Time Series Econometrics using *Microfit 5.0*

Bahram Pesaran and M. Hashem Pesaran

June 2009

Contents

List of Tables

List of Figures

Preface

Microfit 5.0 represents a major advance over the earlier versions of the package. It contains many new features and covers a number of recent developments in the areas of multivariate time series analysis and multivariate volatility modelling.

This volume describes how to install and run the software, use its various menus, options, formulae and commands. In addition, it contains detailed reviews of the underlying econometric and computing methods, together with 86 tutorial lessons using more than 40 different data sets. It is hoped that this volume can serve as an interactive tool in the teaching of time series econometrics, supplementing recent econometric texts.

Microfit 5.0 is particularly suited for the analysis of macroeconomic and financial time series data at different frequencies: daily, monthly, quarterly and yearly.

This volume is in six parts:

Part I (Chapters 1-2) provides an introduction to the package and shows how to install and run it on personal PCs.

Part II (Chapters 3-5) deals with reading/saving of data and graphic files, management and processing of data, and preliminary data analysis.

Part III (Chapters 6-8) provides an account of the estimation menus and the various single and multiple equation options that are available in *Microfit 5.0*.

Part IV (Chapters 9-20) is devoted to tutorial lessons covering many different issues and problems, ranging from data management and data processing to linear and non-linear regressions, univariate time series analysis, $GARCH$ modelling, Probit and Logit estimation, unrestricted VAR modelling, cointegration analysis, and $SURE$ estimation.

Part V (Chapters 21-23) provides a review of the underlying econometric techniques for the analysis of single and multiple equation models.

Part VI (Appendices A-B) provides information on the size limitations of the package, and tables of critical values.

In developing *Microfit 5.0* we have benefited greatly from comments and suggestions by many Microfit users, students, and colleagues, particularly those at Cambridge University, University of Southern California, and the participants of the 'Working with *Microfit*' courses organized by Cambridge Econometrics and Camfit Data. In particular we would like to thank TengTeng Xu for going through all the tutorial lessons and checking them for consistency and accuracy, Kamiar Mohaddes for helpful suggestions and comments, and five anonymous reviewers for their constructive comments and feedbacks on beta versions of the package. We have also benefited from helpful comments from Adrian Pagan and Ron Smith over the many years that *Microfit 5.0* has been under development.

Finally, we would like to single out Elisa Tosetti for her invaluable help with the current version of the manual. She has been responsible for adding new lessons, checking the manual for consistency, identifying numerous errors and for helping with the preparation of the Manual for publication. Her extensive collaborations is gratefully acknowledged.

Bahram Pesaran London
M. Hashem Pesaran Cambridge

(in collaboration with Elisa Tosetti)

June 2009

Bahram Pesaran is currently a Research Consultant at Wadhwani Asset Management. He has also worked as a Research Analyst at Tudor Investment Corporation, The Bank of England, The National Institute of Economics and Social Research and The Confederation of British Industry."

Hashem Pesaran is Professor of Economics at the University of Cambridge, John Elliott Chair at the University of Southern California, and a Professorial Fellow of Trinity College, Cambridge. Previously he has been the head of the Economic Research Department of the Central Bank of Iran, and Professor of economics at University of California at Los Angeles. Dr Pesaran is the founding editor of the Journal of Applied Econometrics and has served as a Vice President at the Tudor Investment Corporation. He is a Fellow of the Econometric Society, and a Fellow of British Academy.

Dedicated to our children: Bijan, Evaleila, Hassanali, Jamal, Kathleen, and Natasha

Part I

Introduction to Microfit

Chapter 1

Introduction

1.1 What is *Microfit?*

Microfit is an interactive econometric software package designed specifically for the econometric modelling of time series data. It is suitable for classroom teaching of undergraduate and post-graduate courses in applied econometrics. It has powerful features for data processing, file management, graphic display, estimation, hypothesis testing, and forecasting under a variety of univariate and multivariate model specifications. These features make *Microfit 5.0* one of the most powerful menu-driven time series econometric packages currently available.

Microfit generates output in carefully set out tables and graphs, virtually in a matter of seconds. Output from *Microfit* can be sent directly to a printer, saved on a disk file to be printed subsequently, or used in a text file as part of a printed report.

Microfit accepts ASCII and binary data files, *Excel* worksheets, and data files in a variety of formats such as comma delimited (CVS), TXT and AREMOS (TSD) files. It also readily allows for extension, revision, and merging of data files. Data on *Microfit*'s workspace can be exported to spreadsheet packages in the CSV and TSD formats. Other software in the operating system can be accessed easily. For routine and repetitive data processing tasks, *Microfit* employs commands close to conventional algebraic notation.

The strength of the package lies in the fact that it can be used at different levels of technical sophistication. For experienced users of econometric programs it offers a variety of univariate and multivariate estimation methods and provides a large number of diagnostic and non-nested tests not readily available on other packages. The interaction of excellent graphics and estimation capabilities in *Microfit* allows important econometric research to be carried out in a matter of days rather than weeks.

1.2 New features of *Microfit 5.0*

Microfit 5.0 represents a major advance over the earlier versions of the package. It makes more intensive use of screen editors and window facilities for data entry, model specification, and for easy storage and retrieval of data and result files. Using the new version you can run regressions up to 102 regressors with an almost unlimited number of data points (5,000,000 observations as compared to the 3,000 limit imposed in *Microfit 4.0*). The time series dimension of the observations can be adjusted dynamically. You can also move readily between drives, directories, and sub-directories for retrieving and saving data input and output files. Scrolling within a result screen is also possible. Almost all files created using *Microfit 4.0* can be used in *Microfit 5.0*.

The new options in *Microfit 5.0* include:

1. Virtually no limits on the data sizes being analyzed – given available PC memory.
2. Importing and exporting of *Excel* files.
3. Much enhanced graphic module with the possibility of many types of graphs and an almost unrestricted number of plots per screen.
4. Revamped interface – much more transparent instruction screens.
5. Enhanced help files – online documentation.
6. Additional unit roots test Phillips-Perron, ADF-GLS, ADF-WS, and ADF-MAX.
7. Analysis of cointegrating $VARX$ models. This extends the popular cointegration module to the case where the model contains weakly exogenous variables, essential for modelling of small open economies, for example, used in modelling of global economy. This option is also used for Global VAR ($GVAR$) modelling. See Pesaran, Schuermann, and Weiner (2004) and Garratt, Lee, Pesaran, and Shin (2006).
8. Forecasting, impulse response analysis, persistence profiles and error variance decomposition for $VARX$ models.
9. Bootstrapped critical values for tests of over-identifying restrictions in cointegrated models (very important in practical uses of the cointegrating options).
10. Small sample simulation of the critical values of unit roots and cointegration tests.
11. Bootstrapped error bounds for the impulse responses, persistence profiles and error variance decompositions for VAR, $VARX$, and cointegrated VAR and $VARX$ options.
12. Multivariate $GARCH$ models – this option allows modelling of volatilities of many assets jointly based on Pesaran and Pesaran (2007). This option allows estimation with Gaussian and multivariate t-distributed shocks (important for the analysis of fat-tailed distributions) and would be particularly helpful in empirical finance.

For data analysis, *Microfit 5.0* has a large number of additional time series and econometric features, including new functions and commands, new single-equation options, and new multivariate time series techniques.

1.2.1 New functions and commands

New functions included in *Microfit 5.0* are:

- **NONPARM** computes nonparametric density estimation using Gaussian and Epanechnikov kernels with Silverman rule of thumb and least squares cross-validation band widths.
- **REC_MAX**(X) and **REC_MIN**(X) compute the maximum and minimum of X recursively over a specified sample period.
- **ROLL_MAX**(X, h) and **ROLL_MIN**(X, h) compute the maximum and minimum of X using rolling windows of size h over a specified sample period.
- **MAV**(X, p) function, which computes a p-th order moving average of X.
- **GDL**(X,λ) is the geometric distributed lag function with the lag coefficient λ.

New commands in *Microfit 5.0* are:

- **DF_PP** is the Phillips-Perron (due to Phillips and Perron (1988)) unit roots test.
- In addition to standard ADF statistics, it is also possible to compute **ADF_GLS** (due to Elliott, Rothenberg, and Stock (1996)), **ADF_MAX** (due to Leybourne (1995)), and **ADF_WS** (due to Park and Fuller (1995)) statistics. Unit root tests can also be carried out in the case of ADF regressions subject to known breaks.
- **CCA** performs canonical correlation analysis on two sets of variables, after controlling for a third set of variables.
- **FILL_FORWARD** replaces current missing values by the last available observations.
- **FILL_MISSING** replaces current missing values with a value specified by the user.
- **PCA** performs principal components analysis on a set of variables after filtering out the effects of another set of variables.

1.2.2 Single equation estimation techniques

The single-equation options in *Microfit 5.0* include:

- **Linear and non-linear OLS and Instrumental Variables (IV) regressions**.
- **Recursive and Rolling Regressions**.
- **Estimation of Regression Models with Autoregressive and Moving Average Errors Cochrane-Orcutt, Maximum Likelihood and IV Procedures**.

- **Estimation of Conditionally Heteroscedastic Models.** Maximum likelihood estimation of regression models under a variety of conditionally heteroscedastic error specifications, such as $ARCH$, $GARCH$, $GARCH$ in mean,Absolute value $GARCH$, absolute value $GARCH$ in mean, exponential $GARCH$, exponential $GARCH$ in mean. The $ARCH$ and $GARCH$ models can be estimated for two different specifications of the conditional distribution of the errors, namely normal and the Student's t-distributions.

- **Logit and Probit Estimation.**

- **Phillips-Hansen's Fully Modified OLS Estimation.** This procedure provides single-equation estimates of the cointegrating relations.

- **Autoregressive-Distributed Lag (ARDL) Estimation Procedure.** This procedure provides estimates of a single cointegrating relation on the basis of an $ARDL$ model selected by means of model selection procedures such as Akaike, Schwarz, Hannan and Quinn, and $\bar{R}^2$. This approach also readily allows for inclusion of time trends, seasonal dummies and other deterministic/exogenous regressors in cointegrating relation. See Pesaran and Shin (1999) and Pesaran, Shin, and Smith (2001).

- **Diagnostic and misspecification test statistics.**

- **Non-nested tests.** Tests of linear versus log-linear models, level-differenced versus log-differenced models, and other non-linear specifications of the dependent variable.

1.2.3 System equation estimation techniques

Microfit 5.0 provides an integrated tool-box for the analysis of multivariate time series models. The estimation and testing procedures cover the following models:

- **Estimation of Unrestricted VAR Models.** This option provides estimates of the coefficients in the VAR model, together with various diagnostic test statistics for each equation in the VAR model, separately. It allows automatic order selection in the VAR using Akaike, Schwarz, and likelihood-ratio procedures, Granger (1969) block non-causality tests, orthogonalized and generalized impulse response in VAR models (Sims (1980), Koop, Pesaran, and Potter (1996), Pesaran and Shin (1998)). Orthogonalized and Generalized Forecast Error Variance Decomposition in unrestricted and cointegrating VAR models. The generalized impulse responses are new and, unlike the orthogonalized responses, do not depend on the ordering of the variables of the VAR model.

- **Estimation of Seemingly Unrelated Regression Equations.** Estimation and hypothesis testing in systems of equations by the Seemingly Unrelated Regression Equations ($SURE$) method (Zellner 1962). ML estimation and hypothesis testing in systems of equations subject to parametric restrictions. The restrictions could be homogeneous or non-homogeneous, and could involve coefficients from different relations (such as cross-equation restrictions).

- **Two-Stage and Three-Stage Least Squares.**

- **Two-Stage and Three-Stage Restricted Least Squares.**

- **New Cointegration Tests in VAR and $VARX$ Models.** These tests allow for deterministic linear trends in the underlying VAR model both with and without restrictions on the trend coefficients, and enable the user to carry out the Cointegration tests when one or more of the $I(1)$ variables are exogenously determined. *Microfit 5.0* allows the users to simulate critical values at any desired level of significance for the cointegration tests.

- **Long-Run Structural Modelling.** This estimation procedure allows you to estimate and test more than one cointegrating relations subject to identifying and over-identifying restrictions on the long-run (or cointegrating) relations. The restrictions can be homogeneous or non-homogeneous, and can involve coefficients from different cointegrating relations. The long-run structural modelling also allows analysis of sub-systems where one or more of the $I(1)$ variables are exogenously determined. See Pesaran, Shin, and Smith (2000).

- **Impulse Response Analysis and Forecast Error-Variance Decomposition.** The program now allows computations of orthogonalized and generalized impulse response functions, and forecast error variance decomposition in the case of cointegrating VAR models. It also produces estimates of the persistence profiles for the effect of system-wide shocks on the cointegrating relations. See Pesaran and Shin (1996) and Pesaran and Shin (1998).

- **Computation of Multivariate Dynamic Forecasts.** Multivariate dynamic forecasts for various horizons can be readily computed using *Microfit 5.0*, both for unrestricted and cointegrating VAR models, and for Seemingly Unrelated Regression Equations with and without parametric restrictions.

1.3 Tutorial lessons

Important features of *Microfit* are demonstrated throughout this manual by means of 82 tutorial lessons, using data files supplied with the program. There are lessons in data management; data transformation; graphics (plotting time series, scatter diagrams, histograms); displaying; printing and saving results; estimation; and hypothesis testing and forecasting, using a variety of univariate and multivariate econometric models. These lessons and the details of econometric methods provided in Chapters 21, 22 and 23 are intended to complement the more traditional econometric texts used in quantitative economics courses, and help further establish the concept of interactive econometric teaching in the profession.

1.4 Other features of *Microfit 5.0*

Many useful features of *Microfit 4.0* have been either retained or have been greatly enhanced, particularly as far as data inputs, interface and graphics are concerned. A summary of these features follows.

1.4.1 Data management

Microfit can be used to input data directly from the keyboard, from raw ASCII data files, csv, *Excel* spreadsheet files, or from special *Microfit* files prepared and saved previously by the package. The data can be input as undated series or as daily, monthly, quarterly, half-yearly or yearly frequencies. Integral parts of the data management system of *Microfit* are the facilities provided for the extension, revision and merging of the data files. These data management facilities allow the user to extend an existing data file by adding more observations and/or more variables to it. It is also possible to input and output raw data files in ASCII, CSV and *Excel* spreadsheets to and from *Microfit*.

1.4.2 Data transformations

Microfit allows the user to compute new series as algebraic transforms of existing data using standard arithmetic operators, such as + - / *, and a wide range of built-in functions including **MAX**, **MIN**, **SIGN**, **RANK**, and **ORDER**. Time trends, seasonal dummies, and random numbers from **UNIFORM** and **NORMAL** distributions can also be generated easily by *Microfit*.

1.4.3 High-resolution graphics

Microfit can be used to produce high-quality scatter diagrams and graphs of variables plotted against time or against another variable, with the option of adding headings to the graph. In *Microfit 5.0* there are no effective limits on the number of time series plots that can be shown on the same screen. A hard copy of the graphs can be produced on Laser and Laserjet printers in black and white and in colour when available. *Microfit* can also automatically produce graphs of actual and fitted values, residuals, histograms with superimposed normal and t-distribution in the case of the *ARCH* and *GARCH* options, as well as graphs of forecasts and concentrated log-likelihood functions.

Microfit 5.0 allows the graphs to be saved for importation into word-processing programs such as *Microsoft Word* and *Scientific Word*.

1.4.4 Batch operations

Microfit allows the user to run batch jobs containing formulae, samples, and simulation commands. This facility is particularly useful when the same transformations of different or revised data sets are required.

1.4.5 General statistics

Microfit allows the user to compute:

- Means and standard deviations.
- Skewness and kurtosis measures.
- The coefficient of variation.
- Correlation coefficients of two or more variables.
- Minimum and maximum of a series.
- The autocorrelation coefficients and their standard errors.
- Estimates of the spectral density function and their standard errors using Bartlett, Tukey and Parzen windows.

1.4.6 Dynamic simulation

Important facilities in *Microfit* are the **SIM** and **SIMB** commands, which allow the user to simulate dynamically any non-linear difference equation both forwards and backwards.

1.4.7 Other single equation estimation techniques

Microfit 5.0 estimates regression equations under a variety of stochastic specifications. The estimation techniques carried over from *Microfit 3.0* include:

- Ordinary least squares.
- Generalized instrumental variables.
- Two-stage least squares.
- Recursive and rolling estimation by the least squares and instrumental variables methods.
- Non-linear least squares and non-linear two-stage least squares.
- Cochrane-Orcutt iterative technique.
- Maximum likelihood estimation of regression models with serially correlated errors (both for AR and MA error processes).
- Instrumental variable estimation of models with serially correlated errors (both for AR and MA error processes).
- Maximum likelihood estimation of $ARMA$ or $ARIMA$ processes.

- Maximum likelihood estimation of cointegrated systems.

Models with autocorrelated errors of up to order 12 can be estimated, both when the pattern of residual autocorrelation is unrestricted and when it is subject to zero restrictions. The estimation results are compactly tabulated and provide parameter estimates and other statistics of interest including t-statistics, standard errors, Durbin-Watson statistic, Durbin's h-statistic (when relevant), $\overline{R}^2$, and more.

Alternative estimates of the variance-covariance matrix of the parameter estimates, namely White's heteroscedasticity-consistent estimates and Newey-West adjusted estimates with equal weights, Bartlett weights, and Parzen weights can also be obtained with *Microfit*, for the cases of linear and non-linear least squares and instrumental variables methods.

Microfit enables the user to list and plot the actual and fitted values, as well as the residuals. The fitted values and the residuals can be saved for use in subsequent econometric analysis.

1.4.8 Model respecification

The specification of equations in *Microfit* can be changed simply by using a screen editor to add and/or delete regressors, or to change the dependent variable. The equations and variable lists can be saved to a file for later use. It is possible to re-estimate the same regression equation over different time periods and under different stochastic specifications simply by selecting the relevant items from the menus.

1.4.9 Diagnostic tests and model selection criteria

Microfit supplies the user with an array of diagnostic statistics for testing residual autocorrelation, heteroscedasticity, autoregressive conditional heteroscedasticity, normality of regression disturbances, predictive failure and structural stability. It automatically computes:

- Lagrange multiplier tests for serially correlated residuals in the case of OLS and IV estimation methods.
- Ramsey's $RESET$ test of functional form mis-specification.
- Jarque-Bera's test of normality of regression residuals.
- A heteroscedasticity test.
- Predictive failure test.
- Chow's test of stability of regression coefficients.
- Unit roots tests.
- $ARCH$ test.
- The $CUSUM$, and the $CUSUM$ of Squares tests for structural stability.

- $\bar{R}^2$, Akaike, Schwarz, and Hannan and Quinn model selection criteria.
- Generalized $\bar{R}^2$ for models estimated by Instrumental Variables method (see Pesaran and Smith (1994)).

1.4.10 Variable addition and variable deletion tests

Microfit has options for carrying out variable addition and variable deletion tests. These facilities are very helpful in the process of model constructions, and enable users to follow either of the two basic modelling strategies, namely specific-to-general or general-to-specific. The former facility is also particularly useful as it allows the user to carry out further diagnostic tests, such as higher-order $RESET$ or $ARCH$ tests, or to test for the independence between the disturbances and the regressors of the regression equation.

1.4.11 Cointegration tests

Microfit provides a user-friendly method of testing for cointegration among a set of at most 12 variables by the Johansen's Maximum Likelihood method. Both versions of Johansen's tests, namely the Maximal eigenvalue and the trace tests, are computed. These features are extensively enhanced in *Microfit 5.0*.

1.4.12 Testing for unit roots

Microfit automatically computes a *variety* of Augmented Dickey-Fuller statistics and allows the users to simulate the appropriate critical values at any desired level of significance.

1.4.13 Tests of linear and non-linear restrictions

Tests of linear and non-linear restrictions on the parameters of the regression model (both the deterministic and the stochastic parts of the model) can be carried out using *Microfit*. It is also possible to compute estimates of functions (possibly non-linear ones) of the parameters of the regression model, together with their asymptotic standard errors for all the estimation methods. This facility is particularly useful for the analysis of long-run properties, such as estimation of long-run responses, mean lags, and computation of persistence measures.

1.4.14 Non-nested tests

Microfit provides a number of non-nested statistics proposed in the literature for tests of non-nested linear regression models. These include Godfrey and Pesaran's small sample-adjusted Cox statistic, Davidson and MacKinnon's J-test, and the encompassing F-statistic based on a general model. It also contains options for testing linear versus log-linear models, and for testing ratio models versus log-linear and linear models. A number of important model selection criteria such as Akaike's information criterion, and Schwarz's Bayesian criterion are also computed in the case of non-nested models.

1.4.15 Static and dynamic univariate forecasts

Microfit generates one-period-ahead (static) and dynamic forecasts of single-equation regression models. It gives forecast errors and a number of useful summary statistics. If lagged dependent variables are included in the regression, *Microfit* automatically computes dynamic forecasts, otherwise it generates static forecasts. The plot of actual and forecast values is also provided, with the possibility of saving forecast values and forecast errors on a file for later analysis.

Missing values are fully supported by *Microfit*, and when a transformation (including leading and lagging) results or involves undefined values, the undefined values are set to blank. At the Estimation/Testing/Forecasting stage *Microfit* looks for blank fields and adjusts the specified sample period automatically so that beginning and end periods with missing values are discarded. If missing values are encountered in the middle of the estimation and/or forecast periods the Estimation/Testing/Forecasting will use as sample the set of observations from observation 1 to the observation preceeding the first encountered missing value. The value *NONE* is displayed whenever the computations involve operations that cannot be carried out, such as taking the square root or the logarithm of a negative number.

1.5 Installation and system configuration

Microfit can be easily installed both on personal computers and can be configured to suit the taste and the needs of the user. It allows the user to alter the colour of text displayed on the screen by choosing a combination of foreground and background colours. *Microfit* needs to be configured only once, at the time the package is installed on the PC, but can easily be reconfigured at a later date.

1.6 System requirements for *Microfit 5.0*

Microfit 5.0 is available for Microsoft Windows 2000, XP and Vista operating systems.

- 1MB of Random Access Memory (RAM).
- At least 45MB of free hard disk.
- *Microsoft* mouse or compatible (optional).
- A printer for producing hard copies of graphs and regression results (optional).

Microfit 5.0 allows running regressions with up to 102 regressors and up to 5,000,000 observations. In the case of the unrestricted VAR option *Microfit* allows for up to 12 variables in the VAR model and a maximum order of 24. In the cointegration option it allows a maximum of 12 endogenous $I(1)$ variables and 5 exogenous $I(1)$ variables, and 18 deterministic and/or exogenous $I(0)$ variables. See Appendix A for further information on the size limitations of the program.

Chapter 2

Installation and Getting Started

2.1 Installation

To install *Microfit 5.0,* follow the installation instructions outlined below for a single user (see Section 2.1.1) and for network installation (see Section 2.1.2).

2.1.1 Single user installation

Microfit 5.0 is distributed on a single CD-ROM. Close all applications before inserting the CD into your computer's drive and wait for the setup program to launch automatically. If the setup does not start automatically, navigate to the CD drive and click on the Setup icon (SETUP.EXE). Follow the simple instructions indicated by the program.

2.1.2 Network installation

This involves the following two steps:

1. Server Installation by the System Operator. The network version of *Microfit* 5.0 is distributed on a single CD-ROM. The machine used for the server installation must be a client machine and have internet access. Close all applications before inserting the CD into a client computer's drive and wait for the setup program to launch automatically. If the setup does not start automatically, navigate to the CD drive and click on the Setup icon (NETWORK_SETUP.EXE). The only action required is to browse to a folder to be used for installation on the server.

2. Installation by a Client. The client browses to the server folder specified in Step 1 above and starts the setup program (MICROFIT5_SETUP.EXE) and follows the simple instructions (identical to the single user setup, see Section 2.1.1) indicated by the program.

2.2 Starting and ending a session

2.2.1 Running *Microfit*

When you have successfully installed *Microfit 5.0* on your system, you can start it by double-clicking the icon for Mfit5 on your desktop or from the Programs menu. You can also upload a special *Microfit* data file directly into the program by double-clicking on the *Microfit* data file.

2.2.2 Quitting *Microfit*

To quit *Microfit*, do one of the following:

- Choose Exit from the File Menu.
- Double click on the left of the title bar.
- Click on the left of the title bar once and choose 

.
- Press ALT+F4.

You are warned that all unsaved data will be lost. If you are sure you have saved all your data, choose Yes. If not, click No, save any unsaved data, and try quitting again.

2.3 Using windows, menus and buttons

You work your way through a *Microfit 5.0* session using a combination of windows, menus, and buttons. At the heart of the application are several menus and sub-menus for processing your data. Each menu or sub-menu contains a selection of up to 11 options with one of the options (usually the most common) already selected.

A simple method of familiarizing yourself with *Microfit* and what it can do is to learn about its main menus and their interrelationships.

The rest of this volume describes the various options in these menus, show you how to use them to input and process data, do preliminary data analysis, and estimate/test/forecast using a number of univariate and multi-variate time series models.

2.3.1 The main window

The main window is your workspace. From here you can access all the main functions of the program: the Variable window, the Data Editor, the Process window, the Single Equation Estimation window, and the System Estimation window.

2.3.2 Main Menu bar

Many of the program's functions are controlled using the menus of the main menu bar.

File Menu

Open file Opens an existing data file saved in any of the following formats: ASCII (.DAT), *Microfit* (.FIT), *Excel* (.XLS) comma delimited values (.CSV), AREMOS time series data (.TDS).

Open file from tutorials data files Opens an existing data file saved in the tutorial directory.

Input new data from the keyboard Allows you to enter a new data set from the keyboard (see Section 3.2.1).

Input new data from clipboard Enables you to copy and paste a data set from the clipboard (see Section 3.2.8).

Add 2 special *Microfit* files Allows you to load two *Microfit* files (see Section 3.3).

Add a special *Microfit* file to workspace Adds a *Microfit* file to existing data in *Microfit* (see Section 3.2.4).

Paste data from clipboard to workspace Enables you to paste data from the clipboard to existing data in *Microfit* workspace (see Section 3.2.9).

Save as Saves your data in a new file in special *Microfit*, ASCII, *Excel*, CSV, or AREMOS format (see Section 3.5).

Change the data dimension Allows you to change the dimension of the data (see Section 3.1).

View a file Opens a file and allows you to examine its content (see Section 4.2.1).

Exit Quits the program.

Edit Menu

Cut, Copy, and **Paste** allow you to perform standard Windows editing functions.

Constant (intercept) term Creates a constant (see Section 4.1.1). Its button equivalent is 

.

Time trend Creates a time trend (see Section 4.1.2). Its button equivalent is 

.

Seasonal dummies Creates seasonal dummies (see Section 4.1.3). Its button equivalents are the 

, 

, and 

buttons.

Options Menu

European/US date format Allows you to set the format of the date, when reading CSV/*Excel* files with daily data or calendar dates in the first column.

Result printing Allows you to make choices about how results are printed out.

Editor defaults Specifies the size, style, and colour of default.

Location of Acrobate PDF file viewer enables you to find (automatically or manually) the location of Adobe Acrobate reader.

Location of default files allows you to choose the location of default files.

Location of temporary files allows you to choose the location of temporary files.

Univariate Menu

This is the Single Equation Estimation Menu from which all the sigle-equation estimation options may be accessed. These are discussed in detail in Chapter 6.

Multivariate Menu

This is the System Estimation Menu from which all the multiple equation estimation options may be accessed. These are discussed in detail in Chapter 7.

Volatility Modelling Menu

From this menu all univariate and multivariate $GARCH$ options may be accessed. These are discussed in details in Chapter 8.

Help Menu

Accesses the program's help functions (see Section 2.5.2).

2.3.3 Buttons

The buttons in the main window control the most common functions:

Button	**Function**
Variables	Switches to the Variable window where the current list of variables and their descriptions are displayed and may be edited.
Data	Moves to the Data window where you can edit your data.
Process	Moves to the Process window where you can process your data.
Single	Switches to the Single Equation Estimation window. The Single Equation Estimation options can also be accessed from the univariate menu.
Multi	Switches to the System Equation Estimation window. The System Estimation options can also be accessed from the Multivariate menu.
Volatility	Switches to the Volatility Modelling window. The Volatility options can also be accessed from the Volatility Modelling menu.
Batch	Runs a Batch file you have saved earlier.
Help	Displays the on-line help facility.
Save	Saves the contents of the current editor to disk.
Retrieve	Loads a previously save LST or EQU file into the current box editor.
Clear	Clears the contents of the current box editor.
Font	Changes the font of the contents in the current box editor.
Find	Finds a word in the current box editor.
Print	Prints the contents of the current box editor.
Load Graph	Loads into *Microfit* a graph in Olectra Chart format.

2.4 The Variables window

The Variables window displays the name of the variables and their descriptions in the workspace, and can be accessed by clicking the Variables button (see Figure 2.1 below). Variable descriptions can be up to 80 characters long. If you wish to add or change a description for one variable, move to the corresponding description field and type in a title, click Save and then Close. Note that you cannot undo the changes you have made to the variable description if you click the Save button.

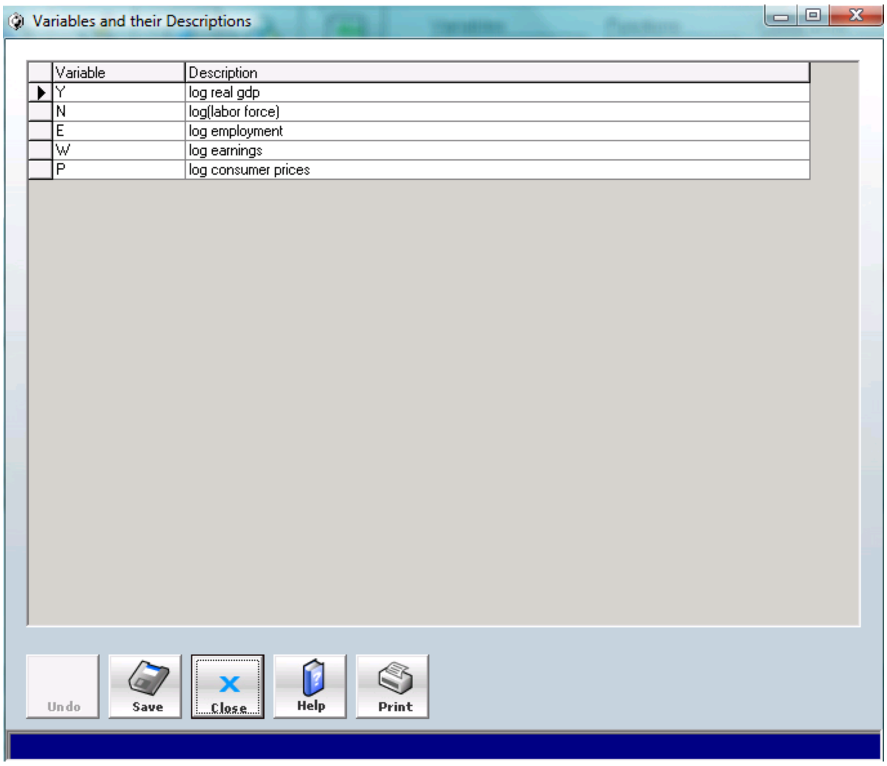


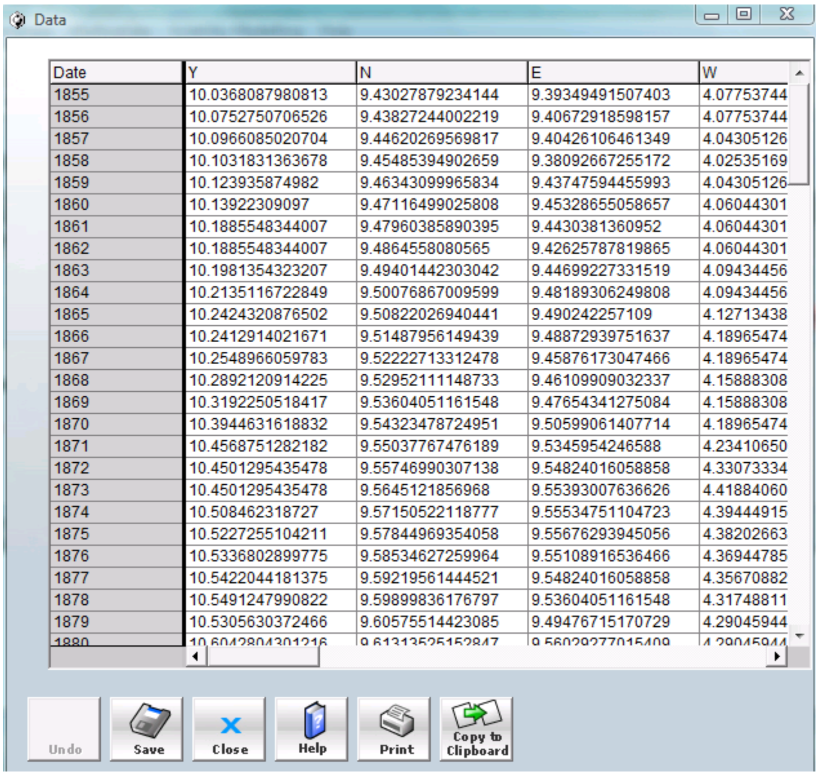
Figure 2.1: The Variable window

2.5 The Data window

The Data window displays data in the workspace and can be accessed by clicking the  button (see Figure 2.2). In this window you can inspect and edit your data. Use the horizontal and vertical scroll panes to move through the contents of the data windows. When you have finished editing your data, click  and then . Once you press the  button, changes you have made to your data are permanent and cannot be undone. You can also copy to clipboard a set of cells, by selecting the cells and pressing the  button. *Microfit* will automatically add dates and column names to copied data.

2.5.1 Program options

Various aspects of the program may be customized to personal preference using Edit, Options, and Window Menus.



Date	Y	N	E	W
1855	10.0368087980813	9.43027879234144	9.39349491507403	4.07753744
1856	10.0752750706526	9.43827244002219	9.40672918598157	4.07753744
1857	10.0966085020704	9.44620269569817	9.40426106461349	4.04305126
1858	10.1031831363678	9.45485394902659	9.38092667255172	4.02535169
1859	10.123935874982	9.46343099965834	9.43747594455993	4.04305126
1860	10.13922309097	9.47116499025808	9.45328655058657	4.06044301
1861	10.1885548344007	9.47960385890395	9.4430381360952	4.06044301
1862	10.1885548344007	9.4864558080565	9.42625787819865	4.06044301
1863	10.1981354323207	9.49401442303042	9.44699227331519	4.09434456
1864	10.2135116722849	9.50076867009599	9.48189306249808	4.09434456
1865	10.2424320876502	9.50822026940441	9.490242257109	4.12713438
1866	10.2412914021671	9.51487956149439	9.48872939751637	4.18965474
1867	10.2548966059783	9.52222713312478	9.45876173047466	4.18965474
1868	10.2892120914225	9.52952111148733	9.46109909032337	4.15888308
1869	10.3192250518417	9.53604051161548	9.47654341275084	4.15888308
1870	10.3944631618832	9.54323478724951	9.50599061407714	4.18965474
1871	10.4568751282182	9.55037767476189	9.5345954246588	4.23410650
1872	10.4501295435478	9.55746990307138	9.54824016058858	4.33073334
1873	10.4501295435478	9.5645121856968	9.55393007636626	4.41884060
1874	10.508462318727	9.57150522118777	9.55534751104723	4.39444915
1875	10.5227255104211	9.57844969354058	9.55676293945056	4.38202663
1876	10.5336802899775	9.58534627259964	9.55108916536466	4.36944785
1877	10.5422044181375	9.59219561444521	9.54824016058858	4.35670882
1878	10.5491247990822	9.59899836176797	9.53604051161548	4.31748811
1879	10.5305630372466	9.60575514423085	9.49476715170729	4.29045944
1880	10.6042804301216	9.61313525152847	9.56029277015409	4.29045944

Figure 2.2: The Data Editor

Printing of results

To specify how your results are printed out, choose Result Printing from the Options Menu. When you first install *Microfit* you are prompted to enter the name of the current researcher (user). To edit or replace the name, click on the Name of Researcher field and edit (or delete) the name as necessary.

Page numbers are added to your result printout by default. The date of your results and the name of the researcher are added to results printouts by default. To exclude them, uncheck the 'Print date and researcher name' option.

Changing font

You have two choices when specifying how the contents of *Microfit* editors and results windows are displayed. You can either change the font in the current window, or change the font for all editors and/or results windows. To change the font style, size or colour in which the current editor or result window is displayed, click the 

button and choose an alternative font, size, and/or colour from the Font dialogue. To change the default fonts, choose 'Editor defaults' from the Options Menu. The default font used in all Editor boxes is Courier New.

The default colour and size are blue 16-point. To change the default font, click the button 'All Editor Boxes' and choose an alternative font, size, and/or colour from the Font dialogue box. Then click . The changes you have made will be immediately implemented.

The default font used in all result windows is black Courier New 9-point. To change the default font, click the button 'All Result Windows' and choose an alternative font, size and/or colour from the Font dialogue box. Then click . (To ensure the correct display of results, only the fixed fonts available on your computer are listed).

To save the new font as default in *Microfit*, check the 'Save as defaults' checkbox in the Options Menu.

2.5.2 Help

Microfit has an extensive on-line help facility. For help on the part of the application you are currently using, click the  button or press F1. Alternatively, use the Help Menu. This has several options:

Overview Shows general information on using *Microfit* help. Click on a topic highlighted in green to move there.

Contents Shows a list of help topics. Click on a topic highlighted in green to move there.

Help on functions Displays help on various functions available in *Microfit*.

Help on commands Displays help on commands.

About Provides copyright information about *Microfit* for Windows.

To access the help options in *Microfit 5.0* you need to have Acrobat Reader 6 or above installed on your PC. Also before using the help options in *Microfit* we recommend that you ensure that your Acrobat Reader (or Adobe Acrobat if you have one installed) allows you to access the internet from PDF files.

Part II

Processing and Data Management

Chapter 3

Inputting and Saving Data Files

Data can be input into *Microfit* directly from the keyboard, from ASCII files, CSV files, or *Excel* files, from special *Microfit* files saved previously, or from AREMOS (TSD) files.[1] Data on the workspace of *Microfit* can be saved as ASCII/text files, *Excel* sheet, as special *Microfit* files, as TSD files or as comma delimited (CSV) files with descriptions of variables in the second row. It is also possible to copy data sets from the clipboard to *Microfit*.

To input and save data, use the options in the File Menu. You can also use the 'Add a Special *Microfit* File to Workspace' option in the File Menu to add new variables (already saved in a special *Microfit* file) to your current data set, or select 'Add 2 Special *Microfit* Files' to combine two *Microfit* files containing the same variables.

3.1 Change data dimension

Before inputting data, make sure that the dimension of your data set does not exceed the size limits of *Microfit*. *Microfit 5.0* has a limit of 200 variables, and can run regressions up to 102 regressors, while it has no limits on the number of observations in the data. The default maximum number of observations is 6,000, but it may be changed upwards or downwards.

You can change the default data/variable dimensions by clicking the 'Change Data Dimension' option in the File Menu. You are presented with a window where you can set the maximum number of observations, variables and regressors (see Figure 3.1). If you set the option 'Save as Defaults', these numbers will be remembered and become the default setting when you open *Microfit*.

3.2 Inputting data

When you start a new session with *Microfit*, you can either input a new data set directly from the keyboard or load an existing data set.

[1] *Excel* is a trademark of *Microsoft*.

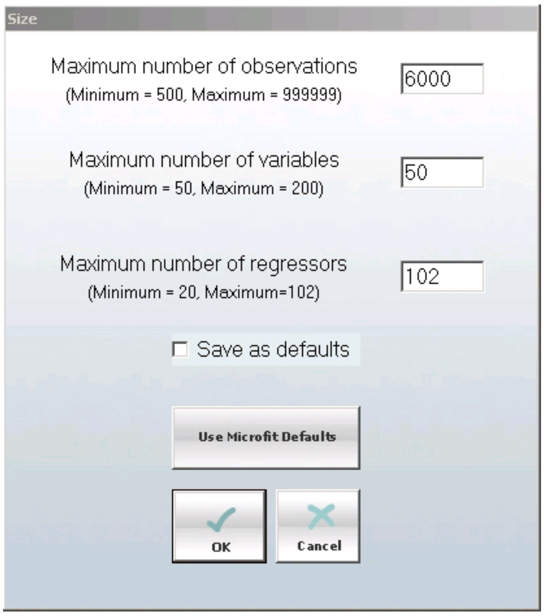


Figure 3.1: Change the data dimension

3.2.1 Inputting data from the keyboard

Inputting data directly from the keyboard is the most basic method of entering data. Before entering new data, make sure you know the frequency of your data (that is, whether your data are undated, or have annual, quarterly, or monthly frequencies), the number of variables in your data set, and the sample period of your observations.

To input a new data set, choose 'Input new data from the keyboard' in the File Menu. This opens a new data set dialogue. The data frequency fields are at the top of the dialogue, with fields for start and end dates, and number of variables below. The following data frequency options are available. To choose one, click the appropriate radio button

Undated
Annual
Half-yearly
Quarterly
Monthly

Entering undated observations: This option is often relevant for entering cross-sectional observations, and when it is chosen *Microfit* assumes that the observations are unordered, and asks how many observations you have. An observation refers to the individual unit on which you have data. For example, if you have cross-section data containing variables such as employment, output, and investment on a number of firms, then each firm represents an observation, and the number of observations will be equal to the number of firms in your data set. If you have time series data covering the period from 1960 to 1970 inclusive and for one reason or another you wish to enter them as undated data; then the number of observations in your data set will be equal to 11.

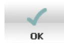
Entering annual, half-yearly, quarterly, or monthly observations: If any of these options are chosen, the program will supply the dates, and you will not need to type them in. You are, however, asked to type the dates for the start and end of your data in the appropriate field. For example, if your data are annual and cover the period 1960-1985 inclusive, when asked, you need to enter 1960 in the Start field and 1985 into the End field. You can type the year in its full (1960) or abbreviated (60) form. However, if your data go beyond the year 1999 you must enter the dates in their full forms, namely 2000, 2025 and so on.

If your data are quarterly and cover the period from the first quarter of 1990 to the last quarter of 2000 inclusive, you need to enter 1990 in the Start field and 1 in the adjacent quarter field, then 2000 in the End field, specifying 4 as its quarter.

Similar responses will be required if your data are half-yearly or monthly.

Note: It is not possible to enter daily data directly from the keyboard. You can only input them from an *Excel* file or copy them from the clipboard. See Section (3.2.10) for further information on how to input daily data.

Entering number of variables: This refers to the number of data items that you have for each observation. Set as appropriate.

When you have finished entering the information, click . This opens the Variables window.

Entering variable names: The Variables window contains the default variable names $X1$, $X2$, $X3$, ... For example, if you specify that you have 10 variables the screen editor appears with the following default variable names

$$X1\ X2\ X3\ X4\ X5\ X6\ X7\ X8\ X9\ X10$$

You can enter your own choice of variable names and/or add a description if you wish. Move to the appropriate field and edit or add text using standard Windows editing functions.

A valid variable name is alphanumeric, can be at most 9 characters long, and must begin with a letter. Lower- and upper-case letters are treated as equivalent. '_' underscore is also allowed anywhere in a variable name. Examples of valid variable names are

$$GDPUK \quad OUTPUT \quad X2Y3\ DATA261 \quad Y\_1$$

Variable names such as $\$GDP, 123, 2X, W\#$ are not allowed. Also, function and command names used in the data processing stage cannot be used as variable names. The list of function and command names can be found in Chapter 4.

Variable descriptions can be up to 80 characters long. You can return to the Variables window at any stage, by clicking the  button. Variables can also be given an optional description of up to 80 characters in the Process window by means of the command **ENTITLE**. See Chapter 4 on how to use this and other commands. When you are satisfied with the changes you have made, click .

Warning: Note that until the **Close** button is pressed, the **Undo** button will restore the variables or their descriptions to their original values. But once the **Close** button is activated the **Undo** button will no longer function.

Entering data: When you have completed listing your variable names, you will be presented with the Data window. This is where you enter the observations on your first variable. Initially, all cells on this screen are set to blank, indicating missing values.

To enter your data, move to each cell in turn and type in your data. Continue until all the observations are entered. If the observations do not fit on one screen use the PgUp and PgDn keys to move between screens. To move to the top or bottom table, press Ctrl+Home and Ctrl+End. When you have finished entering your observations, click .

Warning: Note that until the **Close** button is pressed, the **Undo** button will restore the data to their original values. But once the **Close** button is activated the **Undo** button will no longer function.

3.2.2 Loading an existing data set

You can input data from an existing file in any of the following formats: special *Microfit*, *Excel* sheet, *Excel* 4.0, ASCII, comma delimited (CSV), or AREMOS (TSD).

To load an existing file, choose 'Open file' from the File Menu. This displays the Open dialogue. Select one of the file types from the drop-down list at the base of the dialogue, choose your file types from the appropriate drive and directory in the usual way. and click . If the data are in *Microfit*, *Excel*, CSV, or TDS file formats, the data are loaded automatically. If the data are in another file format, you will be asked to confirm how the data are structured before they are loaded.

You are then presented with the Process window. To view or edit the variable names or descriptions, click the  button. To edit the data, click the  button.

3.2.3 Inputting data from a raw data (ASCII) file

An ASCII (plain text) file may be extracted from an existing data bank, or typed in directly using spreadsheets or other data processing packages, or after transfer from another computer package. The file is expected to contain *only* numbers. A missing value is represented by the number 8934567.0.

Once you have specified a filename successfully (see Section 3.1), you will need to specify the frequency of your data, their coverage, and the names for the variables that they contain. These specifications are the same as those described above in Section 3.2.1, and require similar action.

In addition, you need to specify whether its format is free or fixed.

Data is organized by variable: This option should be chosen if all the observations on the first variable appear in the file before the observations on the second variable are entered, and so on.

Data is organized by observations: This option should be chosen if the first observations on all of the variables appear in the file before the second observations are entered, and so on.

As an example, suppose you have the following data on variables $VAR1$ and $VAR2$, over the period 1980 to 1983 inclusive on your file

Period	$VAR1$	$VAR2$
1980	23	45
1981	26	50
1982	30	52
1983	40	60

If your data appear on the file as

23	26	30	40
45	50	52	60

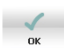
then your data are organized by variable. But if your data appear on the file as

23	45
26	50
30	52
40	60

then your data are organized by observation.

Free format: Choose this option if the data appear in the file with one item separated from another by a space, a comma, or end of line.

Fixed format: The only time you need to choose this option is when data are packed into the file without any such spacing, or have a particular layout specified according to one of the FORTRAN format statements. Users who are not familiar with formatted data are advised to consult a FORTRAN manual.

Once you have made all your choices click . The program starts reading the data from the file and, if successful, presents you with the Process window. However, if the information on the file does not match what has been supplied, then you will get an error message. Click  to return to the Open dialogue and start again.

3.2.4 Inputting data from a special *Microfit* file saved previously

You are likely to choose this option on the second or subsequent time that you use *Microfit*, assuming that you have previously saved a file as a special (non-text) *Microfit* file. If a correct filename is specified, the file is loaded and you are presented with the Process window.

Warning: A special *Microfit* data file must have the file extension FIT. Special FIT files created on earlier versions of *Microfit* can be read in *Microfit 5.0*. Special *Microfit* files created using the current version of the package can be read into *Microfit* version 4.0 but not into *Microfit* version 3.0 to 3.24.

3.2.5 Inputting data from an *Excel* file

Microfit 5.0 can directly read *Excel* workbook files (one sheet of the book at a time). Before loading an *Excel* file into *Microfit*, make sure that you know in which sheet of the *Excel* workbook data are saved. Data should be organized so that the observations for the different variables are arranged in columns, with variable names (up to nine characters) on the first row. The variable names can be followed by their optional descriptions (up to 80 characters for each variable), separated by spaces. Alternatively, descriptions of the variables can be provided on the second row in the cells below the variable names.

The first column of the file must contain dates or, if undated, the observation numbers. Acceptable examples of dates are

1990, 70H1, 1983Q2, 76M12, 30/7/83, 30-7-2001, May-90, 951103

Once you have selected the file, you will need to specify the sheet where data are saved. Select the relevant sheet and click the  button.

3.2.6 Inputting data from CSV files

These files are in ASCII (text) format and are usually created by spreadsheet packages. Before they can be read into *Microfit* they must have the following structure:

1. The file should be organized so that columns are variables and rows are observations.

2. The first row of the file must contain the variable names followed by optional descriptions and separated by spaces. Alternatively, the first row could contain the variable names, with their descriptions given on a second row immediately below the variable names.

3. The first column of the file must contain dates or, if undated, the observation numbers.

Note that the separator for the values/observations for each row in the above files can be a comma, a tab, or spaces.

3.2.7 Inputting data from AREMOS (TSD) files

Files created by the AREMOS package are in the time series data (TSD) format.

Note that only TSD files containing data with the same frequency, namely annual, quarterly, monthly and daily observations, can be read into *Microfit*.

3.2.8 Input new data from the clipboard into *Microfit*

Data copied from a standard Windows spreadsheet package, such as Microsoft *Excel*, can be pasted from the clipboard into *Microfit*. There are two possibilities: You can paste new data sets from the clipboard into an empty Microfit workspace, or you could augment existing data on Microfit workspace with additional data from the clipboard. For the former choose the option 'Input New Data from Clipboard' from the File Menu. The information on the clipboard may contain the variable names on the first row, the descriptions of the variables in the second row, and dates (or observation numbers in the case of undated data) in the first column. See Figure 3.2.

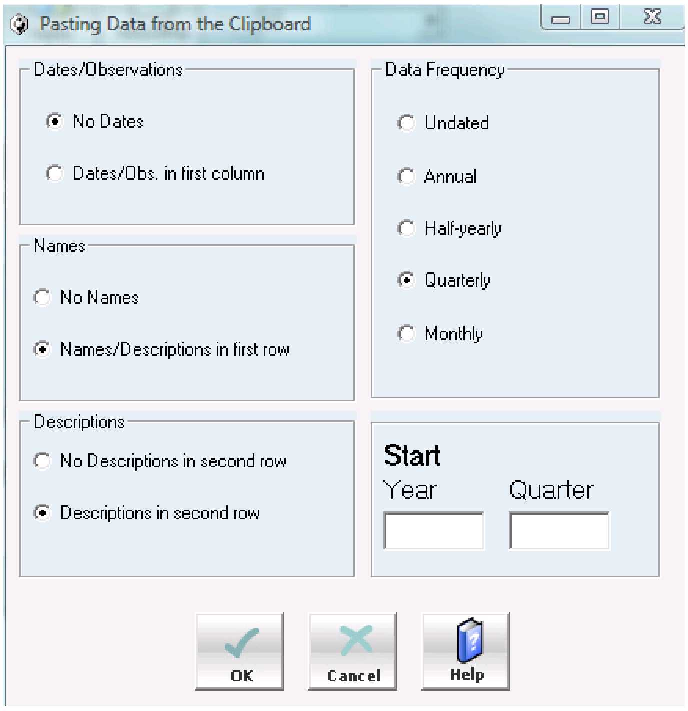


Figure 3.2: Pasting data from the clipboard

When pasting daily data, the information on the clipboard must contain dates in the first column.

Note: *Microfit* sets the precision of data equal to the number of places after the decimal point displayed in the *Excel* worksheet. It is always possible to switch back and forth to the spreadsheet application in order to inspect or change the contents of the clipboard so that the necessary information can be supplied to *Microfit*.

Data are pasted into the Data window. Move to the Data window (click 

) to check that data have been correctly formatted. If you are satisfied, click 

. If not, try to paste data again.

3.2.9 Adding data from the clipboard into *Microfit* workspace

Provided a data set is already loaded into *Microfit*, it is possible to copy data (typically copied from *Excel*) into clipboard and then add the contents of the clipboard to an existing *Microfit* data set. For this you need to select option 'Add Data from Clipboard to Workspace'. For this option the first row of the data to be added must contain variable names. each variable name can be followed by a space and an optional description. Alternatively, the second row can contain the descriptions. The first column of the data on the clipboard must contain valid dates recognized by *Microfit*. The periodicity of the added data must match that of the data already loaded into *Microfit*.

Note that this option is particularly useful for adding daily data. You can also use this option to load a new data set from clipboard into *Microfit*. In this case first create a data set using the 'Input New Data from the Keyboard' from the File Menu and specify the number of variables to be zero. You can then use this option to add data of the same frequency which has been copied to the clipboard.

3.2.10 Inputting daily data

One important advance of *Microfit 5.0* over earlier versions of the package is that it can handle daily observations.

You can only load daily data from a *Microfit* special file or from an *Excel* file.

The first column of your data must contain dates, which can be expressed either in European or in US format. Acceptable examples of dates in European dd/mm/yyyy format are

30/5/05, 30-7-01, 02-May-99

while examples of US dates are

3/14/01, 01/25/2005

When *Microfit* is installed it automatically detects the Windows regional settings of your computer, and sets the appropriate date format as the default. To change the default date format in *Microfit*, go to the Options Menu and choose 'European/US date format'. Then select the preferred date format (either 'European date format' or 'US date format') and click 

.

Once the appropriate date format is selected, you can load in your data by selecting your data file (see Sections 3.2.2 for a description of how to load an existing data set). Once the program has successfully read the data, it creates three new columns containing for each observation the corresponding information on the day, month and year. If your data contain more than 6,000 observations per variable you need to increase the maximum number of observations after selecting the option 'Change Data Dimensions' under File Menu. Make sure to tick the box 'Save as Defaults' to ensure a permanent change in the data dimensions of *Microfit* on your PC.

Note: There is no need to change the date format when data are saved in special *Microfit* files. Dates contained in special *Microfit* files are automatically converted to the European format, regardless of the settings of your computer.

3.3 Adding two data files

You can add two *Microfit* files containing the same variables, or add the variables from one *Microfit* file to another.

To add two files containing the same variables, select the 'Add 2 Special Microfit Files' option from the File Menu. Choose the first file, click OK, and then choose the second file from the Open dialogue. If the files contain the same variables, data frequency, or any number of undated observations, data are combined. The program appends the observations from the second data set to those of the first data set and, when observations can be ordered, sorts them.

To add two files containing different variables but the same data frequency, load the first file into *Microfit* in the usual way. Then select 'Add a Special Microfit File to Workspace' from the File Menu, and choose the file you want to add to your existing file from the Open dialogue.

If the data are incompatible, a warning message is displayed.

3.3.1 Adding two special *Microfit* files containing the same variables

Adding two data files containing different observations on the same variables is particularly useful for extension and/or revision of data in either direction (backward or forward), and for stacking of undated (cross-sectional) data. In the case of files containing dated and overlapping observations the second file that you specify should contain the most recent information. The content of the first file which overlaps with the second file will be overwritten.

As an example suppose you have a special *Microfit* file (say, SET1.FIT) which contains annual observations for the period 1970-1978 on the variables, C, S, and Y.

Obs	C	S	Y
1970	57814.0	0.0908	63585.0
1971	59724.0	0.0747	64544.0
1972	63270.0	0.0989	70214.0
1973	66332.0	0.1163	75059.0
1974	65049.0	0.1215	74049.0
1975	60000.0	0.1429	70000.0
1976	60000.0	0.1429	70000.0
1977	60000.0	0.1429	70000.0
1978	60000.0	0.1429	70000.0

Consider now a second special *Microfit* file called SET2.FIT which contains revised and updated observations on the same variables C, S, and Y over the period 1975 to 1980.

Obs	C	Y	S
1975	64652.0	74005.0	0.1264
1976	64707.0	73437.0	0.1189
1977	64517.0	72288.0	0.1075
1978	68227.0	78259.0	0.1282
1979	71599.0	83666.0	0.1442
1980	17550.0	84771.0	0.1560

Using the 'Add 2 Special *Microfit* files' option from the File Menu, and choosing SET1.FIT as first file and SET2.FIT as second file creates the following 'combined' data set:

Obs.	C	Y	S
1970	57814.0	63585.0	0.0908
1971	59724.0	64544.0	0.0747
1972	63270.0	70214.0	0.0989
1973	66332.0	75059.0	0.1163
1974	65049.0	74049.0	0.1215
1975	64652.0	74005.0	0.1264
1976	64707.0	73437.0	0.1189
1977	64517.0	72288.0	0.1075
1978	68227.0	78259.0	0.1282
1979	71599.0	83666.0	0.1442
1980	71550.0	84771.0	0.1560

Notice that the observations for the period 1975-1978 in the first file (SET1.FIT) which overlap with the observations in the second file (SET2.FIT) are overwritten by the corresponding observations in the second file. Also note that the order of variables in the combined data set is the same as that of the second file. *Remember to save the combined data set as a special Microfit file!*

In the case of data files with non-overlapping observations, the data gaps (if any) will be shown by a blank field, indicating missing observations. For example, combining the files

First File			Second File		
Obs	$X1$	$X2$	Obs	$X1$	$X2$
1960	2.0	10.0	1965	10.0	25.0
1961	3.0	20.0	1966	20.0	35.0
1962	4.0	30.0	1967	22.0	45.0

produces the combined data set

Obs	$X1$	$X2$
1960	2.0	10.0
1961	3.0	20.0
1962	4.0	30.0
1963		
1964		
1965	10.0	25.0
1966	20.0	35.0
1967	22.0	45.0

You can optionally fill in the missing values in the Process window by means of the commands **FILL_MISSING** and **FILL_FORWARD**. See Chapter 4 on how to use these and other commands. In the case of data files containing undated observations (cross-sectional data), the use of this option has the effect of stacking the observations in the two special *Microfit* files. This facility is particularly useful for pooling cross-section and time series data. For example, combining the following two special *Microfit* files containing *undated* observations

First File

Obs	PU	PS
1	3.0	66.0
2	3.0	66.0
3	9.0	62.0
4	9.0	64.0

Second File

Obs	PU	PS
1	16.0	71.0
2	25.0	64.0
3	24.0	64.0
4	22.0	64.0
5	12.0	70.0
6	13.0	66.0

results in the data set which appends the observations in the second file at the end of the observations in the first file.

Only *Microfit* files with the same data frequencies can be combined. For example, a data set containing annual observations cannot be combined with a data set containing quarterly or monthly observations.

3.3.2 Adding two special *Microfit* files containing different variables

Combining two files containing different variables but the same data frequency allows you to add new variables to your current data set. The new variables should already have been stored in a special *Microfit* file.

When using this option the following points are worth bearing in mind:

1. The current data set and the special *Microfit* file to be added to it should have the same data frequencies, otherwise an error message will be displayed
2. The current data set and the special *Microfit* file need not cover the same time period

As an example, suppose your current data set contains

Obs	X	Y
1960	34.0000	76.0000
1961	25.0000	84.0000
1962	76.0000	90.0000

and you have a special *Microfit* file containing variables A and B over the period 1959 to 1963. The special *Microfit* file to be added to the current data set contains

Obs	A	B
1959	20.0000	72.0000
1960	40.0000	98.0000
1961	50.0000	76.0000
1962	30.0000	45.0000
1963	56.0000	87.0000

If you now add the files the above special *Microfit* file is added to your current data set, and your new current data set is

Obs	X	Y	A	B
1959			20.0000	72.0000
1960	34.0000	76.0000	40.0000	98.0000
1961	25.0000	84.0000	50.0000	76.0000
1962	76.0000	90.0000	30.0000	45.0000
1963			56.0000	87.0000

Note: If you wish to add variables to your data set from the keyboard, you should use the **ADD** command followed by the name of the variable in the Process window and then press 

to add the variable name to the workspace. Once the new variable is added to the workspace use the 

button to input values for the new variable just added to the workspace.

3.4 Using the Commands and Data Transformations box

At various stages during the processing of your data you will need to enter text into an on-screen editor box, such as the Commands and Data Transformation box. Text can be edited in the usual way, using the Cut, Paste, and Copy options in the Edit Menu.

To scroll through the contents of the current editor, use the mouse or the cursor keys. To scroll up or down a screen page, press PgUp or PgDn. To scroll to the top or bottom of the editor text, press Ctrl+Home, or Ctrl+End.

When you have finished using the current editor, click the 

button. You can change the font of the text displayed in the box editors using the 

button. To clear a box editor completely, click 

.

3.5 Saving data

You can save your current data set (the data in the *Microfit* workspace), in several different formats: in an ASCII (text) format, in a special *Microfit* format for use in subsequent *Microfit* sessions, in a comma delimited (CSV) file, in an AREMOS file or in an *Excel* format.

To save your current data file, click 'Save as' from the File Menu and select in the sub-menu the type of file in which you want to save your data.

A 'Save as' dialogue appears, choose the drive and the directory in which you want to save the data, and type in a filename in the usual way. Click 

.

3.5.1 Save as a special *Microfit* file

You can save your data in a special *Microfit* file, for use in subsequent sessions. In addition to data, special *Microfit* files also contain other important information, namely data frequency, time periods, and variable names and their description (if any).

Notes

1. If you specify a file that already exists, you will be prompted to confirm that you wish to overwrite it.

2. *Microfit* automatically affixes the file extension .FIT to files saved as special *Microfit* files. An alternative cannot be used.

3. Special *Microfit* files saved using the earlier version of *Microfit* (versions 4.0 and lower) can still be used in *Microfit 5.0*. Special *Microfit* files created in this version of *Microfit* can be used in *Microfit 4.0* but not in earlier versions.

3.5.2 Save as an *Excel* sheet

You can save your data in *Excel* format. The *Excel* file created by *Microfit 5.0* contains in the first column the dates or the observation numbers, and in the first row the variable names and their optional descriptions.

3.5.3 Save as a comma separated values (CSV) file

This file format is useful when you wish to export the data from a *Microfit* workspace to spreadsheet packages such as Microsoft *Excel*.

If you select the option 'Comma Separated Values File' from the Save As Menu, *Microfit* saves data as a CSV file. The CSV file created by *Microfit 5.0* contains information in ASCII (raw text) on data frequency (undated, yearly, quarterly, and so on), the variable names and their optional descriptions given in the first row.

Alternatively, if you select the option 'CSV File with Descriptions in 2nd Row', data are saved with the description of the variables given in the second row, immediately below the variable names.

Microfit 5.0 also allows you to save data as a CSV file, *excluding rows with missing values*. This is particularly useful when you want to estimate a regression model, but there are missing values in the middle of the estimation period, since in this case *Microfit* does not carry out the estimation (see Chapter 11). You can create a new file that excludes the rows with missing values, by choosing the option 'CSV, Descriptions in the 2nd Row, Excludes Rows with Missing Values, Undated and Daily Data Only'. Note that this option is only valid for undated and daily data.

After selecting the format in which you want to save your data, type in the filename and click OK. *Microfit* automatically gives the file an extension .CSV. An alternative extension cannot be used.

Notes

1. The start and the finish of the subset of observations to be saved should fall within the specified range (between the minimum and the maximum values).

2. A raw data file does not contain any information on the frequency of your data, variable names or their descriptions. It only contains numbers, or ‘raw’ data.

3. When saving raw data files *Microfit* replaces missing data, namely a blank field, by the number 8934567.0

3.5.4 Save as an AREMOS (TSD) file

You can save data from *Microfit's* workspace as a Time Series Data (TSD) file for export into the *AREMOS* package. The file extension .TSD is assigned by default. An alternative extension cannot be used.

3.5.5 Save as a raw data (numbers only) file

You can save your data in an ASCII (text) file for the purpose of exporting it to other programs or computers that are not capable of reading CSV files (see Section 3.5.3). If you choose this option you are asked to type the start and finish of the subset of observations to be saved, whether you want to save your data variable by variable or by observations, and whether you want to save your data in fixed of free format.

Notes

1. The start and the finish of the subset of observations to be saved should fall within the specified range (between the minimum and the maximum values).

2. *Microfit* will automatically affix the extension .DAT to files saved as raw data files.

3. A raw data file does not contain any information on the frequency of your data, variable names or their descriptions. It only contains numbers, or ‘raw’ data.

4. When saving raw data files *Microfit* replaces missing data, namely a blank field, by the number 8934567.0

3.6 Starting with a new data set

Suppose you wish to abandon your current data set and start with a new data set without exiting *Microfit*. To enter your data from keyboard, choose ‘Input New Data from the Keyboard’ from the File Menu. To open an existing file, choose ‘Open File’ or ‘Open File from Tutorials Data Files’ from the File Menu.

Warning: Before starting with a new data set make sure that your current data set is saved. To save your data, use the ‘Save as’ option from the File Menu.

Chapter 4

Data Processing and Preliminary Data Analysis

When your data has been successfully read in *Microfit*, the program opens the Process window. This is *Microfit*'s gateway to data transformations and preliminary data analyses. To return to the Process window if this is not displayed on screen, click the 

button. Various buttons appear along the top and the base of the Process window.

The rectangular buttons across the top of the Process window are used to access other parts of the application.

To view your variables, and edit their names and/or descriptions, click the 

button. To return to the Data window to edit your data, click the 

button.

To access the Single Equation Estimation window, click the 

button. The Single Equation Estimation options can also be accessed directly using the Univariate Menu on the main menu bar (for more information on these estimation options see Chapter 6).

To access the System Estimation window, click the 

button. The System Estimation functions can also be accessed directly using the Multivariate Menu on the main menu bar (for more information on these estimation options see Chapter 7).

To access the Volatility Modelling window, click the 

button. These functions can also be accessed directly using the Volatility Modelling Menu on the main menu bar (for more information on these estimation options see Chapter 8).

The buttons along the base of the screen, on the right, allow you to create constants, time trends, and seasonal dummies. These functions can also be accessed via the Edit Menu; they are described in Section 4.1.

The Process window is divided into two panes the Variables box, which lists the variable names in the workspace, and the Commands and Data Transformations box, where you enter your commands. You can move up, down and side ways in these boxes by using the PgUp and PgDn and the cursor keys ↑→↓←.

You can type formulae and commands directly in the Commands and Data Transformations box (see Section 4.2). The different formulae need to be separated by semicolons (;). You can see lists of available functions and commands using the drop-down lists above the

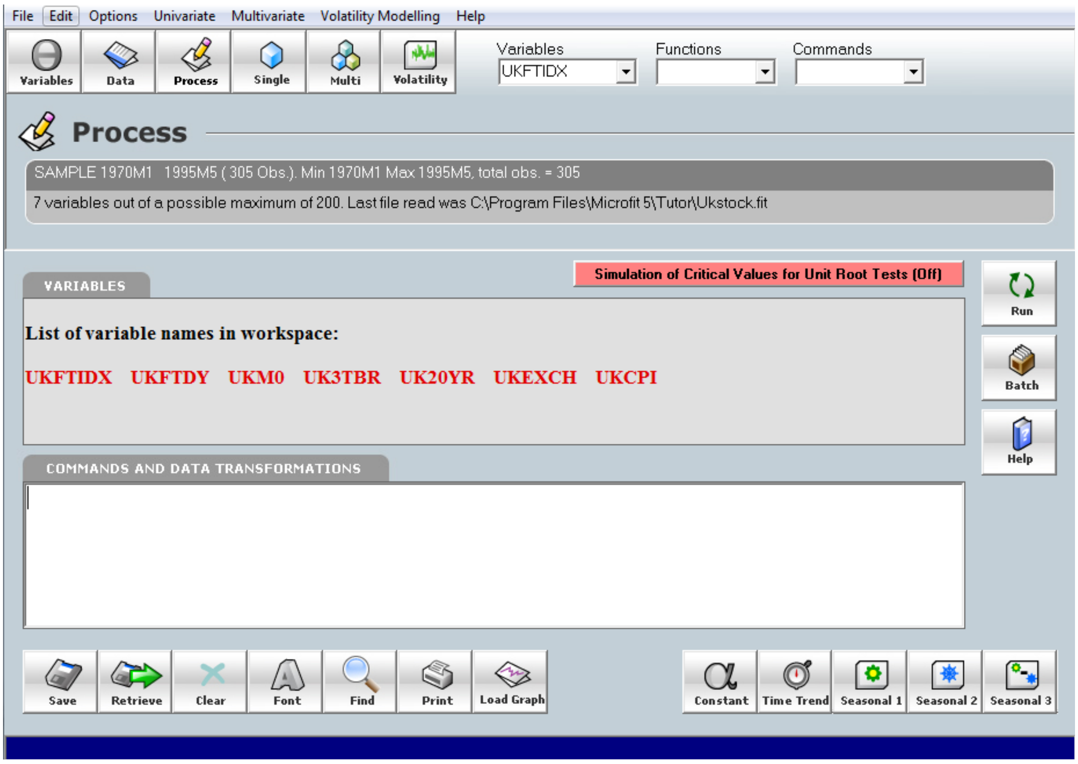


Figure 4.1: The Process window

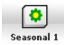
Process window. Click on the appropriate box to view its list. To select a function or a command, click on it in the list. Functions and Commands are discussed in Sections 4.3 and 4.4.

4.1 Creating constant terms, time trends and seasonal dummies

The buttons along the base of the Process window allow you to create constants, time trends, and seasonal dummies. Equivalent options are found in the Edit Menu. These options enable you to create a constant (or an intercept) term, a linear time trend, or seasonal dummies with the frequency of your choice, for use with quarterly or monthly observations.

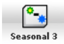
Click here to create a constant.

To create a time trend, use this button.



Click this button to create (0,1) seasonal dummies.

Creates centred seasonal dummies.



Creates seasonal dummies relative to the last period.

4.1.1 Creating a constant (intercept) term

To create a constant (intercept) term click the  button or choose Constant (intercept) term from the Edit Menu. *Microfit* creates a constant term (a variable with all its elements equal to unity) and asks you to supply a name for it.

4.1.2 Creating a time trend

If you click  or choose 'Time trend' from the Edit Menu, *Microfit* creates a time trend and asks you to supply a name for it. The time trend variable begins with the value of 1 at the start of the sample and increases in steps of 1.

4.1.3 Creating (0,1) seasonal dummies

To create traditional seasonal dummies, click the  button or choose 'Seasonal dummies (0,1)' from the Edit Menu. Each seasonal dummy created will have the value of unity in the season under consideration and zeros elsewhere. In the case of quarterly observations, seasonal dummies created by this option will be

Obs.	S1	S2	S3	S4
80q1	1	0	0	0
80q2	0	1	0	0
80q3	0	0	1	0
80q4	0	0	0	1
81q1	1	0	0	0
81q2	0	1	0	0
81q3	0	0	1	0
81q4	0	0	0	1

When you choose this option, you will be asked to specify the periodicity of your seasonal variables. When your data are undated, or are ordered annually, you can choose any periodicity. But for other data frequencies the program automatically creates seasonal dummies with periodicity equal to the frequency of your data, and presents you with a screen editor containing default names for the seasonal dummies. For example, for half-yearly data the periodicity will be 2 and the default variable names $S1$ and $S2$; for quarterly data the periodicity will be 4 and the default variable names $S1, S2, S3$, and $S4$; and for monthly data the periodicity will be 12, and the default variable names, $S1, S2, ..., S12$.

In the case of daily data you will be asked to specify the periodicity of your seasonal variables.

Notice that in each case the sum of the seasonal dummies will add up to unity, and their inclusion in a regression equation containing an intercept (constant) term will cause regressors to be perfectly multicollinear. To avoid this problem one of the seasonal dummies can be dropped, or use the seasonal dummies created using the 

button (see Section 4.1.5).

4.1.4 Creating centred seasonal dummies

Clicking the 

button, or choosing the 'Seasonal Dummy, Centred' option from the Edit Menu, generates seasonal dummies centred at zero. For example, in the case of quarterly observations the centred seasonal dummies will be

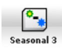
Obs.	SC1	SC2	SC3	SC4
80q1	0.75	-0.25	-0.25	-0.25
80q2	-0.25	0.75	-0.25	-0.25
80q3	-0.25	-0.25	0.75	-0.25
80q4	-0.25	-0.25	-0.25	0.75
81q1	0.75	-0.25	-0.25	-0.25
81q2	-0.25	0.75	-0.25	-0.25
81q3	-0.25	-0.25	0.75	-0.25
81q4	-0.25	-0.25	-0.25	0.75

When you choose this option you will be presented with a screen editor containing the default names SC1, SC2, and so on. Click 

to accept or edit the default variable names and then click 

.

4.1.5 Creating seasonal dummies relative to the last season

Clicking the 

button, or choosing 'Seasonal Dummies Relative to Last Season' from the Edit Menu, creates seasonal dummies relative to the last season, which are transformations of $(0, 1)$ seasonal dummies. For example, in the case of quarterly observations it generates the following seasonal dummies

$$\begin{aligned} SR1 &= S1 - S4 \\ SR2 &= S2 - S4 \\ SR3 &= S3 - S4 \end{aligned}$$

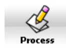
where $S1, S2, S3, S4$ are the (0,1) dummies, created using the 

button. These relative seasonal dummies can be included along with an intercept (constant) term in a regression equation, and their coefficients provide estimates of the first three seasonal effects (say

α_1, α_2,and α_3). The effect of the last season can be computed as $-(\alpha_1 + \alpha_2 + \alpha_3)$. This procedure restricts the sum of the seasonal effects to be zero. A similar logic also applies to monthly or half-yearly observations. For monthly observations, 11 relative seasonal dummies defined as $SR1 = S1 - S12$, $SR2 = S2 - S12, ...$ will be created. Once again $S1, S2, ...$ are (0,1) monthly seasonal dummies, which can be created using, for example, the 

button (see Section 4.1.3).

4.2 Typing formulae in *Microfit*

The Process window is automatically displayed when a data set is opened. To return to the Process window from another part of the application, click the 

button. One or more formulae or commands can be typed in the Commands and Data Transformations box. The different formulae need to be separated by semi-colons (;).

When entering information in the Commands and Data Transformations box you can quickly add one or more variables by *highlighting their names and dragging them from the Variables box* into the editor while holding down the mouse button.

The formula can be as complicated as you wish and, with the help of nested parentheses, you can carry out complicated transformations on one line, using standard arithmetic operators such as +, −, /, ∗, and a wide range of built-in functions. For example, to create a new variable (say, $XLOG$) which is the logarithm of an existing variable (say, X) you need to type

$$XLOG = \mathbf{LOG}(X)$$

in the Functions box. You can then execute the formula by clicking the 

button. This operation places the natural logarithm of X in $XLOG$. The program will show the extra variable ($XLOG$) in the list of variables. The program adds the extra variable ($XLOG$) to the list of variables in the Variables box shown above the Commands and Data Transformations box, if there is enough space. Otherwise you need to move to the Variables box and use the cursor keys or the PgUp and PgDn keys to see all your variables. To view the values for $XLOG$, click the 

button. Scroll through the data to find the new variable if necessary.

In this version of *Microfit* it is also possible to enter two or more formulae/commands in the Functions box before carrying out the operations. Suppose you have annual observations on US output over the period 1950-1994, but you wish to compute the percentage rate of change of the variable $USGDP$, and compute its mean and standard deviations over the period 1970 to 1990. You need to type the following operations in the Functions box:

$$\begin{aligned} G &= 100 * (USGNP - USGNP(-1))/USGNP(-1); \\ & \mathbf{SAMPLE} \quad 1970 \quad 1985; \quad \mathbf{COR} \quad G \end{aligned}$$

and then press the 

button to carry out the required operations. Notice that mistakes can be readily corrected at little cost. *Microfit* does not automatically clear the Functions box and so you can edit any mistakes that you find without having to type all the formulae or commands. Also you can save the content of the Commands and Data Transformations box in a file for later use by clicking the  button, at the base of the screen.

The following points are worth bearing in mind when entering formulae/command:

1. The upper- and lower-case letters are equivalent. So, the above operation could be put into effect by typing
$$XLOG = \mathbf{LOG}(X)$$

2. The new variable $XLOG$ is added to the list of the variables in the workspace, but needs to be saved as a special *Microfit* file if you wish to use it in subsequent sessions (see Section 3.5.1 on how to save a file as a special *Microfit* file).

3. If one or more observations of X are negative the program still creates the new variable $XLOG$ but enters the missing value indicated by a blank field for negative observations.

4. If you attempt to take the logarithm of a non-existent variable you will see the error message

   Error in formula or command or ';' missing

   Examples of other data transformation are as follows:

$$Y = 2 * X + (Z/3) - 5$$

This creates a new variable called Y which is equal to twice X plus a third of Z minus 5. An error message is generated if X and/or Z do not exist.

$$Y = X\hat{}3.25$$

This raises X to the power 3.25 and places the result in Y (on some keyboards the symbol $\uparrow$ is used in place of ˆ).

$$X1 = X(-1)$$

This generates first-order lagged values of X and places the results in the variable $X1$. The initial value, if undefined, will be set to blank, i.e. missing. For higher-order lagged values of X, you need to specify the order in parenthesis. For example, to create third-order lagged values of X in $X3$, you should type

$$X3 = X(-3)$$

It is also possible to generate lead values of a variable of any arbitrary order. For example,

$$XF2 = X(+2)$$

creates the second-order lead of X and places the results in variable $XF2$. In this example the undefined final two observations of $XF2$ will be set to blank.

$$INPT = 1$$

This creates an intercept (or a constant) term called $INPT$ with all its elements equal to 1.

You can also create linear or quadratic time trends in the Process window. Suppose you have a set of monthly observations over the period 1965(1)-1995(12), and you wish to create a linear trend variable, say T, starting with the value of 1 in January 1974. Type

$$\mathbf{SAMPLE}\ 1974m1\ \ 1995m12;\ T = \mathbf{CSUM}(1)$$

The variable T now contains the values of $1, 2, 3, ...$ for the months $74m1$, $74m2$, $74m3$, and so on. The values in T for the months prior to 1974 will be set to blank, unless you specify otherwise. **CSUM**$(\cdot)$ is the 'cumulative sum function' described below in Section 4.3.4.

4.2.1 Printing, saving, viewing, and copying files

You can save the content of the Commands and Data Transformations box in a file for later use by clicking the button. Such a file will be saved with the extension .EQU or .LST. To load a file saved previously, click the . You are presented with an Open file dialogue displaying the files with the extensions .EQU and .LST. Choose the location and the name of the file in the usual way. To view, copy, or save the contents of a saved file (without loading it into the editor) choose the 'View a File' option from the File Menu. Choose the file you want from the dialogue in the usual way. It is displayed in the File View window.

You can copy text from the File View window to the clipboard or to a new file of the same type. First, highlight the text you want to copy. To save it to a file, click the button and specify a name and a location for your saved file. To copy it to the clipboard, click the button.

To edit the font of the text displayed in the window, click the button. A standard windows font dialogue is displayed. Make your choices in the usual way and click .

4.3 Using built-in functions in *Microfit*

Standard mathematical functions, namely LOG (logarithm function), EXP (exponential function), COS (cosine function), SIN (sine function), $SQRT$ (square roots operator), and ABS (absolute value operator) can also be used in a formula.

In addition to the above standard functions, several other functions can also be used in a formula. A brief description of these functions is given in Sections 4.3.1 to 4.3.30. For a list of available functions, click the Functions field above the Process window and scroll through the drop-down list. To copy a function from the list to the Functions box in the Process window, click on it.

4.3.1 Function ABS

Absolute value function. For example,

$$Y = 2 + 3 * \mathbf{ABS}(X)$$

4.3.2 Function COS

The cosine function. Example:

$$Y = 2 + 3.5 * \mathbf{COS}(5 + 2 * X)$$

4.3.3 Function CPHI

This is the cumulative standard normal function so that $\mathbf{CPHI}(X)$ represents the integral between minus infinity to X of the standard normal distribution. For examples,

$$Y = \mathbf{CPHI}(0.0)$$

returns the value of 0.5 for Y,

$$Y = \mathbf{CPHI}(W + 2 * (Z1/Z2))$$

first computes the expression inside the brackets, and then returns the values of the integral of the standard normal distribution from minus infinity to $W + 2 * (Z1/Z2)$.

4.3.4 Function CSUM

This function, when applied to a variable X, calculates the cumulative sum of X. For example, if $X = (6, 2, -1, 3, 1)$, then typing the formula

$$Y = \mathbf{CSUM}(X)$$

will result in $Y = (6, 8, 7, 10, 11)$.

The argument X can itself be a function of other variables as in the following example:

$$Y = \mathbf{CSUM}((z - (\mathbf{SUM}(z)/\mathbf{SUM}(1))^2)$$

Here $\mathbf{SUM}(\cdot)$ is the function **SUM** described below. See Section 4.3.29.

Warning: Before using this command you need to make sure that there are no missing values in the variable X.

4.3.5 Function EXP

This function takes the exponential of the expression that follows in the brackets. For example,

$$Y = 2 + 3.4 * \mathbf{EXP}(1.5 + 4.4 * X)$$

The general form of $\mathbf{EXP}(\cdot)$ is given by

$$\mathbf{EXP}(x) = 1 + x + x^2/2! + x^3/3! + ...$$

4.3.6 Function GDL

This is the geometric distributed lag function and has the form

$$Y = \mathbf{GDL}(X, \lambda)$$

It computes Y as

$$\mathbf{SIM}\ Y = \lambda * Y(-1) + X$$

over the sample in which it is in effect. λ is the parameter of the distributed lag function, and **SIM** is the **SIM** command (see Section 4.4.25). The initial value is set equal to the value of X at the start of the selected sample period. For example, suppose you wish to compute geometric distributed lag of X over the period 1950-1980, with $\lambda = 0.8$, you need to type

$$\mathbf{SAMPLE}\ 1950\ 1980;\ Y = \mathbf{GDL}(X, 0.8)$$

The value of Y in 1950 will be set equal to the value of X in 1950.

Warning: Before using this command you need to make sure that there are no missing values in the variable X.

4.3.7 Function HPF

This function has the form

$$Y = \mathbf{HPF}(X, \lambda)$$

and runs the variable X through a Hodrick-Prescott (HP) filter using the parameter λ. In this function X is a vector, and λ is a non-negative scalar (a vector with all its elements equal to $\lambda \geq 0$). This filter is used extensively in the real business cycle literature as a de-trending procedure (see Hodrick and Prescott 1997)

The choice of λ depends on the frequency of the time series, X. For quarterly observations Hodrick and Prescott set $\lambda = 1,600$. The argument X could also be specified to be a function of other variables in the variable list. Harvey and Jaeger (1993) show that for $\lambda = 1,600$ the transfer function for the HP filter peaks around 30.14 quarters (approximately 7.5 years). This suggests using the value of $\lambda = 7$ for annual observations, and $\lambda = 126,400$ for monthly observations.[1] But, in general, the optimal choice of λ will depend on the particular time series under consideration.

For example, suppose $USGNP$ contains quarterly observations on US aggregate output. The trend series (in logarithms) are given by

$$YT = \mathbf{HPF}(\mathbf{LOG}(USGNP), 1600)$$

To compute the filtered, or de-trended, series you need to type

$$YD = \mathbf{LOG}(USGNP) - YT$$

[1] We are grateful to Micheal Binder for the estimates of λ in the case of annual and monthly observations.

4.3.8 Function INVNORM

This function computes the inverse of the cumulative standard normal distribution, so that for a given probability p, $Y =$**INVNORM**(p) computes Y such that the area under the standard normal curve between minus infinity to Y is equal to p:

$$Y = \textbf{INVNORM}(0.975)$$

In this example, Y will be set to 1.9600, the 95 per cent critical value of the standard normal distribution. Note that $0 < p < 1$.

4.3.9 Function LOG

The function takes logarithm to the base e (natural logarithm) of the expression that follows in brackets. For example

$$Y = 2.4 + 3.5 * \textbf{LOG}(X + 3)$$

For negative or zero values of X, this function returns missing values.

4.3.10 Function MAX

This function has the form

$$Y = \textbf{MAX}(X, Z)$$

and places the maximum of X and Z in Y. For example, if $X = (1, 7, 2, 3, 6)$ and $Z = (6, 2, -1, 3, 1)$, then Y will be set to $(6, 7, 2, 3, 6)$.

4.3.11 Function MAV

This function has the form

$$Y = \textbf{MAV}(X, p)$$

and places the pth order moving-average of the variable X in Y, namely

$$Y_t = \frac{1}{p}(X_t + X_{t-1} + ... + X_{t-p+1})$$

Variable X could be any of the variables on the workspace or a function of them. If p is not an integer *Microfit* chooses the nearest integer to p in order to carry out its computations. If p is negative this function returns a missing value (a blank field).

4.3.12 Function MEAN

This function, when applied to a variable X, calculates the mean of X over the specified sample period. For example,

$$\textbf{SAMPLE}\ 1970\ 1995;\ Y = (X - \textbf{MEAN}(X))$$

generates deviations of variable X from its mean computed over the sample period 1970-1995, inclusive.

Note that the value of $\mathbf{MEAN}(X)$ will be set to missing if one or more values of X are missing over the specified sample period.

4.3.13 Function MIN

This function has the form

$$Y = \mathbf{MIN}(X, Z)$$

and places the minimum of X and Z in Y. The arguments X and Z themselves could be functions of other variables, as in the following example:

$$y = \mathbf{MIN}((G1/G2) + 1, (H1/H2) - 1)$$

4.3.14 Function NORMAL

This function can be used to generate independent standardized normal variates (i.e. with zero means and unit variances). The function should be used in the form of $\mathbf{NORMAL}(j)$ within a formula, where j represents the 'seed' for the quasi-random numbers generated, and it must be an integer in the range $0 < j < 32000$. By changing the value of j, different quasi-random number series can be generated. Examples of the use of this function are:

$$X = \mathbf{NORMAL}(1); \quad Y = 2 + 3.5 * \mathbf{NORMAL}(124) + Z$$

Warning: The function **UNIFORM** and **NORMAL** must not be used in **SIM** or **SIMB** commands!

4.3.15 Function ORDER

This function has the form

$$Y = \mathbf{ORDER}(X, Z)$$

and orders X according to the sorting implied by Z, where Z is sorted in an ascending order. The results is placed in Y. For example, if $X = (1, 7, 2, 3, 6)$ and $Z = (6, 2, -1, 3, 1)$, then Y will be set to $(2, 6, 7, 3, 1)$.

Clearly, as in the case of other functions, the arguments of the function, namely X and Z, could themselves be functions of other variables.

4.3.16 Function PHI

This function gives the ordinates of the standard normal distribution for the expression that follows in brackets. For example,

$$Y = 0.5 * (1/\mathbf{PHI}(0))\,\hat{}\,2$$

$$Y = \mathbf{PHI}(1)$$

$$Z = \mathbf{PHI}(1 + 0.5 * W)$$

The general formula for the PHI function is given by

$$\mathbf{PHI}(X) = ((2 * \pi)\hat{}(-0.5)) * \mathbf{EXP}(-0.5 * X\hat{}2)$$

where **EXP** is the exponential function (see Section 4.3.5), and π=3.14159.

4.3.17 Function PTTEST

This function has the form

$$T = \mathbf{PTTEST}(Y, X)$$

and returns the Pesaran and Timmermann (1992) statistic for a non-parametric test of association between the variables Y and X. Under the null hypothesis that X and Y are distributed independently, $\mathbf{PTTEST}(X, Y)$ has a standard normal distribution in large samples. For example, for a two-sided test, the hypothesis that Y and X are statistically independent will be rejected if **PTTEST** is larger than 1.96 in absolute value.

4.3.18 Function RANK

This function, when applied to a variable X, gives the ranks associated with the elements of X, when X is sorted in an ascending order. For example, if $X = (6, 2, -1, 3, 1)$ then typing the formula

$$Y = \mathbf{RANK}(X)$$

will give $Y = (5, 3, 1, 4, 2)$.

4.3.19 Function RATE

This function has the form

$$PIZ = \mathbf{RATE}(Z)$$

It computes the percentage rate of change of Z and places the result in PIZ. More specifically, PIZ will be computed as

$$PIZ = 100 * (Z - Z(-1))/Z(-1)$$

with its initial value set to blank.

An alternative approximate method of computing rate of change in a variable would be to use changes in logarithms, namely

$$PIZX = 100 * (\mathbf{LOG}(Z/Z(-1))$$

It is easily seen that both approximations are reasonably close to one another for values of PIZ and $PIZX$ around 20 per cent or less.

The argument of this function, namely Z, can itself be a function of other variables, as in the following example

$$Y = \mathbf{RATE}(W + U/V\,)$$

4.3.20 Function REC_MAX

This function, when applied to a variable X, returns for each observation j the maximum value of X over the interval that goes from the start of the sample to the j^{th} observation. For example, if $X = (2, 1, 3, 0, 4)$, then typing

$$Y = \mathbf{REC\_MAX}(X)$$

will return $Y = (2, 2, 3, 3, 4)$.

Warning: Before using this command you need to make sure that there are no missing values in the variable X.

4.3.21 Function REC_MIN

This function, applied to a variable X, returns for each observation j the minimum value of X over the interval from the start of the sample to the j^{th} observation. If, for example, $X = (2, 1, 3, 0, 4)$, then typing

$$Y = \mathbf{REC\_MIN}(X)$$

will set $Y = (2, 1, 1, 0, 0)$.

Warning: Before using this command you need to make sure that there are no missing values in the variable X.

4.3.22 Function ROLL_MAX

This function has the form

$$Y = \mathbf{ROLL\_MAX}(X, h)$$

where h is the window length. It returns the maximum value of X over successive rolling periods of a fixed length. For each observation j, this function computes the maximum value of X over the interval that goes from the $(j-h)^{th}$ to the j^{th} observation. The first $h-1$ observations are set to missing. If for example $X = (2, 1, 3, 0, 4, 3)$, then

$$Y = \mathbf{ROLL\_MAX}(X, 2)$$

will set $Y = (\quad, 2, 3, 3, 4, 4)$.

Warning: Before using this command you need to make sure that there are no missing values in the variable X.

4.3.23 Function ROLL_MIN

This function takes the form

$$Y = \textbf{ROLL\_MIN}(X, h)$$

It returns, for each observation j, the minimum value of X over the sample from the $(j-h)^{th}$ to the j^{th} observation, with the first $h-1$ observations set to missing. If, for example, $X = (2, 1, 3, 0, 4, 3)$, then

$$Y = \textbf{ROLL\_MIN}(X, 2)$$

will set $Y = (\ \ , 1, 1, 0, 0, 3)$.

Warning: Before using this command you need to make sure that there are no missing values in the variable X.

4.3.24 Function SIGN

This function, when applied to a variable X, returns the value of 1 when X is positive and 0 when X is zero or negative. For example, if $X = (3, -4, 2, 0, 1.5)$, then typing

$$Y = \textbf{SIGN}(X)$$

will return $Y = (1, 0, 1, 0, 1)$. Another example is

$$Y = \textbf{SIGN}(X - 2)$$

which will return $Y = (1, 0, 0, 0, 0)$.

4.3.25 Function SIN

This function takes sine of the expression that follows in the brackets. For example,

$$Y = 2 + 3 * \textbf{SIN}(5 + 7 * X)$$

4.3.26 Function SORT

This function, when applied to a variable X, will sort X in an ascending order. For example, if $X = (6, 2, -1, 3, 1)$ then typing

$$Y = \textbf{SORT}(X)$$

will set $Y = (-1, 1, 2, 3, 6)$, while

$$Z = -\textbf{SORT}(-X)$$

will sort X in descending order in Z so that $Z = (6, 3, 2, 1, -1)$.

In the example

$$Y = \textbf{SORT}(2 + w/z)$$

the expression $2 + w/z$ will be first computed, and the resultant expression will be sorted in Y as above.

4.3.27 Function SQRT

This function takes the square-root of the expression that follows in brackets. For example:

$$Y = 3 + 5 * \mathbf{SQRT}(X)$$

For negative values of X, this function returns the missing values.

4.3.28 Function STD

This function, when applied to a variable X, calculates the standard deviation of X over the specified sample period. For example,

$$\begin{array}{l}\mathbf{SAMPLE}\ 1970\ 1995;\\ Z = (X - \mathbf{MEAN}(X))/\mathbf{STD}(X)\end{array}$$

places the standardized values of X over the period 1970-1995 (inclusive) in the variable Z.

Warning: The value of $\mathbf{STD}(X)$ will be set to missing (blank) if one or more values of X are missing over the specified sample period.

4.3.29 Function SUM

This function first calculates the expression specified within closed brackets immediately following it, and then computes the sum of the elements of the result over the relevant sample period. Examples of the use of this function are:

$$\begin{array}{l}\mathbf{SAMPLE}\ 1960\ 1970; \quad XBAR = \mathbf{SUM}(X)/\mathbf{SUM}(1);\\ XD = X - XBAR;\ YBAR = \mathbf{SUM}(Y)/\mathbf{SUM}(1);\ YD = Y - YBAR;\\ BYX = \mathbf{SUM}(XD * YD)/\mathbf{SUM}(XD * XD)\end{array}$$

In the above examples, $\mathbf{SUM}(X)$ is a vector with all its elements equal to the sum of the elements of X over the period 1960-70. $\mathbf{SUM}(1)$ is equal to the number of observations in the sample period (namely, 11). $XBAR$ is, therefore, equal to the arithmetic mean of X, computed over the specified sample period. XD and YD are deviations of X and Y from their respective means and BYX is the ordinary least squares (OLS) estimates of the coefficients of X in the regression of Y on X (including an intercept term).

4.3.30 Function UNIFORM

This function can be used to generate independent random numbers from a uniform distribution within the range 0 and 1. The function should be used in the form of $\mathbf{UNIFORM}(j)$ within a formula, where j represents the 'seed' for the quasi-random numbers generated, and must be an integer in the range $0 < j < 32000$. By changing the value of j, different quasi-random number series can be generated.

Examples of the use of this function are:

$$
\begin{aligned}
X &= \mathbf{UNIFORM}(1) \\
Y &= 2 + 3.5 * \mathbf{UNIFORM}(124) + Z
\end{aligned}
$$

Warning: The functions **UNIFORM** and **NORMAL** must not be used in **SIM** or **SIMB** commands!

4.4 Using commands in *Microfit*

For a list of available commands, click the Commands field above the Process window and scroll through the drop-down list. To select a command from the list and copy it to the Functions box, click on it.

4.4.1 Command ADD

This command enables you to add a new variable to the list in the Variables box. The form of this command is

$$\mathbf{ADD}\ XNEW$$

where $XNEW$ is the name of the new variable.

The variable is added to the variables list, and a new empty column appears on the extreme right-hand side of the Data window. To add a description to your variable, click Variables to move to the Variables window and type in your description. To insert the values for your variable, click the Data button to move to the Data window, then click on each of the relevant cells in turn and type in the value you want.

Note that this command allows you to add one new variable at a time. If you want to add the variables from an existing *Microfit* file to your current data set, you should use the 'Add' option from the File Menu instead.

Remember to save the data set with the added variable(s) as a special *Microfit* file if you wish to use it in subsequent *Microfit* sessions.

4.4.2 Command ADF

This command, when followed by a variable name, displays the Dickey-Fuller (DF) and the augmented Dickey-Fuller (ADF) statistic for testing the unit root hypothesis together with the associated critical values. See Dickey and Fuller (1979), and Lesson 12.1.

For example,

$$\mathbf{SAMPLE}\ 75Q1\ 87Q1;\quad \mathbf{ADF}\ X$$

computes the DF and the ADF test statistics of up to order 4 (the periodicity of the data) for the variable X, and displays the statistics together with their 95 per cent critical values on screen for the following models:

Case I: No intercept and no trends. The ADF test statistic is computed as the t-ratio of ϕ in the $ADF(p)$ regression

$$\Delta X_t = \phi X_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta X_{t-i} + u_t \tag{4.1}$$

where $\Delta X_t = X_t - X_{t-1}$, and p is the order of augmentation of the test.

Case II: With intercept but without a trend. The p^{th} order ADF test statistic is given by the t-ratio of ϕ in the $ADF(p)$ regression

$$\Delta X_t = a_0 + \phi X_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta X_{t-i} + u_t \tag{4.2}$$

Case III: With an intercept and a linear time trend. The p^{th} order ADF test statistic is the t-ratio of ϕ in the regression

$$\Delta X_t = a_0 + a_1 T_t + \phi X_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta X_{t-i} + u_t \tag{4.3}$$

where T_t is a linear time trend.

Microfit computes the ADF statistic for cases I, II and III, and also provides Akaike information and Schwarz Bayesian criteria for selecting the order of augmentation in the ADF tests. The 95 per cent critical values for the test computed using the response surface estimates given in (MacKinnon 1991, Table 1), are provided at the foot of the result tables only for $p = 0$, for cases I and II.

Microfit 5.0 also presents the possibility of computing simulated critical values for unit roots tests by bootstrapping. Click on the rectangular button 

. You will be presented with a window that allows you to set the number of replications, the maximum number of observations used for simulating critical values, and the significance level. Check the 'Simulate Critical Values' checkbox and click 

. The rectangular button will turn green, indicating that, when running unit roots tests, the result table will display a new column, CV, containing the simulated critical values of ADF tests for each order of augmentation, for cases I, II and III.

Note that simulation of critical values could take a long time if your sample size is large (for example, more than 400 units) and you use all observations in the simulation. To control the computational time, reduce the maximum number of observations used for simulations (for example, to 400).

The **ADF** command can also be used to compute the augmented Dickey-Fuller test statistics, up to an order of augmentation specified by the user. The desired order should be specified in parentheses immediately after the variable name. For example,

ADF $X(12)$ 

gives the ADF test statistics for the variable X up to the order 12, assuming, of course, that there are enough observations.

Finally, ADF tests can be applied to a series after controlling for a set of deterministic/exogenous variables. This can be achieved by using the **ADF** command in combination with &. For example

$$\textbf{ADF} \quad Y \quad \& \quad Z1 \quad Z2$$

allows performing unit roots tests on residuals computed from a regression of Y on $Z1$ and $Z2$. Notice that in this case *Microfit* only reports ADF statistics and associated critical values for Case 1, the no intercept and no trends case. Intercept and/or trends can be included in the set of exogenous variables after &. Different orders of augmentation can also be computed as before by issuing the command

$$\textbf{ADF} \quad Y(4) \quad \& \quad Z1 \quad Z2$$

in the case of a 4th order ADF test. This option is particularly useful in testing for unit roots in the presence of structural breaks. For example, to allow for an intercept shift at a known point in time first construct a dummy variable, say $DUM1$, that takes the value of zero before the break and unity after, and then issue the following command

$$\textbf{ADF} \quad Y(4) \quad \& \quad INPT \quad DUM1$$

Also see Lessons 12.5 and 12.6.

Warning (sample selection for computation of ADF statistics): The actual sample period used by the **ADF** command is based on the sample period specified by the user and the lag order, p, of the ADF test. If data on X is available over the sample 1 to 100 (inclusive) and the sample is specified by the user is from 20 to 100 and $p = 2$, ADF regressions are run over the period 2 to 100 and observations 18 and 19 are used as initial values. But if the selected sample is 1 to 100 (and no other data exists before the first observation), then ADF regressions are run using observations 3 to 100, with X_1 and X_2 used as initial values. The same also applies to the other ADF statistics set out below.

4.4.3 Command ADF_GLS

This command, applied to a variable X, takes the form

$$\textbf{ADF\_GLS} \quad X$$

It computes the GLS augmented Dickey-Fuller test statistic due to Elliott, Rothenberg, and Stock (1996) of up to order p (the periodicity of the data) for testing the unit root hypothesis for cases I, II and III. The 95 per cent critical values for the test have been computed by Pantula, Gonzalez-Farias, and Fuller (1994) and by Elliott, Rothenberg, and Stock (1996), and are given at the base of the Result Tables for cases II and III. If you choose the option 'Simulate Critical Values', *Microfit* also provides the simulated critical values for the test for each order of augmentation (see Section 4.4.2).

The **ADF_GLS** command can also be used to compute the *GLS* augmented Dickey-Fuller test statistics up to an order of augmentation specified by the user, putting the order in parentheses immediately after the variable name, namely, by issuing the command

ADF_GLS $X(p)$

where p is the order of augmentation ($p \leq 12$).

This command *cannot* be used in conjunction with & to include additional variables when carrying out the test. To control for breaks or other exogenous effects when carrying out unit root tests use the ADF, ADF_MAX or ADF_WS commands. Also see the discussion and the warning regarding the sample period used for computation of *ADF* statistics in Section 4.4.2.

4.4.4 Command ADF_MAX

This command computes the Maximum Augmented Dickey-Fuller statistic for testing the unit root hypothesis in models of type I to III. The test is due to Leybourne (1995), in which further details can be found.

This test is given by the maximum between the usual *ADF* statistic and the *ADF* statistic computed on the reverse time series, in a regression with an intercept and a time trend. When applied to a variable X this command has the form

ADF_MAX X

Switch on the button Simulation of Critical Values for Unit Root Tests, to obtain the simulated critical values for the test for an arbitrary significance level (see Section 4.4.2).

The **ADF_MAX** command can also be used to compute the Maximum augmented Dickey-Fuller test statistics up to an order of augmentation specified by the user, putting the order in parentheses immediately after the variable name. This command can also be used in combination with & (see Section 4.4.2 for more information).

Examples of this command are

ADF_MAX $X(p)$

ADF_MAX $X(p)$ & $Z1$ $Z2$

where p ($p \leq 12$) is the order of augmentation and $Z1$ and $Z2$ are deterministics such as intercepts, time trends of dummy variables representing breaks in intercepts or the trend coefficients.

Also see the discussion and the warning regarding the sample period used for computation of *ADF* statistics in Section 4.4.2.

4.4.5 Command ADF_WS

This command computes the Weighted Symmetric *ADF* (*WS-ADF*) statistic, for testing the unit root hypothesis advanced by Park and Fuller (1995), in which further details can be found. Critical values of the test statistic for a given level of significance (to be selected by the user) can be obtained via stochastic simulations. For this purpose the option 'Simulate Critical Values' need to be switched on. see also Section 4.4.2.

The **ADF_WS** command can also be used in combination with & to control for effect of a set of exogenous variables (see Section 4.4.2 for more information). It can also be used to compute the weighted-symmetric augmented Dickey-Fuller test statistics up to an order of augmentation specified by the user, putting the order in parentheses immediately after the variable name. Examples of this command are

$$\textbf{ADF\_WS} \quad X(p)$$

$$\textbf{ADF\_WS} \quad X(p) \quad \& \quad Z1 \quad Z2$$

where p ($p \leq 12$) is the order of augmentation and $Z1$ and $Z2$ are deterministics such as intercepts, time trends of dummy variables representing breaks in intercepts or the trend coefficients.

Also see the discussion and the warning regarding the sample period used for computation of *ADF* statistics in Section 4.4.2.

4.4.6 Command BATCH

This command has the form

$$\textbf{BATCH}$$

or

$$\textbf{BATCH} < filename >$$

If you enter the command **BATCH** on its own, the names of the batch files in your default directory will appear on the screen. Select the appropriate batch file, and when prompted press OK.

When the batch command is followed by a file name the instructions in the file will be carried out on the variable in the workspace. A batch file must have the extension .BAT as a part of its name.

This command allows you to place a number of commands in a file so that they are subsequently obeyed in batch mode. The legitimate instructions can either be one or more mathematical formulae and/or commands: **SAMPLE**, **DELETE**, **KEEP**, **ENTITLE**, **SIM**, **SIMB**, **REORDER**, **RESTORE**, and $. When you use the command **ENTITLE** in batch mode you need to enter the descriptions of the variables on separate lines in exactly the same order that the variables are typed after the command. You can also write in comments in your BATCH file by starting your comments with the dollar sign, $. Anything entered on the same line after the $ sign in the BATCH file will be ignored.

An example of a simple BATCH file is given below:

$ Space for comments
SAMPLE $70M1$ $78M5$
$INPT = 1$
$Z = X +$ **LOG**(Y)
$W = X - Y$
SAMPLE $75M1$ $78M5$
ENTITLE X Y
Consumption Expenditures
Labour Income
$ The end of **BATCH** file

Running this file with the **BATCH** command creates variables $INPT$, Z, and W from the data series X and Y, and assigns the title 'Consumption Expenditures' and 'Labour Income' to the variables X and Y, respectively.

This is a useful command enabling you to carry out the same operations on different data sets or on revisions of the same data set.

Notice that in *Microfit*, the commands **COR**, **LIST**, **SPECTRUM**, **HIST**, **PLOT**, **XPLOT**, **SCATTER**, **ADD**, **ADF**, **EDIT**, and **TITLE** cannot be included in the BATCH file.

4.4.7 Command CCA

This command enables you to perform a canonical correlation analysis on two sets of variables, controlling for a third set of variables. Suppose in your workspace you have T observations on three sets of variables named $(Y1, Y2, ..., Yn)$; $(X1, X2, ..., Xs)$ and $(Z1, Z2, ..., Zg)$. Then to obtain the canonical correlations and the associated canonical variates between Y and X, controlling for Z, type

CCA $Y1$ $Y2$ .... Yn & $X1$ $X2$ .... Xs & $Z1$ $Z2$ ... Zg RUN

Microfit reports the squared canonical correlations $\rho_1^2 \geq \rho_2^2 \geq .... \geq \rho_k^2 \geq 0$, with $k = \min(n, s)$, and the canonical variates, u_{it} and v_{it} for $t = 1, 2, ..., T$ and $i = 1, 2, ..., k$, for the two sets of variables $Y1, Y2, ..., Yn$, and $X1, X2, ..., Xs$, once these have been filtered by the variables $Z1$ $Z2$ ... Zg. The filtering is carried out by running regressions of Yi and Xi on $Z1, Z2, ..., Zg$ and then using the residuals from these regressions to compute the canonical correlations. *Microfit* also reports the trace statistic for testing the independence of the two sets of variables.

Under the null hypothesis of independence and certain regularity conditions, the trace statistic is distributed as a chi-squared variate with $(n-1)(s-1)$ degrees of freedom. See Section 22.13 for further details on canonical correlation analysis and references to the literature. For tutorial lessons see 10.17 and 16.8.

4.4.8 Command COR

This command has different effects depending on whether it is followed by one variable or more. When only one variable is specified after **COR**, as in the example

SAMPLE 1 20; **COR** X

Summary statistics for X (mean, standard deviation, coefficient of variation, skewness, kurtosis, minimum, and maximum values) and its auto-correlation coefficients of up to the order of a third of the number of specified observations will be shown on the screen. If you have a graphics adaptor, the plot of the auto-correlation function will also be displayed.

The **COR** command can also be used to compute auto-correlation coefficients up to an order specified by the user. The desired order should be specified in parentheses immediately after the variable. For example,

COR $X(12)$

gives the auto-correlation coefficients for the variable X up to the order of 12 (assuming, of course, that there are enough observations). When the **COR** command is followed by two or more variables, as in the example

COR X Y Z

then summary statistics and the correlation coefficients for these variables, over the specified sample period, will be provided.

For the relevant formulae and appropriate references to the literature, see Section 21.1.

4.4.9 Command DELETE

This command enables you to delete one or more variables from the list of your existing variables in your workspace. The names of the variables to be deleted should follow the command, separated by spaces. For example,

DELETE X Y Z

deletes the variables X, Y, and Z from the list of your existing variables. If you wish to delete a single variable, you can also type

$X =$

This operation has the effect of deleting variable X.

4.4.10 Command DF_PP

This command, applied to a variable X, takes the form

DF_PP $X(h)$

It computes the Phillips-Perron unit roots test due to Phillips and Perron (1988) using a window of length h. This test attempts to correct for the effect of residual serial correlation in a simple DF regression with an intercept both with and without a time trend, using non-parametric estimates of the long-run variance. If you do not specify the window length in parentheses, *Microfit* automatically select it as in the command **SPECTRUM** (see Section 4.4.27). The command **DF_PP** can also be used in combination with & to control for a set of exogenous variables. For example,

$$\textbf{DF\_PP} \quad Y \quad \& \quad INPT \quad Z1 \quad Z2$$

allows performing unit roots tests on residuals from a regression of Y on the variables $INPT$, $Z1$, $Z2$ (see Section 4.4.2 for more information).

The 95 per cent critical values for the test are the same as DF critical values, and are provided at the foot of the Result table. Critical values for an arbitrary significance level can also be obtained via simulation, setting the option 'Simulate Critical Values' (see Section 4.4.2).

See Phillips and Perron (1988) for further information on this test.

4.4.11 Command ENTITLE

This command allows you to enter or change the description of one or more of the variables in your workspace. For example, if you type

ENTITLE

the Variable window opens and you can add the titles (or descriptions) of the variables you require. Note that the description of a variable can be at most 80 characters. If you type in a title which is more than 80 characters long, only the first 80 characters will be saved.

When a new variable is generated using the data transformation facilities, the first 80 characters after the equality sign will be automatically used as the title of the generated variable. Also, when a variable $XNEW$ is created by the formula

$$XNEW = XOLD$$

the title of the variable $XOLD$, if any, will be passed on to the new variable, $XNEW$.

4.4.12 Command FILL_FORWARD

This command allows you to replace missing values of a given variable or all the variables in the workspace. For example, if you type

FILL_FORWARD X

the program replaces any existing missing value of the variable X by the last available observation in X. If you type the command on its own, namely

FILL_FORWARD

for each variable in the workspace the program replaces the missing values by the last available observation for that variable.

Note: It makes sense to use this command only when observations can be ordered. In case of undated observations, the use of **FILL_MISSING** command is preferable (see Section 4.4.13).

4.4.13 Command FILL_MISSING

This command, when applied to a variable X, replaces any existing missing value of X with a value specified by the user. For example,

FILL_MISSING X 10

replaces all missing values in the X variable with 10.

4.4.14 Command HIST

This command will only work on computers with a graphics facility. When followed by a variable name, this command displays the histogram of the variable. For example,

SAMPLE 1 20; **HIST** X

The number of bands is automatically chosen between 6 and 15 according to the formula

$$Min\{15,\ Max(n/10,\ 6)\}$$

where n is the total number of observations.

This command can also be used to plot the histograms for any numbers of intervals chosen by the user. The desired number of classes should be specified in parentheses immediately after the variable. For example,

HIST $X(12)$

In the Graph window you can specify a different period over which you wish to see the histogram of the variable. Click the Start and Finish fields and scroll through the drop down lists to select the sample period, and then press the button 'Refresh graph over the above sample period'.

Use the Help Functions at the foot of the displayed graph for information about the various options available for saving and printing the displayed graph. Also see Section 5.2 on how to add text, print, save and retrieve graphs.

To exit the graphic routine, click the button.

4.4.15 Command KEEP

This command deletes all the variables in the workspace except those specified. For example, suppose you have 10 variables named $X1, X2, ..., X10$ in your workspace, and you wish to keep only the variables $X1$ and $X2$, then type

KEEP $X1$ $X2$

by which only the variables $X1$ and $X2$ will be kept. Also see the command **DELETE**.

4.4.16 Command KPSS

This command allows you to compute the stationarity test, developed by Kwiatkowski, Phillips, Schmidt, and Shin (1992), for a simple DF regression with an intercept and a linear time trend. Note that the null hypothesis in this test is that the time series is stationary. The command takes the form

KPSS $X(h)$

where the number in parentheses is the window size. If you type

KPSS X

a window size is automatically selected as in the command **SPECTRUM** (see Section 4.4.27). You can obtain simulated critical values for this test by setting on the option 'Simulate Critical Values' (see Section 4.4.2).

For further information on this test see Kwiatkowski, Phillips, Schmidt, and Shin (1992), in which appropriate critical values are also provided.

4.4.17 Command LIST

This command allows you to inspect your data on screen and/or to save them in a file to be printed out later. If the command **LIST** is typed on its own followed by , then the values of all the variables will be displayed over the current sample period set by the **SAMPLE** command. If the command **LIST** is followed by one or more variable names, then only the values of the specified variables will be listed.

For example,

SAMPLE 1940 1980;
LIST

displays all the existing variables over the period 1940-80.

SAMPLE $80Q1$ $85Q2$;
LIST X Y Z

displays the observations on the variables X, Y, Z over the period from the first quarter of 1980 to the second quarter of 1985, inclusive.

See Chapter 5 on how to print or save the displayed observations in a result file.

4.4.18 Command NONPARM

The command **NONPARM** provides non-parametric estimates of the density function of a set of n observations set out as $\mathbf{x} = (x_1, x_2, ..., x_n)$. The command has the general form

$$\textbf{NONPARM} \quad 1 \quad 2 \quad 3 \quad 4 \quad X \quad Z \quad h$$

where the integers 1 to 4 indicates the combination of kernel functions and band width to be used from the following choices:

$k = 1$: Gaussian kernel and Silverman rule of thumb band width
$k = 2$: Gaussian kernel and least squares cross-validation band width
$k = 3$: Epanechnikov kernel and Silverman rule of thumb band width
$k = 4$: Epanechnikov kernel and least squares cross-validation band width

We refer to Section (21.2) for the description of the above kernel functions and band width, and references to the literature. The vector Z contains the values at which the nonparametric function is to be evaluated and h gives the choice of the band width imposed by the user.

When h is set to zero the band width is selected automatically according to Silverman rule of thumb if $k = 1$ or 3, and by the least squares cross-validation procedure if $k = 2$ or 4. If h is set to a (small) positive number the badwith will be fixed at that value and only the choice of the kernel is governed by the specified value(s) of k. The optimum value of h under the cross-validation procedure is computed on a grid covering 101 values of h in the range $[0.25h_{silverman},\ 1.5h_{silverman}]$, where $h_{silverman}$ is given by (21.1).

The command computes the values of fitted density at n points uniformly distributed over the range $\left[x_{\min} - \hat{h}, x_{\max} + \hat{h}\right]$, where $x_{\min} = Min(x_1, x_2, ..., x_n)$, $x_{\max} = Max(x_1, x_2, ..., x_n)$, and $\hat{h}$ is the value of the bandwith (either specified by the user or automatically automatically).

In applications where n is relatively large (larger than 1000), the computation of the least squares cross-validation band width could take considerable amount of time. In such cases the user has the option of specifying

$$\textbf{NONPARM} \quad 1 \quad 3 \quad X \quad Z \quad h$$

or

$$\textbf{NONPARM} \quad 1 \quad X \quad Z \quad h$$

In the case of these commands the nonparametric densities will be computed only for the values of k specified.

Also it is not necessary to specify the variable Z. For example, the command can be issued as

$$\textbf{NONPARM} \quad 1 \quad X \quad h$$

In this case the density of x evaluated at n points uniformly distributed over the range $\left[x_{\min} - \hat{h}, x_{\max} + \hat{h}\right]$, where $\hat{h}$ is either the value of h specified by the user or automatically computed by *Microfit*.

If the user wishes to use an automatically computed band width the following simple form of the command can be used.

$$\textbf{NONPARM} \quad 1 \quad X$$

This will have the same effect as issuing the command

$$\textbf{NONPARM} \quad 1 \quad X \quad 0$$

Finally, if the integer values $k = 1, 2, 3, 4$ are dropped from the command, *Microfit* assumes that all the four options is to be computed. Namely the following commands will have the same effects

$$\textbf{NONPARM} \quad 1 \quad 2 \quad 3 \quad 4 \quad X \quad 0$$

$$\textbf{NONPARM} \quad X \quad 0$$

See Section (21.2) for the mathematical details and the references to the literature. See 10.15 for a tutorial lesson.

4.4.19 Command PCA

This command takes the form

$$\textbf{PCA} \quad X1 \quad X2 \quad .... \quad Xn \quad \& \quad Z1 \quad Z2 \quad ... \quad Zs$$

and computes the principal components of the variables $X1 \quad X2 \quad .... \quad Xn$ after filtering out the effects of $Z1 \quad Z2 \quad ... \quad Zs$. The filtering is carried out by regressing Xi, for each i, on $Z1 \quad Z2 \quad ... \quad Zs$ (if any specified), with the residuals from these regressions used in the principal component analysis. Typically Zi, $i = 1, 2, ..., s$, would include intercept or linear trends, although other variables can also be included amongst the Zi variables.

The **PCA** command generates the eigenvalues of the correlation matrix of filtered variables and the associated eigenvectors and principal components. The eigenvalues are reported in descending order together with the list of cumulative and percent cumulative eigenvalues, and the list of eigenvectors associated with non-zero eigenvalues. If you close the output screen, you are presented with the Principal Components Analysis Menu, where you can decide to plot eigenvalues and percent cumulative eigenvalues. You can also save a selected number of principal components as CSV or FIT files, or into workspace. Eigenvectors (or factor loadings) can also be saved, but only in a CSV file.

Notice that there is no need to standardize the variables before carrying out the principal component analysis. However, if you do not standardize them, it is advisable to set $Z1 = 1$. For technical details and references to the literature see Section 22.12. For a tutorial lesson see 10.16.

4.4.20 Command PLOT

This command produces a line graph of up to a maximum of 100 variables against time. You must specify at least one variable name. For example

SAMPLE 1950 1970;
PLOT X

produces a plot of variable X against time, over the period 1950-70.

PLOT $X1$ $X2$ $X3$ $X4$ $X5$ $X6$ $X7$ $X8$ $X9$ $X10$

produces a plot of the ten variables $X1, X2, ..., X10$ against time.

If you type

PLOT $X1\,X2\,X3$ & $Z1\,Z2\;Z3$

Microfit shows the Y axis of the first set of variables $X1\,X2\,X3$ on the left of the screen, and those of $Z1\,Z2\;Z3$ to the right of the screen. Specifically, left Y-axis reports the values of $X1\,X2\,X3$, while the right Y-axis reports the values of $Z1\,Z2\;Z3$.

In the Graph window you can specify a different period over which you wish to see the plots. Click the Start and Finish fields and scroll through the drop-down lists to specify the sample period and then press the 'Refresh graph over the above sample period' button.

See Section 5.2 on how to alter the display of graphs.

4.4.21 Command REORDER

This command enables a complete reordering of *all* the observations in the workspace according to the ordering of the variable that follows the command. For example,

REORDER X

produces a reordering of observations according to the ordering of the observations in variable X. This command is particularly useful when analyzing cross-sectional observations where the investigator wishes to carry out regression analysis on a sub-set of the observations. The exact nature of the particular sub-set of interest is defined by the ordering of the observations in the variable X.

As an example, suppose that the *undated* observations on the workspace refer to both male and female indexed by 0 and 1, respectively, stored in the variable SEX. Issuing the command

REORDER SEX

reorders the observations in the workspace so that observations referring to females appear first. The number of such observations is equal to **SUM**(SEX). See the **SUM** function above.

4.4.22 Command RESTORE

This command should be used after the **REORDER** command, and restores the ordering of the observations to their *original* state (before **REORDER** was used).

4.4.23 Command SAMPLE

This command can be used to change the sample period for subsequent data analysis in the data processing stage, but does not carry over to the other parts of the program. An example of the use of this command for undated observations is

SAMPLE 3 26

For annual observations

SAMPLE 1972 1986

For half-yearly data

SAMPLE $50H2$ $72H1$

For quarterly data

SAMPLE $75Q1$ $78Q2$

For monthly data

SAMPLE $70M1$ $80M11$

For daily data

SAMPLE 03-May-85 25-May-85

4.4.24 Command SCATTER

This command can be used to produce a scatter diagram of one variable against another. When issuing this command, you must specify exactly two variable names. For example,

SCATTER X Y

produces a scatter plot of the variable X against the variable Y.

See Section 5.2 for details concerning adding text, saving, retrieving and printing graphs.

4.4.25 Command SIM

This is a simulation command, and enables you to solve numerically any general univariate linear or non-linear difference equation. For example, to solve the non-linear difference equation

$$X(t) = 0.2X(t-1) + 0.7Log(X(t-2)) + Z$$

for $t = 3, 4, ..., 20$, with initial values $X(1) = 0.05$ and $X(2) = 0.10$, you need to issue the following commands

SAMPLE 1 1; $X = 0.05$;
SAMPLE 2 2; $X = 0.10$;
SAMPLE 3 20; **SIM** $X = 0.2 * X(-1) + 0.7 * LOG(X(-2)) + Z$;
SAMPLE 1 20; **PLOT** X

The first four commands in the above example set the initial values for X, which are used to simulate the values of X for observations 3, 4, ..., 20.

The following points should be borne in mind when using the **SIM** command:

1. When the **SIM** command is used, the values of the simulated variable will be overwritten. To avoid this problem, one possibility would be to create a new variable called, say $XNEW$, which contains the appropriate values, but is otherwise undefined for other periods. The **SIM** command can then be applied to $XNEW$ over the sample period for which $XNEW$ is undefined. A typical example of this procedure would be (assuming that the specified sample period runs from 1950 to 1980)

   SAMPLE 1950 1950; $XNEW = 0.05$;
   SAMPLE 1951 1980;
   SIM $XNEW = 4 * XNEW(-1) * (1 - XNEW(-1))$

   The above will solve the well-known chaotic bifurcation equation

   $$X_t = 4X_{t-1}(1 - X_{t-1})$$

   starting with the initial value $X_{1950} = 0.05$ over the period 1951-1980.

2. Choose your sample period carefully and make sure that well-defined initial values exist for the simulation, otherwise all the values of the variable being simulated will be set to missing.

3. In the case of unstable difference equations, the use of the **SIM** command may cause an overflow. When the value of the simulated variable exceeds 10 to the power 50, to prevent the program from crashing, all the subsequent values will be set to missing.

4.4.26 Command SIMB

This is a simulation command which allows you to solve numerically any general univariate linear or non-linear difference equation involving lead (not lagged) values of the left-hand side variable. The difference equation is solved backwards over the specified sample period. For example, to solve the linear difference equation:

$$X(t) = 1.2 * X(t+1) + Z(t), \quad \text{for} \quad t = 20, 19, ..., 1$$

with a terminal value of 30.5 at observation 20, the following commands should be issued:

SAMPLE 20 20; $X = 30.5$;
SAMPLE 1 19; **SIMB** $X = 1.2 * X(+1) + Z$;
SAMPLE 1 20; **PLOT** X

For more information, see the description of the **SIM** command in Section 4.4.25.

4.4.27 Command SPECTRUM

This command, when followed by a variable name, displays the estimates of the standardized spectral density function of the variable and their estimated standard errors using Bartlett, Tukey and Parzen lag windows as in the following example:

SAMPLE 1 120; **SPECTRUM** X

The window size will be taken to be twice the square root of the number of specified observations. If you have a graphics adaptor, the plot of the different spectral density functions and the associated standard error bands will also be displayed.

The **SPECTRUM** command can also be used to estimate the spectrum for a window size specified by the user. The desired window size should be specified in parentheses after the variable. For example,

SPECTRUM $X(12)$

See Section 5.2 on how to alter the display of graphs.

The algorithms used to compute the different estimates of the spectral density and the relevant references to the literature are given in Section 21.3.

4.4.28 Command TITLE

This command generates a list of the names of your variables, together with their description, if any. If you type

TITLE

the variable names and the descriptions (if any) of all your variables will be displayed.

4.4.29 Command XPLOT

This command can be used to plot up to a maximum of 100 variables against another variable. When issuing this command you must specify at least two variable names. For example,

XPLOT X Y

produces a plot of the variable X against the variable Y.

XPLOT $X1$ $X2$ $X3$ $X4$ $X5$ $X6$ Y

produces a plot of the variables $X1, X2, X3, X4, X5$, and $X6$ against the variable Y.

See Section 5.2 for details concerning how to add text, save, retrieve and print graphs.

Chapter 5

Printing/Saving Results and Graphs

Output from *Microfit* appears on your screen in the form of texts and graphs. These can be output to a variety of printers attached to your PC, or can be saved as a file to be printed at a later stage or for importation into Word-processing packages such as Microsoft Word or Scientific Word. *Microfit 5.0* also allows to save regression results in equation format, suitable for use with modelling or simulation packages.

5.1 Result screens

Text output of *Microfit* is displayed inside a result window. You can scroll through the window in the usual way using the mouse and scroll bar and/or the PgUp/PgDn, Ctrl+Home/End keys.

Use standard Windows editing functions to edit the contents of the results window if you wish. To copy text to the clipboard, highlight the text you want by clicking and dragging with mouse, and click the 

button.

To edit the font of text displayed in the window, highlight the text you want to change, and click the 

button. A standard Windows font dialogue is displayed. Make your choices in the usual way and click 

.

To exit the result window, click 

.

5.1.1 On-line printing of results

The content of each result screen can be printed, separately, by clicking on the 

icon. Make any choice about the number of copies and so on. and click .

5.1.2 Saving results

When saving results you can save them either in 'report file' format, or in 'model file' format. The former saves the content of the result screen in a text (ASCII) file for use with word-processing editors. The latter saves only the estimated coefficients in the form of a linear/non-linear regression equation.

Saving results to a report file. To save your results in a result file, click the 

button and select the 'Save to Existing/New Result File' option. When the 'Save as' dialogue appears, specify a filename, drive and directory. Result files in *Microfit* are given the extension .OUT, and if you have any such files in your default directory you should see them in the list.

Suppose now that you have already opened a result file called RESULT.OUT. To add the results displayed on the screen to this existing report (or output) file, click the 

button and select the 'Add to Current Result File' option, or choose the 'Save to Existing/New Result File' option and select the RESULT.OUT file.

To view the contents of a result file, use the 'View a File' option from the File Menu. The result files created in *Microfit* are in ASCII (text) format and can be edited/printed using text editing or word-processing packages.

Saving equation specification to a model or an equation file. This function applies only when the displayed results contain coefficients of an estimated relation/model.

To save your results in a model or an equation file, click the 

button and select the option 'Save to Existing/New Model File'. A menu appears giving you a choice of model file types; choose the type most suited to the package into which you want to import the file. Four different model file formats are allowed:

1. *Microfit* model file format. This is internal to *Microfit* and may not be compatible with model specification formats used by other packages.
2. Winsolve model file format. This is the format used in the model solver program, Winsolve, developed by Richard Pierse of the Department of Economics, University of Surrey, England.
3. National Institute model file format. This is the equation format used by the National Institute of Economic and Social Research (NIESR), London.
4. London Business School model file format. This is the equation format currently used in the London Business School (LBS) forecasting model.

5.2 Print/save/retrieve graphs

The graphic facilities in *Microfit 5.0* are considerably enhanced in comparison with earlier versions of the package. In order to take full advantage of these facilities you need at least a laser-jet or a PostScript printer.

5.2.1 Altering the display of graphs

The default graph display may be edited using the Chart Control facility. Select the chart options from the Edit Menu or click the 

button below the graph to access it. Chart Control contains numerous options for adjusting the background, colour, axes, style, title, and so on, of your graph. Each option (such as 'Background') has its own property page; click the appropriate page tab to view it. Page tabs contain one or more inner tabs that group related properties together. Some tabs also contain a list that selects a specific object to edit. The property changes are immediately applied unless the option 'IsBatched' in the Control page tab is set. When this option is set click the Apply button to display property changes. To exit Graph Control without implementing our changes, click CANCEL. To minimize the graph window and reduce it to an icon, click the MINIMIZE button. To close the window click CLOSE.

The most common functions you may want to use with *Microfit* are described here.

For more information, visit the web site
http://helpcentral.componentone.com.

Titles: A graph can have two titles, called the header and footer. You can use this page to set the text alignment, positioning, colours, border style and font used for the header and/or footer. For example, to change or insert a title select Titles from the Edit and then select Label in the '2D Chart Control Properties' screen that follows. To edit the 'Title at the Header' select Header in the left window insert. You can also change the location of the titles on the graph, add Border to them or change their fonts.

Legends. In this screen you can decide the positioning, border, colours and font used for the legend. Use the Anchor property in the General tab to specify where to position the legend relative to the ChartArea. When the IsDefault property is used, the chart automatically positions the legend. You can also remove the legends by switching off IsShowing. Use the Title property to specify the legend title.

Colours. Using this option you can set the line thickness, pattern and colour for each line in the graph, separately. For example, to change the pattern/colour of the first line in the graph select Colours under Edit and in the '2D Chart Control Properties' select Style1 under ChartGroup 1 then change Pattern, Width and Colour using the right panel of this screen insert. To do the same for the second line in the graph select Style 2 in ChartGroup 1 and repeat the process.

Variable names. This option allows you to change the variable names of the displayed graphs. For example, if you have plotted the variables Y and P and you wish to change their names in the graph select Variable Names under Edit and in the '2D Chart Control Properties' screen select the variable name that you wish to change and then type in the new name in the text box provided.

Other options of the graph menu. To access the other options of '2D Chart Control Properties' select 'All Chart Options' from the Edit Menu. For example you can change the Y (vertical) and X (horizontal) axes, decide whether or not to display them, show them in logarithmic scales or reverse the way they are displayed.

Axes. This page allows you to modify the annotation method, the scale and the style of axes and to give a title to axes. The Scale tab in this page allows you to frame the graph at specific data values (using 'Data Min' and 'Data Max' options) and/or at specific axis values (using 'Max' and 'Min' options), and to control the placement of the origin point. Use the TitleRotation property to rotate the axis title to either 90 or 270 degrees counterclockwise. In the Axis/Grid Lines tab you can change the properties of axes lines or display a grid on a graph. Use the tick length property to choose the length of the tick marks on the axis, or use the IsStyleDefault axis line property to allow the graph to set it automatically. Use the gridlines Spacing property to change the grid spacing for an axis. In the AxisStyle and GridStyle tabs you can control the line pattern, thickness, and colour properties of the axis lines and ticks and of the grid, respectively.

View 3D. Graphs can be enhanced with a 3D effect. Use the Depth property to set the visual depth of the 3D effect, as a percentage of the chart width. The maximum value is 500. Use the Elevation property to set the distance above the X-axis for the 3D effect, in degrees. This can be from -45 to 45 degrees. Use the Rotation property to set the distance right of the Y-axis for the 3D effect, in degrees. This cannot be higher than 45 or lower than -45. Use the Shading property to set the shading method for the 3D areas of the chart.

5.2.2 Printing graphs

To obtain a hard copy of the displayed graph on the default printer click the 

button. Make any choices about the number of copies and so on, and click 

.

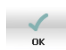
5.2.3 Saving graphs

A displayed graph can be saved in a variety of formats: as an Olectra Chart (OC2), as a Bitmap (BMP), Windows metafile (WMF), Enhanced metafile (EMF), JPEG or Portable Network Graphics (PNG) file.

The Olectra Chart format is useful if you want to load the graph into *Microfit* at a later stage for further editing (see Section 5.2.4). To save the graph as an OC2 file, select the 'Save the Chart (Olectra Chart Format)' option from the File Menu in the Graph window. Enter the filename and location for your file and click 

.

To save the graph in BMP, WMF, EMF, JPEG or PNG format, select the 'Save as' option from the File Menu or click the 

button. Choose the file type you want from the drop-down list and enter the filename and location for your file before clicking 

. The graph's image only will be saved.

5.2.4 Retrieval of graphic files

To retrieve a graphic file, click the 

button at the base of the screen. In the Open dialogue, select Graph files from the List of types box, find the location and name of the file you want, and click 

. Only graphs saved in Olectra Chart (OC2) format can be loaded into *Microfit*.

5.2.5 Capturing graphs onto the clipboard

It is possible to capture the displayed graph onto the Windows clipboard. Click the 

button. From the clipboard the graph may be pasted into another application in the usual way, using the special Paste option available in word-processing packages such as Microsoft Word or Scientific Word.

5.3 Exercises using graphs

5.3.1 Exercise 5.1

Carry out Lesson 10.5, copy the plot of C and Y to the clipboard and then past the graph into Microsoft Word or Scientific Word. When you see the graph on the screen type some text around it, resize and print.

Note that any text that you wish to add to the graph must be done in *Microfit*. Once a graph is imported into a word-processing package you cannot add text inside the graphic box. You can only give it a title.

5.3.2 Exercise 5.2

Carry out Lessons 10.9 and 10.10, and save the results from both lessons in *one* file. Add titles and other descriptions to this result file by editing it, and then print the file.

Part III

Estimation Menus

Chapter 6

Single-Equation Options

In this chapter we show how *Microfit* can be used to estimate a large number of single-equation econometric models, compute diagnostic statistics for them, carry out tests of linear or non-linear restrictions on their parameters, and use them in forecasting. First we review briefly the classical linear regression model and the likelihood approach that underlie the various estimation options in *Microfit*. The more technical details of the econometric methods and the computational algorithms used are given in Chapter 21, where further references to the literature can also be found.

6.1 The classical normal linear regression model

The econometric model underlying the linear regression estimation options in *Microfit* is the classical linear regression model. This model assumes that the relationship between y_t (the dependent variable) and x_{1t}, x_{2t}, ..., x_k (the k regressors) is linear

$$y_t = \sum_{i=1}^{k} \beta_i x_{it} + u_t, \qquad t = 1, 2, ..., n \tag{6.1}$$

where u_t's are unobserved 'disturbance' or 'error' terms, subject to the following assumptions.

A1 Zero mean assumption. The disturbances u_t have zero means

$$E(u_t) = 0, \quad t = 1, 2, ..., n$$

A2 Homoscedasticity assumption. The disturbances u_t have a constant conditional variance

$$V(u_t \,|x_{1t}, x_{2t}, ..., x_{kt}) = \sigma^2, \quad \text{for all} \quad t$$

A3 Non-autocorrelated error assumption. The disturbances u_t are serially uncorrelated

$$Cov(u_t, u_s) = E(u_t u_s) = 0, \qquad \text{for all} \quad t \neq s$$

A4 Orthogonality assumption. The disturbances u_t and the regressors x_{1t}, x_{2t}, ... x_{kt} are uncorrelated

$$E(u_t\,|x_{1t}, x_{2t}, ..., x_{kt}) = 0, \quad \text{for all} \quad t$$

A5 Normality assumption. The disturbances u_t are normally distributed.

Adding the fifth assumption to the classical model yields the classical linear normal regression model. The latter model can also be derived using the *joint* distribution of y_t, x_{1t}, x_{2t}, ..., x_{kt}, and by assuming that this distribution is a multivariate *normal* with constant means, variances and covariances. In this setting the regression of y_t on x_{1t}, x_{2t}, ..., x_{kt}, is defined as the mathematical expectation of y_t conditional on the realized values of the regressors, will be linear in the regressors. The linearity of the regression equation follows from the joint normality assumption and need not hold if this assumption is relaxed. Both of the above interpretations of the classical normal regression model have been used in the literature (see, for example, Spanos 1989).

In time-series analysis, the critical assumptions are A3 and A4. Assumption A3 is particularly important when the regression equation contains lagged values of the dependent variable, namely y_{t-1}, y_{t-2},... . However, even if lagged values of y_ts are not included among the regressors, the breakdown of assumption A3 can lead to misleading inferences; a problem recognized as early as the 1920s by Yule (1926), and known in the econometrics time-series literature as the spurious regression problem.[1] The orthogonality assumption, A4, allows the empirical analysis of the relationship between y_t and x_{1t}, x_{2t},...,x_{kt} to be carried out without fully specifying the stochastic processes generating the regressors, or the 'forcing' variables. Assumption A1 is implied by A4, if a vector of ones is included among the regressors. It is therefore important that an intercept is always included in the regression model, unless it is found to be statistically insignificant. Assumption A2 specifies that u_ts have constant variances both conditionally and unconditionally. The assumption that the error variances are constant unconditionally is likely to be violated when dealing with cross-sectional regressions. The assumption that the conditional variance of u_t is constant is often violated in analysis of financial and macroeconomic time-series, such as exchange rates, stock returns and interest rates. The normality assumption A5 is important in small samples, but is not generally required when the sample under consideration is large enough. All the various departures from the classical normal regression model mentioned here can be analysed using the options that are available in *Microfit.*

6.1.1 Testing the assumptions of the classical model

Microfit enables the user to test the assumptions that underlie the classical model in a simple and straightforward manner. This type of diagnostic testing is an essential component of any applied econometric research. However, it is important that the outcomes of such diagnostic testing exercises are properly interpreted and acted upon.

[1] Champernowne (1960), and Granger and Newbold (1974) provide Monte Carlo evidence of the spurious regression problem, and Phillips (1986) establishes a number of theoretical results.

Guidelines

You may find the following broad guidelines useful when working with *Microfit*:

1. Rejection of an hypothesis against an alternative does not necessarily imply that the alternative hypothesis is acceptable or that it should be necessarily adopted. Rejection of a given hypothesis could be due to a number of different interlocking factors, and it is therefore important that a variety of nested and non-nested alternative explanations are considered before a firm conclusion, as to the appropriate choice of the alternative hypothesis, is reached. For example, when assumption A3 (the non-autocorrelated error assumption) is rejected it may be due to any one or a mixture of the following model mis-specifications: omitted variables, structural change, mis-specified dynamics, or aggregation across heterogenous groups. Rejection of the normality assumption may be due to the presence of outliers or non-linearities. Rejection of the orthogonality assumption could arise because of simultaneity, expectational effects, omitted variables and/or mis-specified dynamics.

2. A regression equation that passes all the diagnostic tests generated by *Microfit* is not necessarily a statistically adequate model, and should not be regarded as the 'true' data-generating process! It is quite possible for two rival (or non-nested) models to pass all the diagnostic tests produced by *Microfit*, but yet one of the models could be rejected by the other and not vice versa.[2] The non-nested test options in *Microfit* can be used to deal with such eventualities. Even then there could be other possible models that may not have been considered or thought out by the investigator. A satisfactory econometric model should satisfy a number of quantitative and qualitative criteria, and can at best represent a reasonable approximation of one or more aspects of the reality that the investigator is interested in analyzing. Pesaran and Smith (1985) summarize these different criteria under the heading 'relevance', 'consistency' and 'adequacy'.

3. The t-tests on individual regression coefficients should be carried out with great care, particularly when the regressions exhibit a high degree of collinearity. It is good practice to combine the t-tests on individual coefficients with F-tests of *joint* restrictions on the coefficients. It is important that the results of the individual t -tests (also known as separate-induced tests) and the joint tests are not in conflict. Otherwise, inferences based on individual t-tests can be misleading. For a demonstration of this point see Lesson 10.4 on the multicollinearity problem.

6.1.2 Estimation of the classical linear regression model

Under the classical assumptions (A1 to A4), the estimation of the regression coefficients, β_1, β_2,..., β_k is carried out by minimizing the sum of squares of the errors, u_t, with respect to β_1, β_2,...,β_k. Writing (6.1) in matrix notations we have

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{6.2}$$

[2] As an example, see the comparative empirical analysis of the Keynesian and the Neo-Classical explanations of US unemployment in Pesaran (1982b), Pesaran (1988b) and Rush and Waldo (1988).

where

$$\underset{n\times 1}{\mathbf{y}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \underset{n\times 1}{\mathbf{u}} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

$$\underset{n\times k}{\mathbf{X}} = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{k1} \\ x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & \dots & x_{kn} \end{pmatrix}$$

and $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_k)'$ is a $k \times 1$ vector of unknown coefficients. The sum of squares of the errors can now be written in matrix notations as

$$Q(\boldsymbol{\beta}) = \sum_{t=1}^{n} u_t^2 = \mathbf{u}'\mathbf{u} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{6.3}$$

The first-order conditions for minimization of $Q(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ are given by

$$\frac{\partial Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\mathbf{X}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{6.4}$$

The Ordinary Least Squares (OLS) estimator of $\boldsymbol{\beta}$ is obtained by solving the normal equations in $\widehat{\boldsymbol{\beta}}_{OLS}$

$$\mathbf{X}'(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_{OLS}) = \mathbf{0}$$

For these equations to have a unique solution it is necessary that $\mathbf{X}'\mathbf{X}$ has a unique inverse, $(\mathbf{X}'\mathbf{X})^{-1}$. When $\mathbf{X}$ is rank deficient, any one of the columns of $\mathbf{X}$ can be written as exact linear combinations of the other columns, and it is said that the regressors are *perfectly multicollinear*. In what follows we make the following assumption:

A6: The observation matrix $\mathbf{X}$ has a full column rank: $Rank(\mathbf{X}) = k$.

Under this assumption the OLS estimator of $\boldsymbol{\beta}$ is given by

$$\widehat{\boldsymbol{\beta}}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{6.5}$$

Under the classical assumptions (A1 to A4), $\widehat{\boldsymbol{\beta}}_{OLS}$ is unbiased ($E(\widehat{\boldsymbol{\beta}}_{OLS}) = \boldsymbol{\beta}$), and among the class of linear unbiased estimators it has the least variance. (This result is known as the Gauss Markov Theorem.).

Under the normality assumption A5, the Maximum Likelihood (ML) estimator of $\boldsymbol{\beta}$ is identical to the OLS estimator, and the log-likelihood function is given by

$$L(\boldsymbol{\beta}, \sigma^2) = -\tfrac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{6.6}$$

where σ^2 denotes the variance of u_t. The ML estimator of σ^2 is given by

$$\tilde{\sigma}^2 = n^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_{OLS})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_{OLS}) \tag{6.7}$$

and is biased. In fact

$$E(\tilde{\sigma}^2) = \left(\frac{n-k}{n}\right)\sigma^2 \tag{6.8}$$

The unbiased estimator of σ^2, which we denote by $\hat{\sigma}^2$, is defined by

$$\hat{\sigma}^2 = (n-k)^{-1}\,(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_{OLS})'\,(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_{OLS}) \tag{6.9}$$

The variance matrix of $\widehat{\boldsymbol{\beta}}_{OLS}$, together with a number of useful summary statistics for regression analysis, are given in Section 21.6.1.

6.1.3 Testing zero restrictions and reporting probability values

Consider the problem of testing the 'null' hypothesis that

$$H_0 : \beta_i = \beta_i^0$$

against

$$H_1 : \beta_i \neq \beta_i^0$$

where β_i is the ith element of $\boldsymbol{\beta}$ in (6.2). The relevant test statistic is given by the t-ratio

$$t_i = \frac{\widehat{\beta}_i - \beta_i^0}{\sqrt{\widehat{V}(\widehat{\beta}_i)}} \tag{6.10}$$

where $\widehat{\beta}_i$ is the ith element of $\widehat{\boldsymbol{\beta}}_{OLS}$, and $\widehat{V}(\widehat{\beta}_i)$ is the estimator of the variance of $\widehat{\beta}_i$ and is given by the ith diagonal element of the variance matrix defined by (21.6). Since the alternative hypothesis, H_1, is two-sided, the absolute value of t_i should be compared with the appropriate critical value of the Student-t distribution with $n-k$ degrees of freedom. *Microfit* reports the probability of falsely rejecting the null hypothesis that $\beta_i = 0$ against $\beta_i \neq 0$, in square brackets next to the t-ratios. These probability values are valid under assumptions A1 to A5 for two-sided tests, and provide an indication of the level of significance of the test. For example, if the probability value reported for β_i is equal to 0.025, it means that the probability of falsely rejecting $\beta_i = 0$ is at most equal to 0.025. Therefore, the null hypothesis that $\beta_i = 0$ against $\beta_i \neq 0$ is rejected at the 2.5 per cent significance level. The probability values are applicable even if the normality assumption is violated, provided that the sample is large enough.

6.2 The maximum likelihood approach

Many of the estimation options in *Microfit* compute estimates of the regression coefficients when one or more of the classical assumptions are violated. For example, the AR and the

MA options, discussed in Sections 6.8-6.11 below, compute estimates of $\boldsymbol{\beta}$ under a variety of assumptions concerning the autocorrelation patterns in the disturbances. To deal with such departures, *Microfit* makes use of two general principles: the Likelihood Principle, and the Instrumental Variables Method which is a special case of the Generalized Method of Moments (GMM). Here we briefly review the likelihood principle.

Let $L(\boldsymbol{\theta})$ be the likelihood function of the $k \times 1$ vector of unknown parameters, $\boldsymbol{\theta}$, associated with the joint probability distribution of $\mathbf{y} = (y_1, y_2, ..., y_n)'$, conditional (possibly) on a set of predetermined variables or regressors. Assume that $L(\boldsymbol{\theta})$ is twice differentiable and satisfies a number of regularity conditions. See, for example, Chapter 8 in Davidson and MacKinnon (1993).

The Maximum Likelihood (ML) estimator of $\boldsymbol{\theta}$ is that value of $\boldsymbol{\theta}$ which globally maximizes $L(\boldsymbol{\theta})$. Let $\widetilde{\boldsymbol{\theta}}$ be the ML estimator of $\boldsymbol{\theta}$, then it must also satisfy the first-order condition

$$\left.\frac{\partial \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}} = \mathbf{0}$$

6.2.1 Newton-Raphson algorithm

The computation of the ML estimators in *Microfit* is generally carried out using the Newton-Raphson algorithm. Denote the estimator of $\boldsymbol{\theta}$ in the jth iteration by $\widetilde{\boldsymbol{\theta}}_{(j-1)}$, then the iterative algorithm used is given by

$$\widetilde{\boldsymbol{\theta}}_{(j)} = \widetilde{\boldsymbol{\theta}}_{(j-1)} + d\left[-\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right]^{-1}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_{(j-1)}} \left[\frac{\partial \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_{(j-1)}} \tag{6.11}$$

where d is a scalar 'dumping factor'. In cases where convergence of the numerical algorithm may be problematic, *Microfit* allows the user to start the iterations with different choices for the initial estimates, $\widetilde{\theta}_{(0)}$, and the value of the damping factor in the range $[0.01 - 2.00]$. The iterations are terminated if

$$\sum_{i=1}^{k} \left|\widetilde{\theta}_{i,(j)} - \widetilde{\theta}_{i,(j-1)}\right| < k\epsilon$$

where $\widetilde{\theta}_{i,(j)}$ is the ith element of $\widetilde{\boldsymbol{\theta}}_{(j)}$, and ϵ is a small positive number usually set equal to $1/10,000$. In some cases the program also checks to ensure that at termination the maximized value of the log-likelihood function is at least as large as the log-likelihood values obtained throughout the iterations.

6.2.2 Properties of maximum likelihood estimators

The optimum properties of ML estimators are asymptotic; that is, they are valid in large samples. Assuming that certain regularity conditions are satisfied, and in particular $\mathbf{y}_t$, is a stationary process, then $\widetilde{\boldsymbol{\theta}}$, the ML estimator of $\boldsymbol{\theta}$, has the following properties:[3]

[3] Many of these properties continue to hold even if the stationary assumption is relaxed. For general results in the case of the ML estimation of models with unit root processes see Chapters 7 and 22.

1. $\widetilde{\boldsymbol{\theta}}$ is a consistent estimator of $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_k)'$ that is

$$\lim_{n\to\infty} \Pr\left\{\left|\widetilde{\theta}_i - \theta_{i0}\right| < \epsilon\right\} = 1, \qquad \text{for} \quad i = 1, 2, ..., k$$

   where θ_{i0} is the true value of θ_i and $\epsilon(> 0)$ is a small positive number.

2. Asymptotically (as $n \to \infty$), $\sqrt{n}\left(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)$ has a multivariate normal distribution with zero means and the variance matrix $\sum$, where

$$\sum\nolimits^{-1} = E\left\{-\frac{1}{n} \cdot \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right\} \qquad \text{or} \qquad \sum\nolimits^{-1} = \operatorname*{plim}_{n\to\infty}\left\{-\frac{1}{n} \cdot \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right\}$$

3. $\widetilde{\boldsymbol{\theta}}$ attains the Cramer-Rao lower bound asymptotically.

4. $\widetilde{\boldsymbol{\theta}}$ is an asymptotically unbiased estimator of $\boldsymbol{\theta}$, that is

$$\lim_{n\to\infty} E(\widetilde{\boldsymbol{\theta}}) = \boldsymbol{\theta}_0$$

5. $\widetilde{\boldsymbol{\theta}}$ is an asymptotically efficient estimator. That is, $\widetilde{\boldsymbol{\theta}}$ has the lowest asymptotic variance in the class of all asymptotically unbiased estimators.

6.2.3 Likelihood-based tests

There are three main likelihood-based test procedures that are commonly used in econometrics for testing linear or non-linear parametric restrictions on a maintained model. They are:

1. The Likelihood Ratio (LR) approach.
2. The Lagrange Multiplier (LM) approach.
3. The Wald (W) approach.

All these three procedures yield asymptotically valid tests, in the sense that they will have the correct size (the type I error) and possess certain optimal power properties in large samples. They are asymptotically equivalent, although they can lead to different results in small samples. The choice between them is often made on the basis of computational simplicity and ease of use.

The Likelihood Ratio test procedure

Suppose that the hypothesis of interest to be tested can be written as a set of r independent restrictions (linear and/or non-linear) on $\boldsymbol{\theta}$. Denote these r restrictions by[4]

$$\begin{array}{rcl}\phi_1(\theta_1, \theta_2, ..., \theta_k) & = & 0 \\ \phi_2(\theta_1, \theta_2, ..., \theta_k) & = & 0 \\ \vdots & & \vdots \\ \phi_r(\theta_1, \theta_2, ..., \theta_k) & = & 0\end{array}$$

which can be written compactly in vector notations as

$$H_0 : \boldsymbol{\phi}(\boldsymbol{\theta}) = 0$$

where $\boldsymbol{\phi}(\cdot)$ is an $r \times 1$ twice differentiable function of the $k \times 1$ parameter vector, $\boldsymbol{\theta}$. Consider the two-sided alternative hypothesis

$$H_1 : \boldsymbol{\phi}(\boldsymbol{\theta}) \neq 0$$

The log-likelihood ratio (LR) statistic for testing H_0 against H_1 is defined by

$$LR = 2\left\{\log\left[L(\widetilde{\boldsymbol{\theta}}_U)\right] - \log\left[L(\widetilde{\boldsymbol{\theta}}_R)\right]\right\} \tag{6.12}$$

where $\widetilde{\boldsymbol{\theta}}_U$ is the *unrestricted* ML estimator of $\boldsymbol{\theta}$, and $\widetilde{\boldsymbol{\theta}}_R$ is the *restricted* ML estimator of $\boldsymbol{\theta}$. The latter is computed by maximizing $L(\boldsymbol{\theta})$ subject to the r restrictions $\boldsymbol{\phi}(\boldsymbol{\theta}) = 0$. Under the null hypothesis, H_0, and assuming that certain regularity conditions are met, the statistic LR is asymptotically distributed as a chi-squared variate with r degrees of freedom. The hypothesis H_0 is then rejected if the log-likelihood ratio statistic, LR, is larger than the relevant critical value of the chi-squared distribution.

The LR approach requires that the maintained model, characterized by the likelihood function $L(\theta)$, be estimated both under the null and the alternative hypotheses. The other two likelihood-based approaches to be presented below require the estimation of the maintained model either under the null or under the alternative hypothesis, but not under both hypotheses.

The Lagrange Multiplier test procedure

The Lagrange Multiplier (LM) procedure uses the restricted estimators, $\widetilde{\boldsymbol{\theta}}_R$, and requires the computation of the following statistic:

$$LM = \left\{\frac{\partial \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right\}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_R} \left\{-\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right\}^{-1}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_R} \left\{\frac{\partial \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right\}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_R} \tag{6.13}$$

[4] The assumption that these restrictions are independent requires that the $r \times k$ matrix of the derivatives $\partial\boldsymbol{\phi}/\partial\boldsymbol{\theta}'$ has a full rank, namely that $Rank(\partial\boldsymbol{\phi}/\partial\boldsymbol{\theta}') = r$.

where $\left\{\frac{\partial \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right\}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_R}$ and $\left\{\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right\}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_R}$ are the first and the second derivatives of the log-likelihood function which are evaluated at $\widetilde{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}}_R$, the *restricted* estimator of $\boldsymbol{\theta}$. Recall that $\widetilde{\boldsymbol{\theta}}_R$ is computed under the null hypothesis, H_0, which defines the set of restrictions to be tested. The LM test was originally proposed by Rao (1948) and is also referred to as Rao's score test, or simply the 'score test'.

The Wald test procedure

The Wald (W) test makes use of the unrestricted estimators, $\widetilde{\boldsymbol{\theta}}_U$, and is defined by

$$W = \boldsymbol{\phi}'(\widetilde{\boldsymbol{\theta}}_U)\left\{\widehat{V}\left[\boldsymbol{\phi}(\widetilde{\boldsymbol{\theta}}_U)\right]\right\}^{-1}\boldsymbol{\phi}(\widetilde{\boldsymbol{\theta}}_U) \tag{6.14}$$

where $\widehat{V}\left[\boldsymbol{\phi}(\widetilde{\boldsymbol{\theta}}_U)\right]$ is the variance of $\boldsymbol{\phi}(\widetilde{\boldsymbol{\theta}}_U)$ and can be estimated consistently by

$$\widehat{V}\left[\boldsymbol{\phi}(\widetilde{\boldsymbol{\theta}}_U)\right] = \left\{\frac{\partial \boldsymbol{\phi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right\}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_U}\left\{-\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right\}^{-1}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_U}\left\{\frac{\partial \boldsymbol{\phi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right\}_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_U} \tag{6.15}$$

Asymptotically (namely as the sample size, n, is allowed to increase without a bound), all the three test procedures are equivalent. Like the LR statistic, under the null hypothesis, the LM and the W statistics are asymptotically distributed as chi-squared variates with r degrees of freedom. We can write

$$LR \overset{a}{\sim} LM \overset{a}{\sim} W$$

where $\overset{a}{\sim}$ denotes 'asymptotic equivalence' in distribution functions.

Other versions of the LM and the W statistics which replace

$$\left\{\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right\}$$

in (6.13) and (6.15) by a consistent estimate of

$$n \operatorname*{plim}_{n\to\infty}\left\{n^{-1}\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right\}$$

are also used in *Microfit*. This would not affect the asymptotic distribution of the test statistics, but in some cases could simplify the computation of the statistics.

The various applications of the Likelihood Approach to single equation econometric models are reviewed in Chapter 18.

6.3 Estimation menus in *Microfit*

Microfit's gateway to econometric analysis consists of the Single Equation Estimation Menu (shortened to Multivariate Menu on the menu bar), the System Estimation Menu (shortened to Multivariate Menu on the menu bar), and the Volatility Modelling Menu. The Univariate Menu opens the Single Equation Estimation Menu (see Section 6.4) which provides a large number of options for estimation of linear and non-linear single equation models.

The Multivariate Menu opens the System Estimation Menu (see Section 7.3) which allows you to estimate unrestricted vector autoregressive (VAR), cointegrating VAR models with exactly identifying and over-identifying restrictions on the long-run relations, cointegrating $VARX$, and system of seemingly unrelated equations ($SURE$), with and without restrictions.

The Volatility Modelling Menu allows you to estimate univariate and multivariate $GARCH$ models (see Chapter 8).

Alternatively, use the 

, 

or 

buttons to move to Single Equation Estimation window, the System Estimation window, or the Volatility Modelling window respectively. The window opens with the last menu option chosen (or the first option in the menu) selected by default.

Before any of these estimation options are used, it is important that the data are correctly entered, and that all the variables to be included in the regression equation, such as the intercept (the constant term), time trends, seasonal dummies, or transformations of your existing variables (for example, their first differences or logarithms). See the data processing functions described in Chapter 4.

6.4 Single Equation Estimation Menu

The Single Equation Estimation Menu (abbreviated to Univariate Menu in the menu bar) contains the following options:

1. Linear Regression Menu.
2. Recursive Linear Regression Menu.
3. Rolling Linear Regression Menu.
4. Non-linear Regression Menu.
5. Phillips-Hansen Estimation Menu.
6. $ARDL$ Approach to Cointegration.
7. Logit and Probit Models.

In *Microfit* each of these options is regarded as a menu on its own.

When you choose any of these options, or their submenu options, you will be asked to enter the specification of your econometric model in the editor window on the screen.

Option 1 allows you to estimate linear regression models by a variety of methods: ordinary least squares (OLS), instrumental variables (IV) or two-stage least squares ($TSLS$), maximum likelihood (ML) estimates for regression models with serially correlated errors

(*AR*, *CO*, *MA*), and *IV* estimates of regression models with serially correlated errors (*IV*/*AR* and *IV*/*MA* options)

Options 2 and 3 compute recursive and rolling regressions estimated by the *OLS* and the *IV* methods.

Option 4 enables you to estimate non-linear regression equations by the least squares or the two-stage least squares methods.

Option 5 can be used to obtain fully-modified *OLS* (*FM-OLS*) estimators of a single cointegrating relation proposed by Phillips and Hansen (1990).

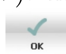
Option 6 implements the Autoregressive-Distributed Lag (*ARDL*) approach to estimation of a single long-run relationship advanced by Pesaran and Shin (1999), with automatic order selection using any one of the four model selection criteria, namely $\bar{R}^2$, the Akaike information criterion (*AIC*), the Schwartz Bayesian criterion (*SBC*), and the Hannan and Quinn criterion (*HQC*).

Option 7 can be used to estimate univariate binary quantitative response models for normal and logistic probability distributions (namely, the Probit and Logit models).

6.5 The Linear Regression Menu

This is the main menu for estimation of single equation linear regression models. It contains the following options

1. Ordinary Least Squares.
2. Gen. Instr. Var. Method.
3. *AR* Errors (Exact ML) J<=12.
4. *AR* Errors (Cochran Orcutt) J<=12.
5. *AR* Errors (Gauss-Newton).
6. *IV* with *AR* Errors (Gauss-Newton).
7. *MA* errors (Exact *ML*) J<=12.
8. *IV* with *MA* Errors J<=12.

The options in this menu can be used to compute estimates of a linear regression equation under a number of different stochastic specifications of the disturbances. To start your calculations once they have been set up, click . All the options in this menu assume that the observation matrix of the regressors has a full rank (that is, that Assumption A6 is satisfied and the specified regressors are not *perfectly multicollinear*). If this condition is not satisfied *Microfit* gives an error message and invites you to click .

To access the Linear Regression Estimation Menu choose option 1 in the Single Equation Estimation Menu (see Section 6.4), and then follow the instructions below to specify your regression equation and the estimation period.

6.5.1 Specification of a linear regression equation

When you choose options 1 to 3 in the Single Equation Estimation Menu (see Section 6.4) you will be asked to type the name of your dependent variable followed by the list of your regressors, separated by spaces, in the box editor that appears on the screen. For example, to specify the regression

$$YFOOD_t = a_0 + a_1 XL_t + a_2 XC_t + u_t$$

you need to type

$$YFOOD \ INPT \ \ XL \ \ XC$$

where $YFOOD$ is the dependent variable and $INPT$, XL, and XC are the regressors. The variable $INPT$ here denotes for the intercept (or constant) term and can be created either by using the  button (see Section 4.1), or by typing the formula

$$INPT = 1$$

into the Functions box. Before running this regression you must ensure that all the four variables $YFOOD$, $INPT$, XL, and XC are in the variable list by clicking the  button (see Chapter 4).

In specifying the regression equation the following points are worth bearing in mind:

1. It is possible to specify lagged or lead values of the dependent variable or other variables as regressors by including the order of lags or leads enclosed within brackets immediately after the relevant variables. For example, to specify the regression equation

   $$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t+1} + \beta_0 x_t + \beta_1 x_{t-1} + u_t$$

   when asked to list your regression equation you can type

   $$Y \quad INPT \quad Y(-1) \quad Y(-2) \quad Y(+1) \quad X \quad X(-1)$$

   where $INPT$ stands for an intercept term (a vector of ones), $Y(-1)$ and $Y(-2)$ represent the first and second-order lags of the dependent variable (Y), respectively, $Y(+1)$ stands for the first-order lead of Y, and $X(-1)$ is the first-order lag of X. The variables $Y(-1)$, $Y(-2)$, $Y(+1)$, and $X(-1)$ are created temporarily for use only in the estimation/testing/forecasting stage of the program. This is a useful facility and allows you to include lags of variables in the regression equation without having to create them explicitly as new variables in the Process window. When the specified equation contains lagged variables, the information in the order of lags will also be used in the calculation of dynamic forecasts (see the forecast option in the Post Regression Menu in Section 6.20). However, if lagged values of the dependent variable are created in the Process window, before entering the estimation/testing/forecasting stage, these lagged values will be treated like any other regressors, and static forecasts will be calculated. For example, suppose $Y1$, $Y2$, and $X1$ are created in the Process window as

   $$Y1 = Y(-1); \quad Y2 = Y(-2); \quad X1 = X(-1)$$

The regression of Y on

$$INPT \quad Y1 \quad Y2 \quad X1$$

will generate the same results as the regression of Y on

$$INPT \quad Y(-1) \quad Y(-2) \quad X(-1)$$

except for the forecasts; which will be static (see options 8 and 9 in Section 6.20).

2. Also, in the regression of Y on

$$INPT \quad Y(-1) \quad X$$

the program recognizes that only the first-order lag of the dependent variable, namely $Y(-1)$, is specified amongst the regressors, and automatically includes Durbin's h-statistic in the OLS regression results. But if the regression is specified as Y on $INPUT$ $Y1$ X the program treats $Y1$ like any other regressor and does not report the h-statistic.

3. In specifying distributed lag functions it is often convenient to use the facility that allows the user to include a number of lagged values of a variable without having to type all of their names in full. For example to include the variables

$$x_t, x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, z_{t-10}, z_{t-11}, z_{t-12}$$

among your regressors you simply need to type

$$X\{0-5\} \;\; Z\{10-12\}$$

As another example, if you wish to include the following regressors in your model

$$w_t, w_{t-2}, w_{t-5}, w_{t-8}, w_{t-9}, w_{t-10}$$

you need to type

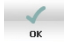
$$W\{0 \; 2 \; 5 \; 8-10\}$$

4. Note that except for Phillips-Hansen's Fully Modified OLS estimator (item 5 in the Single Equation Estimation Menu), *Microfit* does not automatically include an intercept term in the regression equation, and you need to include it explicitly amongst your regressors.

6.5.2 Specification of the estimation period

You need to specify the estimation period once you have set up the model (see Section 6.5.1). By default, all available observations will be chosen, and the start and finish for the estimation period will be the same as the minimum and maximum dates (or observations) displayed on the screen.

You may, however, wish to choose a subset of available observations for estimation, perhaps saving some of the observations for the predictive failure and structural stability tests, or for forecasting. In this case you should enter the start and finish of your estimation period, by clicking on the Start and End boxes in turn and choosing a date from the drop-down list.

If there are observations at the end of the sample period which have not been included in the estimation period, in the case of the *OLS* option you will also be asked to specify the number of observations to be used in the predictive failure/structural stability tests. You will be presented with a window stating: 'Number of observations for predictive failure/structural stability tests(s) (Min 0 Max < >)' . Enter your desired number of observations between zero and the maximum number specified, and press .

In specifying the estimation period the following points are worth bearing in mind:

1. The estimation period cannot fall outside the period defined by the minimum and maximum dates (or observations).

2. The program automatically adjusts the chosen estimation period to allow for missing observations (blank fields) at the beginning and at the end of the sample period. For example, if the available observations run from 1960 to 1980, but the observations on the dependent variable and/or one of the regressors are missing for the years 1960, 1961, and 1980, then the default estimation period will be 1962-1979.

3. If one or more observations on the dependent variable and/or on the regressors are missing in the middle of the specified estimation period, estimation will be carried out on a shorter sample period with no missing values (if possible).

6.6 Ordinary Least Squares option

This option computes the Ordinary Least Squares (OLS) estimates of $\boldsymbol{\beta}$ together with the corresponding standard errors, t-ratios, and probability values. (See Sections 6.1.2 and 6.1.3). It also computes a number of summary statistics and diagnostic test statistics (with probability values) aimed at checking for possible deviations from the classical normal assumptions (A1 to A5). The summary statistics include R^2, $\bar{R}^2$, Akaike information criterion (AIC), Schwartz Bayesian criterion (SBC), residual sum of squares, standard error of regression, and the maximized value of the log-likelihood function. The formulae used for the computation of these and other statistics are given in Sections 21.6 and 21.7.

The diagnostic statistics included in the OLS regression results are for testing the following hypotheses:

- Residual serial correlation.

- Functional form misspecification.

- Normality of residuals.

- Heteroscedasticity.
- Predictive failure.
- Structural stability.

For each of these hypotheses the program computes two test statistics: a Lagrange multiplier (LM), or score statistic, and an F statistic. The LM statistic is asymptotically distributed as a chi-square (χ^2) variate. For a comprehensive review of the use of LM tests in econometrics, see Godfrey (1988). The F-statistic, also known in the literature as 'LM F' or 'modified LM' statistic, is taken approximately to have the F distribution: see Harvey (1981), p. 277. The LM and the F statistics have the same distribution asymptotically. But, on the basis of Monte Carlo results, Kiviet (1986) has shown that in small samples the F version is generally preferable to the LM version. In what follows we provide a brief account of these diagnostic tests. For further details of the econometric methods involved and the relevant references to the literature, see Section 21.6.2.

6.6.1 Tests of residual serial correlation

The program provides the following tests of the non-autocorrelated error assumption, A3:

- Durbin-Watson test (Durbin and Watson (1950), Durbin and Watson (1951)).
- Durbin's h-test (Durbin (1970))[5].
- Lagrange multiplier (LM) tests[6].

The program always supplies the DW statistic, but reports the h-statistic only when the regression equation is explicitly specified to include a single, one-period lag of the dependent variable. The LM statistic is included in the diagnostic tests table, and is applicable to models with and without lagged dependent variables (Godfrey 1978b, 1978c). It is applicable in testing the hypothesis that the disturbances, u_t, are serially uncorrelated against the alternative hypothesis that they are autocorrelated of order p (either as autoregressive or moving average processes). In the diagnostic tests table the following values are chosen for p:

$p = 1$	for undated, annual and daily data
$p = 2$	for half-yearly data
$p = 4$	for quarterly data
$p = 12$	for monthly data

Other values for p can be specified using option 1 in the Hypothesis Testing Menu (see Section 6.23).

[5] See also Godfrey (1978a).

[6] For example, see Godfrey (1978b), Godfrey (1978c), Breusch and Pagan (1980), and Breusch and Godfrey (1981).

6.6.2 Ramsey's RESET test for functional form misspecification

The RESET test (Ramsey 1969) reported in the diagnostic tests table refers to the simple case where only the square of fitted values ($\hat{y}_t^2$) are included in the extended regression of $e_t = y_t - \mathbf{x}_t'\widehat{\boldsymbol{\beta}}$ (or $\hat{y}_t$) on $\mathbf{x}_t$ and $\hat{y}_t^2$. A pth order RESET test can be carried out by using Option 6 in the Hypothesis Testing Menu (Section 6.23), with $\hat{y}_t^2$, $\hat{y}_t^3$, ..., $\hat{y}_t^p$ specified as additional variables. Notice that to carry out such a test, you first need to save the fitted values of the regression of y on $\mathbf{X}$ by means of Option 7 in the Display/Save Residuals and Fitted Values Menu (see Section 6.21).

6.6.3 The normality test

This is the test proposed by Jarque and Bera (1980) for testing the normality assumption, A5, and is valid irrespective of whether or not the regression equation includes an intercept term.

6.6.4 Heteroscedasticity test

This is a simple test of the (unconditional) homoscedasticity assumption, A4, and provides an LM test of $\gamma = 0$ in the model

$$E\left(u_t^2\right) = \sigma_t^2 = \sigma^2 + \gamma\left(\mathbf{x}_t'\boldsymbol{\beta}\right)^2, \qquad t = 1, 2, ..., n$$

See Koenker (1981), where it is also shown that such a test is robust with respect to the non-normality of the disturbances.

6.6.5 Predictive failure test

This is the second test discussed in Chow (1960), and is applicable even if the number of available observations for the test is less than the number of unknown parameters. As shown in Pesaran, Smith, and Yeo (1985) the predictive failure test can also be used as a general specification error test.

6.6.6 Chow's test of the stability of regression coefficients

This is the first test discussed in Chow (1960), and tests the equality of regression coefficients over two sample periods conditional on the equality of error variances. In the statistics literature this test is known as the analysis of covariance test: see Scheffe (1959). The program computes this test if the number of observations available after the estimation period is greater than k, the number of regressors included in the model.

6.6.7 Measures of leverage

In the classical linear regression model (6.2), the leverage (or the influence) of points in the regression design is measured by the diagonal elements of the matrix[7]

$$\mathbf{A} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = (a_{ij}) \tag{6.16}$$

The relevance of the leverage measures, a_{ii}, $i = 1, 2, ..., n$ in regression analysis has been discussed in detail by Belsley, Kuh, and Welsch (1980) Chapter 2, and Cook and Weisberg (1982) Chapter 2.

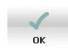
The program provides plots of a_{ii}, $i = 1, 2, ..., n$, and allows you to save them for subsequent analysis. In the plot of the leverage measures, the average value of a_{ii}, which is equal to k/n, is also displayed.[8]

The leverage measures also provide important information on the extent of small sample bias that may be present in the heteroscedasticity-consistent estimators of the covariance matrix of $\hat{\boldsymbol{\beta}}_{OLS}$ (Section 21.22). As shown by Chesher and Jewitt (1987), substantial downward bias can result in the heteroscedasticity-consistent estimators of the variance of the least squares estimators, if regression design contains points of high leverage.

6.7 Generalized instrumental variable method option

Option 2 (the IV or $2SLS$ option) in the Linear Regression Estimation Menu (see Section 6.5) enables you to obtain consistent estimates of the parameters of the regression model when the orthogonality assumption A4 is violated.[9] The breakdown of the orthogonality assumption could be due to a variety of problems, such as simultaneity, measurement errors or sample selection bias, or could be because actual values are used as a proxy for expectational variables under the rational expectations hypothesis. For example, see Sargan (1958), McCallum (1976), Wickens (1982), and Pesaran (1987b). A unified account of the IV method can be found in Pesaran and Smith (1990).

The IV option can also be used to compute two-stage least squares ($2SLS$) estimates of a single equation from a simultaneous equation system. Notice, however, that the computations of the $2SLS$ estimates require that all the predetermined variables of the simultaneous equation model be specified as instrumental variables.

When you choose this option, you will be asked to list your instrumental variables separated by spaces. The number of instruments should be at least as large as the number of regressors. In the case of exact collinearity amongst the instruments and/or the regressors, the program displays an error message and invites you to click  to continue.

If you specify fewer instruments than regressors, the program shows the minimum number of required instruments (i.e. the number of regressors) and asks you to try again.

[7] Since matrix $\mathbf{A}$ maps $\mathbf{y}$ into $\hat{\mathbf{y}} = \mathbf{A}\mathbf{y}$, the matrix $\mathbf{A}$ is also known in the literature as the 'hat' matrix.

[8] Note that since $Tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} = k$, then the simple average of a_{ii}, $i = 1, 2, ..., n$ will be equal to k/n.

[9] A test of the orthogonality assumption can be carried out by computing Wu-Hausman statistic T_2, (Wu 1973 and Hausman 1978), using the variable addition test option in the Hypothesis Testing Menu (see Section 6.23). See also Lesson 11.10.

The estimation results for the IV option are summarized in two tables. The first table gives the parameter estimates, their estimated asymptotic standard errors, and t-ratios, as well as Sargan's statistic for a general test of misspecification of the model and the instruments. This test statistic is asymptotically distributed as χ^2 with $s-k$ degrees of freedom, where s represents the number of instruments and k the number of the regressors ($s > k$). (See Section 21.10.3). This table also reports probability values, the values of the IV minimand, R^2, $\overline{R}^2$, GR^2, $\overline{GR}^2$, F, and DW statistics, and a few other summary statistics. But, note that these statistics in the case of the IV option do not have the usual OLS interpretations. For example, R^2, $\overline{R}^2$ are not valid for regressions estimated by the IV method, and can produce misleading results. Appropriate measures of overall fit for IV regressions are given by the Generalized R-Bar-Squared statistics (GR^2, and $\overline{GR}^2$) proposed in Pesaran and Smith (1994) (see Section 21.10.2). The same also applies to the DW statistic. For tests of residual serial correlation the appropriate statistics is the LM statistic reported in the Diagnostic Tests Table. Finally, note that the probability values reported are only valid asymptotically.

The second table supplies diagnostic statistics (with the associated probability values) for the tests of residual serial correlation, functional form misspecification, non-normal errors, and heteroscedasticity. The tests of residual autocorrelation and functional form misspecification are both based on the statistics in equation (21.69), originally due to Sargan (1976), using different specifications for the $\mathbf{W}$ matrix. In the case of the test of residual autocorrelation, the $\mathbf{W}$ matrix is defined by equation (21.70), with p, the order of the test, being

$p = 1$	for undated, annual and daily data
$p = 2$	for half-yearly data
$p = 4$	for quarterly data
$p = 12$	for monthly data

Other values for p can be specified using option 1 in the Hypothesis Testing Menu (see Section 6.23). The statistic for the test of functional form misspecification is computed using (21.69) with the $\mathbf{W}$ matrix specialized to

$$\mathbf{W} = (\hat{y}^2_{1,IV}, \hat{y}^2_{2,IV}, ..., \hat{y}^2_{n,IV})'$$

where $\hat{y}_{t,IV} = \mathbf{x}'_t\widehat{\boldsymbol{\beta}}_{IV}$ are the IV fitted values.

The statistics for normality and the heteroscedasticity tests are computed as in the OLS case, with the difference that the IV fitted values and the IV residuals are used in place of the OLS ones (see Section 21.6.2).

The details of the algorithms used to compute the IV estimators and the associated test statistics are given in Section 21.10.

6.8 AR errors (exact ML) option

Option 3 in the Linear Regression Estimation Menu (see Section 6.5) computes exact maximum likelihood estimators of the regression equation (6.1) under the assumption that the disturbances, u_t, follow a stationary autoregressive process with stochastic initial values. This option differs from the Cochrane-Orcutt option (option 4 in the Linear Regression Menu), which estimates the AR error regression model under the assumption of fixed initial values. The idea of allowing for initial values in the estimation of $AR(1)$ error models was first put forward in econometrics by Hildreth and Lu (1960). The method was then subsequently extended to higher-order AR error models by Pesaran (1972), and Beach and MacKinnon (1978). See Section 21.11 for more details.

When you click the button to begin your calculation you need to specify the order of the AR error process. You can either choose the $AR(1)$ specification

$$AR(1): \quad u_t = \rho u_{t-1} + \epsilon_t$$

or the $AR(2)$ specification

$$AR(2): \quad u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \epsilon_t$$

For example, to choose the $AR(2)$ specification, when prompted, type 2 and click . The estimation results are displayed in a table in two parts. At the top are the estimates of the regression coefficients, their (asymptotic) standard errors, and other summary statistics such as R^2, $\bar{R}^2$, standard errors of regression ($\hat{\sigma}_\epsilon$), are given. At the bottom (use the scroll bar if necessary) is the second part of the results, which gives the parameter estimates of the AR error process, together with the associated t-ratios computed on the basis of the (asymptotic) standard errors (see Sections 21.11.1 and 21.11.2 for the relevant formulae). The program also reports the log-likelihood ratio statistics for the test of $AR(1)$ against the non-autocorrelated error hypothesis, and for the test of the $AR(2)$ error specification against the $AR(1)$ error specification. The latter statistic is reported only in the case of the $AR(2)$ option. These statistics are computed according to the formulae set out in Section 21.11.4

In case of the $AR(1)$ option, you will also be given the opportunity to see the plot of the concentrated log-likelihood function in terms of the parameter of $AR(1)$ error process over the range $|\rho| < 1$ (see equation (21.77) in Section 21.11.1). The plot of the concentrated log-likelihood function is particularly useful for checking the possibility of multiple maxima.

Notes

1. In the case of this option the standard errors (and hence t-ratios) reported for the estimates of the regression coefficients and the parameters of the AR-error process are valid (asymptotically) if the regression equation does not contain lagged dependent variables. When your equation includes lagged dependent variables try other AR options, namely options 4 to 6 in the Linear Regression Menu.

2. The iterations, if convergent, always converge to a stationary solution. This is a particular feature of the exact ML/AR option and does not apply to the other AR options.

3. If the estimation method fails to converge within 40 iterations, a sub-menu will be displayed. The options in this sub-menu allow you to terminate the iterations and

start with a different set of initial parameter estimates, or to increase the number of iterations in steps of 20 until convergence is reached. If you choose option 0 (abandon estimation), you will be presented with another menu with which to specify a new set of initial parameter estimates and another chance to try the iterations (see Section 6.13.1 for more details). In situations where the convergence cannot be attained even after, say, 100 iterations, and for different sets of initial parameter estimates, it is perhaps best to terminate the iterations and try other AR options in the Linear Regression Menu. Notice, however, that in the case of option 3 where a first-order error process is specified, the iterations are certain to converge (see Section 21.11.1).

6.9 AR errors (Cochrane-Orcutt) option

Option 4 in the Linear Regression Menu computes estimates of the regression equation (6.1) under the following $AR(m)$ error process ($m \leq 12$)

$$AR(m): \quad u_t = \sum_{i=1}^{m} \rho_i u_{t-i} + \epsilon_t \tag{6.17}$$

using a generalization of the Cochrane and Orcutt (1949) iterative method. This method assumes that the initial values $u_1, u_2, ..., u_m$ are given (or fixed). Notice, however, that if the $AR(m)$ process is stationary, the Cochrane-Orcutt (CO) option yields estimates that are asymptotically equivalent to the exact ML estimators that explicitly allow for the distribution of the initial values.

The results for the CO option are summarized in a table, the top half giving the estimates of the regression equation (6.1), and the bottom half giving the estimates of the error process (6.17). The details of the computations can be found in Section 21.12.

Notes

1. For the case where $m = 1$, the program provides you with the option of seeing the plot of concentrated log-likelihood function, given by equation (21.102) in Section 21.12.

2. The estimated standard errors computed under the CO option are valid (asymptotically) even if the regression equation contains lagged values of the dependent variable.

3. The program displays a warning if the estimates of $\rho_1, \rho_2, ..., \rho_m$ result in a non-stationary error process. In such a case, inferences based on the reported standard errors can be misleading.

4. The program allows you to increase the number of iterations interactively or to change the choice of the initial parameter estimates for the start of the iterations, if the method fails to converge within 40 iterations (for details see note 3 in Section 6.8.)

6.10 AR errors (Gauss-Newton) option

Option 5 in the Linear Regression Menu (see Section 6.5) provides estimates of equations (6.1) and (6.17) when the $AR(m)$ process is subject to zero restrictions. For example, it allows estimation of equation (6.1) under

$$u_t = \rho_4 u_{t-4} + \epsilon_t, \tag{6.18}$$

or under

$$u_t = \rho_1 u_{t-1} + \rho_4 u_{t-4} + \rho_{15} u_{t-14} + \epsilon_t. \tag{6.19}$$

When this option is chosen you will be prompted to type the non-zero lags in the AR error process (6.17) in an ascending order separated by space(s). To choose, for example, specification (6.18), you need to type 4 and then click on the  button. To choose specification (6.19) you need to type

1 4 15

then click . In the case of example (6.18), the order of the AR error process is $m = 4$, but there is only one unknown parameter in the AR error process. Similarly, in example (6.19), $m = 15$, but the number of unknown parameters of the AR error process is equal to $r = 3$. The following restrictions apply:

$$r \leq 12$$

and

$$n > m + k + r$$

where

n	$\equiv$	the number of observations in the chosen sample period.
k	$\equiv$	the number of regressors in the regression equation.
m	$\equiv$	the order of the AR-error process.
r	$\equiv$	the number of non-zero coefficients in the AR-error process.

See Section 21.13 for more details. Notice, however, that in the case of this option, the plot of the concentrated log-likelihood function can be obtained if $r = 1$, irrespective of the value specified for m.

6.11 IV with AR errors (Gauss-Newton) option

Option 6 in the Linear Regression Menu is appropriate for the estimation of a regression equation with autocorrelated disturbances when one or more of the regressors are suspected of being correlated with the disturbances. The estimation method which is due to Sargan (1959) is, however, applicable if there exists a sufficient number of instrumental variables that are uncorrelated with the current and past values of the transformed disturbances, ϵ_t in equation (6.17), but at the same time are asymptotically correlated with lagged disturbances, $u_{t-1}, u_{t-2}, .., u_{t-m}$. This option also enables you to compute IV/AR estimates when the AR-error process is subject to zero restrictions. See Section 21.13.1 for details. The econometric methods underlying this option are briefly described in Section 21.14. Other relevant information can be found in Sections 21.14.1 and 21.14.2.

When you click 

you will be asked first to type the non-zero lags in the AR process (as in Section 6.10). You will then be presented with a screen editor to type the list of your instruments. At least $k+m$ instruments are needed for this option. The program provides a number of useful error messages if the instruments and/or regressors are exactly collinear or if the number of instruments supplied is insufficient (see also Section 6.7). You can retrieve a list of instruments previously saved as an .LST file using the 

button, or save your instrument lists for use in subsequent sessions using the 

button.

Notes

1. In the absence of adequate initial observations, the program automatically adjusts the estimation period to allow for the specification of lagged values of the dependent variable and/or the regressors as instruments. In the case of the IV regressions only the Generalized R^2 measures are appropriate for this option.

2. The Sargan misspecification test statistic reported in the case of this option is computed using (21.111), and is useful as a general test of misspecification. It is asymptotically distributed as a chi-squared variate with $s-k-r$ degrees of freedom, where s is the number of specified instruments, k is the number of regressors, and r is the number of unknown parameters of the AR error process (see Section 21.14.1).

3. The R^2, $\bar{R}^2$, GR^2, and $\overline{GR}^2$ statistics reported for this option are based on adjusted residuals and prediction errors, respectively. The relevant formulae are given in Section 21.14.2. Notice that in the case of the IV regressions only the Generalized R^2 measures are appropriate for this option.

4. When $r=1$, the program gives you the option of plotting the minimized values of the IV minimand (21.109), in terms of the unknown parameter of the AR process. This is useful for checking the possibility of multiple minima.

5. The program enables you to increase the number of iterations interactively if the method fails to converge within 40 iterations. (For details see note 3 in Section 6.8).

6. The program gives a warning if the method converges to a non-stationary AR process (see note 3 in Section 6.9).

7. In the case of this option the estimated standard errors are valid (asymptotically) even if the regression equation contains lagged values of the dependent variable.

6.12 MA errors (exact ML) option

Option 7 in the Linear Regression Menu allows you to estimate the regression equation (6.1) under the following $MA(q)$ error specification

$$u_t = \sum_{i=0}^{q} \gamma_i \epsilon_{t-i}, \qquad \gamma_0 = 1 \tag{6.20}$$

Like option 5, this allows you to impose zero restrictions on the MA parameters, γ_i. The estimation is carried out by the exact ML method described in Pesaran (1988a), and does not require the MA process to be invertible. For a description of the method see Section 21.15. The MA option can also be used to estimate univariate $ARMA$ or $ARIMA$ processes.

Notes

1. The estimation of high-order MA error processes (with $q > 6$) can be time-consuming, especially in the case of regression equations with a large number of regressors and observations.

2. The standard errors of the parameter estimates obtained under the MA (or the IV/MA) options are valid asymptotically so long as the estimated MA process is invertible; that is, when all the roots of $\sum_{i=1}^{q} \gamma_i z^i = 0$ fall outside the unit circle. *Microfit* displays a warning if the estimated MA process is non-invertible.

See also the notes in Section 6.11.

6.13 IV with MA errors option

This is the MA version of the IV/AR option described in Section 6.11. It differs from the IV/AR option in two important respects:

1. Following Hayashi and Sims (1983), the IV/MA estimates are computed using 'forward filter' transformation of the regressors and the dependent variable (but not the instruments) to correct for the residual serial correlation. In effect, the IV/MA option is an iterated version of the Hayashi-Sims procedure: see Pesaran (1987b), Section 7.6.2. It is particularly useful for the estimation of linear rational expectations models with future expectations of the dependent variable, where the $u_t's$ may be correlated with the future values of the instruments.

2. The IV minimand for the IV/MA option contains an additional Jacobian term. Although in the case of invertible processes this additional term is asymptotically negligible, our experience suggests that its inclusion in the IV minimand helps the convergence of the iterative process when the roots of the MA part are close to the unit circle. The details of the algorithm and the rationale behind it can be found in Pesaran (1990). A similar procedure has also been suggested by Power (1990) for the first-order case. A description of the underlying econometric method can be found in Section 21.16.

3. The R^2, $\overline{R}^2$, GR^2 and $\overline{GR}^2$ statistics are computed using the formulae in Section 21.16.1. Only the Generalized R^2 statistics are appropriate for this option.

6.13.1 Specification of initial estimates for the parameters of the AR/MA error process

When you choose the AR/MA options in the Linear Regression Menu and click , you will be presented with a menu[10] which gives you a choice between starting the iterations with initial estimates supplied by the program or the initial estimates to be supplied by you. In the case of the AR and MA options with $r = 1$ (when the error process depends only on one unknown parameter), you will also be presented with an option to see the plot of the concentrated log-likelihood function or the IV minimand.

To enter the initial estimate for the first-order lag coefficient, type your choice and move the cursor to the AR lag 2 position. Repeat this process until all the initial estimates are supplied. Then click  to start the iteration.

Since there is no guarantee that the iterative procedures will converge on the global maximum (minimum) of the likelihood function (the IV minimand), we recommend that you check the computations by starting the iterations from a number of different initial values. In the case of error processes with only one unknown parameter, the plot of the concentrated log-likelihood function or the IV minimand can be used to determine whether the global optimum has been achieved.

6.14 Recursive regression options

Option 2 in the Single Equation Menu (See Section 6.4) is the Recursive Linear Regression Menu with the following options:

1. Recursive Least Squares
2. Two-Stage Recursive Least Squares

Option 1 enables you to estimate a linear regression equation recursively by the OLS method.

Option 2 allows you to estimate a linear regression equation recursively by the $2SLS$ (or the IV) method. When you choose this option you will be prompted to list at least as many instruments as there are regressors in your model.

Specify the estimation period and your linear regression equation as described in Sections 6.5.1 and 6.5.2. Set the number of observations you want to use for updating recursive estimation. When the computations are completed you will be presented with the Recursive OLS (or IV) Regression Results Menu described in Section 6.14.1. For the details of the algorithms used in carrying out the necessary computations see Section 21.17.

[10] There is an exception. The $AR(1)$ error specification in Option 3 of the Linear Regression Menu does not give you a choice for the specification of the initial parameter value of the $AR(1)$ process. The iterative method used does not require it, and is always sure to converge.

6.14.1 Recursive OLS Regression Results Menu

This menu has the following options

0. Move to Backtracking Menu
1. Plot recursive coefficients and their standard errors
2. Plot standard errors of recursive regressions
3. Save recursive coefficients
4. Save standard errors of recursive coefficients
5. Save standard errors of the recursive regressions
6. Save standardized recursive residuals
7. Save recursive predictions based on existing regressors
8. Save recursive predictions based on variables to be specified
9. Save adaptive coefficients

Option 0 takes you back to the Commands and Data Transformations box**.**

Option 1 allows you to plot the recursive coefficients, $\hat{\boldsymbol{\beta}}_r$, $r = r^*, r^* + 1, ..., n,$ and their standard error bands (computed as $\hat{\boldsymbol{\beta}}_r$ plus or minus twice their standard errors). To avoid the large uncertainties that are associated with the initial estimates, only the final $\frac{3}{4}$ of the estimates for each coefficient are displayed: namely, r^* is set equal to $\frac{1}{4}n + \frac{3}{4}(k + 1)$. When you choose this option you will be presented with the variable names for your regressors and will be asked for the name of the regressor whose coefficient estimates you wish to see plotted. See Sections 21.17.3 and 21.17.5.

Option 2 plots the standard errors of the recursive regressions, defined by $\hat{\sigma}_r^2$, $r = r^*, r^* + 1, ..., n$, computed using equations (21.138) and (21.143) for the OLS and the IV options, respectively. To avoid the uncertain initial estimates the plots are displayed for $r^* = \frac{1}{4}n + \frac{3}{4}(k + 1)$.

Options 3, **4** and **5** allow you to save all the estimated recursive coefficients, their standard errors, and the standard errors of the recursive regressions as variables in *Microfit*'s workspace.

Option 6 enables you to save standardized recursive residuals defined by equations (21.136) and (21.141) for the OLS and the IV options, respectively.

Option 7 enables you to save recursive predictions and their standard errors. See Section 21.17.8.

Option 8 allows you to save recursive predictions and their standard errors obtained with respect to the variable $\mathbf{w}_t$, which may differ from the regressors, $\mathbf{x}_t$. See equation (21.146) in Section 21.17.8. When you choose this option you will be prompted to list the variable names in $\mathbf{w}_t$ to be used in the calculations of the recursive predictions. You *must* specify exactly the same number of variables as there are regressors in your regression equation.

Option 9 enables you to save adaptive coefficients defined by equation (21.144).

6.15 Rolling Linear Regression Menu

Option 3 in the Single Equation Menu (see Section 6.4) is the Rolling Linear Regression Menu. The menu has the following options

1. Rolling least squares
2. Rolling two-stage least squares

Option 1 allows you to estimate the coefficients of a linear regression equation by the *OLS* method over successive rolling periods of a fixed length.

Option 2 allows you to estimate the coefficients of a linear regression equation by the two-stage least squares (or *IV*) method over successive rolling periods of a fixed length (set using option 1). If you choose this option you will be prompted to list at least as many instruments as there are regressors in your equation.

Specify the estimation and your regression equation as usual. You will be asked to specify the length of the window to be used in the estimation, and to set the number of observations you want to use for updating the estimation.

6.15.1 Rolling Regression Results Menu

This menu has the following options

0. Move to Backtracking Menu (Rolling Regression)
1. Plot rolling coefficients and their standard errors
2. Plot standard errors of rolling regressions
3. Save rolling coefficients
4. Save standard errors of rolling coefficients
5. Save standard errors of rolling regressions
6. Plot one-step-ahead rolling forecasts
7. Save one-step-ahead rolling forecasts
8. Save standard errors of rolling forecasts

Option 0 takes you back to the Commands and Data Transformations box**.**

Option 1 allows you to plot the rolling regression coefficients and their standard errors. See also the description of option 1 in the Recursive Regression Results Menu in Section 6.14.1.

Option 2 allows you to plot the standard errors of the rolling regressions.

Options 3, 4 and **5** enable you to save the rolling coefficients, their standard errors, and the standard errors of the rolling regressions on *Microfit*'s workspace.

Options 6, 7 and **8** allow you to plot and save one-step-ahead rolling forecasts and their standard errors

6.16 Non-Linear Regression Menu

The Non-Linear Regression Menu is option 4 in the Single Equation Estimation Menu (Univariate Menu: see Section 6.4). It contains the following option

1. Non-linear least squares
2. Non-linear 2-stage least squares

Option 1 allows you to estimate your specified non-linear equation by the least squares method.

Option 2 allows you to estimate your specified non-linear equation by the $2SLS$ (or IV) method. When you choose this option you will be prompted to list at least as many instruments as there are unknown parameters in your non-linear model.

Notes

1. See Section 6.16.1 on how to specify/modify a non-linear equation

2. Special care needs to be exercised with respect to the choice of initial parameter estimates. An appropriate choice of initial estimates can hamper the convergence of the iterative process, and may lead to error messages which can be difficult to decipher at first. For example, suppose you are interested in estimating the following non-linear equation:
$$C_t = A_0 + A_1 \exp\left(Y_t/A_2\right)$$
where C is real consumption expenditure, Y is the real disposable income and A_0, A_1, and A_2 are unknown parameters. If you start the iteration with $A_2 = 0$, you will see an error message stating that there are insufficient observations to estimate. This is because with A_2 initially set equal to zero, all the values of Y_t/A_2, being undefined, will be set to blank. Another problem that arises frequently in the estimation of non-linear regression models concerns the scaling on the regressors. In the case of the above examples, unless the Y_ts are reasonably small, exponentation of Y_t can result in very large numbers, and the computer will not be able to handle the estimation problem. When this arises you will see an error message on the screen.

3. A similar problem can arise when initial values chosen for A_2 are very small, even if the Y_ts are reasonably small. In some applications the use of zero initial values for the parameters can result in an error message. It is always advisable to think carefully about the scale of your regressors and the choice of the initial parameter values before running the non-linear regression option.

4. When estimating a linear regression equation via the non-linear regression option, it is acceptable to use zeros as initial values for the parameters

5. The least squares and the IV options in the Non-Linear Regression Menu allow you to estimate a linear or a non-linear regression, subject to linear or non-linear parametric restrictions. For example,, to estimate the $ARDL$ model

$$y_t = \alpha + \lambda y_{t-1} + \beta_1 x_t + \beta_2 x_{t-1} + u_t$$

subject to the common factor restrictions

$$\beta_1 \lambda + \beta_2 = 0$$

you need to type the following formula.

$$Y = A0 + A1 * Y(-1) + A2 * X - A1 * A2 * X(-1)$$

6. The non-linear $2SLS$ option can also be used to estimate Euler equations, namely the first-order conditions for intertemporal optimization problems under uncertainty. For an example, see Lesson 13.2.

6.16.1 Specification of a non-linear regression equation

Enter the specification of the non-linear equation, and type the formula for the equation in the box editor provided on the screen. You can type your formula using standard arithmetic operators such as $+$, $-$, $/$, and $*$, and any one of the built-in functions set out in Section 4.2. For example, suppose you are interested in estimating the following Cobb-Douglas production function with additive errors:

$$Y_t = AL_t^{\alpha} K_t^{\beta} + u_t$$

where Y_t is output, L_t and K_t are labour and capital inputs, and u_t is a disturbance term. The unknown parameters are represented by A, α and β. Then you need to type

$$Y = A0 * (L\hat{\ } A1) * (K\hat{\ } A2)$$

As another example, suppose you wish to specify the following non-linear regression:

$$z_t = \alpha_1 e^{\beta_1 x_{1t}} + \alpha_2 e^{\beta_2 x_{2t}} + u_t$$

In the box editor that appears on the screen you need to type

$$Z = a1 * \mathbf{EXP}(b1 * X1) + a2 * \mathbf{EXP}(b2 * X2)$$

When specifying a non-linear regression, the following points are worth bearing in mind:

1. In the case of the above two examples, it is assumed that the variables (Y, K, and L), and (z, $x1$ and $x2$) are in the variable list and that $A0, A1, A2, B1$ and $B2$ are parameters to be estimated, and are not, therefore, in your list of variables. (Note that in *Microfit* upper- and lower-case letters are treated as identical.)

2. You need to watch for two important types of mistake: using an existing variable name to represent a parameter value, and including a non-existent variable in the specification of the non-linear equation. *Microfit* is not capable of recognizing these types of mistake. But you should be able to detect your mistakes at a later stage when you will be asked to provide initial estimates for the unknown parameters of your equation (see Section 6.16.2). If, by mistake, you use a variable name to present a parameter, you will not be asked to supply an initial estimate for the parameter in question, and most likely the computations will fail to converge. In the opposite case, where a non-existent variable is included in the regression equation, *Microfit* treats the non-existent variable as an unknown parameter and asks you to supply an initial estimate for it! *To reduce the probability of making such mistakes we recommend that you reserve the names* $A0, A1, A2...$ *and* $B0, B1, B2, ...$ *for parameter values, and try not to use them as names for the variables on the workspace.*

3. Mistakes in typing the regression formula are readily detected by *Microfit*. But you need to fix the problem by carefully checking the non-linear formula that you have typed, and by ensuring that you have not inadvertently mixed up variable names and parameter values!

4. The list of variables specified under the various estimation options, including the non-linear equation specified under option 4, can be saved in a file for use at a later stage using the 

button. The variable lists are saved in files with extension .LST, and the non-linear equations are saved in files with the extension .EQU. To retrieve a file you saved earlier, click 

.

6.16.2 Specification of initial parameter estimates

The non-linear estimation option, and the other estimation methods that use iterative techniques, will require the users to supply initial values for the unknown parameters in their specified econometric model. In such cases you will be presented with a screen containing the names of the known parameters, all of which are initially set equal to zero. You can change these initial settings by moving the cursor to the desired position and typing your own choice of the initial estimate. These initial parameter values can be readily changed if the estimation method fails to converge.

6.16.3 Estimation results for the non-linear regression equation

Once the estimation of the non-linear equation is successfully complete, you will be presented with the estimation results in a format similar to that given for the *OLS* and the *IV* options in the case of linear regression (see Section 6.5). *Microfit* also automatically computes diagnostic statistics for tests of residual serial correlation, functional form mis-specification, non-normality of disturbances, and heteroscedasticity in the case of non-linear equations estimated by the least squares or the *IV* methods. These statistics are computed using the same procedures as outlined in Sections 21.6.2 and 21.10, with the matrix $\mathbf{X}$ replaced by the matrix of first derivatives (of the non-linear equation with respect to the unknown parameters) evaluated at the parameter estimates obtained on convergence. The non-linear options do not compute statistics for structural stability and predictive failure tests.

The relevant formulae for the non-linear estimation options are given in 21.21.

6.17 Phillips-Hansen Estimation Menu

This menu allows you to estimate the parameters of a *single* cointegrating relation by the fully-modified *OLS* (*FM-OLS*) procedure proposed by Phillips and Hansen (1990). The underlying econometric model is given by

$$y_t = \beta_0 + \boldsymbol{\beta}_1' \mathbf{x}_t + u_t,\ t = 1, 2, ..., n \tag{6.21}$$

where y_t is an $I(1)$ variable, and $\mathbf{x}_t$ is a $k \times 1$ vector of $I(1)$ regressors, assumed not to be cointegrated among themselves.[11] It is also assumed that $\mathbf{x}_t$ has the following first-difference stationary process

$$\Delta \mathbf{x}_t = \boldsymbol{\mu} + \mathbf{v}_t,\ \ t = 2, 3, ..., n \tag{6.22}$$

where $\boldsymbol{\mu}$ is a $k \times 1$ vector of drift parameters, and $\mathbf{v}_t$ is a $k \times 1$ vector of $I(0)$, or stationary variables. It is also assumed that $\boldsymbol{\xi}_t = (u_t, \mathbf{v}_t')'$ is strictly stationary with zero mean and a finite positive-definite covariance matrix, $\boldsymbol{\Sigma}$.

The *OLS* estimators of $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}_1')'$ in (6.21) are consistent even if $\mathbf{x}_t$ and u_t (equivalently $\mathbf{v}_t$ and u_t) are contemporaneously correlated: see, for example, Engle and Granger (1987), and Stock (1987). But in general the asymptotic distribution of the *OLS* estimator involves the unit-root distribution and is non-standard, and carrying out inferences on $\boldsymbol{\beta}$ using the usual t-tests in the *OLS* regression of (6.21) will be invalid. To overcome these problems, appropriate corrections for the possible correlation between u_t and $\mathbf{v}_t$ and their lagged values is required. The Phillips-Hansen fully-modified *OLS* (*FM-OLS*) estimator takes account of these correlations in a semi-parametric manner. But it is important to recognize that the validity of this estimation procedure critically depends on the assumption that $\mathbf{x}_t$s are $I(1)$ and are not themselves cointegrated. For Monte Carlo evidence on small sample properties of the *FM-OLS* estimators see Pesaran and Shin (1999). For details of the computational algorithms see Section 21.18.

To access this menu, select option 5 in the Single Equation Estimation Menu (see Section 6.4). It contains the following options

1. None of the regressors has a drift
2. At least one regressor is $I(1)$ with drift

You need to choose option 1 if $\boldsymbol{\mu}$ in (6.22) is zero. Otherwise you should select option 2. List the dependent variable, y_t, followed by the $I(1)$ regressions, x_{1t}, x_{2t},..., x_{kt} in the box

[11] A variable is said to be $I(1)$ if it must be differenced once before it can be rendered stationary. A random walk variable is a simple example of an $I(1)$ variable. A set of $I(1)$ variables are said to be cointegrated if there exists a linear combination of them which is $I(0)$, or stationary. For further details see, for example, Engle and Granger (1991).

editor on the screen. Do not include intercept or time trends among the regressors. Specify your estimation period and the length of your lag window. When you click , you will be presented with the following menu for selecting the lag window for the estimation of the long-run variances used in the estimation procedure

0. Move to Backtracking Menu
1. Equal weights lag window
2. Bartlett lag window
3. Tukey lag window
4. Parzen lag window

The use of the equal weights (or uniform) lag window may result in negative standard errors, and when this happens you need to choose one of the other three lag windows. We recommend the Parzen lag window.

Once the lag window is chosen you will be asked to specify the length of the lag window, and are then presented with the estimation results. You can also carry out tests of linear and non-linear restrictions on the cointegrating coefficients $\boldsymbol{\beta} = \left(\beta_0, \boldsymbol{\beta}_1'\right)'$, using the options in the Post Regression Menu (see Section 6.20).

6.18 ARDL approach to cointegration

Option 6 in the Single Equation Estimation Menu (Univariate Menu: see Section 6.4) allows you to estimate the following $ARDL(p, q_1, q_2, ..., q_k)$ models

$$\phi(L,p)y_t = \sum_{i=1}^{k} \beta_i(L,q_i)x_{it} + \boldsymbol{\delta}'\mathbf{w}_t + u_t \tag{6.23}$$

where

$$\begin{aligned} \phi(L,p) &= 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p, \\ \beta_i(L,q_i) &= 1 - \beta_{i1} L - \beta_{i2} L^2 - ... - \beta_{iq_i} L^{q_i}, \text{ for } i = 1,2,...,k \end{aligned} \tag{6.24}$$

L is a lag operator such that $Ly_t = y_{t-1}$, and $\mathbf{w}_t$ is a $s \times 1$ vector of deterministic variables such as the intercept term, seasonal dummies or time trends, or exogenous variables with fixed lags. *Microfit* first estimates (6.23) by the OLS method for *all* possible values of $p = 0, 1, 2, ..., m$, $q_i = 0, 1, 2, ..., m$, $i = 1, 2, ..., k$; namely a total of $(m+1)^{k+1}$ different $ARDL$ models. The maximum lag, m, is chosen by the user, and all the models are estimated on the same sample period, namely $t = m+1, m+2, ..., n$.

In the second stage the user is given the option of selecting one of the $(m+1)^{k+1}$ estimated models using one of the following four model selection criteria: the $\bar{R}^2$ criterion, Akaike information criterion (AIC), Schwartz Bayesian criterion (SBC) and the Hannan and Quinn criterion (HQC).[12] The program then computes the long-run coefficients and

[12] These model selection criteria are described in Section (21.7).

their asymptotic standard errors for the selected $ARDL$ model. It also provides estimates of the error correction model (ECM) that corresponds to the selected $ARDL$ model. For further details and the relevant formulae for the computation of the long-run coefficients and the associated Error Correction Model (ECM) see Section 21.19.

6.18.1 Specification of an ARDL regression equation

Specify the estimation period and maximum lag order m ($m < 24$) for your $ARDL$ specification before specifying your equation. The list of the variables to be included in the $ARDL$ model should be typed followed by deterministic regressors such as the intercept term, time trends, and regressors with fixed lags; separating the two sets of variables by &. The dependent variable should be the first variable in the list. The first set of variables should not appear in lagged or lead form, and it should not contain an intercept term or time trends.

As an example, suppose you wish to specify the following $ARDL$ model:

$$\begin{aligned}\phi(L,s)y_t &= \alpha_0 + \alpha_1 T_t + \alpha_2 z_t + \beta_1(L,s)x_{1t} \\ &\quad + \beta_2(L,s)x_{2t} + u_t\end{aligned}$$

where $\phi(L,s)$, $\beta_i(L,s)$, $i = 1,2$ are polynomial lag operators of the maximum order equal to s, T_t is a deterministic time trend, and z_t is an exogenous regressor. Once presented with the box editor you need to type

$$Y \quad X1 \quad X2 \quad \& \quad INPT \quad T \quad Z$$

It is important to note that even in the case of $ARDL$ models with a small number of regressors (say $k = 2$), the number of $ARDL$ models to be estimated could be substantial, if m is chosen to be larger than 6. In the case where $k = 2$, and $m = 6$, the total number of $ARDL$ models to be estimated by the program is equal to $(6+1)^3 = 343$. If this number of $ARDL$ models to be estimated exceeds 125 you will be presented with a warning that computation may take a long time to complete. If you choose to go ahead, *Microfit* carries out the necessary computations and presents you with the $ARDL$ Order Selection Menu (see Section 6.18.2).

6.18.2 ARDL Order Selection Menu

This menu has the following options

0. Move to Backtracking menu
1. Choose maximum lag to be used in model selection
2. R-BAR Squared
3. Akaike information criterion
4. Schwartz Bayesian criterion
5. Hannan-Quinn criterion
6. Specify the order of the $ARDL$ model yourself

Option 1 allows you to change the maximum lag order, s, to be used in the computations

Option 2 selects the orders of the $ARDL(p, q_1, q_2, ..., q_k)$ model, namely the values of $p, q_1, q_2..., q_k$ using the $\overline{R}^2$ criterion.

Option 3 selects the orders of the $ARDL$ model using the Akaike information criterion.

Option 4 selects the orders of the $ARDL$ model using the Schwartz Bayesian criterion.

Option 5 selects the orders of the $ARDL$ model using the Hannan and Quinn criterion.

Option 6 allows you to specify your own choice of the lag-orders, $p, q_1, q_2, ..., q_k$. When you choose this option you will be asked to specify exactly $k + 1$ integers representing the order of the lag on the dependent variable, followed by the order(s) of the lag(s) on the k regressor(s). *Microfit* works out the maximum value of these orders that can be chosen by the user given the sample size available.

Once the orders $p, q_1, q_2, ..., q_k$ are selected either by one of the model selection criteria (options 2 to 5) or by specifying them yourself (option 6) you will be presented with the Post ARDL Selection Menu (see Section 6.18.3).

6.18.3 Post ARDL Model Selection Menu

This menu has the following options

0. Return to $ARDL$ Order Selection Menu
1. Display the estimates of the Selected $ARDL$ regression
2. Display long run coefficients and their asymptotic standard errors
3. Display Error Correction Model
4. Compute forecasts from the $ARDL$ model

Option 0 returns you to the $ARDL$ Order Selection Menu (see Section 6.18.2).

Option 1 gives the estimated coefficients of the $ARDL$ model, together with the associated summary and diagnostic statistics. This option also allows you to make use of all the options available under the OLS method, for hypothesis testing, plotting fitted values, residuals, leverage measures, and so on (See Section 6.6).

Option 2 presents you with a table giving the estimates of the long-run coefficients, their asymptotic standard errors, and the associated t-ratios. The orders $\hat{p}, \hat{q}_1, \hat{q}_2, ..., \hat{q}_k$ selected for the underlying $ARDL$ model are also specified at the top of the table.

Option 3 displays a result table containing the estimates of the error correction model (ECM) associated with the selected $ARDL$ model. These estimates are computed using the relations in Section 21.19, and allow for possible parametric restrictions that may exist across the long-run and the short-run coefficients. The estimated standard errors also take account of such parametric restrictions.

Option 4 computes forecasts based on the selected $ARDL$ model, and asks you whether you wish to see forecasts of the levels or the first-differences of y_t. You will then be presented with the $ARDL$ Forecast Menu (see Section 6.18.4).

6.18.4 ARDL Forecast Menu

This menu appears on the screen if you choose option 4 in the Post *ARDL* Model Selection Menu (see Section 6.18.3). It contains the following options

0. Choose another variable
1. Display forecasts and forecast errors
2. Plot of in-sample fitted values and out of sample forecasts
3. Save in-sample fitted values and out of sample forecasts

Option 0 enables you to alter your choice of the levels or the first-differences of y_t that you may wish to forecast.

Option 1 displays the forecasts and the forecast errors for y_t (or Δy_t) computed on the basis of the selected *ARDL* model. It also provides a number of summary statistics computed both for the estimation and the prediction periods.

Option 2 plots the actual values of y_t (or Δy_t), and the fitted and forecast values of y_t (or Δy_t) over the estimation and the forecast periods, respectively.

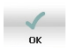
Option 3 allows you to save the fitted values of y_t (or Δy_t) over the estimation period, and the forecast values of y_t (or Δy_t) over the forecast period.

6.19 Logit and Probit models

The Logit and Probit options are appropriate when the dependent variable, y_i, $i = 1, 2, ..., n$ takes the value of 1 or 0. In econometrics such models naturally arise when the economic agents are faced with a choice between *two* alternatives. (For example, whether to use public transportation, or to purchase a car), and their choice depends on a set of k explanatory variables or factors. The models are also referred to as 'qualitative' or ' limited dependent' variable models. In the biological literature they are known as 'quantal variables', or as 'stimulus and response models'.

Comprehensive surveys of the literature on binary response models can be found in McFadden (1976) and Amemiya (1981). Other useful references are Maddala (1983), Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) Chapter 18, Cramer (1991), and Greene (2002) Chapter 19.

6.19.1 Specification of the Logit/Probit model

To access the Logit and Probit estimation options choose option 7 in the Single Equation Estimation Menu (see Section 6.4). You will then be asked to list the dependent variable, y_i, followed by the regressors (or the explanatory) variables, $x_{i1}, x_{i2}, ..., x_{ik}$. The dependent variable must contain only ones and zeros. The explanatory variables could contain both continuous and discrete variables. Once you have completed the specification of the model you will be asked to select a sample period for estimation. If the dependent variable in your model contains values other than ones and zeros you will be presented with an error message to that effect; click 

to continue. This will take you to the Backtracking Menu for the Logit/Probit Estimator.

6.19.2 Logit/Probit Estimation Menu

The Logit/Probit Estimation Menu contains the following options

1. Logit
2. Probit

Option 1 computes ML estimates of the coefficients assuming the logistic probability model

$$\Pr(y_i = 1) = \Lambda(\boldsymbol{\beta}'\mathbf{x}_i) = \frac{e^{\boldsymbol{\beta}'\mathbf{x}_i}}{1 + e^{\boldsymbol{\beta}'\mathbf{x}_i}}, \quad i = 1, 2, ..., n \tag{6.25}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_k)'$ is the $k \times 1$ vector of unknown coefficients, and $\mathbf{x}_i$ is a $k \times 1$ vector of explanatory variables, possibly containing a vector of ones (the intercept term). The effect of a unit change in the jth element of $\mathbf{x}_i$ on $\Pr(y_i = 1)$ is given by

$$\frac{\partial \Pr(y_i = 1)}{\partial x_{ij}} = \beta_j \Lambda_i (1 - \Lambda_i), \quad \text{for} \quad j = 1, 2, ..., k \quad \text{and} \quad i = 1, 2, ..., n \tag{6.26}$$

where $\Lambda_i = \Lambda(\boldsymbol{\beta}'\mathbf{x}_i)$. The ML estimation is carried out using the iterative method of Scoring (see (21.182)).

Option 2 computes ML estimates of the coefficients assuming the normal probability model

$$\Pr(y_i = 1) = \Phi(\boldsymbol{\gamma}'\mathbf{x}_i) = \int_{-\infty}^{\boldsymbol{\gamma}'\mathbf{x}_i} \frac{1}{\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2}t^2\right\} dt, \tag{6.27}$$

where as in option 1 above, $\boldsymbol{\gamma}$ is a $k \times 1$ vector of unknown coefficients and $\mathbf{x}_i$ is a $k \times 1$ vector of explanatory variables. In the case of this option the effect of a unit change in the jth element of $\mathbf{x}_i$ on $\Pr(y_i = 1)$ is given by

$$\frac{\partial \Pr(y_i = 1)}{\partial x_{ij}} = \beta_j \phi(\boldsymbol{\beta}'\mathbf{x}_i), \quad \text{for} \quad j = 1, 2, ..., k, \quad \text{and} \quad i = 1, 2, ..., n, \tag{6.28}$$

where $\phi(\cdot)$ stands for the standard normal density

$$\phi(\boldsymbol{\beta}'\mathbf{x}_i) = (2\pi)^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2}(\boldsymbol{\beta}'\mathbf{x}_i)^2\right\}$$

6.19.3 Estimation results for Logit and Probit options

The estimation results for both the Logit and Probit options are set out in a table with two parts. The top part gives the ML estimates of coefficients together with their (asymptotic) standard errors and the t-ratios. The bottom part of the table gives the factor

$$\Lambda\left(\hat{\boldsymbol{\beta}}'\bar{\mathbf{x}}_i\right)\left(1 - \Lambda(\hat{\boldsymbol{\beta}}'\bar{\mathbf{x}})\right)$$

or

$$\phi\left(\hat{\boldsymbol{\beta}}'\bar{\mathbf{x}}_i\right)$$

needed to compute the marginal effects (6.26) and (6.28) for different coefficients evaluated at sample means, a number of summary statistics, test statistics and model selection criteria. Here $\bar{\mathbf{x}}$ refers to the sample mean of the regressors. The details of these are given in Section 21.20.3. Since under both probability models the log-likelihood function is concave, the computations usually converge very quickly to the unique ML estimators (when they exist). Also note that the variances of the logistic and normal distributions that underlie the Logit and Probit options differ and are given by $\pi^2/3$ and 1, respectively. As a result, to ensure comparability of the ML estimates obtained under these two options the ML estimates using the Probit option must be multiplied by $\pi/\sqrt{3} \approx 1.814$ to make them comparable with those computed using the Logit option.

The problem of choosing between the Probit and Logit models can be approached either by application of the model selection criteria such as Akaike information criterion or the Schwartz Bayesian criterion, or by means of non-nested hypothesis testing procedures. In cases where the Logit and Probit options are used on the same set of regressors, $\mathbf{x}_i$ the applications of the various model selection criteria reduced to a simple comparison of the maximized log-likelihood values. In practice these log-likelihood values will be quite close, particularly if the estimates of $\boldsymbol{\beta}'\mathbf{x}_i$ (or $\boldsymbol{\gamma}'\mathbf{x}_i$) lie in the range $(-1.6, 1.6)$. In these circumstances the application of the non-nested testing methods is more appropriate: see Pesaran and Pesaran (1993).

6.19.4 Logit/Probit Post Estimation Menu

You will be presented with this menu when you have finished with the estimation results tables (see above). This menu has the following options:

0. Quit Logit/Probit Estimation
1. Display results again
2. List actual and fitted values, and fitted probabilities
3. Plot actual values and fitted probabilities
4. Save fitted probabilities (and forecasts if any)
5. Wald test of linear/non-linear restrictions
6. Estimate/test functions of parameters of the mode
7. Compute forecasts

Option 0 takes you back to the Commands and Data Transformations box (see Section 6.19.2).

Option 1 enables you to see the ML estimation results again (see Section 6.19.3).

Option 2 lists actual values (y_i), the fitted values ($\hat{y}_i$), and the fitted probability values $\Phi(\widehat{\boldsymbol{\beta}}'\mathbf{x}_i)$ and $\Lambda(\widehat{\boldsymbol{\beta}}'\mathbf{x}_i)$ for the Probit and Logit options, respectively. (See Section 21.20.2).

Option 3 plots actual values (y_i) and the fitted values, $\Phi(\widehat{\boldsymbol{\beta}}'\mathbf{x}_i)$ and $\Lambda(\widehat{\boldsymbol{\beta}}'\mathbf{x}_i)$ for the Probit and Logit options, respectively.

Option 4 allows you to save fitted probability values and their forecasts (if any). Use Option 7 below to compute forecasts of probability values.

Option 5 enables you to carry out Wald tests for linear and non-linear restrictions on the coefficients $\boldsymbol{\beta}$. Also see Option 7 in the Hypothesis Testing Menu (see Section 6.23). For the relevant formulae see Section 21.25.

Option 6 allows you to estimate linear and non-linear functions of the coefficients $\boldsymbol{\beta}$. Also see Option 5 in the Post Regression Menu in Section 6.20. For the relevant formulae see Section 21.24.

Option 7 computes forecasts of the probability values and the associated forecasts of y using the formulae in Section 21.20.4. This option also gives a number of summary statistics computed over the estimation and forecast periods.

6.20 Post Regression Menu

This menu appears on the screen immediately after the estimation results for the single equation linear and the non-linear estimation options. The Probit and Logit estimation option has its own Post Estimation Menu. (See Section 6.19.4). The Post Regression Menu contains the following options

0. Move to Backtracking Menu
1. Display regression results again
2. Move to Hypothesis Testing Menu
3. List/plot/save residuals and fitted values
4. White and Newey-West adjusted variance menu
5. Estimate/Test (possibly non-linear) functions of parameters
6. Plot the leverage measures of the regression (OLS)
7. Save the leverage measures of the regression (OLS)
8. Forecast
9. Plot of forecast values only

These options enable you to study the properties of your specified regression equation in more detail. The highlighting in this menu is initially placed on option 2 in the case of the least squares and the IV methods; otherwise the highlighting is placed on option 1.

Option 0 takes you back to the Commands and Data Transformations box, where a new estimation period and/or a regression equation can be specified.

Option 1 enables you to see your regression results again.

Option 2 takes you to the Hypothesis Testing Menu (see Section 6.23)

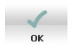
Option 3 takes you to the Display/Save Residuals and Fitted Values Menu (see Section 6.21).

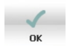
Option 4 takes you to the menu for the computation of alternative estimators of the variance matrices (see Section 6.22).

Option 5 allows you to estimate linear or non-linear functions of the parameters of your regression model. In the case of linear regressions *Microfit* assigns $A1$, $A2$, ..., to the regression coefficients and $B1$, $B2$,... to the parameters of the AR/MA error processes. For non-linear regression *Microfit* works directly in terms of the parameter names that you have specified. When you choose this option you will be asked to type your functions one at a time in the box editor that appears on the screen, separating the functions by a semicolon (;). The program computes and displays the estimates of the functions and the estimates of their (asymptotic) variance-covariance matrix. The relevant formula for the variance-covariance matrix of the parameter estimates is given in Section 21.24.

Option 6 provides plots of the measures of the leverage (or the influence) of points in the regression design, together with a horizontal line representing the average value of the leverage measures (see Section 6.6.7 for more details and relevant references to the literature.)

Option 7 allows you to save the leverage measures for subsequent analysis

Option 8 computes static or dynamic forecasts of the dependent variable conditional on the observed values of the regressors over the forecast period, if any, together with forecast errors, and the standard errors of the forecast. Dynamic forecasts will be computed if the lagged value of the dependent variable are explicitly included among the regressors (see note 4 in Section 6.5.1). When you choose this option you will be asked to specify the final observation in your forecast period. Type in the observation number, or the relevant date, and click 

. To choose all the available observations in the forecast period you only need to click 

(see Section 21.26 for details of the computations).

Option 9 provides a plot of actual and forecast values. The emphasis in this plot is on the forecast values, and in contrast to the plot provided under Option 8, it does not, in general, cover the whole of the estimation period.

6.21 Display/Save Residuals and Fitted Values Menu

This is a sub-menu of the Post Regression Menu and contains the following options

0. Return to Post Regression Menu
1. List residuals and fitted values
2. Plot actual and fitted values
3. Plot residuals
4. Plot the autocorrelation function and the spectrum of residuals
5. Plot the histogram of residuals
6. Save residuals (and forecast errors if any)
7. Save fitted values (and forecasts if any)

Option 0 takes you back to the Post Regression Menu (see Section 6.20).

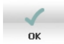
Option 1 displays on the screen the residuals and fitted values together with the actual values of the dependent variable. In the case of the AR and MA options (options 3 and 8 in the Linear Regression Menu), adjusted residuals and fitted values are reported.

Option 2 provides a plot of actual and (adjusted) fitted values.

Option 3 provides a plot of (adjusted) residuals, together with a standard error band. The band represents $\pm 2\hat{\sigma}_\epsilon$, where $\hat{\sigma}_\epsilon$ is the estimated standard error of the regression.

Option 4 displays graphs of the autocorrelation function and the standardized spectral density function of the residuals estimated using the Parzen window. To obtain estimates that utilize other windows or to compute the standard errors of these estimates, you can apply the **COR** and the **SPECTRUM** commands to the residuals, after saving them using option 6 in this menu.

Option 5 displays a histogram of the residuals. If you wish to produce a histogram with a different number of bands, save the residuals using Option 6 in this menu and then apply the **HIST** command to the saved residuals at the data processing stage.

Option 6 allows you to save the residuals and forecast errors (if any) in a variable for use in subsequent analyses. When you select this option, you will be asked to specify a variable name for the residuals to be saved. Type in the variable name followed by an optional description and click .

Option 7 allows you to save fitted and forecast values (if any) in a variable for use in subsequent analyses (see previous option above in this menu for more details).

6.22 Standard, White and Newey-West Adjusted Variance Menu

This menu has the following options

0. Return to Post Regression Menu
1. Standard variance-covariance matrix
2. White heteroscedasticity adjusted
3. Newey-West adjusted with equal weights
4. Newey-West adjusted with Bartlett weights
5. Newey-West adjusted with Tukey weights
6. Newey-West adjusted with Parzen weights

Option 0 returns you to the Post Regression Menu.

Option 1 displays the conventional variance-covariance matrix of the estimated coefficients. This option applies to all the methods available for the estimation of linear and non-linear regression models. It also computes and displays the covariance matrix of the regression coefficients and the parameters of the AR and MA error processes.

Option 2 computes and displays a 'degree of freedom adjusted' version of White (1980) and White (1982) heteroscedasticity-consistent estimates of the variance-covariance matrix of the parameter estimates in the case of the OLS, the IV, the non-linear least squares and the non-linear IV options. (See Section 21.22 for the relevant formulae). If you click , you will be presented with the following choices:

0 Return to White and Newey-West Adjusted Variance Menu
1. Display regression results for the adjusted covariance matrix
2. Display the adjusted covariance matrix
3. Wald test of restrictions based on adjusted covariance matrix
4. Estimate/test functions of parameters based on adjusted matrix

The options in this sub-menu allow you to test hypotheses on the regression coefficients using the heteroscedasticity-consistent estimates of the variance-covariance matrices.

Options 3 to **6** compute a 'degree of freedom adjusted' version of the Newey and West (1987) heteroscedasticity and autocorrelation consistent estimates of the variance-covariance matrix of the parameter estimates in the case of the linear and non-linear least squares and IV options for different choices of lag windows (see Section 21.23). Newey and West use the Bartlett weights, but in general the Parzen weights are preferable. The 'equal weights' options is relevant when the residual serial correlation can be approximated by a finite order MA process. When you choose any of these options you will be prompted to specify the size of the lag window. We recommend that you do not specify a window size which is in excess of one third of the available observations. *Microfit* then computes and displays the estimates of the Newey-West adjusted variance-covariance matrices. If you click you will be presented with the same choices as in the case of option 2 (White heteroscedasticity adjusted variance-covariance matrix) set out above.

Note: The formula for the White standard errors is a special case of the Newey-West formula, and can also be obtained using Options 3 to 6 by setting the window size equal to zero.

6.23 Hypothesis Testing Menu

This menu contains the following options

0. Return to Post Regression Menu
1. LM tests for serial correlation (OLS, IV, NLS, & $IV - NLS$)
2. Autoregressive conditional heteroscedasticity tests (OLS & NLS)
3. Unit root tests for residuals (OLS & NLS)
4. $CUSUM$ and $CUSUMSQ$ tests (OLS)
5. Variable deletion test (OLS & IV)
6. Variable addition test (OLS & IV)
7. Wald test of linear/non-linear restrictions
8. Non-nested tests against another linear regression (OLS)
9. Non-nested tests by simulation for log-linear ratios etc (OLS)

and allows you to subject your chosen linear and non-linear regression model to additional tests.

Option 0 takes you back to the Post Regression Menu (see Section 6.20)

Option 1 allows you to carry out a pth order test of residual serial correlation ($p \leq 12$). In the case of the least squares option (OLS and NLS) it provides (asymptotic) t-ratios for individual coefficients of the AR error process as well as the LM, the F, and the log-likelihood ratio statistics. When this option is chosen you will be asked to specify the order of the test. Type in your answer (an integer between 1 and 12) and click . In the case of the least squares options, the program computes the LM and the F-version of Godfrey's test statistic given respectively by equations (21.19) and (21.21). For the IV options, the program computes Sargan's test statistic given by (21.69).

Option 2 allows you to compute the autoregressive-conditional heteroscedasticity ($ARCH$) test statistic due to Engle (1982). For an ($ARCH$) test of order p, the program computes the LM statistic for the test of $\delta_i = 0$, $i = 1, 2, ..., p$ in the auxiliary regression

$$e_t^2 = \text{intercept} + \sum_{i=1}^{p} \delta_i e_{t-i}^2 + \text{Error}$$

estimated over the period $t = p + 1,\ \ p + 2, ..., m$, where e_t are the OLS residuals. See also Section 23.1.7.

Option 3 allows you to carry out the Dickey-Fuller and Augmented Dickey-Fuller tests of the unit root hypothesis in the residuals (see the **ADF** command for more details).[13] This test has been discussed in Engle and Granger (1987) and Engle and Yoo (1987) as a test of cointegration. The program also displays the 95 per cent critical values, using the results in MacKinnon (1991). Note that these critical values differ from those supplied with the **ADF** command, and depend on the number of $I(1)$ variables in the underlying regression (excluding the intercept term and the time trend), and whether or not the regression model includes a time trend. *Microfit* checks the regressions and reports the correct critical values. If an intercept term is not included in the original regression, a warning is displayed. It is important to note that in view of the above considerations, a direct application of the ADF command to saved residuals will generate incorrect critical values for the test, and must be avoided.

Option 4 enables you to carry out the cumulative sum ($CUSUM$) and the $CUSUM$ of squares ($CUSUMSQ$) tests of structural stability proposed by Brown, Durbin, and Evans (1975). When you choose this option, *Microfit* displays two graphs, one giving the plot of the $CUSUM$ statistic (21.131), and the other giving the plot of the $CUSUMSQ$ statistic (21.133). Each graph also displays a pair of straight lines drawn at the 5 per cent level of significance defined by equations (21.132) and (21.134), respectively. If either of the lines is crossed, the null hypothesis that the regression equation is correctly specified must be rejected at the 5 per cent level of significance. The $CUSUM$ test is particularly useful for detecting systematic changes in the regression coefficients, and the $CUSUMSQ$ test is useful in situations where the departure from the constancy of the regression coefficients is haphazard and sudden.

Option 5 enables you to test for the statistical significance of deleting one or more regressors from your linear regression model.

Option 6 enables you to test the statistical significance of adding one or more regressors to your linear regression model.

Option 7 allows you to carry out a Wald test of linear or non-linear restrictions on the parameters of your model. When you choose this option you will first be prompted to specify the number of the restrictions that you wish to test, and then the restrictions themselves.

Notes

[13] Unit roots tests can also be applied to the residuals directly by making use of the ADF commands in the Processing Window. See Section 4.4.2 for more information.

1. Restrictions must be linearly independent and should not exceed the number of the unknown parameters in your model.

2. In the case of linear regressions, *Microfit* assigns $A1$, $A2$, ... to the regression coefficients and $B1$, $B2$, ... to the parameters of the AR/MA error processes. For example, to test the hypothesis that in the regression of C on $INPT$, Y, $Y(-1)$, $C(-1)$ the long run response of C to Y is equal to 1, you need to specify either $(A2 + A3) / (1 - A4) = 1$, or $A2 + A3 + A4 = 1$. Both are mathematically equivalent, and in large samples give the same results. In small samples, however, they could give very different results (see Gregory and Veall 1985, 1987). The linear form of the restriction is preferable and should be used in practice.

3. Another method of testing restrictions would be to use option 5 in the Post Regression Menu (see Section 6.20).

4. The relevant expression for the test statistic is given by equation (21.198).

Option 8 enables you to compute a number of test statistics proposed in the literature for the test of non-nested linear regression models. This option also computes a number of useful summary statistics, including Akaike and Schwartz Bayesian information criteria (see Section 21.8 for details, and for relevant references to the literature).

Option 9 allows you to carry out non-nested tests of the following linear regression models :

$$\begin{aligned} M_1 &: \mathbf{f}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}_1 + \mathbf{u}_1, \qquad \mathbf{u}_1 \sim N(0, \sigma^2\mathbf{I}_n), \\ M_2 &: \mathbf{g}(\mathbf{y}) = \mathbf{Z}\boldsymbol{\beta}_2 + \mathbf{u}_2, \qquad \mathbf{u}_2 \sim N(0, \omega^2\mathbf{I}_n), \end{aligned}$$

where $\mathbf{f}(\mathbf{y})$ and $\mathbf{g}(\mathbf{y})$ are known transformations of the $n \times 1$ vector of observations on the underlying dependent variable of interest, $\mathbf{y}$, and $\mathbf{X}$ and $\mathbf{Z}$ are $n \times k_1$and $n \times k_2$, observation matrices for the models M_1 and M_2, respectively. In what follows we refer to $\mathbf{f}(\mathbf{y})$ and $\mathbf{g}(\mathbf{y})$ as the right-hand-side (RHS) variables.

When you choose this option, you will be prompted to list the regressors of model M_2. The currently specified regression equation will be treated as model M_1. Then you will be presented with the following menu, asking you to specify the nature of the transformation of the dependent variable for model M_1:

0. Move to Hypothesis Testing Menu
1. Linear form
2. Logarithmic form
3. Ratio form
4. Difference form
5. Log-difference form
6. General non-linear form to be specified by you

Options 1 to **5** allow you to specify the following transformations of the dependent variable under model M_1:

Linear form:	$\mathbf{f}(\mathbf{y}) = \mathbf{y}$
Logarithmic form:	$\mathbf{f}(\mathbf{y}) = \log(\mathbf{y})$
Ratio form:	$\mathbf{f}(\mathbf{y}) = \mathbf{y}/\mathbf{z}$
Difference form:	$\mathbf{f}(\mathbf{y}) = \mathbf{y} - \mathbf{y}(-1)$
Log-difference form:	$\mathbf{f}(\mathbf{y}) = \log \mathbf{y} - \log \mathbf{y}(-1)$

where $\mathbf{z}$ is another variable on the workspace. Notice that $\log(\mathbf{y})$ refers to a vector of observations with elements equal to $\log(y_t)$, $t = 1, 2, ..., n$. Also $\mathbf{y} - \mathbf{y}(-1)$ refers to a vector with a typical element equal to $y_t - y_{t-1}$, $t = 1, 2, ..., n$.

Option 6 in this sub-menu allows you to specify your own particular functional form for $\mathbf{f}(\mathbf{y})$. See note 4 below for more details.

Once one of these options is chosen, the program presents you with a similar menu to identify the functional form, $\mathbf{g}(\mathbf{y})$, for the RHS variable under model M_2. Having specified the functional forms for the dependent variables of the two models, the program asks you to give the number of replications, R, to be used in the simulations, and computes the following test statistics: P_E test statistic due to tciteNMacKinnon1983, the Bera and McAleer (1989) test statistic, the double-length regression test due to Davidson and MacKinnon (1984), and the simulated Cox test statistic, SC_c, proposed in Pesaran and Pesaran (1995). Furthermore, it reports Sargan (1964) and Vuong (1989) likelihood criteria for the choice between the three models. See Section 21.9 for the details. Using this option it is possible, for example, to test linear versus log-linear models, first-difference models versus models in log-differences, models in ratio forms against models in logarithms.

Notes

1. In the case of testing linear versus log-linear models (or first-difference versus log-difference models) the program first computes the probability of drawing a negative value of y under the linear model, and displays a warning if this probability is larger than 0.0001. In such an event, the Cox statistic for testing the log-linear versus the linear model cannot be computed, and the program only computes the Cox statistic for testing the linear versus the log-linear model. See Pesaran and Pesaran (1995).

2. The results are displayed in two separate screens. The first screen gives the *OLS* estimates of models M_1 and M_2, and the quasi-maximum likelihood estimators of the parameters of model M_1 and M_2, and vice versa. The different test statistics are displayed in the subsequent screen.

3. Our experience suggests that for most problems 150 to 200 replications should be enough for achieving accuracies of up to two decimal places in the computation of the simulated Cox statistic. Nevertheless, we recommend that you try different numbers of replications to check the robustness of the results.

4. You need to choose option 6 when the functional forms $\mathbf{f}(\mathbf{y})$ or $\mathbf{g}(\mathbf{y})$ are not among the menu choices. When you choose this option you will be presented with a box editor,

asking you to specify the functional forms for the RHS variable, their inverse functions, and their derivatives, for each of the non-nested models separately. You need to provide the required information first for model M_1 and then for model M_2, separating them by a semicolon (;). For example, suppose you wish to specify the functions

$$\begin{aligned} f(y_t) &= \frac{y_t - y_{t-1}}{z_t} \\ g(y_t) &= \log(y_t/z_t) \end{aligned}$$

for the non-nested regression models M_1 and M_2, respectively. You need to type

$$\begin{aligned} F &= (Y-(Y(-1))/Z; \quad Y = Z*F+Y(-1); \quad DFY = 1/Z; \\ G &= \mathbf{LOG}(Y/Z); \quad Y = Z*\mathbf{EXP}(G); \quad DGY = 1/Y \end{aligned}$$

Notice that the variables F, Y, G, Z should exist on *Microfit*'s workspace, but the variables DFY and DGY should not exist.

Chapter 7

Multiple Equation Options

This chapter deals with the multiple (system) equation options in *Microfit*, and covers estimation, hypothesis testing and forecasting in the context of unrestricted Vector Autoregressive (VAR) models, Seemingly Unrelated Regression Equations ($SURE$), and cointegrating VAR models. The chapter also shows how to compute/plot orthogonalized and generalized impulse response functions, forecast error variance decompositions, and persistence profiles for the analysis of the effect of system-wide shocks on the cointegrating relations. There are also important new options for long-run structural modelling, enabling the user to estimate and test models with multiple cointegrating relations subject to general linear (possibly) non-homogeneous restrictions. The details of the econometric methods and the computational algorithms that underlie the multivariate options are set out in Chapter 22, where references to the literature can also be found.

7.1 The canonical multivariate model

The multivariate estimation options in *Microfit* are all based on the following augmented vector autoregressive model of order p, or $AVAR(p)$ for short:

$$\mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \sum_{i=1}^{p} \mathbf{\Phi}_i \mathbf{z}_{t-i} + \mathbf{\Psi}\mathbf{w}_t + \mathbf{u}_t \tag{7.1}$$

where $\mathbf{z}_t$ is an $m \times 1$ vector of jointly determined (endogenous) variables, t is a linear time trend, $\mathbf{w}_t$ is a $q \times 1$ vector of exogenous variables, and $\mathbf{u}_t$ is an $m \times 1$ vector of unobserved disturbances assumed to satisfy the following assumptions:[1]

B1 Zero Mean Assumption. The $m \times 1$ vector of disturbances, $\mathbf{u}_t$, have zero means

$$E(\mathbf{u}_t) = \mathbf{0}, \qquad \text{for} \quad t = 1, 2, ..., n$$

[1] These assumptions are the multivariate generalizations of those underlying the univariate classical linear regression model described in Section 6.1.

B2 Homoscedasticity Assumption. The $m \times 1$ vector of disturbances, $\mathbf{u}_t$, has a time-invariant conditional variance matrix

$$E\left(\mathbf{u}_t \mathbf{u}_t' \,| \mathbf{z}_{t-1}, \mathbf{z}_{t-2}, ..., \mathbf{w}_t, \mathbf{w}_{t-1},\ ...\right) = \mathbf{\Sigma}$$

where $\mathbf{\Sigma} = (\sigma_{ij})$ is an $m \times m$ symmetric positive definite matrix.

B3 Non-autocorrelated Error Assumption. The $m \times 1$ vector of disturbances, $\mathbf{u}_t$, are serially uncorrelated

$$E\left(\mathbf{u}_t \mathbf{u}_s'\right) = 0 \qquad \text{for all} \qquad t \neq s$$

B4 Orthogonality Assumption. The $m \times 1$ vector of disturbances, $\mathbf{u}_t$, and the regressors $\mathbf{w}_t$ are uncorrelated

$$E\left(\mathbf{u}_t \,| \mathbf{w}_t\right) = 0 \qquad \text{for all } t$$

B5 Stability Assumption. The augmented $VAR(p)$ model (7.1) is stable. That is, all the roots of the determinantal equation

$$|\ \mathbf{I}_m - \mathbf{\Phi}_1 \lambda - \mathbf{\Phi}_2 \lambda^2 - \cdots - \mathbf{\Phi}_p \lambda^p\ | = 0 \tag{7.2}$$

fall outside the unit circle.

For maximum likelihood estimation, the following normality assumption is also needed

B6 Normality Assumption. The $m \times 1$ vector of disturbances has a multivariate normal distribution.

The VAR specification is chosen for its flexibility and computational ease, and it is 'hoped' that by choosing p (the order of the VAR) high enough, the residual serial correlation problem can be avoided. The conditional homoscedasticity assumption, B3, is likely to be violated in the case of financial time series at monthly or higher frequencies. See Chapter 8 for the use of multivariate $GARCH$ models. Assumption B2 allows for contemporaneous correlation across the errors in different equations, and therefore also accommodates instantaneous feedbacks between the different variables in $\mathbf{z}_t$.

The canonical multivariate model (7.1) also forms the basis of the Seemingly Unrelated Regression Equations ($SURE$) and the restricted $SURE$ options in *Microfit*. The general $SURE$ model results when one allows for different lag orders and/or exogenous variables in different equations in (7.1) (see Section 7.7).

Finally, the cointegrating VAR options discussed in Section 7.5 are based on equation (7.1) but allows one or more roots of the determinantal equation (7.2) to fall on the unit circle.

7.1.1 The log-likelihood function of the multivariate model

The various multivariate estimation options in *Microfit* compute maximum likelihood (ML) estimators of the parameters of (7.1) subject to appropriate parametric restrictions. The log-likelihood function of (7.1), conditional on $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n$ and the initial values, $\mathbf{z}_0, \mathbf{z}_{-1}, ..., \mathbf{z}_{-p+1}$, is given by

$$\ell_n(\boldsymbol{\varphi}) = \frac{-nm}{2} \log 2\pi - \frac{n}{2} \log |\boldsymbol{\Sigma}| - \tfrac{1}{2} \sum_{t=1}^{n} \mathbf{u}_t' \boldsymbol{\Sigma}^{-1} \mathbf{u}_t, \tag{7.3}$$

where $\boldsymbol{\varphi}$ stands for all the unknown parameters of the model. Stacking the n observations on the m equations in (7.1), the log-likelihood in (7.3) can also be written in matrix form as

$$\ell_n(\boldsymbol{\varphi}) = \frac{-nm}{2} \log 2\pi - \frac{n}{2} \log |\boldsymbol{\Sigma}| - \tfrac{1}{2} Tr\left(\boldsymbol{\Sigma}^{-1} \mathbf{U}\mathbf{U}'\right), \tag{7.4}$$

where $Tr(\cdot)$ denotes the trace of a matrix, and

$$\mathbf{U} = \mathbf{Z} - \boldsymbol{\tau}_n \mathbf{a}_0' - \mathbf{t}_n \mathbf{a}_1' - \sum_{i=1}^{p} \mathbf{Z}_{-i} \boldsymbol{\Phi}_i' - \mathbf{W}\boldsymbol{\Psi}' \tag{7.5}$$

$$\underset{n \times m}{\mathbf{U}} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n)', \qquad \underset{n \times m}{\mathbf{Z}} = (\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n)' \tag{7.6}$$

$$\underset{n \times m}{\mathbf{Z}_{-i}} = (\mathbf{z}_{-i+1}, \mathbf{z}_{-i+2}, ..., \mathbf{z}_{-1+n})', \qquad \underset{n \times q}{\mathbf{W}} = (\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n)' \tag{7.7}$$

$$\underset{n \times 1}{\boldsymbol{\tau}_n} = (1, 1, ..., 1)', \qquad \underset{n \times 1}{\mathbf{t}_n} = (1, 2, ..., n)' \tag{7.8}$$

The particular computational algorithm used to carry out the maximization of the above log-likelihood function depends on the nature of the restrictions on the parameters of the model, and are set out in detail in Chapter 22. In this chapter our focus will be on how to use the various multiple estimation options.

7.2 General guidelines

Before proceeding any further, the following points are worth bearing in mind when using the multiple equation options in *Microfit*:

1. The order of the VAR, p, often plays a crucial role in the empirical analysis, and in selecting it special care must be taken to ensure that it is high enough so that the disturbances $\mathbf{u}_t$ in (7.1) are not serially correlated and, for the p chosen, the remaining sample for estimation is large enough for the asymptotic theory to work reasonably well. This involves a difficult balancing act. For VAR order selection *Microfit* automatically generates the Akaike information criterion (AIC) and the Schwarz Bayesian criterion (SBC), as well as a sequence of log-likelihood ratio statistics. In practice, their use often leads to different choices for p, and the user must decide on the best choice of p for the problem in hand.

2. The multivariate techniques are often highly data intensive, particularly when m, the number of jointly determined variables, is large. For example, when $m = 10$, $p = 4$, and $q = 2$, each equation in the $AVAR$ model (7.1) contains $s = mp + q + 2 = 44$ unknown coefficients. For such a specification our experience suggests that sample sizes of 200 or more is often needed if any sensible results are to be obtained. It is, therefore, important that the multivariate options in *Microfit* are applied in cases where n is sufficiently large.[2]

3. Cointegrating VAR options presume that the variables $\mathbf{z}_t$ are $I(1)$, *and* that the user already knows the nature of the *unconditional* mean of the variables in the underlying VAR model, namely whether the variables $\mathbf{z}_t$ have non-zero means or are trended, and whether the trend is linear. Therefore, it is important that the variables in $\mathbf{z}_t$ are tested for unit roots (for example using the **ADF** command in the Process window), and that the nature of the trends in $\mathbf{z}_t$ is ascertained, for example, by plotting each elements of $\mathbf{z}_t$ against time! Econometric techniques are often not powerful enough to identify the nature of the trends in the variables being modelled.

4. Before using the long-run structural modelling options described in Section 7.5.3, you need to specify the number of cointegrating (or long-run) relations of your model. The maximum eigenvalue and the trace statistics advanced by Johansen (1988) and Johansen (1991) can be employed for this purpose (see Section 7.5). However, the results of these tests are often ambiguous in practice, and it may be necessary that the number of cointegrating relations is chosen on the basis of *a priori* information; for example from the long-run predictions of a suitable economic model: see Pesaran (1997).

5. Another important issue in the use of long-run structural modelling options concerns the nature of the just-identifying restrictions on the long-run relations. This invariably requires an explicit formulation of the long-run economic model that underlies the empirical analysis. *Microfit* invites you to specify the cointegration or the long-run relations of your model at two different stages. In the first stage you will be asked to specify the cointegrating relations that are *just-identified*. Once such a just-identified model is successfully estimated, you will be prompted to specify your over-identifying restrictions (if any). For the ML estimation procedure to converge it is important that the over-identifying restrictions are introduced one at a time, starting with those that are less likely to be rejected. The asymptotic standard errors computed for the estimated coefficients of the exactly-identified cointegrating relations can be used as a guide in deciding which over-identifying restrictions to impose first, and which one to impose second, and so on.

6. Since the cointegration analysis focuses on the long-run properties of the economic model, it is important to combine it with some additional information on how the long-run relations of the model respond to shocks. For example, it may be of interest

[2] Size limitations in the case of the multivariate estimation options are set out in Appendix A.

to know whether there are over-shooting effects, and how long, on average, it will take for the economy to settle back into its long-run state after being shocked. To shed light on these and other related issues we recommend the use of the generalized impulse response functions for characterizing the time-profiles of the effects of variable-specific shocks on the long-run relations, and the persistence profiles for characterizing the effects of system-wide shocks on the cointegrating relations. See Sections 22.9.4, 22.9.5 and 22.9.6.

7.3 System Estimation Menu

All the multivariate options in *Microfit* can be accessed from the System Estimation Menu (the Multivariate Menu) or by clicking the 

button. When you use this button the currently selected menu option (Unrestricted VAR) is automatically selected. The Multivariate Menu options open the System Estimation Menu which contains the following options

1. Unrestricted VAR
2. Cointegrating VAR Menu
3. Cointegrating $VARX$
4. $SURE$ method
5. Restricted $SURE$ method
6. $2SLS$
7. Restricted $2SLS$
8. $3SLS$
9. Restricted $3SLS$

Option 1 enables you to estimate unrestricted VAR models, test a number of restrictions on their parameters and compute multivariate, multi-step ahead forecasts. You can also use this option to estimate univariate AR models. But if you wish to estimate univariate ARMA models you need to choose the MA option in the Linear Regression Estimation Menu. See Sections 6.5 and 6.12.

Option 2 enables you to carry out cointegration analysis in a VAR framework; distinguishing between $I(1)$ jointly determined variables, $I(1)$ exogenous variables, and $I(0)$ exogenous variables. You can choose among five different specifications of intercepts and/or trends in the underlying VAR model. This option also allows you to perform impulse response analysis and trend/cycle decompositions. See Sections 22.10.1, 22.11 and 22.9.4.

Option 3 enables you to estimate vector error correction models with weakly exogenous $I(1)$ variables ($VARX$). This option can also be used to perform impulse response analysis and trend/cycle decompositions. See Sections 22.10.1, 22.11 and 22.9.4.

Option 4 allows you to estimate a system of Seemingly Unrelated Regression Equations ($SURE$) by the full maximum likelihood method (see Zellner 1962).

Option 5 provides an important extension of the $SURE$ estimation method of option 3, and allows estimation of $SURE$ models subject to linear restrictions, possibly involving coefficients from different regression equations in the model. This option can be used, for example, to estimate systems of demand equations subject to the homogeneity and symmetry restrictions.

Options 6 and 7 enable you to estimate $SURE$ models and $SURE$ models subject to linear restrictions by two-stages least squares (see Zellner 1962).

Options 8 and 9 allow three-stages least squares estimation of unrestricted and restricted $SURE$ models. This method accounts for possible serial correlation of disturbances, and leads to more efficient estimates of regression coefficients than two-stages least squares when regression errors are autocorrelated (see Kmenta and Gilbert (1970) and Parks (1967)).

7.4 Unrestricted VAR option

Option 1 in the System Estimation Menu (Multivariate Menu: see Section 7.3) enables you to estimate the augmented $VAR(p)$ model defined by (7.1). When you choose this option you will be presented with Figure 7.1, which prompts you to list the jointly determined variables in the VAR, namely $\mathbf{z}_t$, followed by the deterministic/exogenous variables, namely intercepts, trend terms (if any) and possibly exogenous variables determined outside the VAR model, denoted by $\mathbf{w}_t$ in equation (7.1). It is possible to specify only one variable in $\mathbf{z}_t$. In this case an augmented *univariate* autoregressive model will be estimated. The two sets of variables must be separated by &.

For example, to specify a VAR model in the three variables

C	Real consumption expenditure
I	Real investment expenditure
Y	Real output

including in it an intercept ($INPT$) and a linear trend (T), you need to type

$$C \quad I \quad Y \quad \& \quad INPT \quad T$$

in the box editor shown in Figure 7.1.

Type in the start and the finish of your estimation period, and the order of the VAR model ($p \leq 24$), and then click 

. *Microfit* carries out the necessary computations and presents you with the Unrestricted VAR Post Estimation Menu. See Section 7.4.1.

Notes

1. The ordering of the variables in the VAR is important only as far as computation of the orthogonalized impulse responses and orthogonalized error variance decompositions are concerned. See Sections 22.5 and 22.6.

2. Lagged values cannot appear among the set of jointly determined variables, $\mathbf{z}_t$. Although they could be included in the second set, namely $\mathbf{w}_t$s, so long as these are not lagged values of the first set!

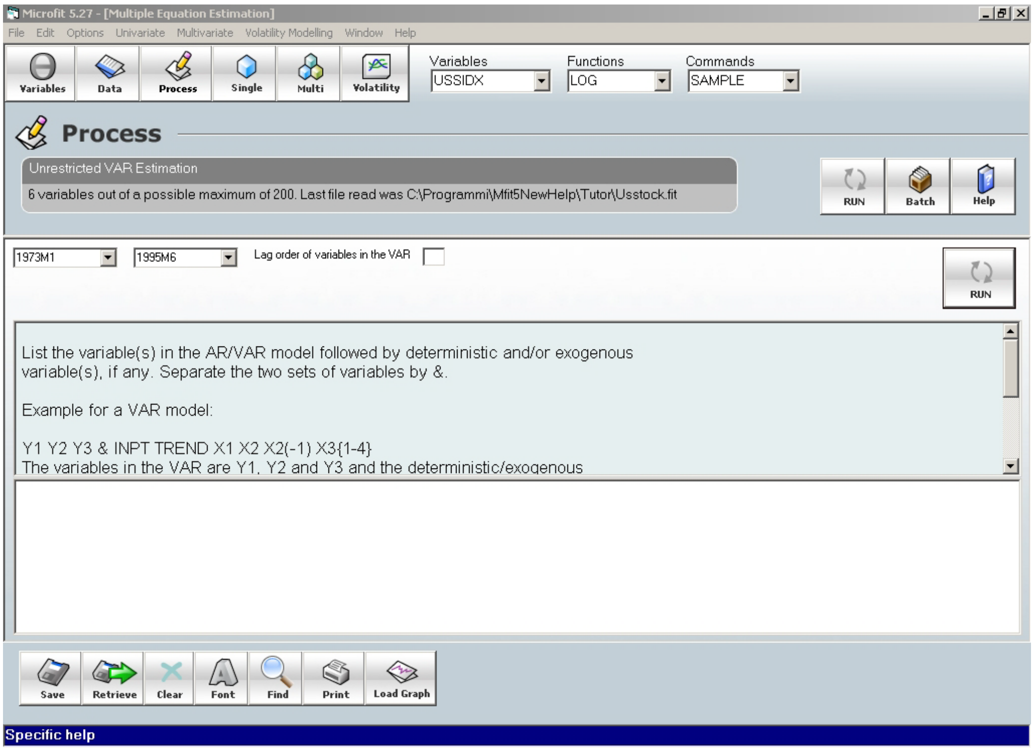


Figure 7.1: The System Estimation window with the unrestricted VAR option

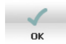
3. If a lagged variable, say $C(-1)$, is included as one of the variables in the VAR, *Microfit* will present you with an error message

4. If you include lagged values of the jointly determined variables as an exogenously determined variable, initially no error messages will appear on the screen, but at the stage of computations the program will encounter a perfect multicollinearity problem and issue an error message.

7.4.1 Unrestricted VAR Post Estimation Menu

The Unrestricted VAR Post Estimation Menu appears on the screen when the estimation of the VAR model is complete. It contains the following options

0. Move to Backtracking Menu
1. Display single equation estimation results
2. Display system covariance matrix of errors
3. Impulse Response and Forecast Error Variance Decomposition
4. Hypothesis testing and lag order selection in the VAR
5. Compute multivariate dynamic forecasts

Option 1 enables you to see estimation results for individual equations in the VAR model. When you choose this option you will be asked to select the equation to be displayed. Initially, the highlighting will be on the first variable in the VAR. Move the cursor to the variable whose equation you wish to inspect, and click . The estimation results for the selected equation should now appear on the screen. Since all the equations in the unrestricted VAR model have the same variables in common, the ML estimates of the VAR model is the same as the OLS estimates. See Section 22.4. Similarly, the summary and diagnostic statistics supplied with the estimation results are computed using the same formula as in the case of the OLS option described in Section 6.6. The only additional information provided is the maximized value of the system log-likelihood function given at the bottom right-hand corner of the first result table that appears on the screen. See equation (22.38). As in the case of the OLS option, after the estimation results you will be presented with the Post Regression Menu, which provides you with a number of options for further analysis of the residuals and hypothesis testing. For example, you can compute White and Newey-West adjusted standard errors for the parameter estimates or carry out tests of linear/non-linear restrictions on the coefficients of the chosen equation in the VAR. But to carry out tests of restrictions that involve parameters of different equations in the VAR model you need to choose the $SURE$ and restricted $SURE$ options described in Section 7.7.1 and 7.7.2.

Option 2 presents you with the unbiased estimates of $\mathbf{\Sigma}$, given by $\widehat{\mathbf{\Sigma}}$ in equation (21.12).

Option 3 takes you to the Unrestricted VAR Dynamic Response Analysis Menu, from where you can compute and plot orthogonalized and generalized impulse responses and forecast error variance decompositions for unit shocks to the ith equation in the VAR model (7.1). See Section 7.4.2.

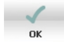
Option 4 moves you to the VAR Hypothesis Testing Menu, where you can select the order of the VAR, test for the statistical significance of the exogenous/deterministic variables, $\mathbf{w}_t$, and finally test for non-causality of a sub-set of jointly determined variables, $\mathbf{z}_t$, in the VAR model defined by (7.1). See Section 7.4.3.

Option 5 enables you to compute forecasts of the jointly determined variables, $\mathbf{z}_t$, in (7.1), for given values of the exogenous/deterministic variables, $\mathbf{w}_t$, over the forecast period. When you select this option you will be first asked to specify the forecast period, and then to choose the variable you wish to forecast. For each variable that you choose you will be given a choice of forecasting the levels of the variable or its first-differences. The program then computes the forecasts and presents you with the Multivariate Forecast Menu (see Section 7.4.4).

7.4.2 Unrestricted VAR Dynamic Response Analysis Menu

This menu has the following options

0. Return to VAR Post Estimation Menu
1. Orthogonalized IR of variables to shocks in equations
2. Generalized IR of variables to shocks in equations
3. Orthogonalized forecast error variance decomposition
4. Generalized forecast error variance decomposition

When you choose any one of the above options 1 to 4, you will be asked to choose the equation to be shocked. Each equation is designated by its left-hand-side variable. Move the cursor to the desired variable name (equation) and click . You will now be asked to specify the horizon (denoted in Chapter 22 by N) for the impulse responses (or forecast error variance decomposition). The default value is set to 50, otherwise you need to type your desired value of N and then click  (with $N \leq 150$). Once you have specified the horizon the program carries out the computations and presents you with a list of impulse responses (or forecast error variance decompositions) at different horizons. To plot or save the results click  to move to the Impulse Response Results Menu. This menu has the following options

0. Move back/bootstrap confidence intervals
1. Display results again
2. Graph
3. Save in a CSV file

Option 0 allows you to compute the empirical distribution of impulse responses for one or more variables by applying the bootstrap method. Choose the desired number of replications and the confidence level $(1-\alpha)$ and click , then select the variable you want to inspect. You will be presented with the list of impulse responses at different horizons for the selected variable, together with their bootstrapped $1-\frac{\alpha}{2}$ and $\frac{\alpha}{2}$ percentiles, their median and mean. Click  to return to the Impulse Response Results Menu.

Option 1 enables you to see the results of the impulse response analysis and forecast error variance decompositions again.

Option 2 enables you to plot the impulse responses (or the forecast error variance decompositions) for one or more of the variables in the VAR at different horizons. If you have previously used option 0 to compute the bootstrapped confidence intervals at a confidence level $(1-\alpha)$, you can also plot the mean, median, $1-\frac{\alpha}{2}$ and $\frac{\alpha}{2}$ percentiles of the impulse responses bootstrapped empirical distributions.

Option 3 allows you to save the impulse responses (or the forecast error variance decompositions) for all the variables in a CSV file for subsequent analysis.

It is worth noting that the orthogonalized impulse responses and the orthogonalized forecast error variance decompositions usually depend on the ordering of the variables in the VAR, but their generalized counterparts do not. The orthogonalized and the generalized impulse responses exactly coincide either for the first variable in the VAR or if $\mathbf{\Sigma}$ is diagonal. An account of these concepts and the details of their computation are set out in Sections 22.5 and 22.6.

7.4.3 VAR Hypothesis Testing Menu

This menu appears on the screen when option 4 in the Unrestricted VAR Post Estimation Menu is chosen. (see Section 7.4.1), and has the following options

0. Return to VAR Post Estimation Menu
1. Testing and selection criteria for order (lag length) of the VAR
2. Testing for deletion of deterministic/exogenous variables in the VAR
3. Testing for block non-causality of a subset of variables in the VAR

Option 0 returns you to the Unrestricted VAR Post Estimation Menu (see Section 7.4.1).

Option 1 computes Akaike information and Schwarz Bayesian model selection criteria for selecting the order of the $VAR(p)$, for $p = 0, 1, 2, ..., P$, where P represents the maximum order selected by the user (see Section 7.4). The selection procedure involves choosing the $VAR(p)$ model with the highest value of the AIC or the SBC. In practice, the use of SBC is likely to result in selecting a lower order VAR model, as compared to the AIC. But in using both criteria it is important that the maximum order chosen for the VAR is high enough, so that high-order VAR specifications are given a reasonable chance of being selected, if they happen to be appropriate. This option also computes log-likelihood ratio statistics and their small sample adjusted values which can be used in the order-selection process. The log-likelihood ratio statistics are computed for testing the hypothesis that the order of the VAR is p against the alternative that it is P, for $p = 0, 1, 2, .., P - 1$. Users interested in testing the hypothesis that the order of the VAR model is p against the alternative that it is $p + 1$, for $p = 0, 1, 2, ..., P - 1$, can construct the relevant log-likelihood statistics for these tests by using the maximized values of the log-likelihood function given in the first column of the result table corresponding to this option. For example, to test the hypothesis that the order of the VAR model is 2 against the alternative that it is 3, the relevant log-likelihood ratio statistic is given by

$$LR(2:3) = 2\left(LL_3 - LL_2\right) \tag{7.9}$$

where LL_p, $p = 0, 1, 2, ..., p$ refers to the maximized value of the log-likelihood function for the $VAR(p)$ model. Under the null hypothesis, $LR(2:3)$ is distributed asymptotically as a chi-squared variate with $m^2(3-2) = m^2$ degrees of freedom, where m is the dimension of $\mathbf{z}_t$ in equation (7.1). For further details and the relevant formulae see Section 22.4.1.

Option 2 computes the log-likelihood ratio statistic for testing zero restrictions on the coefficients of a sub-set of deterministic/exogenous variables in the VAR. For example, to test the hypothesis that the VAR specification in (7.1) does not contain a deterministic trend the relevant hypothesis will be $\mathbf{a}_1 = \mathbf{0}$. In general, this option can be used to test the validity of deleting one or more of the exogenous/deterministic variables from the VAR. When you choose this option you will be asked to list the deterministic/exogenous variable(s) to be dropped from the VAR model. Type in the variable name(s) in the box editor and click  to process. The test results should now appear on the screen; they give the maximized values of the log-likelihood function for the unrestricted and the restricted model, and the log-likelihood ratio statistic for testing the restrictions. The degrees of freedom and the rejection probability of the test are given in round ( ) and square [ ] brackets, respectively. For further details and the relevant formulae see Section 22.4.2.

Option 3 computes the log-likelihood ratio statistic for testing the null hypothesis that the coefficients of a sub-set of jointly determined variables in the VAR are equal to zero.

This is known as Block Granger Non-Causality test and provides a statistical measure of the extent to which lagged values of a set of variables (say $\mathbf{z}_{2t}$) are important in predicting another set of variables, (say $\mathbf{z}_{1t}$) once lagged values of the latter set are included in the model.

More formally, in (7.1), let $\mathbf{z}_t = (\mathbf{z}'_{1t}, \mathbf{z}'_{2t})$ where $\mathbf{z}_{1t}$ and $\mathbf{z}_{2t}$ are $m_1 \times 1$ and $m_2 \times 1$ sub-sets of $\mathbf{z}_t$, and $m = m_1 + m_2$. Consider now the following block decomposition of (7.1)

$$\mathbf{z}_{1t} = \mathbf{a}_{10} + \mathbf{a}_{11}t + \sum_{i=1}^{p} \mathbf{\Phi}_{i,11}\mathbf{z}_{1,t-i} + \sum_{i=1}^{p} \mathbf{\Phi}_{i,12}\mathbf{z}_{2,t-i} + \mathbf{\Psi}_1\mathbf{w}_t + \mathbf{u}_{1t}, \tag{7.10}$$

$$\mathbf{z}_{2t} = \mathbf{a}_{20} + \mathbf{a}_{21}t + \sum_{i=1}^{p} \mathbf{\Phi}_{i,21}\mathbf{z}_{1,t-i} + \sum_{i=1}^{p} \mathbf{\Phi}_{i,22}\mathbf{z}_{2,t-i} + \mathbf{\Psi}_2\mathbf{w}_t + \mathbf{u}_{2t}.$$

The hypothesis that the subset $\mathbf{z}_{2t}$ does not 'Granger-cause' $\mathbf{z}_{1t}$ is defined by

$$H_G : \mathbf{\Phi}_{12} = 0,$$

where $\mathbf{\Phi}_{12} = (\mathbf{\Phi}_{1,12}, \mathbf{\Phi}_{2,12} ..., \mathbf{\Phi}_{p,12})$. When you choose this option you will be asked to list the subset of variable(s) on which you wish to carry out the block non-causality test, namely $\mathbf{z}_{2t}$, in the above formulation. The program then computes the relevant log-likelihood ratio statistic and presents you with the test results, also giving the maximized log-likelihood values under the unrestricted ($\mathbf{\Phi}_{12} \neq \mathbf{0}$) and the restricted model ($\mathbf{\Phi}_{12} = \mathbf{0}$). For further details and the relevant formulae see Section 22.4.3. Note that the Granger non-causality tests may give misleading results if the variables in the VAR contain unit roots (namely when one or more roots of (22.34) lie on the unit circle). In such a case one must ideally either use VAR models in first differences, or cointegrating VAR models if the underlying variables are cointegrated. See the discussion in Canova (1995) p. 104, and the references cited therein.

7.4.4 Multivariate Forecast Menu

This menu appears on the screen when option 5 in the Unrestricted VAR Post Estimation Menu is selected. (See Section 7.4.1). It contains the following options

0. Choose another variable
1. Display forecast and forecast errors
2. Plot of in-sample fitted values and out of sample forecasts
3. Save in-sample fitted values and out of sample forecasts

Option 0 enables you to inspect forecasts of the level or first-differences of another variable in the VAR.

Option 1 lists the actual values, multivariate forecasts and the forecast errors. In cases where actual values for the jointly determined variables over the forecast period are not available, it is still possible to generate multi-step ahead forecasts so long as observations on the exogenous/deterministic variables in the VAR (namely $\mathbf{w}_t$, intercepts and trends) are available over the forecast period. In addition to listing the forecasts, this option also computes a number of standard summary statistics for checking the adequacy of the forecasts over the estimation and the forecast periods.

Option 2 enables you to see plots of the actual and forecast values for the selected variable. In the graph window you can specify a different period over which you wish to see the plots. Click the Start and Finish fields and scroll through the drop-down lists to select the desired sample period, and then press the button 'Refresh graph over the above sample period'.

Option 3 allows you to save the fitted and forecast values of the selected variable in the workspace in a new variable to be used in subsequent analysis.

7.5 Cointegrating VAR options

The econometric model that underlies the cointegrating VAR options is given by the following general vector error correction model ($VECM$):

$$\Delta\mathbf{y}_t = \mathbf{a}_{0y} + \mathbf{a}_{1y}t - \mathbf{\Pi}_y\mathbf{z}_{t-1} + \sum_{i=1}^{p-1}\mathbf{\Gamma}_{iy}\Delta\mathbf{z}_{t-i} + \mathbf{\Psi}_y\mathbf{w}_t + \mathbf{u}_{ty} \tag{7.11}$$

where $\mathbf{z}_t = \begin{pmatrix}\mathbf{y}_t \\ \mathbf{x}_t\end{pmatrix}$. This model distinguishes between four categories of variables:

1. $\mathbf{y}_t$, which is an $m_y \times 1$ vector of jointly determined (or endogenous) $I(1)$ variables.[3]

2. $\mathbf{x}_t$, which is an $m_x \times 1$ vector of $I(1)$ exogenous variables.

3. $\mathbf{w}_t$, which is a $q \times 1$ vector of $I(0)$ exogenous variables.

4. Intercepts and deterministic linear trends.

The implicit VAR model for the $I(1)$ exogenous variables, $\mathbf{x}_t$, is given by

$$\Delta\mathbf{x}_t = \mathbf{a}_{0x} + \sum_{i=1}^{p-1}\mathbf{\Gamma}_{ix}\Delta\mathbf{x}_{t-i} + \mathbf{\Psi}_x\mathbf{w}_t + \mathbf{v}_t \tag{7.12}$$

and assumes that $\mathbf{x}_t$s are not themselves cointegrated. Notice also that despite the fact that (7.12) does not contain a deterministic trend, the levels of $\mathbf{x}_t$ will be trended due to the drift coefficients, $\mathbf{a}_{0x}$.

Combining (7.11) and (7.12) we obtain

$$\Delta\mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t - \mathbf{\Pi}\mathbf{z}_{t-1} + \sum_{i=1}^{p-1}\mathbf{\Gamma}_i\Delta\mathbf{z}_{t-i} + \mathbf{\Psi}\mathbf{w}_t + \mathbf{u}_t \tag{7.13}$$

[3] Vector $\mathbf{y}_t$ is said to be $I(1)$ if all its elements *must* be differenced to achieve stationarity.

where

$$\mathbf{z}_t = \begin{pmatrix} \mathbf{y}_t \\ \mathbf{x}_t \end{pmatrix}, \quad \mathbf{u}_t = \begin{pmatrix} \mathbf{u}_{ty} \\ \mathbf{v}_t \end{pmatrix}, \quad \mathbf{a}_0 = \begin{pmatrix} \mathbf{a}_{0y} \\ \mathbf{a}_{0x} \end{pmatrix}, \quad \mathbf{a}_1 = \begin{pmatrix} \mathbf{a}_{1y} \\ \mathbf{0} \end{pmatrix}$$

$$\mathbf{\Pi} = \begin{pmatrix} \mathbf{\Pi}_{0y} \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{\Gamma}_i = \begin{pmatrix} \mathbf{\Gamma}_{iy} \\ \mathbf{\Gamma}_{ix} \end{pmatrix}, \quad \mathbf{\Psi} = \begin{pmatrix} \mathbf{\Psi}_y \\ \mathbf{\Psi}_x \end{pmatrix}$$

which is a restricted vector error correction form of (7.1).

In estimating (7.11) *Microfit* distinguishes between five cases depending on whether (7.11) contains intercepts and/or time trends, and on whether the intercepts and the trend coefficients are restricted. Ordering these cases according to the importance of deterministic trends in the model we have

Case I. $\mathbf{a}_{0y} = \mathbf{a}_{1y} = \mathbf{0}$ (no intercepts and no deterministic trends).

Case II. $\mathbf{a}_{1y} = \mathbf{0}$, and $\mathbf{a}_{0y} = \mathbf{\Pi}_y \mu_y$ (restricted intercepts and no deterministic trends).

Case III. $\mathbf{a}_{1y} = \mathbf{0}$, and $\mathbf{a}_{0y} \neq \mathbf{0}$ (unrestricted intercepts and no deterministic trends).

Case IV. $\mathbf{a}_{0y} \neq \mathbf{0}$, and $\mathbf{a}_{1y} = \mathbf{\Pi}_y \gamma_y$ (unrestricted intercepts and restricted deterministic trends).

Case V. $\mathbf{a}_{0y} \neq \mathbf{0}$, and $\mathbf{a}_{1y} \neq \mathbf{0}$ (unrestricted intercepts and trends).

The intercept and the trend coefficients, $\mathbf{a}_{0y}$ and $\mathbf{a}_{1y}$ are $m_y \times 1$ vectors, $\mathbf{\Pi}_y$ is the long-run multiplier matrix of order $m_y \times m$, where $m = m_x + m_y$, $\mathbf{\Gamma}_{1y}, \mathbf{\Gamma}_{2y}, ..., \mathbf{\Gamma}_{p-1,y}$ are $m_y \times m$ coefficient matrices capturing the short-run dynamic effects, and $\mathbf{\Psi}_y$ is the $m_y \times q$ matrix of coefficients on the $I(0)$ exogenous variables.

Firstly (7.11) allows for a sub-system approach in which the m_x-vector of random variables, $\mathbf{x}_t$, are treated structurally *exogenous*, in the sense that there are no error correction feedbacks in the equations explaining $\Delta \mathbf{x}_t$. Models of this type naturally arise in empirical macroeconomic analysis of small open economies where for the purpose of modelling the domestic macroeconomic relations, foreign prices, interest rates and foreign incomes can often be treated as exogenous $I(1)$ variables. Secondly, the model (7.11) explicitly allows for the possibility of deterministic trends, which again could be an important consideration in macroeconomic applications.

The importance of distinguishing among the above five cases is discussed in Section 22.7. In the case where $\mathbf{a}_0$ and $\mathbf{a}_1$ are both unrestricted (Case V), $\mathbf{y}_t$ will be trend-stationary when the rank of $\mathbf{\Pi}_y$ is full, but when $\mathbf{\Pi}_y$ is rank deficient, the solution of $\mathbf{y}_t$ will contain quadratic trends, unless $\mathbf{a}_{1y}$ is restricted as in Case IV. Similarly, in Case III, when Π_y is rank deficient then y_t will contain a linear deterministic trend, unless $\mathbf{a}_{0y}$ is restricted as in Case II. Case I is included for completeness and is unlikely to be of relevance in economic applications.

It is also worth noting that in Case IV where $\mathbf{a}_{1y} \neq 0$, the cointegrating vectors contain a deterministic trend, and in Case II where $\mathbf{a}_{0y} \neq 0$, the cointegrating vectors contain intercepts.

Under the assumption that Rank($\mathbf{\Pi}_y$) $= r$, that is, when there exists r cointegrating relations among the variables in $\mathbf{z}_t$, we have

$$\mathbf{\Pi}_y = \boldsymbol{\alpha}_y \boldsymbol{\beta}' \tag{7.14}$$

where $\boldsymbol{\alpha}_y$ and $\boldsymbol{\beta}$ are $m_y \times m$ and $m \times r$ matrices each with full column rank, r. The r cointegrating relations are then given by

$$\boldsymbol{\beta}'\mathbf{z}_t = \boldsymbol{\beta}'\mathbf{z}_0 + \left(\boldsymbol{\beta}'\mathbf{C}^*(1)\mathbf{a}_1\right) t + \boldsymbol{\eta}_t \tag{7.15}$$

where $\mathbf{z}_0$ is the initial value of $\mathbf{z}_t$, $\mathbf{C}^*(1) = \mathbf{C}_0^* + \mathbf{C}_1^* + \mathbf{C}_2^* + \cdots$, defined by the recursive relations (22.83), and $\boldsymbol{\eta}_t$ ($\boldsymbol{\eta}_0 = \mathbf{0}$) is an $r \times 1$ vector of mean-zero, $I(0)$ disturbance vector, representing the covariance stationary components of the cointegrating relations.[4] In the case where the trend coefficients $\mathbf{a}_{1y}$ in the underlying $VECM$ are restricted (Case IV), we have $\boldsymbol{\beta}'\mathbf{C}^*(1)\mathbf{a}_1 = \boldsymbol{\beta}'\boldsymbol{\gamma}$, where $\boldsymbol{\gamma}$ is an $r \times 1$ vector of unknown coefficients. In this case the trend coefficients enter the cointegrating relations and we have

$$\Delta\mathbf{y}_t' = \mathbf{a}_{0y} - \boldsymbol{\alpha}_y\boldsymbol{\beta}'\left(\mathbf{z}_{t-1} - \boldsymbol{\gamma}t\right) + \sum_{i=1}^{p-1}\mathbf{\Gamma}_{iy}\Delta\mathbf{z}_{t-i} + \mathbf{\Psi}_y\mathbf{w}_t + \mathbf{u}_{ty} \tag{7.16}$$

$\boldsymbol{\beta}'\left(\mathbf{z}_{t-1} - \boldsymbol{\gamma}t\right)$ will be a covariance stationary process with a constant mean. In this case the presence of a deterministic trend in the cointegrating relations can be empirically tested by testing the r 'co-trending' restrictions

$$\boldsymbol{\beta}'\boldsymbol{\gamma} = 0 \tag{7.17}$$

In Case V where $\mathbf{a}_{1y}$ is unrestricted, the deterministic trends in the error correction model (7.11) are specified outside the cointegrating relations. One could test for their presence in the error correction model by means of standard t-tests. Similar considerations also apply when comparing models with restricted and unrestricted intercepts, but with no deterministic trends (Cases II and III). Under Case II, the intercepts in the underlying error correction model will appear as a part of the cointegrating relations, while under case III, the unrestricted intercepts appear as a part of the error correction specification.

In most macroeconomic applications of interest, where $\mathbf{y}_t$ and $\mathbf{x}_t$ contain deterministic trend components, and $\mathbf{\Pi}_y$ is rank deficient, the appropriate vector error correction model is given by Case IV, where the trend coefficients are restricted. In cases where $(\mathbf{y}_t, \mathbf{x}_t)$ do not contain deterministic trends, Case II is the appropriate error correction model.

7.5.1 Specification of the cointegrating VAR model

The Cointegrating VAR Menu is option 2 in the System Estimation Menu (See Section 7.3). The menu contains the following choices for the inclusion of intercept/trends in the VAR model

[4] For a derivation and an explicit expression for $\boldsymbol{\eta}_t$ see Section 22.7.1. Also see Park (1992) and Ogaki (1992).

1. No intercepts or trends included in VAR model
2. Restricted intercepts and no trends in VAR model
3. Unrestricted intercepts and no trends in VAR model
4. Unrestricted intercepts and restricted trends in VAR model
5. Unrestricted intercepts and unrestricted trends in VAR model

Options 1 to 5 refer to Cases I to V set out above, and correspond to different specifications of intercept/trend in the underlying VAR model. Option 2 is appropriate when the jointly determined variables do not contain a deterministic trend. Option 4 (that corresponds to Case IV) is appropriate when the jointly determined variables in the VAR have a linear deterministic trend.

Options 3 to 5 can lead to error correction models with different trend properties depending on the number of cointegrating relations.

You need to list the $I(1)$ variables in your model, namely variables $\mathbf{z}_t$, followed by the list of your $I(0)$ variables, $\mathbf{w}_t$, if any, separating them by **&**. The division of $I(1)$ variables into the jointly determined variables, $\mathbf{y}_t$, and the exogenously determined variables, $\mathbf{x}_t$, is done by using the semicolon character (**;**) as a separator. For example, suppose you wish to estimate a cointegrating VAR model containing the following variables:

P	Domestic price level
PF	Foreign price level
E	Exchange rate
R	Domestic interest rate
RF	Foreign interest rate
DPO	Changes in real oil prices
$DPO(-1)$	Lagged changes in real oil prices

where P, E and R are endogenous $I(1)$ variables, PF and RF are exogenous $I(1)$ variables, and DPO is the exogenous $I(0)$ variable in the model. Then you need to type

$$P \quad E \quad R\ ; \quad PF \quad RF \quad \& \quad DPO \quad DPO(-1)$$

You should not include an intercept or a deterministic trend term among these variables. You need to specify the order of the VAR model ($p \leq 24$), and then click RUN to begin calculations. *Microfit* presents you with the trace and maximum eigenvalue statistics for testing a number of hypotheses concerning the rank of the long-run matrix $\mathbf{\Pi}_y$ in (7.11), together with the relevant 90 and 95 per cent critical values (see Sections 22.8, 22.8.1 and 22.8.2).

The program also gives the maximized values of the log-likelihood function of the cointegrating VAR model, Akaike, Schwarz, and Hannan and Quinn model selection criteria, for the different values of r, the rank of the long-run matrix, $\mathbf{\Pi}_y$ (see Section 22.8.3).

The above test results and model selection criteria can be used to determine the appropriate number of cointegrating relations that are likely to exist among the $I(1)$ variables. Before moving to the next stage of the cointegration analysis, you *must* choose a value for r. It is only meaningful to continue with the cointegration analysis if your choice of r lies strictly between 0 and m_y.

7.5.2 Cointegrating VAR Post Estimation Menu

This menu appears on the screen after the cointegration test results obtained choosing either option 2 or 3 in the System Estimation Menu (see Section 7.6 for further details on option 3, the cointegrating $VARX$ option). This menu has the following options

0. Move to Backtracking Menu
1. Display cointegration tests again
2. Specify r, the number of cointegrating vectors (CVs)
3. Display CVs using Johansen's just-identifying restrictions
4. Display system covariance matrix of errors
5. Display matrix of long-run multipliers for the specified r
6. Long-run structural modelling, IR Analysis and Forecasting
7. Compute multivariate dynamic forecasts

Option 0 takes you back to the Commands and Data Transformations box.

Option 1 enables you to see the cointegration test results again.

Option 2 allows you to specify your choice of r, the number of cointegrating or long-run relations among the $I(1)$ variables. Notice that r cannot be zero. If r is chosen to be equal to m_y, the dimension of $\mathbf{y}_t$, the estimation results will be the same as using the unrestricted VAR option.

Option 3 displays the ML estimates of the cointegrating vectors under Johansen's exact identifying restrictions. These estimates lack any meaningful economic interpretations when $r > 2$. In the case where $r = 1$, Johansen's estimates (when appropriately normalized) will be the same as those obtained using option 6 in this menu.

Option 4 displays the estimates of $\mathbf{\Sigma}$, the variance matrix of the errors in the cointegrating VAR model, assuming rank of $\mathbf{\Pi}_y$ is equal to r.

Option 5 displays the ML estimates of $\mathbf{\Pi}_y$, the matrix of the long-run coefficients defined by (7.14), subject to the cointegrating restrictions. Notice that by construction rank of $\widehat{\mathbf{\Pi}}_y$ is equal to r.

Option 6 moves you to the Long-Run Structural Modelling Menu and enables you to estimate the cointegrating vectors subject to general linear restrictions, possibly involving parameters across the cointegrating vectors, and to test over-identifying restrictions (if any). This option is also the gateway to computation of impulse response functions, forecast error variance decompositions, persistence profile analysis and multivariate forecasting, with the cointegrating vectors, $\boldsymbol{\beta}$, being subject to (possibly) over-identifying restrictions (see Section 7.5.3).

Option 7 enables you to compute multivariate dynamic forecasts of $\mathbf{y}_t$, the jointly determined variables, for given values of $\mathbf{x}_t$ and $\mathbf{w}_t$, over the forecast period, and assuming that $rank\,(\mathbf{\Pi}_y) = r$. The forecasts obtained under this option implicitly assume that the cointegrating vectors, $\boldsymbol{\beta}$, are *exactly* identified. To compute multivariate forecasts when $\boldsymbol{\beta}$ is subject to over-identifying restrictions you must use option 5 in the Impulse Response Analysis and Forecasting Menu (see Section 7.5.4). When you choose this option you will be first asked to specify the forecast period, and then to choose the variable you wish to forecast. For each variable that you choose you will be given a choice of forecasting the levels of the variable or its first differences. The program then computes the forecasts and presents you with the Multivariate Forecast Menu (see Section 7.4.4).

7.5.3 Long-Run Structural Modelling Menu

This menu can be accessed selecting option 6 in the Cointegrating VAR Post Estimation Menu (see Section 7.5.2). It contains the following options

0. Move to Cointegrating VAR Post Estimation Menu
1. Likelihood ratio test of fixing some cointegrating vectors (CVs)
2. Likelihood ratio test of imposing same restriction(s) on all CVs
3. Likelihood ratio test of imposing restriction(s) on only one CV
4. Likelihood ratio test of imposing general restrictions on CVs
5. Use CVs obtained under Johansen's just-identifying restrictions
6. Fix all the cointegrating vectors

Option 0 returns you to the Cointegrating VAR Post Estimation Menu (see Section 7.5.2).

Options 1 to **3** represent different ways of testing simple homogenous restrictions on the cointegrating vectors. Since the same restrictions can be imposed and tested using option 4 we will not discuss these options here. The interested user should consult the manual for *Microfit 3.0* where these tests are described in detail (see Pesaran and Pesaran (1991), pp. 88-89).

Option 4 enables you to estimate the VAR model subject to general (possibly) non-homogenous restrictions on the cointegrating (or the long-run) coefficients, and compute log-likelihood ratio statistics for testing over-identifying restrictions on the long-run coefficients. However, when you first choose this option you will be asked to specify exactly r restrictions on each of the r cointegrating vectors to just-identify the long-run restrictions. For example, if the number of cointegrating relations, r, is equal to 4, a typical set of just-identifying restrictions could be

$$\begin{array}{llll} A1 = 1; & A2 = 0; & A3 = 0; & A4 = 0; \\ B1 = 0; & B2 = 1; & B3 = 0; & B4 = 0; \\ C1 = 0; & C2 = 0; & C3 = 1; & C4 = 0; \\ D1 = 0; & D2 = 0; & D3 = 0; & D4 = 1 \end{array}$$

The files CO2.EQU, CO3.EQU, ..., CO10.EQU in the *Microfit* Tutorial Directory contains such just-identifying restrictions for $r = 2, 3, .., 10$, respectively. The above type of just-identifying restrictions could be made relevant to your particular application, by a suitable choice of the ordering of the variables in the VAR.[5] Once you have successfully estimated the model subject to the just-identifying restrictions, you can add over-identifying restrictions

[5] Notice that except for the results on orthogonalized impulse response functions, and the orthogonalized forecast error variance decomposition, the cointegration tests and the ML estimates are invariant to the ordering of the variables in the VAR.

at a later stage (see option 0 in Section 7.5.4). The econometric and computational details are set out in Section 22.9.

Option 5 sets the cointegrating vectors equal to Johansen's estimates, $\widehat{\boldsymbol{\beta}}_J$, which are obtained subject to the just-identifying restrictions defined by (22.106) and (22.107).

Option 6 enables you to fix the cointegrating vectors by specifying all their elements.

Options 1 to 6 in the menu, once successfully implemented, take you to the Impulse Response Analysis and Forecasting Menu (see Section 7.5.4).

7.5.4 Impulse Response Analysis and Forecasting Menu

This menu appears on the screen after a successful implementation of options 1 to 6 in the Long-Run Structural Modelling Menu (see Section 7.5.3). It contains the following option:

0. Return to identify/test cointegrating vectors
1. Impulse Response of variables to shocks in equations
2. Forecast Error Variance Decomposition analysis
3. Impulse Response of CVs to shocks in equations
4. Persistence Profile of CVs to system-wide shocks
5. Trend/Cycle Decomposition
6. Compute multivariate dynamic forecasts
7. Display restricted/fixed CVs again
8. Display error correction equations
9. Save error correction terms
10. Display system covariance matrix of errors
11. Save the cointegrating VAR model in a CSV file

Option 0 enables you to estimate/test (further) over-identify restrictions on the cointegrating or long-run coefficients. The restrictions could involve parameters from different long-run relations (see Section 22.7 and option 4 in the Long-Run Structural Modelling Menu described in Section 7.5.3). When you choose this option you will be asked to confirm whether you wish to test over-identifying restrictions on the long-run relations. If you say 'No', you will be returned to the Long-Run Structural Modelling Menu (see Sections 22.9.1-22.9.3). If your answer is in the affirmative you will be presented with a box-editor to specify your over-identifying restrictions. *Our recommendation is to introduce these restrictions gradually (ideally one by one), starting from those that are less likely to be rejected.* The asymptotic standard errors reported below the just-identified estimates could provide a good guide as to which over-identifying restrictions to impose first, second and so on. Once your over-identifying restrictions are added successfully to the existing set of restrictions (including the just-identifying ones), you will be presented with a screen containing initial values for all the long-run coefficients. These are the estimates obtained under the previous set of restrictions. We recommend that you accept these initial estimates.[6] If you

[6] You can, of course, edit these initial estimates if you experience difficulties with the convergence of the iterative algorithm.

now click  to accept the initial values you will be presented with a small menu giving you a choice of the 'Back substitution algorithm (A) as in *Microfit* 4', the 'Back substitution algorithm (B) new to *Microfit* 5', and the 'Modified Newton-Raphson algorithm'. The highlighting is always on the 'Back substitution algorithm (B) new to *Microfit* 5', which is the one that we recommend. If you choose the modified Newton-Raphson algorithm[7] you will also be given a choice of a damping factor in the range [0.01 to 2.0]. We recommend starting with the value of 0.01, unless you experience difficulties with getting the algorithm to converge. Once you have chosen the algorithm, the program starts the computations and, if the iterative process converges successfully, presents you with the ML estimates of the long-run relations subject to the over-identifying restriction, together with their asymptotic standard errors in round brackets. *Microfit* also generates log-likelihood ratio statistics for testing the over-identifying restrictions, which are asymptotically distributed as χ^2 variates with degrees of freedom given by $k - r^2$, where k is the total number of restrictions and r^2 is the number of just-identifying restrictions (see Section 22.9.3). If you click , you will be presented with a window asking whether you want bootstrapped critical values of overidentifying restrictions on long-run relationships. You can choose the number of replications and two different significance levels. If you click *Microfit* starts the computation and presents you with an output window which reports the bootstrapped critical values of the log-likelihood ratio statistics. If you click  you return to the Impulse Response Analysis and Forecasting Menu.

Option 1 computes and displays orthogonalized and generalized impulse responses of variable-specific shocks on the different variables in the cointegrating VAR model, (possibly) subject to over-identifying restrictions on the long-run coefficients. Once the results are obtained, it is also possible to compute bootstrapped confidence intervals of impulse responses, for any desired confidence level.

Option 2 computes and displays orthogonalized and generalized forecast error variance decompositions for the cointegrating VAR model, (possibly) subject to restrictions on the long-run relationships. You can then obtain bootstrapped confidence intervals for the error variance decomposition at a given confidence level.

Option 3 computes and displays orthogonalized and generalized impulse responses of the effect of variable-specific shocks on the r cointegrating relations.

Option 4 computes and displays the time profile of the effect of system-wide shocks on the cointegrating relations, referred to as 'persistence profiles'. Selecting options 3 or 4 allows you to obtain bootstrapped confidence intervals for persistence profiles, for a given confidence level. The algorithms used to carry out the computations for options 3 and 4 are set out in Section 22.9.5 and 22.9.6, where references to the literature can also be found.

Option 5 allows you to perform the multivariate Beveridge Nelson trend/cycle decomposition (see Section 22.11).

Option 6 enables you to compute multivariate, multi-step ahead forecasts (of levels and of first-differences) of $\mathbf{y}_t$ conditional on values of $\mathbf{x}_t$ and $\mathbf{w}_t$. The forecasts obtained using

[7] For a detailed account of the Back substitution algorithm (A) and of the Modified Newton-Raphson algorithm see Section 22.9.2.

this option and those obtained using option 7 in the Cointegrating VAR Post Estimation Menu will be identical under just-identifying restrictions on the cointegrating relations, and differ only when there are over-identifying restrictions on $\boldsymbol{\beta}$ (see Section 7.4.4, and option 7 in Section 7.5.2). In the case of cointegrating $VARX$ (option 3 from the System Estimation Menu), you can choose between conditional and unconditional or *ex ante* forecasts, depending on whether you wish to use the realized values of the exogenous variables or their forecast values. For the conditional forecasts the values of the exogenous variables during the forecast period are treated as known. The unconditional forecasts use forecasts of the exogenous variables, obtained using the marginal model.

Option 7 displays the ML (or fixed) estimates of the cointegrating vectors again**.** This option also allows you to obtain bootstrapped confidence intervals for ML estimates, for two different significance levels.

Option 8 displays error correction equations for each of the jointly determined $I(1)$ variables in the model. These estimates are followed by diagnostic statistics and the other options available after the OLS option. See section 6.6.

Option 9 saves error correction terms in the workspace.

Option 10 displays the degrees-of-freedom adjusted system covariance matrix of the errors in the underlying VAR model, (7.11). The adjustments are made by dividing the cross-product of residuals from different equations by $n - s$, where s is the total number of coefficients estimated for each equation in the VAR. This adjustment does not take account of the cross-equation restrictions on the long-run coefficients, $\mathbf{\Pi}_y$, implicit in the cointegrating restrictions. These estimates will be identical to those obtained using option 4 in the Cointegrating VAR Post Estimation Menu, if the cointegrating vectors, β, are not subject to over-identifying restrictions. See Section 7.5.2.

Option 11 allows you to save the estimated cointegrating VAR model as a CSV file.

7.5.5 Beveridge-Nelson Trend/Cycle Decomposition

Option 5 from the Impulse Response Analysis and Forecasting Menu allows you to compute the multivariate Beveridge-Nelson trend/cycle decomposition (see Section 22.11). When choosing this option you will be presented with a screen containing the names of the endogenous and exogenous variables and the default initial values for the corresponding parameters and t-ratios on the intercept and trend. If you click , you are then asked to select the variable for which you wish to perform the trend/cycle decomposition.

7.5.6 Trend/Cycle Decomposition Results Menu

This menu appears on screen after you have performed a trend cycle decomposition, using option 5 from the Impulse Response Analysis and Forecasting Menu. It contains the following options

0. Move back
1. Display results again
2. Graph
3. Save the decomposition for all variables in a CSV file
4. Save the trend component for all variables in a CSV file
5. Save the stochastic component for all variables in a CSV file
6. Save the permanent component for all variables in a CSV file
7. Save the cyclical component for all variables in a CSV file
8. Save all the detrended variables in a CSV file

Option 0 allows to move back to the Impulse Response Analysis and Forecasting Menu.

Options 1 and **2** enables to display results or to plot the various components.

Option 3 to 8 save the decomposition or some components of the Beveridge-Nelson decomposition for all variables in a CSV file.

7.6 Cointegrating VARX option

Option 3 in the System Estimation Menu (Multivariate Menu: see Section 7.3) enables you to estimate vector error correction models ($VECM$) with weakly exogenous $I(1)$ variables. When you choose this option you will be presented with the screen in Figure 7.2 (below), which prompts you to list the jointly determined (endogenous) variables, followed by the list of the exogenous variables, and the deterministic variables such as intercepts and trend terms (if any). The division of variables into endogenous and exogenous is done using the semicolon character (;) as a separator, and the deterministic variables are separated from the rest by &.

Type in the start and the finish of your estimation period, and the order of the VAR model which could differ between endogenous and exogenous variables (in both cases must be smaller than 24) and then click RUN. *Microfit* asks you to select the variables you wish to restrict, and then presents you with the screen in Figure 7.3.

Here you are asked to specify the equations for the marginal models for your exogenous variables. You need to list first the lag order of the endogenous variables, then the lag order of exogenous variables, followed by the deterministic variables (if any). You must specify one equation for each exogenous variable and separate the equation specifications using the semicolon (;).

Once you have completed the specification of the marginal models, click RUN. *Microfit* carries out the necessary computations and produces an output window with the results on cointegration tests. If you close this output window, you will be then presented with the Cointegrating VAR Post Estimation Menu (see Section 7.5.2).

7.7 SURE options

There are six options in *Microfit* for estimation of Seemingly Unrelated Regression Equations ($SURE$) models. These options can be accessed from the System Estimation Menu (see

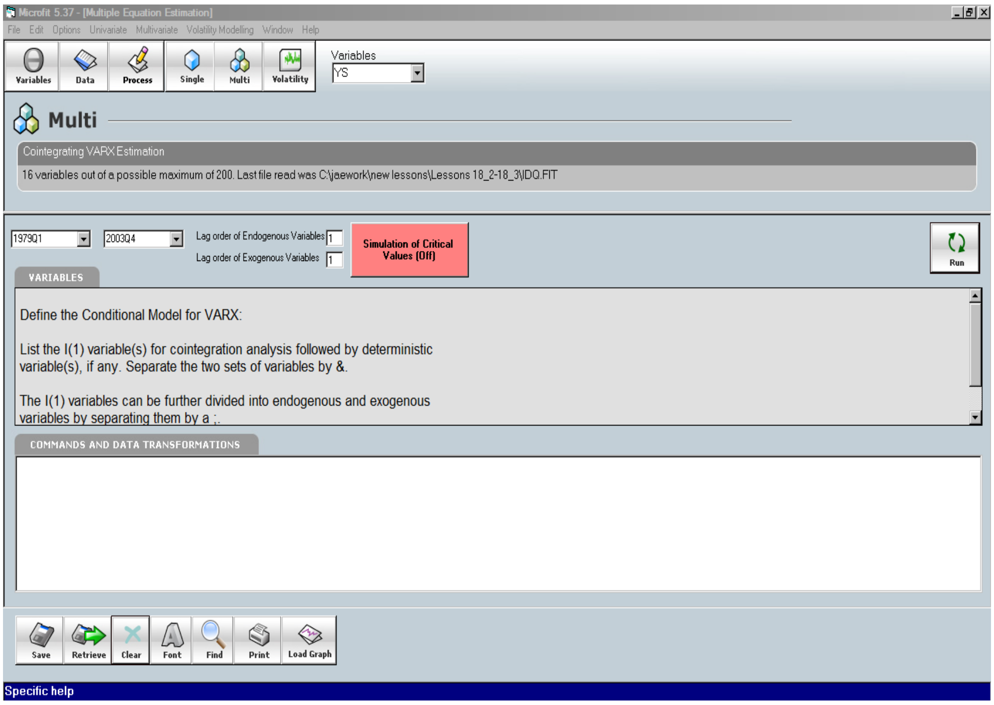


Figure 7.2: The System Estimation window with the $VARX$ option

Section 7.3).

Option 4, 6 and **8** in this menu enable you to compute, respectively, maximum likelihood (ML), two-stages least squares ($2SLS$) and three-stages least squares ($3SLS$) estimates (see Section 22.1 and 22.2) of the parameters of the following $SURE$ model:

$$y_{it} = \boldsymbol{\beta}_i' \mathbf{x}_{it} + u_{it}, \qquad i = 1, 2, ..., m \tag{7.18}$$

where y_{it} is the i^{th} dependent variable in the model composed of an equation, $\mathbf{x}_{it}$ is the $k_i \times 1$ vector of regressors in the ith equation, $\boldsymbol{\beta}_i$ is the $k_i \times 1$ vector of unknown coefficients of the i^{th} equation, and u_{it} is the disturbance term. The disturbances (shocks) u_{it}, $i = 1, 2, ..., m$ are assumed to be homoscedastic and serially uncorrelated, but are allowed to be contemporaneously correlated ($\text{Cov}\,(u_{it}, u_{jt}) = \sigma_{ij}$, need not be zero for $i \neq j$).

Options 5, 7 and **9** in the System Estimation Menu allow you to compute, respectively, maximum likelihood, two-stages least squares and three-stages least squares estimates of (7.18) when β_is are subject to the following general linear restrictions:

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{b} \tag{7.19}$$

where $\mathbf{R}$ and $\mathbf{b}$ are $r \times k$ matrix and $r \times 1$ vector of known constants, and $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', ...\boldsymbol{\beta}_m'\right)'$ is a $k \times 1$ vector of unknown coefficients, $k = \sum_{i=1}^{m} k_i$. When there are no cross-equation restrictions, we have

$$\mathbf{R}_i \boldsymbol{\beta}_i = \mathbf{b}_i \tag{7.20}$$

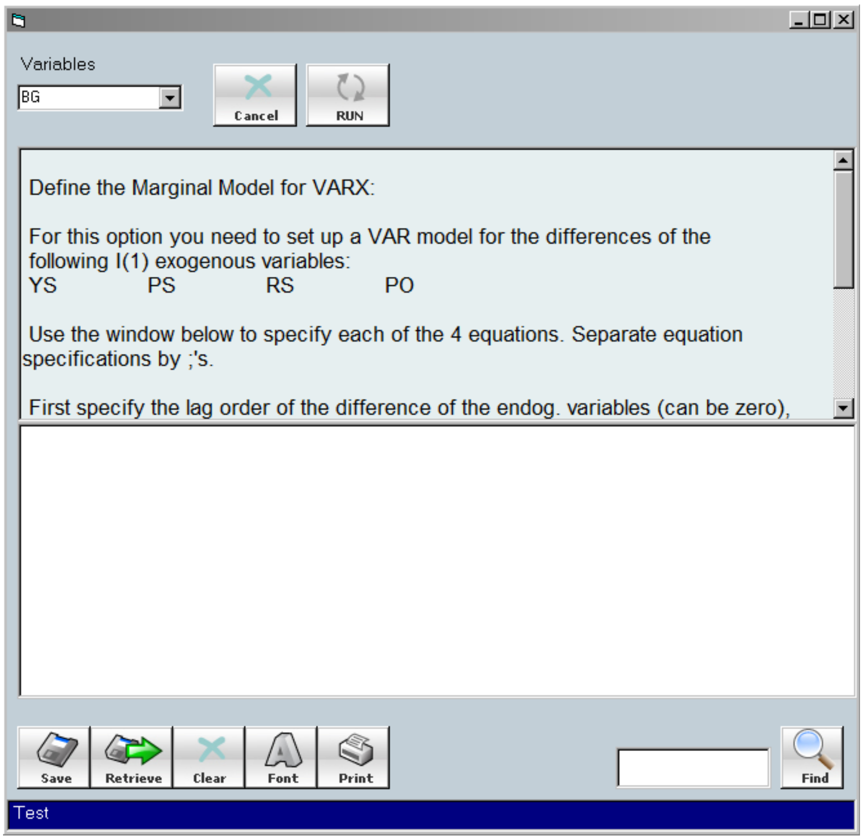


Figure 7.3: Box editor for the specification of the marginal model

where $\mathbf{R}_i$ and $\mathbf{b}_i$ are the $r_i \times k_i$ and $r_i \times 1$ matrix vector of restrictions applicable only to the coefficients of the ith equation in the model. See Section 22.3 for more details.

7.7.1 Unrestricted SURE options

These are options 4, 6 and 8 in the System Estimation Menu (see Section 7.3). When you choose one of these options you will be presented with a box editor for specifying the equations in the $SURE$ model. You should list the endogenous variable, y_{it}, followed by its regressors, $\mathbf{x}_{it}$, for each equation, separating the different equations by a semicolon (;). The program then automatically works out the number of equations in the $SURE$ model, namely m ($m \leq 10$). A simple example, of a $SURE$ specification is

$$\begin{array}{lllllll} W1 & INPT & W1(-1) & LP1 & LP2 & LP3 & LRY; \\ W2 & INPT & W2(-1) & LP1 & LP2 & LP3 & LRY \end{array}$$

where $INPT$ is an intercept term, $W1$ and $W2$ could be budget shares of two different commodity groups, $LP1$, $LP2$ and $LP3$, the logarithm of price indices and LRY, the logarithm of real income. It is also possible to estimate restricted VAR models using the $SURE$ option. For example, suppose you wish to estimate the following restricted $VAR(4)$ model:

$$x_{1t} = a_1 + \sum_{j=1}^{4} b_{1j} x_{1,t-j} + c_{11} x_{2,t-1} + d_{11} x_{3,t-1} + u_{1t}$$

$$x_{2t} = a_2 + \sum_{j=1}^{4} b_{2j} x_{2,t-j} + c_{21} x_{1,t-1} + d_{21} x_{3,t-1} + u_{2t}$$

$$x_{3t} = a_3 + \sum_{j=1}^{4} b_{3j} x_{3,t-j} + c_{31} x_{1,t-1} + d_{31} x_{2,t-1} + u_{3t}$$

then you need to specify the $SURE$ model as

$$\begin{array}{lllll} X1 & INPT & X1\{1-4\} & X2(-1) & X3(-1); \\ X2 & INPT & X2\{1-4\} & X1(-1) & X3(-1); \\ X3 & INPT & X3\{1-4\} & X1(-1) & X2(-1) \end{array}$$

For specification of VAR models with linear non-homogeneous and/or cross-equation parametric restrictions you need to use the restricted $SURE$ option (see Section 7.7.2).

Once the unrestricted $SURE$ model is successfully specified you will be prompted to specify the period over which you wish to estimate the model. *Microfit* then starts to compute ML, $2SLS$ or $3SLS$ estimators of the parameters (depending on whether you selected option 4, 6 or 8 from the System Estimation Menu) and, when successful, presents you with the $SURE$ Post Estimation Menu (see Section 7.7.3). Maximum likelihood estimates are computed using the back-substitution algorithm described in Section 22.1.1.

7.7.2 Restricted SURE options

These are options 5, 7 and 9 in the System Estimation Menu (see Section 7.3). When you select one of these options you will be first presented with a box editor to specify the equations in the $SURE$ model (see Section 7.7.1 on how to specify the $SURE$ model). When you have done this you will be presented with *another* box editor for you to specify coefficients of the $SURE$ model. These restrictions must be linear, but can include cross-equation restrictions. For example, suppose you are interested in estimating the following system of equations

$$Y_{it} = \alpha_i + \beta_i X_{it} + \gamma_i W_{it} + u_{it}$$

for $i = 1, 2, ..., 4$, and $t = 1, 2, ..., n$, assuming the homogeneity of the slope coefficients, namely $\beta_i = \beta$, and $\gamma_i = \gamma$, for $i = 1, 2, ..., 4$. In the first box editor that appears on the screen (after you choose the restricted $SURE$ option) type

$$\begin{array}{llll} Y1 & INPT & X1 & W1; \\ Y2 & INPT & X2 & W2; \\ Y3 & INPT & X3 & W3; \\ Y4 & INPT & X4 & W4 \end{array}$$

In the second box editor that appears on the screen type the restrictions

$$A2 = B2; \quad B2 = C2; \quad C2 = D2;$$
$$A3 = B3; \quad B3 = C3; \quad C3 = D3$$

Note that *Microfit* assigns the coefficients $A1$, $A2$, and $A3$ to the parameters of the first equation, $B1$, $B2$, and $B3$ to the parameters of the second equation, $C1$, $C2$ and $C3$ to the parameters of the third equation, and so on. Therefore,

$$\alpha_1 = A1, \quad \beta_1 = A2, \quad \gamma_1 = A3;$$
$$\alpha_2 = B1, \quad \beta_2 = B2, \quad \gamma_2 = B3$$

and so on.

Once the restrictions are specified successfully you will be asked to specify the period over which you wish to estimate the model. When this is done, *Microfit* starts the task of computing ML, $2SLS$ or $3SLS$ estimators (depending on whether you selected option 5, 6 or 9 from the System Estimation Menu) of the parameters of the $SURE$ model subject to the restrictions. You will be then presented with the $SURE$ Post Estimation Menu, with options for displaying the estimates and their standard errors, carrying out tests on the parameters of the model and computing multivariate forecasts. The technical details and the relevant formulae are given in Section 22.3.

7.7.3 SURE Post Estimation Menu

This menu appears on the screen after a $SURE$ model or a restricted $SURE$ model is successfully estimated either by maximum likelihood, two-stages least squares, or three-stages least squares (see Sections 7.7.1 and 7.7.2). It contains the following options

0. Move back to System Estimation Menu
1. Edit the model and estimate
2. Display individual equation estimation results
3. Display system covariance matrix of errors
4. Wald tests of hypotheses on the parameters of the model
5. Estimate/test functions of parameters of the model
6. Compute multivariate dynamic forecasts

Option 0 returns you to the System Estimation Menu (see Section 7.3).

Option 1 allows you to edit the equations in the $SURE$ model and estimate the revised model.

Option 2 enables you to see the ML estimates of the coefficients of the equations in the $SURE$ model. When you choose this option you will be asked to select the equation in the model that you wish to inspect. You will then be presented with estimation results together with a number of summary statistics, including the values of the AIC and SBC for the $SURE$ model. (See Sections 22.1 and 22.3 for computational details). If you press the Esc key you will then be presented with the Post Regression Menu, with a number of options including plotting/listing/saving residuals, fitted values, and displaying the estimates of the covariance matrix of the coefficients of the chosen equation. See the *OLS* option in Section 6.6 for further details. Notice, however, that if you wish to test restrictions on the coefficients of the *SURE* model, or estimate known functions of the parameters, or compute dynamic forecasts, you need to use options 4 to 6.

Option 3 displays the estimates of the variance matrix of the error, namely $\widehat{\boldsymbol{\Sigma}}$, given by (22.31) and (22.32).

Option 4 enables you to compute Wald statistics for testing the general linear/non-linear restrictions

$$H_0: \quad \mathbf{h}(\boldsymbol{\beta}) = \mathbf{0}$$

against

$$H_1: \quad \mathbf{h}(\boldsymbol{\beta}) \neq \mathbf{0}$$

where $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', ..., \boldsymbol{\beta}_m'\right)'$, and $\mathbf{h}(\cdot)$ is a known $r \times 1$ vector function with continuous partial derivatives. See Section 22.2.1.

Option 5 allows you to compute estimates of known (possibly) non-linear functions of the coefficients, $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', ..., \boldsymbol{\beta}_m'\right)'$.

Option 6 enables you to compute multivariate forecasts of the dependent (left-hand-side) variables in the *SURE* model. When the regressors include lagged values of the dependent variables, the program computes multivariate *dynamic* forecasts. When you choose this option you will be asked to specify the forecast period, and then to choose the variable you wish to forecast. For each variable that you choose you will be given a choice of forecasting the levels of the variables or their first-differences. *Microfit* then computes the forecasts and presents you with the Multivariate Forecast Menu 7.4.4.

Chapter 8

Volatility Modelling Options

8.1 Introduction

This chapter describes how *Microfit* can be used to estimate univariate and multivariate conditionally heteroscedastic models. In the following, we briefly review a variety of univariate and multivariate models with time-varying conditional variance that can be estimated using the volatility modelling options in *Microfit*. More technical details on the econometric methods and computational algorithms used are given in Chapter 23, where further references to the literature can also be found.

Volatility of a series (say asset returns, r_t) is generally measured by its conditional variance, and is denoted by

$$h_t^2 = Var(r_t \,|\mathcal{F}_{t-1}).$$

Volatility can arise due to a number of factors: over-reaction to news, incomplete learning, parameter variations and abrupt switches in policy regimes.

In the case of asset returns, volatility can be estimated either from option prices (when they are available) or from historical observations. The former is referred to as the Implied Volatility (IV) approach and is subject to a number of shortcoming: it depends on the particular option pricing model used; most option pricing models assume that volatility is constant, which is not true; the horizon of the IV is fixed (say 3 months) while the risk manager is often interested in shorter horizons; the number of assets with option prices is not sufficiently comprehensive for most risk management tasks. Measures of volatility based on historical data are broadly known as 'realized volatility' typically using intra-daily observations. Econometric approaches to volatility usually focus on daily observations. Measures of realized volatility are also used in evaluation of econometric models of volatility.

Forecasts of volatility are used in risk management, option pricing and asset portfolio decisions. Most of these applications involve multivariate volatility models. We begin with univariate models of asset return volatility.

8.2 Historical approaches to volatility measurement

8.2.1 RiskMetrics^{TM} (JP Morgan) method

RiskMetrics uses an exponentially weighted moving average model. Let

$$z_t = r_t - \bar{r},$$

The historical volatility of z_t conditional on observations available at time $t-1$ is computed as

$$h_t^2 = (1-\lambda)\sum_{\tau=0}^{\infty} \lambda^{\tau} z_{t-\tau-1}^2 \tag{8.1}$$

where λ is known as decay factor (or $1-\lambda$ the decay rate). The weights

$$w_\tau = (1-\lambda)\lambda^\tau\,, \quad \tau = 0, 1, 2, ...$$

add up to unity, and h_t^2 can be computed recursively

$$h_t^2 = \lambda h_{t-1}^2 + (1-\lambda) z_{t-1}^2$$

which is a special version of the $GARCH$(1,1) model to be discussed below.

Model (8.1) requires the initialization of the process. For a finite observation window, denoted by H, a more appropriate specification is

$$h_{H,t}^2 = \sum_{\tau=0}^{H} w_{H\tau} z_{t-1-\tau}^2$$

where

$$w_{H\tau} = \frac{(1-\lambda)\lambda^\tau}{1-\lambda^{H+1}}, \quad \tau = 0, 1, ..., H \tag{8.2}$$

are weights that add up to unity.

Other weighting schemes have also been considered; in particular the equal weighted specification

$$h_t^2 = \frac{1}{H+1}\sum_{\tau=0}^{H} z_{t-1-\tau}^2,$$

where $w_{H\tau} = 1/(1+H)$, which is a simple moving average specification.

The value chosen for the decay factor, λ, and the size of the observation window, H, are related. For example, for $\lambda = 0.9$, even if a relatively large value is chosen for H, due to the exponentially declining weights attached to past observations only around 110 observations are effectively used in the computation of h_t^2.

8.2.2 Econometric approaches

Consider the regression model

$$r_t = \boldsymbol{\beta}'\mathbf{x}_{t-1} + \varepsilon_t = \beta_0 + \boldsymbol{\beta}_1'\mathbf{x}_{1,t-1} + \varepsilon_t$$

and assume that all the classical assumptions are valid except that $V(\varepsilon_t | \mathcal{F}_{t-1})$ is not constant, but varies over time. In the case of daily asset returns, $\boldsymbol{\beta}_1 = \mathbf{0}$. One possible model capturing such variations over time is the Autoregressive Conditional Heteroscedasticity ($GARCH$) model first proposed by Engle (1982). Other related models where the conditional variance of ε_t is used as one of the regressors explaining the conditional mean of y_t have also been suggested in the literature and are known as $GARCH$-in-Mean and $GARCH$-in-Mean (or $GARCH$-M, for short) models. (See, for example, Engle, Lillien, and Robins (1987)). For a useful survey of the literature on $GARCH$ modelling see Bollerslev, Chou, and Kroner (1992). Shephard (2005) provides selected readings of the literature.

8.3 Univariate GARCH models

The basic econometric model underlying volatility modelling options is the Generalized Autoregressive Conditional Heteroscedastic ($GARCH$) model proposed by Engle (1982) and Bollerslev (1986). This model assumes that

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + u_t \tag{8.3}$$

$$V(u_t | \Omega_{t-1}) = h_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \phi_i h_{t-i}^2 + \boldsymbol{\delta}'\mathbf{w}_t \tag{8.4}$$

where h_t^2 is the conditional variance of u_t with respect to the information set Ω_{t-1}, and $\mathbf{w}_t$ is a vector of predetermined variables assumed to influence the conditional error variances in addition to the past squared errors. In what follows we refer to $\sum_{i=1}^{q} \alpha_i u_{t-i}^2$ and $\sum_{i=1}^{p} \phi_i h_{t-i}^2$ in (8.4) as the MA and the AR parts of the $GARCH(p,q)$, respectively. The $GARCH$ model of Bollerslev, or $GARCH(p,q)$ for short is a special case of (8.4) where $\boldsymbol{\delta} = 0$.

Microfit allows to estimate a number of variants of the $GARCH$ model, such as the $GARCH$ in mean, the absolute $GARCH$ and the exponential $GARCH$ models.

In the $GARCH(p,q)$-in-mean specification (for short, $GARCH(p,q)$-M) the conditional error variance is used as one of the regressors explaining the conditional mean of y_t

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + \gamma h_t^2 + u_t \tag{8.5}$$

where the conditional error variance $h_t^2 = V(u_t | \Omega_{t-1})$ is defined by (8.4).

In the absolute $GARCH(p,q)$ model the conditional standard error of the disturbances u_t in (8.3) are specified by

$$h_t = \sqrt{V(u_t | \Omega_{t-1})} = \alpha_0 + \sum_{i=1}^{q} \alpha_i |u_{t-i}| + \sum_{i=1}^{p} \phi_i h_{t-i} + \boldsymbol{\delta}'\mathbf{w}_t \tag{8.6}$$

A variant of the $GARCH(p,q)$-M is the absolute value $GARCH$-in-Mean model (or $AGARCH(p,q)$-M for short), defined by (8.5) and (8.6). This model has been introduced into the literature by Heutschel (1991).

According to the exponential $GARCH(p,q)$ model ($EGARCH(p,q)$ for short) the logarithm of the conditional variance of the errors in (8.3) has the following specification:

$$\begin{aligned}\log h_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \left(\frac{u_{t-i}}{h_{t-i}}\right) + \sum_{i=1}^{q} \alpha_i^* \left(\left|\frac{u_{t-i}}{h_{t-i}}\right| - \mu\right) \\ &\quad + \sum_{i=1}^{p} \phi_i \log h_{t-i}^2 + \boldsymbol{\delta}'\mathbf{w}_t \end{aligned} \tag{8.7}$$

where $\mu = E\left(\left|\frac{u_t}{h_t}\right|\right)$. The value of μ depends on the density function assumed for the standardized disturbances, $\epsilon_t = u_t/h_t$. This model, which is due to Nelson (1991), allows for the possible asymmetric effects of past errors on the conditional error variances.

A variant of the above model is the $EGARCH(p,q)$-in-mean model, given by (8.5) and (8.7).

See Section 23.1 for further details on the above models and the relevant algorithms.

8.4 Multivariate GARCH models

The literature on multivariate volatility modelling is large and expanding. Bauwens, Laurent, and Rombouts (2006) provide a recent review. A general class of such models is the multivariate generalized autoregressive conditional heteroscedastic (MGARCH) specification (Engle and Kroner (1995)). However, the number of unknown parameters of the unrestricted MGARCH model rises exponentially with m, and its estimation will not be possible even for a modest number of assets. The diagonal-VEC version of the MGARCH model is more parsimonious, but still contains too many parameters in most applications. To deal with the curse of dimensionality the dynamic conditional correlations (DCC) model is proposed by Engle (2002), which generalizes an earlier specification by Bollerslev (1990) by allowing for time variations in the correlation matrix. This is achieved parsimoniously by separating the specification of the conditional volatilities from that of the conditional correlations. The latter are then modelled in terms of a small number of unknown parameters, which avoids the curse of the dimensionality. With Gaussian standardized innovations; Engle (2002) shows that the log-likelihood function of the DCC model can be maximized using a two-step procedure. In the first step, m univariate $GARCH$ models are estimated separately. In the second step using standardized residuals, computed from the estimated volatilities from the first stage, the parameters of the conditional correlations are then estimated. The two-step procedure can then be iterated, if desired, for full maximum likelihood estimation.

DCC is an attractive estimation procedure which is reasonably flexible in modelling individual volatilities, and can be applied to portfolios with a large number of assets. However, in most applications in finance the Gaussian assumption that underlies the two-step procedure is likely to be violated. To capture the fat-tailed nature of the distribution of asset returns, it is more appropriate if the *DCC* model is combined with a multivariate t-distribution, particularly for risk analysis where the tail properties of return distributions are of primary concern. But Engle's two-step procedure will no longer be applicable to such a *t-DCC* specification, and a simultaneous approach to the estimation of the parameters of the model, including the degree-of-freedom parameter of the multivariate t-distribution, would be needed. Pesaran and Pesaran (2007) (PP) develop such an estimation procedure and propose the use of devolatized returns computed as returns standardized by realized volatilities rather than by *GARCH*-type volatility estimates.

Microfit 5 allows you to estimate Engle's *DCC* and PP's *t-DCC* specifications for a relatively large number of asset returns. For an empirical application to risk management see Pesaran, Schleicherc, and Zaffaroni (2009).

8.4.1 DCC and t-DCC Multivariate Volatility Models

Let $\mathbf{r}_t$ be an $m \times 1$ vector of asset returns at close day t assumed to have mean $\boldsymbol{\mu}_{t-1}$ and the non-singular variance-covariance matrix $\boldsymbol{\Sigma}_{t-1}$, which we decompose as follows (see Bollerslev (1990) and Engle (2002)):

$$\boldsymbol{\Sigma}_{t-1} = \mathbf{D}_{t-1}\mathbf{R}_{t-1}\mathbf{D}_{t-1}$$

where

$$\mathbf{D}_{t-1} = \begin{bmatrix} \sigma_{1,t-1} & 0 & \dots & 0 \\ 0 & \sigma_{2,t-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & \sigma_{m,t-1} \end{bmatrix} \tag{8.8}$$

$$\mathbf{R}_{t-1} = \begin{bmatrix} 1 & \rho_{12,t-1} & \rho_{13,t-1} & \cdots & \rho_{1m,t-1} \\ \rho_{21,t-1} & 1 & \rho_{23,t-1} & \cdots & \rho_{2m,t-1} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & & \rho_{m-1,m,t-1} \\ \rho_{m1,t-1} & \cdots & \cdots & \rho_{m,m-1,t-1} & 1 \end{bmatrix} \tag{8.9}$$

In (8.8)-(8.9), $\sigma^2_{i,t-1}$ is the conditional volatility defined by

$$\sigma^2_{i,t-1} = V\left(r_{it} \mid \Omega_{t-1}\right)$$

and $\rho_{ij,t-1}$ are the conditional pair-wise return correlations

$$\rho_{ij,t-1} = \frac{Cov\left(r_{it}, r_{jt} \mid \Omega_{t-1}\right)}{\sigma_{i,t-1}\sigma_{j,t-1}}$$

where Ω_{t-1} is the information set available at close of day $t-1$. Clearly, $\rho_{ij,t-1} = 1$, for $i = j$.

The *MGARCH* option in *Microfit* allows joint estimation, by maximum likelihood, of the following system of dynamic equations, known in the literature as the dynamic conditional correlation (*DCC*) model (Engle (2002)):

$$\sigma^2_{i,t-1} = \bar{\sigma}^2_i\left(1 - \lambda_{1i} - \lambda_{2i}\right) + \lambda_{1i}\sigma^2_{i,t-2} + \lambda_{2i}r^2_{i,t-1}, \quad i = 1, ..., m \tag{8.10}$$

$$\tilde{\rho}_{ij,t-1}(\delta) = \frac{q_{ij,t-1}}{\sqrt{q_{ii,t-1}q_{jj,t-1}}}, \quad i \neq j = 1, ..., m \tag{8.11}$$

In (8.10), $\bar{\sigma}_i^2$ is the unconditional variance of the ith asset return, and $\lambda_{1i}, \lambda_{2i}$ for $i = 1, ..., m$, are unknown parameters. In (8.11), $q_{ij,t-1}$ is

$$q_{ij,t-1} = \bar{\rho}_{ij}(1 - \delta_1 - \delta_2) + \delta_1 q_{ij,t-2} + \delta_2 \tilde{r}_{i,t-1}\tilde{r}_{j,t-1} \tag{8.12}$$

where $\bar{\rho}_{ij}$ is the unconditional pairwise correlation between r_{it} and r_{jt}; δ_1, δ_2 are unknown parameters; and $\tilde{r}_{i,t-1}$ are standardized returns. *Microfit* offers two alternative ways of standardizing returns:

1. Exponentially weighted returns $\tilde{r}_{it}$ (Engle (2002)):

$$\tilde{r}_{it} = \frac{r_{it}}{\sigma_{i,t-1}} \tag{8.13}$$

   with $\sigma_{i,t-1}^2$ given by (8.10).

2. Devolatilized returns (Pesaran and Pesaran (2007)):

$$\tilde{r}_{it} = \frac{r_{it}}{\tilde{\sigma}_{it}^2(p)} \tag{8.14}$$

   with $\tilde{\sigma}_{it}^2(p) = \frac{\sum_{s=0}^{p-1} r_{i,t-s}^2}{p}$.

Further details on the estimation of dynamic conditional correlations models are provided in Section 23.2. See also Engle (2002), Pesaran and Pesaran (2007), and Pesaran, Schleicherc, and Zaffaroni (2009).

8.5 Volatility Modelling Menu

The Volatility Modelling Menu contains the following options

1. Univariate $GARCH$
2. Multivariate $GARCH$ applied to a set of regressors
3. Multivariate $GARCH$ applied to OLS residuals

Option 1 allows you to estimate a variety of conditionally heteroscedastic models, such as $GARCH$, $GARCH$, exponential $GARCH$, absolute $GARCH$, $GARCH$ in mean models both for normally and Student's t-distributed errors. See Section 8.6.

Option 2 enables you to estimate dynamic conditional correlation models for a set of variables which could follow a multivariate normal or a Student's t-distribution. See Section 8.7.

Option 3 enables you to estimate dynamic conditional correlation models on OLS residuals, after having controlled for a set of regressors. See Sections 21.6 and 8.7.

8.6 Univariate GARCH Estimation Menu

The Univariate $GARCH$ Estimation Menu enables you to estimate a variety of univariate conditional heteroscedastic models. When you click on option 1 you will be presented with the following options:

0. Move to Backtracking Menu
1. $GARCH$, Auto-Regressive Conditional Heteroscedasticity
2. $GARCH$-M, $GARCH$ in Mean
3. $AGARCH$, Absolute value $GARCH$
4. $AGARCH$-M, Absolute value $GARCH$ in Mean
5. $EGARCH$, Exponential $GARCH$
6. $EGARCH$-M Exponential $GARCH$ in Mean

Option 0 takes you back to the Commands and Data Transformations box.

Option 1 allows you to estimate the $GARCH$ model (8.3)-(8.4). The variables $\mathbf{x}_t$ and $\mathbf{w}_t$ must be in *Microfit's* workspace and can include lagged values of y_t.

Option 2 enables you to compute ML estimates of the $GARCH(p, q)$-in-mean model.

Option 3 allows you to compute ML estimates of the absolute value $GARCH(p, q)$

Option 4 enables you to estimate the absolute value $GARCH$-in-Mean model.

Option 5 allows you to compute ML estimates of the exponential $GARCH(p, q)$ model.

Option 6 enables you to estimate the $EGARCH(p, q)$-in-Mean model given by (8.5) and (8.7).

8.6.1 Specification of the GARCH, AGARCH and EGARCH models

When you choose any one of the six estimation options in $GARCH$-M Estimation Menu you will be presented with the following sub-menu

0. Return to $GARCH$ Estimation Menu
1. Estimate assuming a normal distribution for conditional errors
2. Estimate assuming a t-distribution for conditional errors

which gives a choice between a conditional normal density and a conditional standardized Student-t distribution for the disturbances.[1] Once you select one of these two conditional distributions you will be asked to specify the orders of the $GARCH(p, q)$ in the box editor that appears on the screen. You need to type the non-zero lags in the AR and the MA parts of the $GARCH$ specification, respectively.

Separate the two sets of numbers by ;. Each set of numbers should be in ascending order.

[1]The use of Student-t distribution for the standardized errors, $\epsilon_t = u_t/h_t$, has been suggested by Bollerslev (1987).

A set can contain only a single 0. Examples:

To specify an $GARCH(1)$ type 0 ; 1

To specify a $GARCH(2, 1)$ type 1 2 ; 1

To specify a restricted $GARCH(4)$ type 0 ; 4

To specify OLS/ML estimation type 0 ; 0

Notice that the same rules apply when you specify $AGARCH$ and $EGARCH$ classes of model.

Having specified the orders of your $GARCH$ model the program asks you to specify the list of the variables $\mathbf{w}_t$ (if any), to be added to the specification of the conditional variance equations (see (8.4), (8.6) and (8.7)). If you do not wish to include any other variables in the equation for the conditional variances simply click the button to move to the next stage of the program, where you will be asked to supply initial estimates for the parameters of the $GARCH$-M models.

8.6.2 Specification of the initial parameter values for GARCH, AGARCH and EGARCH models

Once you have completed the specification of your conditional heteroscedastic model you will be asked to supply initial estimates for the parameters of your model. Type your choice and click the button. An appropriate choice for the initial estimates is often critical for a successful convergence of the numerical algorithm used to compute the ML estimates. This is particularly important in the case of generalized $ARCH$ models with a non-zero AR component. The following points are worth bearing in mind.

1. The algorithm often fails to converge if you try to estimate a $GARCH$ model when there is in fact no statistically significant evidence of an $ARCH$ effect in the data. After running an OLS regression to check for the presence of an $ARCH$ effect in your regression, use option 2 in the Hypothesis Testing Menu (see Section 6.23).

2. It is often advisable not to choose initial values that are on the boundary of the feasible set. For example, in the case of the $GARCH$ specification (8.4), the values of α_i and ϕ_i should be such that $\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \phi_i$ is not too close to unity. It is important that *positive non-zero* values are chosen for these parameters.

3. Make sure that the residuals in the underlying regression model are not serially correlated. Presence of significant residual serial correlation can create problems for the ML estimation of the $GARCH$ model.

4. The algorithm may fail to converge if the observations on the dependent variable are very small. Scale up these observations and re-estimate the $GARCH$ model.

8.6.3 Estimation results for the GARCH-M options

The estimation results for the $GARCH$-M options are summarized in a table. The top half gives the ML estimates of the regression coefficients, $\boldsymbol{\beta}$, and the estimate of γ (when a $GARCH$-in-Mean model is estimated), their estimated asymptotic standard errors, t-ratios, as well as a number of summary statistics, and model selection criteria. The bottom of the table gives the ML estimates of the parameters of the conditional variance model together with their asymptotic standard errors.

After the estimation results you will be presented with the Post Regression Menu, with a number of options described in detail in Section 6.20. In particular, you can plot/save estimates of the conditional standard errors, $\hat{h}_t$, and save their forecasts, if any. (To compute forecasts of h_t you need to choose options 8 or 9 in the Post Regression Menu.) You can also plot the histogram of the standardized (or scaled) residuals, $\hat{\epsilon}_t = \hat{u}_t/\hat{h}_t$. To access these options select option 3 in the Post Regression Menu after the $GARCH$ estimation results.

8.7 Multivariate GARCH Menu

Option 2 and 3 in the Volatility Modelling Menu allow you to estimate dynamic conditional correlation (DCC) models on a set of variables or regression residuals. If you choose option 2, in the screen editor you need to list the variables that you wish to include in your DCC model; if you select option 3 you should list the dependent variable followed by its regressors for each equation, separating the different equations by a semicolon (;).

After you specify your model (using either option 2 or 3 in the Volatility Modelling Menu), you will be presented with the screen shown in Figure 8.1.

The Multivariate $GARCH$ window is divided into two panels. The left panel allows you to set a number of characteristics in your multivariate $GARCH$ model. In particular:

- In the 'Decay Factor for the Variance' field you can impose the parameters λ_{1i} and λ_{2i} in (8.10) to be the same for all assets, that is $\lambda_{1i} = \lambda_1$ and $\lambda_{2i} = \lambda_2$ for $i = 1, ..., m$. Further, you can impose the restriction that $\lambda_{1i} + \lambda_{2i} = 1$; that is that conditional volatilities are non-mean reverting. See Section 23.2 for further details.

- In the 'Decay Factor for the Covariance' field you can impose the restriction on (8.11) that $\delta_1 + \delta_2 = 1$; that is, that conditional correlations are non-mean reverting in the case of all the assets.

- In the 'Distributional assumption' field, you can choose either the multivariate normal distribution or the multivariate Student's t-distribution for conditional returns $\mathbf{r}_t$. If you choose t-distributed returns, a new parameter, the number of degrees of freedom of the t-distribution (df), appears among the parameters to be estimated.

- In the 'Standardized returns' field you can decide the way you standardize your returns, using the exponentially weighted returns given by (8.13), or the devolatilized returns given by (8.14). For devolatilized returns you need to choose the lag-order, p, to compute the realized volatilities in the 'Rolling volatility window' field. Note that

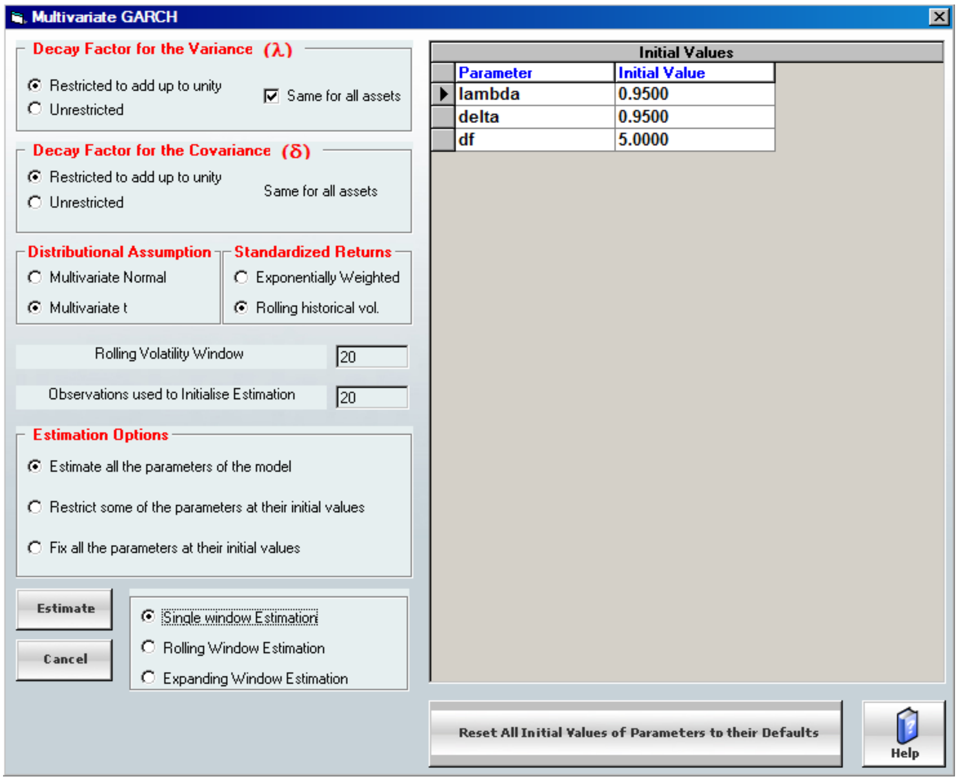


Figure 8.1: The Multivariate *GARCH* window

p should be smaller or equal to the number of observations used for inizialization of estimation.

- In the field 'Observations used to initialized' you need to set the number of observations, T_0, used for the initialization of recursions in (8.10) and (8.11).

- In the 'Estimation options' you may restrict some or all parameters at their initial values. See Section 23.2.1.

- At the bottom of the Multivariate *GARCH* window you can choose to estimate the *DCC* model using a single, expanding or rolling window of observations. The option 'Single window estimation' employs all available observations in the maximum likelihood estimation. The option 'Expanding window estimation' uses an expanding set of observations. If you choose this option, you need to specify the observation from which you wish to start expanding the estimation. The 'Rolling window estimation' option estimates the model over successive rolling samples of a fixed length to be specified. You can choose the size of the rolling window in the field below this option. Note that the number of observations you specify for the expanding or rolling sample must be greater than 20 times the number of parameters to be estimated, and smaller than the total number of observations used for estimation.

The left panel in the *MGARCH* window contains the names of the parameters to be estimated and their initial values. You can change these initial settings by moving the cursor to the desired position and by typing your own choice of the initial estimate. These initial parameter values can be readily changed if the estimation method fails to converge.

8.7.1 Estimation results for the MGARCH

The estimation results for the multivariate *GARCH* options 2 and 3 are set out in a window which consists of two parts. The top part gives the *ML* estimates of the volatility decay parameters λ_1, λ_2, the correlation decay parameters δ_1 and δ_2, and, if the multivariate t-distribution is selected, the degrees of freedom (df) of the t-distribution, together with their standard errors and the associated t-ratios.

The bottom part of the result window gives a table which has on the diagonal the unconditional variances of asset returns, computed over the initialization plus estimation sample (see Section 23.2.1). Namely, for the ith asset

$$\bar{\sigma}_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T} r_{it}^2}$$

where T is the sample size of the initialization and estimation sample. The off-diagonal elements are the unconditional correlations, estimated as

$$\bar{\rho}_{ij} = \frac{\sum_{t=1}^{T} r_{it} r_{jt}}{\sqrt{\sum_{t=1}^{T} r_{it}^2}\sqrt{\sum_{t=1}^{T} r_{jt}^2}}$$

8.8 Multivariate GARCH Post Estimation Menu

This menu appears on the screen after the estimation results for the *DCC* model, using either option 2 or 3 in the Volatility Modelling Menu. The Multivariate *GARCH* Post estimation Menu contains the following options

0. Re-specify *MGARCH* model and estimate or Quit
1. Display estimation results again
2. List/Plot/Save estimated conditional volatilities, correlations and eigenvalues
3. Wald test for linear/non-linear restrictions
4. Estimate/test functions of parameters of the model
5. Test the validity of the *MGARCH* model (*VaR* diagnostics)
6. Calculate the Value at Risk (*VaR*) of a portfolio
7. Compute forecasts of conditional volatilities and correlations

Option 0 returns you to the Multivariate *GARCH* window where you can re-specify your model (see Section 8.7).

Option 1 enables you to see the ML estimation of the $MGARCH$ model again (see Section 8.7.1).

Option 2 allows you to to list/plot/save the estimated conditional volatilities and correlations and the eigenvalues of covariance and correlation matrices.

Option 3 enables you to carry out Wald tests for linear and non-linear restrictions on the coefficients. For the relevant formulae see Section 21.25.

Option 4 allows you to estimate linear and non-linear functions of the coefficients. For the relevant formulae see Section 21.24.

Option 5 allows you to test the validity of the DCC model using a set of diagnostics based on the VaR and on the Probability integral transforms. See Section 23.2.3.

Option 6 enables you to compute the Value at Risk (VaR) of a portfolio. See Section 23.2.3.

Option 7 computes forecasts of conditional volatilities and correlations. Using this option, you can list or plot forecasts, or save them in a special *Microfit* file for later use. See Section 23.2.4.

8.8.1 Testing the Validity of Multivariate GARCH Menu

This menu appears if you choose option 5 from the Multivariate $GARCH$ Post Estimation Menu and contains the following options

0. Quit
1. Tests based on Probability Integral Transforms
2. Testing for VaR Exceptions

Option 0 returns you to the $MGARCH$ Post estimation Menu.

Option 1 allows you to compute the LM test and the KS test of randomness of the probability integral transforms, and to plot or save the histogram of the probability integral transforms. Under a correct specification of the DCC model, these should reproduce the density of a uniform random variable (see Section 23.2.3 for further details).

Option 2 enables you to list, plot and save the VaR for a given tolerance level, over the evaluations period. This option also allows you to compute the mean VaR violations and associated diagnostic test statistics (see Section 23.2.3 for further details).

8.8.2 Compute the VaR of a portfolio

When you select option 6 from the Multivariate $GARCH$ Post Estimation Menu you will be presented with the screen reported in Figure 8.2. The VaR calculator allows you to compute the one-step ahead Value at Risk of a portfolio of your own choice for a given probability level and to compute its probability level for a given VaR.

On the left panel of the screen you can choose the allocation of the assets in your portfolio, as well as the return for each asset. On the right part of the screen you can choose whether asset returns are expressed in percentage or in points, by checking the appropriate checkbox. On the bottom of the screen you can decide whether in the computation of the VaR you want to use the one-step ahead forecasts of variances and covariances or the estimated variances and covariances.

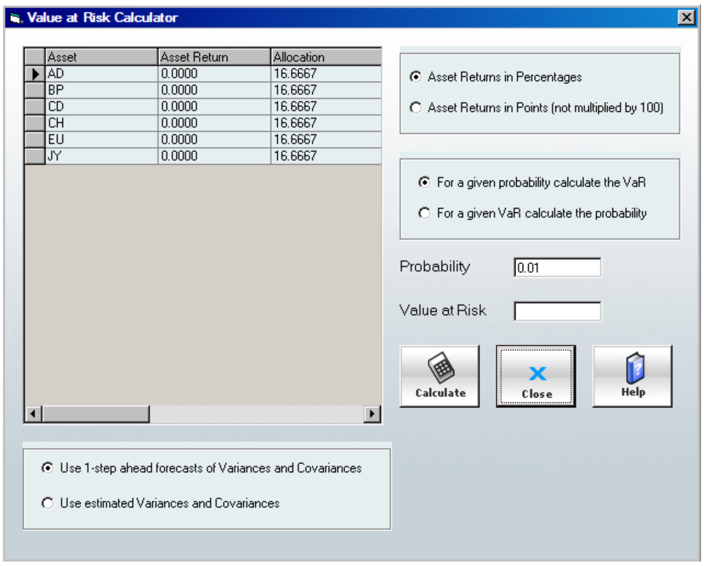


Figure 8.2: Compute the VaR of a portfolio

Part IV

Tutorial Lessons

Chapter 9

Lessons in Data Management

The tutorial lessons in this chapter demonstrate the input/output features of *Microfit*. We start these lessons with an example of how to read ASCII (text) data files, using the data file UKSTOCK.DAT. This file contains the monthly observations on a number of financial series for the UK economy. It is assumed that you have already gone through the steps set out in Chapter 2, that *Microfit 5.0* is properly installed on your system and that the various tutorial data files are on your hard-disk. To open a data file from the tutorial directory, choose 'Open file from tutorial data files' from the file Menu.

9.1 Lesson 9.1: Reading in the raw data file UKSTOCK.DAT

Load *Microfit 5.0*, and choose 'Open file from tutorial data files' from the file Menu. This displays the Open file dialogue with a list of files on the left.

Initially, we suggest that you try to load into *Microfit* the data file UKSTOCK.DAT which is in ASCII format. This file contains seven economic time series for the UK economy organized by observations over the period $1970M1$-$1995M5$, where M denotes that the data are monthly and the integers $1, 2, ..., 12$ denote the months starting with January = 1. All variables refer to the last trading day of the month. The seven variables are arranged in the file in the following order:

1. Financial Times 500 Composite Share Index
2. FT30 Dividend Yield
3. Money Supply (M0)
4. Three Month Treasury Bill Rate (end of period)
5. Average Gross Redemption Yield on 20-year Government Securities
6. Exchange rate: US $ to 1 Pound Sterling (Spot Rate)
7. Retail (Consumer) Price Index

To load (read) the data, choose .DAT data files from the List file of type box, then double click on the file UKSTOCK.DAT. You should now see the New dataset dialogue on the screen.

For frequency of the data choose monthly. For the sample period choose the start year and month as 70 and 1, and then the end year and month as 95 and 5. Choose to organize your data by observation, make sure the Free format radio button is selected, then click 

.

Microfit assigns the variable names $X1$, $X2$, $X3$, $X4$, $X5$, $X6$, and $X7$ to your variables by default. To change these variables, move to each cell of the table in turn and type in the variable names

$$ukftidx \;\; ukftdy \;\; ukm0 \;\; uk3tbr \;\; uk20yr \;\; ukexch \;\; ukcpi$$

and then click 

.

If the data are read unsuccessfully, you will see an error message. Read Section 3.2 carefully, and start the lesson again!

After the raw data file is read in successfully it is good practice to first inspect the data in the Data window to ensure that they have been read in correctly. To check the data click the 

option. You should see the list of monthly observations on all the seven variables on the screen. You can use the scroll bars, and the PgUp, PgDn, Ctrl+Home keys to move around the list.

To save the data in the workspace in binary format for subsequent use with *Microfit*, you need to select the 'Save as' option from the File Menu (see Lesson 9.2).

9.2 Lesson 9.2: Saving your current dataset as a special *Microfit* file

Once you have satisfied yourself that the raw data file is read in correctly, you may wish to save it as a special *Microfit* file for use in subsequent sessions. Special *Microfit* files are saved as binary files and allow you rapid access to your data without any need to supply the details of your data every time you wish to use them with *Microfit*.

To save your current dataset in a special *Microfit* file choose 'Save as' from the File Menu (see Section 3.5). This takes you to the Save as dialogue. Make sure '*Microfit* data files' is selected in the List files of type box, enter the filename UKSTOCK.FIT, and click 

. Since the file UKSTOCK.FIT already exists, you will be asked if you want to replace it. Choose No to return to the Save as dialogue (but to overwrite the file UKSTOCK.FIT choose Yes).

9.3 Lesson 9.3: Reading in the special *Microfit* file UKSTOCK.FIT

The file UKSTOCK.FIT is the special *Microfit* file corresponding to the raw data file UKSTOCK.DAT. To read UKSTOCK.FIT you need to choose 'Open file from tutorial data files' from the File Menu and choose UKSTOCK.FIT by double clicking on it (see Section 3.2). The program starts reading the data from the file, and assuming that the data have been read in successfully, it displays the Process window (see Chapter 4).

9.4 Lesson 9.4: Combining two special *Microfit* files containing different variables

Suppose you wish to add the variables in *Microfit* file UKSTOCK.FIT to the variables in the special *Microfit* file USSTOCK.FIT. First read in the file UKSTOCK.FIT (see Lesson 9.3). Once this is done successfully, choose 'Add a special *Microfit* file to workspace' from the File Menu and when the Open file dialogue appears select USSTOCK.FIT. If the two files are combined successfully, a message confirming this is displayed.

Click OK to return to the Process window. To make sure that the variables in USSTOCK.FIT ($USLGR$, $USCPI$, $USM1$, $US3TBR$, $USSIDX$, and $USDY$) are correctly added to the current variables (namely the variables in UKSTOCK.FIT), click the Variables button to display the list of variables. There should be 13 variables: seven from UKSTOCK.FIT and six from USSTOCK.FIT on your workspace. Use the 'Save as' option to save the combined dataset under a different filename before proceeding further.

9.5 Lesson 9.5: Combining two special *Microfit* files containing the same variables

One of the tutorial files, the special *Microfit* file EJCON1.FIT, contains annual observations (1948-1981) on the following eight variables[1]:

AB	Personal bond holdings
AM	Net liquid assets net of house loans
AS	Personal share holdings
BP	Bond prices
C	Real consumption expenditures
PC	Nominal consumption expenditures
SP	Share prices
Y	Real disposable income

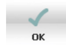
Suppose you wish to update and extend this dataset with observations on the same variables over the period 1970 to 1985, saved in the Special *Microfit* file $EJCON2.FIT$.

Select the option 'Add 2 special *Microfit* files' from the File Menu. First choose the file EJCON1.FIT, click OK, and then choose EJCON2.FIT from the Open dialogue. The program augments the data contained in EJCON1.FIT using the new and additional observations from the file EJCON2.FIT in the manner described in Section 3.3.1. To inspect the combined dataset click the Data button. The dataset displayed on the screen should now contain observations on the eight variables C, PC, Y, AS, AB, AM, SP, and BP over the extended period from 1948 to 1985, inclusive. Notice also that observations for the period 1970 to 1981 contained in the file EJCON1.FIT are now overwritten by the corresponding

[1] For a description and sources of this dataset see Pesaran and Evans (1984) and the manual of *Microfit 3.0*.

observations in the file EJCON2.FIT. To save this revised and extended dataset in a special *Microfit* file, see Lesson 9.2. We have already saved this dataset in the file EJCON.FIT and this file should be in the tutorial directory or on your *Microfit* disks.

9.6 Lesson 9.6: Extending the sample period of a special *Microfit* file

By combining two files you can extend the sample period of an existing special *Microfit* file. Suppose you wish to extend the period of the dataset UKSTOCK.FIT from 1970(1)-1995(5) to 1965(1)-1996(12). First start with a new dataset choosing 'Input data from the keyboard' from the File Menu (see Section 3.2). Choose the monthly data frequency, and enter the start year and month as 1965 and 1, and the end year and month as 1996 and 12. Save your file as a special *Microfit* file with a name of your choice. Choose 'Add a special *Microfit* file to workspace' from the File Menu and find the file UKSTOCK.FIT. Click 

. Back in the Process window, click the 

button. You should now see the observations on the variables, $UKFTIDX$, $UKFTDY$, $UKM0$, $UK3TBR$, $UK20YR$, $UKEXCH$, $UKCPI$ in your workspace, now extended over the period 1965(1) to 1996(12). The values of all these variables over the periods 1965(1) to 1969(12) and 1995(6) to 1996(12) will be set to blank. You can replace some or all missing values by clicking on the relevant cells and typing in new values, or by using the the **FILL_MISSING** and **FILL_FORWARD** commands (see Chapter 4 on how to use these and other commands).

9.7 Lesson 9.7: Reading the CSV file UKCON.CSV into *Microfit*

The file UKCON.CSV is a comma delimited values file containing quarterly observations on the following seven variables obtained from the UK Central Statistical Office (CSO95) data bank:

$AIIWQA$	Personal disposable income £m (seasonally adjusted)
$AIIXQA$	Consumers' expenditure: Total £m CURR SA
$CAABQA$	Consumers' expenditure: Total £m CONS (1990 prices) SA
$CCBHQU$	Consumers' expenditure: Total £m CONS (1990 prices) NSA
$CECOQU$	Real personal disposable income at 1990 prices (seasonally unadjusted)
$CECPQA$	Real personal disposable income at 1990 prices (seasonally adjusted)
$DQABQU$	Tax and prices index (Jan 1987 = 100)

The variable names and their descriptions are the same as those in the CSO95 databank.

If you read this file into *Excel* you will see that the data in the worksheet are arranged in columns with the first column being the dates, while the first two rows contain the variable names and their descriptions, respectively. The last column contains observations on 'Tax and Prices index', but only over the period 1987(1) to 1995(1).

To read this worksheet into *Microfit*, choose 'Open file' from the File Menu. From the Open dialogue select UKCON.CSV by double-clicking on it. *Microfit* attempts to read in the file and, if successful, displays the message

UKCON.CSV imported successfully

Click OK to move to the Process window, then click the Variables button to see the seven variables in the file UKCON.XLS and their descriptions. To see the observations on all the variables click the Data button. Inspect the data carefully and make sure that they are imported into *Microfit* correctly. Note that the missing values of $DQABQU$ (tax and prices index) over the period 1955(1) to 1986(4) are set to blank.

9.8 Lesson 9.8: Reading the Excel file DAILYFUTURES.XLS into *Microfit*

The file DAILYFUTURES.XLS is an *Excel* file containing daily observations on futures returns on the S&P 500 index of the US stock market (SP), the Financial Times Stock Exchange 100 index of the London Stock Exchange ($FTSE$), and the Blue chip index of the German stock market (DAX), over the period from 01-May-02 to the 01-May-07.

To read this worksheet into *Microfit*, choose 'Open file from tutorial data files' from the File Menu. From the Open dialogue select DAILYFUTURES.XLS by double-clicking on it.

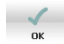
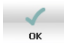
Once *Microfit* has successfully loaded the data, click on the Data button. You will see that *Microfit* has created three new columns, the DAY, $MONTH$ and $YEAR$, containing for each observation the corresponding information on the day, month and year.

For further information on how to input daily data, see Section 3.2.10.

Note: Dates in the file DAILYFUTURES.XLS are expressed in the European format. If the Windows regional settings of your computer are set to English United States, then you will need to change the default date format to the European case in *Microfit* before opening the file. To change the default date format in *Microfit*, go to the Options Menu, choose the 'European/US date format', select the European Date format option and click OK. Alternatively, you can load data in CSV format, contained in the file DAILYFUTURES.CSV in the tutorial directory.

9.9 Lesson 9.9: Saving the DAILYFUTURES.XLS file excluding missing values

Load into *Microfit* the *Excel* file DAILYFUTURES.XLS (see Lesson 9.8), and click on the Data button. Notice that there are jumps in the data, due to the fact that the stock exchange does not trade during weekends, and missing values for each series, generally due to holidays that can vary from country to country. In the presence of missing values, *Microfit* does not allow the use of some functions, such as **MEAN, SUM, CSUM,** and commands, such as **ADF, KPSS, SPECTRUM** (see Chapter 4 for the use of these and other commands). Further, in the presence of missing values in the middle of the estimation period *Microfit* does not carry the estimation of any univariate or multivariate models. To deal with this problem, you can either create a new dataset where observations with missing values are dropped, or impute missing data using the **FILL_FORWARD** and **FILL_MISSING** commands, and save the 'filled-in' data as a new data file for your future use.

To exclude observations with missing values, choose the 'Save as' option in the File Menu, and select 'CSV, descriptions in 2nd row, exclude rows with missing values undated and daily data only'. A 'Save as' dialogue appears; type in a filename in the usual way, and click . You will be asked to select the sample period. Specify the initial and the final period, and click . If you wish to see the new dataset with no missing observation that *Microfit* has created, select the 'Open file' option from the File Menu, and from the Open dialogue double click on the file you have just saved. Once *Microfit* has successfully imported the data, click on the  button, and check that the data do not contain blank cells.

9.10 Exercises in data management

9.10.1 Exercise 9.1

The raw data file USSTOCK.TXT contains monthly observations covering the period 1973M3 to 1995M6 (inclusive) on the following variables

Yield on Long-Term US Government Bonds
US Consumer Price Index
US Money Supply (M1 definition)
Three-month US Treasury Bill Rate
US Share Prices - Standard and Poor 500 (SP500) Composite Index
Dividend Yield on SP500

These observations are arranged 'variable-by-variable' in the file in free format. Input this file into *Microfit*, check that it is correctly read in, and save it as a Special *Microfit* file.

9.10.2 Exercise 9.2

Load the file EJCON.FIT containing the variables AB, AM, AS, BP, C, PC, SP, and Y into *Microfit* and add to it the file EU.FIT, containing annual observations on the variables

E Employees in employment (1000s)
U Unemployed including school leavers

9.10.3 Exercise 9.3

Repeat the steps in Lesson 9.7 and read in the *Excel* worksheet file UKCON.XLS into *Microfit*. Save them in a *Microfit* file and add the five variables in the file UKCON.FIT to the variables in the workspace. Check that the observations on the variables C and $CAABQA$ are in fact identical.

9.10.4 Exercise 9.4

Load the file UKSTOCK.FIT into *Microfit*, and then save the variables in this file as a CSV file. Exit *Microfit* and read the CSV file created by *Microfit* into *Excel*. Are the observations exported correctly?

Chapter 10

Lessons in Data Processing

The Lessons in this chapter show how to carry out data transformations on the data on the workspace by issuing commands and formulae in the Process window. It is assumed that you have already gone through the steps set out in Chapter 2, that *Microfit 5.0* is properly installed on your system and that the various tutorial data files are on your hard-disk. To open a data file from the tutorial directory, choose 'Open file from tutorial data files' from the file Menu.

10.1 Lesson 10.1: Interactive data transformations

You can carry out the transformations which you require on your data either interactively or by executing an already prepared batch/equation file. Suppose you wish to analyse the quarterly movements in aggregate consumption expenditures in the UK. First read in the special *Microfit* file UKCON.FIT (see Lesson 9.3). You should see the Commands and Data Transformation box on the screen in which you can type a formula to carry out data transformations on your existing variables, or issue one of the commands described above. For example, if you type

$$INPT = 1; \quad P = CNOM/C; \quad LY = \mathbf{LOG}(Y); \quad LC = \mathbf{LOG}(C);$$
$$PI = \mathbf{LOG}(P/P(-1)); \quad DLY = LY - LY(-1);$$
$$DLC = LC - LC(-1)$$

the program generates seven new variables:

$INPT$	Intercept term, (a vector with all its elements equal to unity)
P	Implicit price deflator of consumption expenditures (1990=1.00, on average)
LY	Logarithm of Y
LC	Logarithm of C
PI	Inflation rate (measured as the change in log of P)
DLY	Change in log of Y
DLC	Change in log of C

These new variables are now added to the list of your existing variables, and you should see them in the Variables box. If you wish to edit the variables' descriptions, click the 

button.

Note: The content of the Commands and Data Transformation box can be saved as an equation file (with the extension EQU), and retrieved at a later stage. Click the 

button to save the content of the Commands and Data Transformation box, enter the filename and click . We have already saved this file as UKCON.EQU, and should be on the tutorial directory (typically C:\PROGRAM FILES\MFIT5\TUTOR\).

10.2 Lesson 10.2: Doing data transformations using the BATCH command

A convenient method of carrying out data transformations is to first to create a BATCH file (using your preferred text editor before running *Microfit*), containing the instructions (formulae and commands) that you wish carried out, and then running this BATCH file interactively by means of the **BATCH** command. The file UKCON.BAT in the tutorial directory is an example of such a BATCH file. The content of the file UKCON.BAT is reproduced in Table 10.1.

Table 10.1: Content of the BATCH file UKCON.BAT

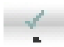
```
$ BATCH file UKCON.BAT, to be used in conjunction
$ with the special Microfit file UKCON.FIT
$
$
$ Generate an intercept term
Inpt=1
$ Generate implicit price deflator of consumer expenditures
p=cnom/c
$ Take (natural) logarithms
ly=log(y)
lc=log(c)
$ Generate rates of change of the variables computed as log-changes
pi=log(p/p(-1))
dly=ly-ly(-1)
dlc=lc-lc(-1)
$ Generate rates of change of the variables computed as
$ percentage change
rp=rate(p)
ry=rate(y)
rc=rate(c)
$ Note that rate (y) is computed as 100*(y-y(-1))/y(-1)
s=(y-c)/y
$ End of the BATCH file
```

You can also see the content of this file on screen by using the option 'View a File' from the File Menu and then double-clicking on UKCON.BAT.

To run the BATCH file UKCON.BAT, make sure that the file UKCON.FIT is loaded into *Microfit* and that the Commands and Data Transformation box is clear. Then click on the 

button on the right-hand side of the screen and select the desired BATCH file from the list of file names by double clicking on the file UKCON.BAT. Wait until the computations are completed and the message

Operations on batch file completed successfully

appears on the screen. If you now click  and then click the 

button, you will see the list of the five original variables together with 11 more variables created by the program in the process of executing the instructions contained in the BATCH file UKCON.BAT. The variable window should now look like the screen shown in Figure 10.1.

The variables in this list will be used in other lessons on preliminary data analysis, estimation, hypothesis testing, and forecasting.

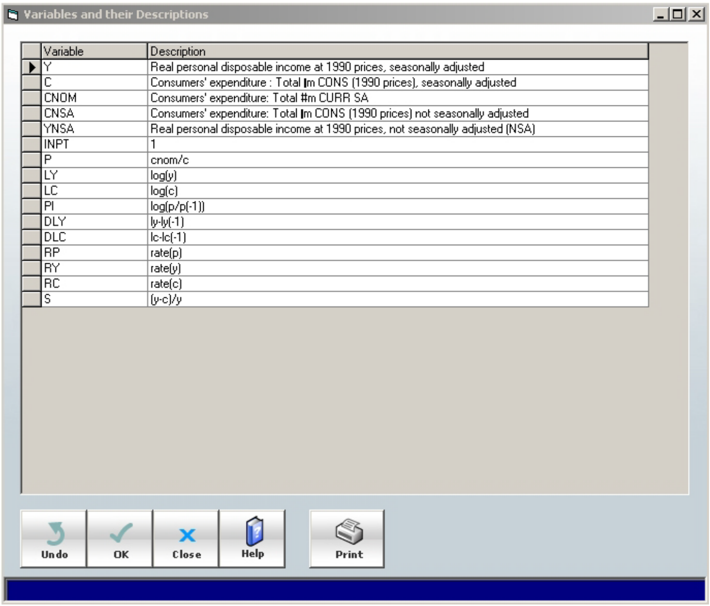


Figure 10.1: The Variable window

10.3 Lesson 10.3: Adding titles (descriptions) to variables

Suppose you wish to give titles to the seven variables in the special *Microfit* file UKSTOCK.FIT (see Lesson 9.3). This can be done either interactively or by means of a BATCH file. Read in the file UKSTOCK.FIT and click the  button. Move to each of the variables' description field in turn and type in a title. Alternatively, in the Commands and Data Transformations box type

ENTITLE

and click . You will be prompted in the Variable window where you can supply a title (a description not more than 80 characters long) to any variable you wish. For example, for the variable $UKFTIDX$ type in the title

Financial Times 500 Composite Share Index

Once you have supplied the description for all variables, click  to save and then  to return to the Process window. Note that you cannot undo the changes you have made to the variable description if you click .

Another alternative is to run the BATCH file UKNAMES.BAT. The content of this file is listed in Table 10.2. To run this BATCH file click the  button in the Process window. You will be presented with an Open file dialogue to help you search for the file UKNAMES.BAT on your PC.

Table 10.2: Content of the BATCH file UKNAMES.BAT

```
ENTITLE ukftidx ukftdy ukm0 uk3tbr uk20yr ukexch ukcpi
Financial Times 500 Composite Share Index
FT30 Dividend Yield
Money Supply M0
Three Month Treasury Bill Rate
Average Gross Redemption Yield on 20-Year Government Securities
Exchange Rate:  US$ to œl
Retail (Consumer) Price Index
```

10.4 Lesson 10.4: Creating dummy variables

In this lesson we will describe how to construct a dummy variable in *Microfit*. Suppose your current sample period is 1948-1981, and you wish to construct the following dummy variable

$$\begin{array}{ll} D_t = 0 & \text{for } 1948, 1949 \\ D_t = 1 & \text{for } 1950, ..., 1955 \\ D_t = 2 & \text{for } 1956, ..., 1960 \\ D_t = 3 & \text{for } 1961, ..., 1970 \\ D_t = 4 & \text{for } 1971, ..., 1981 \end{array}$$

Read in the file EJCON.FIT. In the Commands and Data Transformations box type

SAMPLE 1948 1981; $D = 0$;
SAMPLE 1950 1955; $D = 1$;
SAMPLE 1956 1960; $D = 2$;
SAMPLE 1961 1970; $D = 3$;
SAMPLE 1971 1981; $D = 4$;
SAMPLE 1948 1981; 

The program now creates the variable D in your workspace, with the following values:

OBS	D	OBS	D
1948	0	1965	3
1949	0	1966	3
1950	1	1967	3
1951	1	1968	3
1952	1	1969	3
1953	1	1970	4
1954	1	1971	4
1955	1	1972	4
1956	2	1973	4
1957	2	1974	4
1958	2	1975	4
1959	2	1976	4
1960	2	1977	4
1961	3	1978	4
1962	3	1979	4
1963	3	1980	4
1964	3	1981	4

As another example suppose you wish to create a variable which takes the value of zero over the period 1948-1968 (inclusive), and then increases by steps of unity from 1969 onward. You need to type

SAMPLE 1948 1968; $TD = 0$;
SAMPLE 1969 1981; **SIM** $TD = TD(-1) + 1$;
SAMPLE 1948 1981; **LIST** RUN

The variable TD should now have the following values:

OBS	TD	OBS.	TD
1948	0	1956	0
1949	0	1966	0
1950	0	1967	0
1951	0	1968	0
1952	0	1969	1
1953	0	1970	2
1954	0	1971	3
1955	0	1972	4
1956	0	1973	5
1957	0	1974	6
1958	0	1975	7
1959	0	1976	8
1960	0	1977	9
1961	0	1978	10
1962	0	1979	11
1963	0	1980	12
1964	0	1981	13

Alternatively you can use the cumulative sum function, **CSUM**(•), to construct this trend (see Section 4.3.4). Type

SAMPLE 1948 1968; $TD = 0$;
SAMPLE 1969 1981; $TD =$ **CSUM**(1);
SAMPLE 1948 1981; **LIST** TD

10.5 Lesson 10.5: Plotting variables against time and/or against each other

Suppose you have loaded in the special *Microfit* file UKCON.FIT, and wish to plot the variables C (real consumption expenditures) and Y (real disposable income) against time on the same screen. In the Commands and Data Transformations box type

PLOT C Y

for the graph to appear on the screen (see Figure 10.2 below).

You can alter the display of the graph, print it, or save it. To alter the display click and choose one the options. For more information, see Section 5.2. To print the displayed graph, click . You will be presented with a standard Windows Print dialogue.

You can save the displayed graph in a variety of graphic formats, such as Bitmap (BMP), Windows metafile (WMF), Enhance metafile (EMF), JPEG and PNG, by clicking the button, or choosing the 'Save as' option from the File Menu. You can save in Olectra Chart

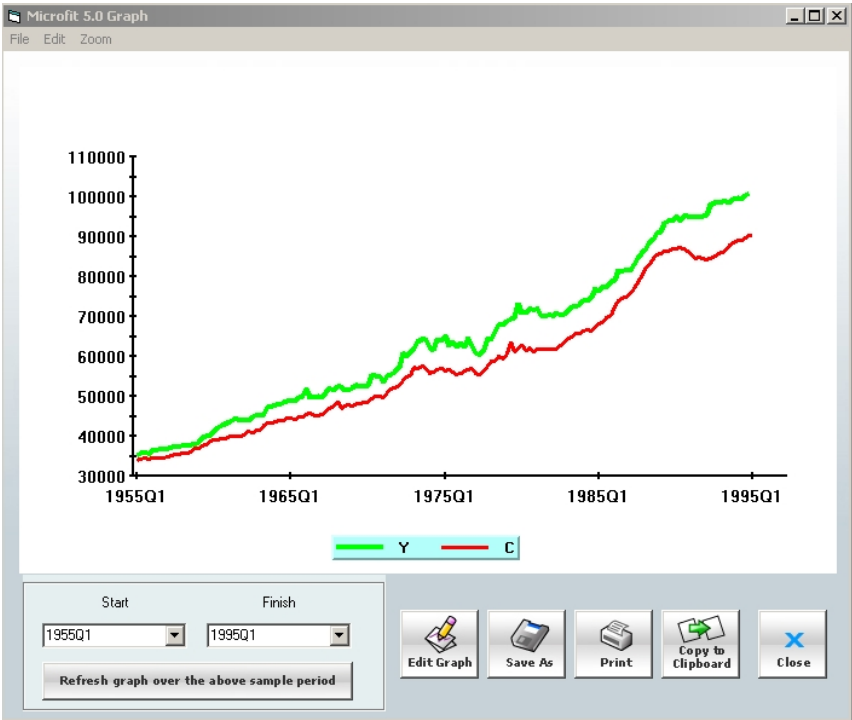


Figure 10.2: Real consumption expenditure and real disposable income in the UK

Format (OC2) by selecting the 'Save the chart (Olectra Chart Format') option from the File Menu. This format is useful if you want to load the graph into *Microfit* at a later stage for further editing. If you click the  button, you can also copy the displayed graph to clipboard for pasting into other programs (see Section 5.2 for further details).

Suppose now that you wish to obtain a scatter plot of the rate of change of real consumption (DLC) against the rate of change of real disposable income (DLY). Type

BATCH *UKCON*; **SCATTER** *DLC* *DLY* 

to see the scatter plot on your screen (see Figure 10.3). (Recall that lower- and upper-case letters have the same effect in *Microfit*.) Clearly there seems to be a high degree of association between the rate of change of consumption expenditure and the real disposable income.

10.6 Lesson 10.6: The use of command XPLOT in generating probability density function

The command **XPLOT** can be used for a variety of purposes, such as for plotting probability distributions, and Lorenz (or concentration) curves. For example, to generate a plot of the

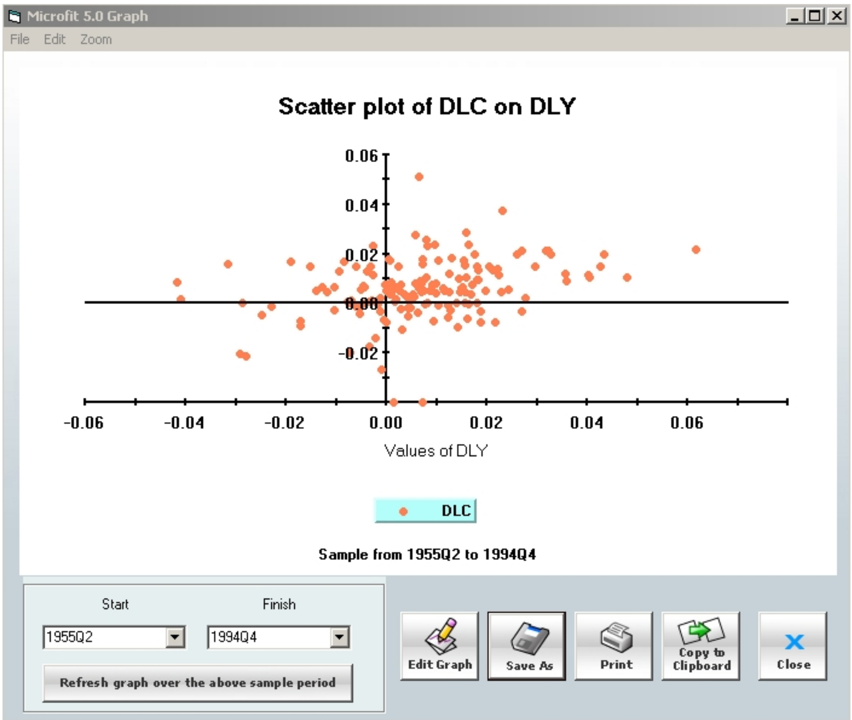


Figure 10.3: Scatter plot of *DLC* and *DLY*

standard normal distribution and the Cauchy distribution on the same graph, read the special *Microfit* file X.FIT. This file should be in the tutorial directory and contains the variable X, $\{x_t = (t - 100)/10, \quad t = 1, 2, ..., 200\}$ (see Lesson 9.1 on how to read in a special *Microfit* file). In the Commands and Data Transformations box type

$$MEU = 0; \qquad ZIG = 1;$$
BATCH *DENSITY* RUN

When the operations in the BATCH file are completed successfully, type

XPLOT *NORM CAUCHY X* RUN

You will see the plot of the standard Normal and Cauchy distributions on the screen (see Figure 10.4).

10.7 Lesson 10.7: Histogram of US stock market returns

The **HIST** command can be used to generate histograms and check the extent to which the empirical distribution function of a given variable deviates from the normal distribution. For instance, suppose you are interested in obtaining the histogram of the return on the US

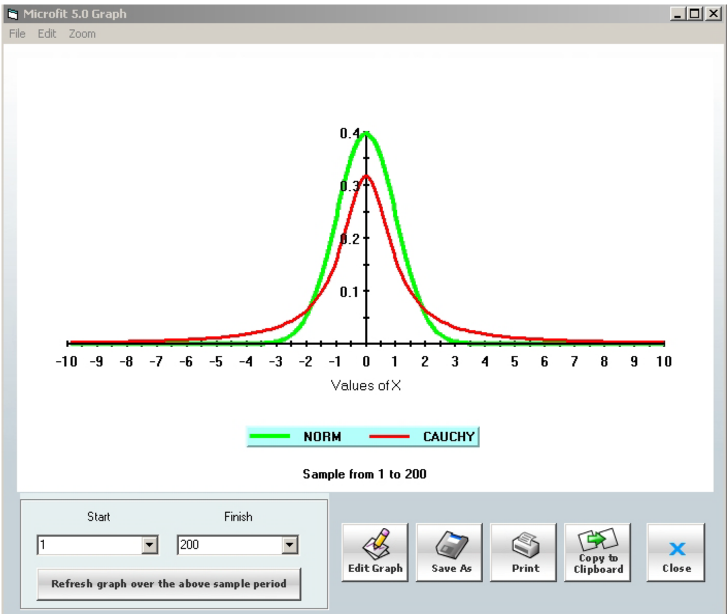


Figure 10.4: Plot of Normal and Chauchy distributions

Stock Market. The Special *Microfit* file USSTOCK.FIT contains 270 monthly observations over the period 1973(1)-1995(6) on the following variables

$US3TBR$	US Three Month Treasury Bill Rate (per cent, per annum)
$USCPI$	Consumer Price Index
$USDY$	Dividend Yield; ratio of dividends to share prices (per cent ,per annum)
$USLGR$	Yield of Long-term US Government Bond (per cent, per annum)
$USM1$	Money Supply M1
$USSIDX$	Share Prices Index-Standard and Poor 500 Composite (beginning of the month)

The monthly rate of return on the Standard and Poor 500 (SP500) share index is defined as the sum of the capital gains/losses $[(P_t - P_{t-1})/P_{t-1}]$ plus the dividend yield (D_{t-1}/P_{t-1}). Since in USSTOCK.FIT observations on the dividend yield variable ($USDY$) are measured in per cents and at annual rates, we first need to compute the dividends paid on SP500 per month. To carry out the necessary computations read the file USSTOCK.FIT into *Microfit*, and in the Commands and Data Transformations box type

$USDIV = (USDY * USSIDX)/1200;$
$USSR = (USSIDX - USSIDX(-1) \ + USDIV(-1))/USSIDX(-1);$
HIST $USSR$

You should see a histogram with 15 bands on the screen. If you wish to draw a histogram with a specific number of bands, say, you need to type

HIST $USSR(20)$

The result should be the same as in Figure 10.5. Compared with the normal distribution, which is given in the background of the histogram, the distribution of $USSR$ is a little skewed and has fat tails, that is, it displays excess kurtosis. There also seems to be an 'outlier', showing a 21.6 per cent decline in monthly returns, which refers to the October 1987 stock market crash.

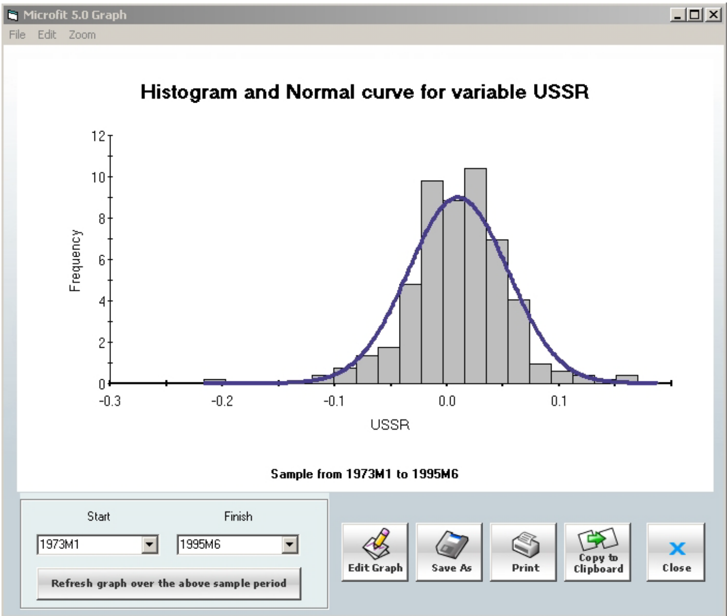


Figure 10.5: Histogram and normal curve for variable $USSR$ (sample from $1973M1$ to $1995M6$)

10.8 Lesson 10.8: Hodrick-Prescott filter applied to UK GDP

The HP filter is a two-sided filter routinely used as a method of detrending aggregate output in the real business cycle (RBC) literature. In this lesson we use the function **HPF**$(\cdot,\cdot)$ described in Section 4.3.7 to detrend the logarithm of the UK GDP.

The Special *Microfit* files GDP95.FIT on the tutorial directory contains the following variables:

$UKGDP$	GDP(A) at constant market prices (1990 prices £ million)
$USGNP$	Gross National Product (BILL.1987$) (T1.10) average

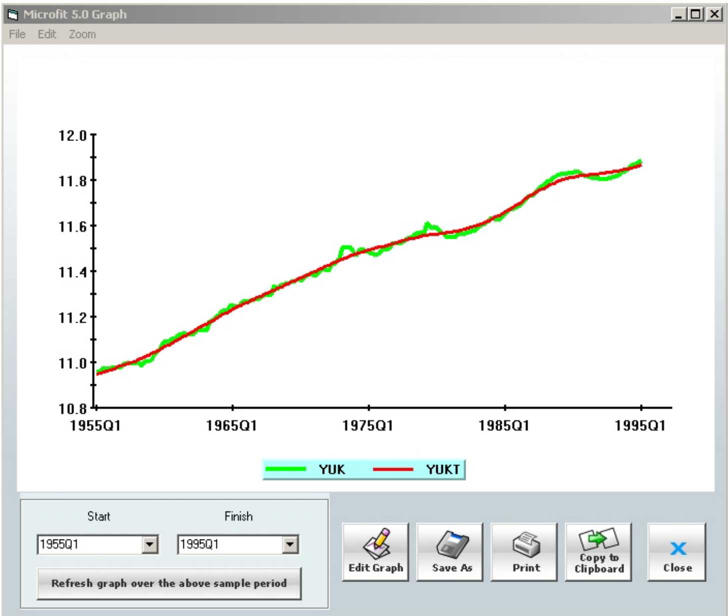


Figure 10.6: Plot of logarithm of the UK GDP and its trend estimated using the Hodrick-Prescott filter with $\lambda = 1600$

The sample period for the US and UK output series are 1960(1)-1995(1), and 1955(1)-1995(1), respectively. Read this file into *Microfit*, and in the Command and Data Transformations box type

$$YUK = \mathbf{LOG}(UKGDP);$$
$$YUKT = \mathbf{HPF}(YUK,\ 1600);$$
$$\mathbf{PLOT}\ YUK\ YUKT$$

RUN

You should now see the plot of the logarithm of UK GDP and its trend computed using the HP procedure with $\lambda = 1600$ on the screen (see Figure 10.6). The detrended series can now be computed as

$$YUKD = YUK - YUKT; \quad \mathbf{PLOT}\ YUKD$$

RUN

You should now see the Figure 10.7 on the screen. To check the sensitivity of the HP detrending procedure to the choice of λ, try other values of λ and plot the results. Notice that for the most part, the trend series are not very sensitive to the value of λ in the range $[600, 3600]$.

Repeat the above exercise with the $USGNP$. But remember to reset the sample to 1960(1)-1995(1), as US output series are not defined outside this period, and the application of the **HPF** function will result in missing values for the trend.

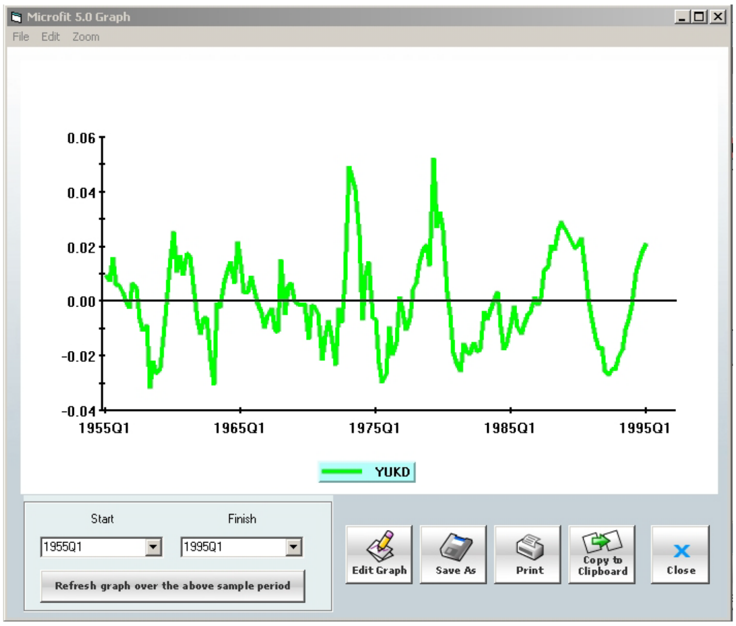


Figure 10.7: Plot of detrended UK output series using the Hodrick-Prescott filter with $\lambda = 1600$

10.9 Lesson 10.9: Summary statistics and correlation coefficients of US and UK output growths

As a part of your preliminary data analysis you may be interested to see the summary statistics and correlation matrix of some of the variables in the variable list. For example, suppose you have read in the file $GDP95.FIT$ and you wish to compute summary statistics and correlation coefficients for the variables, $USGR$ (US output growth), and $UKGR$ (UK output growth). Type in the Command and Variable Transformation box

$$\textbf{SAMPLE}\ 1960Q1\quad 1994Q4;$$
$$USGR = \textbf{RATE}(USGNP);\quad UKGR = \textbf{RATE}(UKGDP);$$
$$\textbf{COR}\quad USGR\quad UKGR$$

RUN

First you should see the summary statistics for the two variables $USGR$ and $UKGR$ on the screen. If you click Close the correlation matrix for these variables will be displayed (see Table 10.3). The result in Table 10.3 shows that the US economy has enjoyed a slightly higher growth than the UK economy over the 1960-1994 period. The US economy has grown around 2.9 per cent per annum as compared to an average annual rate of 2.3 per cent in the UK. Output growth has been relatively more variable in the UK. The coefficients of variations of output growth is 1.24 for the US as compared to 1.95 for the UK.

Finally the correlation coefficient between the two output growth series is 0.22, which is statistically significant at the 5 per cent level. In fact the Pesaran-Timmermann statistic for testing the association between the two growth rates, computed as **PTTEST**($USGR, UKGR$), is equal to 2.82, which is well above 1.96, the 5 per cent critical value of the standard normal distribution (see Section 4.3.17 for an account of the **PTTEST** function).

Table 10.3: Summary statistics for UK and US output growth and estimated correlation matrix of variables

```
Sample period     :1960Q2 to 1994Q4
Variable(s)       :     USGR       UKGR
Maximum           :    2.9980     4.8923
Minimum           :   -2.6335    -2.6096
Mean              :    .72763     .57156
Std. Deviation    :    .90349     1.1160
Skewness          :   -.56162     .54609
Kurtosis - 3      :    1.5262     2.3304
Coef of Variation:    1.2417     1.9526
***********************************
                    USGR       UKGR
 USGR              1.0000     .21979

 UKGR              .21979     1.0000

***********************************
```

10.10 Lesson 10.10: Autocorrelation coefficients of US output growth

Suppose you are interested in computing the autocorrelation coefficients of up to order 14 for the variable $USGR$ (the quarterly growth rate of US and GNP). Carry out the steps in Lesson 10.9, and when presented with the Command and Data Transformation box, type

SAMPLE $1960Q1 \quad 1994Q4$;
$DYUS = \mathbf{LOG}(USGNP/USGNP(-1))$;
COR $DYUS(14)$

The program first computes the logarithmic rate of change of the US real GNP, and then displays the summary statistics (mean, standard deviation, and so on) for the variable $DYUS$. If you now click , the autocorrelation coefficients, the Box-Pierce and Ljung-Box statistics will be displayed (see Table 10.4).

Clicking now yields a plot of the autocorrelation coefficients (see Figure 10.8). The default value for the maximum order of the computed autocorrelation coefficients is equal to $\frac{1}{3}$ of the sample size. For example, if you compute the autocorrelation coefficients over the period 1985(1)-1990(4) only the first 8 autocorrelation coefficients will be computed (see the **COR** command in Section 4.4.8).

The command **COR** applied to a variable, say X, also computes the Q statistic due to (Box and Pierce 1970) and its modified version, the Q^* statistic, due to (Ljung and Box 1978) for X (see Section 21.1). These statistics can be used to carry out general tests of serial correlation. The Ljung-Box Q^* statistic tends to be more reliable in small samples. The figures in square brackets refer to the probability of falsely rejecting the null hypothesis of no serial correlation. A small p-value provides evidence against the null hypothesis that the variable X is serially uncorrelated. In the case of the results in Table 10.4 there is clear evidence of serial correlation in US output growth. The first- and second-order autocorrelation coefficients 0.31864 and 0.23792 are large relative to their standard errors (the t-ratios for these autocorrelation coefficients are 3.76, and 2.56 which are above the critical value of the standard normal distribution at the level of 5 per cent). The remaining autocorrelation coefficients are not statistically significant.

Table 10.4: Summary statistics and autocorrelation coefficients for US output growth

Sample period	:1960Q2 to 1994Q4
Variable(s)	: DYUS
Maximum	: .029539
Minimum	: -.026688
Mean	: .0072099
Std. Deviation	: .0089935
Skewness	: -.60515
Kurtosis - 3	: 1.6251
Coef of Variation:	1.2474

Variable DYUS Sample from 1960Q2 to 1994Q4

Order	Autocorrelation Coefficient	Standard Error	Box-Pierce Statistic	Ljung-Box Statistic
1	.31864	.084819	14.1132[.000]	14.4200[.000]
2	.23792	.093033	21.9812[.000]	22.5178[.000]
3	.044503	.097312	22.2565[.000]	22.8032[.000]
4	.056147	.097458	22.6947[.000]	23.2609[.000]
5	-.057294	.097691	23.1510[.000]	23.7410[.000]
6	.028012	.097932	23.2601[.001]	23.8566[.001]
7	-.071426	.097990	23.9692[.001]	24.6141[.001]
8	-.14831	.098364	27.0268[.001]	27.9051[.000]
9	-.067555	.099960	27.6612[.001]	28.5931[.001]
10	.042191	.10029	27.9086[.002]	28.8636[.001]
11	-.0054286	.10042	27.9127[.003]	28.8681[.002]
12	-.17486	.10042	32.1626[.001]	33.5865[.001]
13	-.15910	.10258	35.6813[.001]	37.5241[.000]
14	-.11420	.10434	37.4941[.001]	39.5689[.000]

10.11 Lesson 10.11: Spectral density function of the US output growth

The **SPECTRUM** command (see Section 4.4.27) can be used to obtain different estimates of the standardized spectral density function. As an example, consider the problem of estimating the spectral density function for the rate of change of the US real GNP.

Use the option 'Open File' from the File Menu to read the GDP95.FIT file into *Microfit* and in the Commands and Data Transformation box create the variables (using the full sample)

$$DYUS = \mathbf{LOG}(USGNP/USGNP(-1))$$

RUN

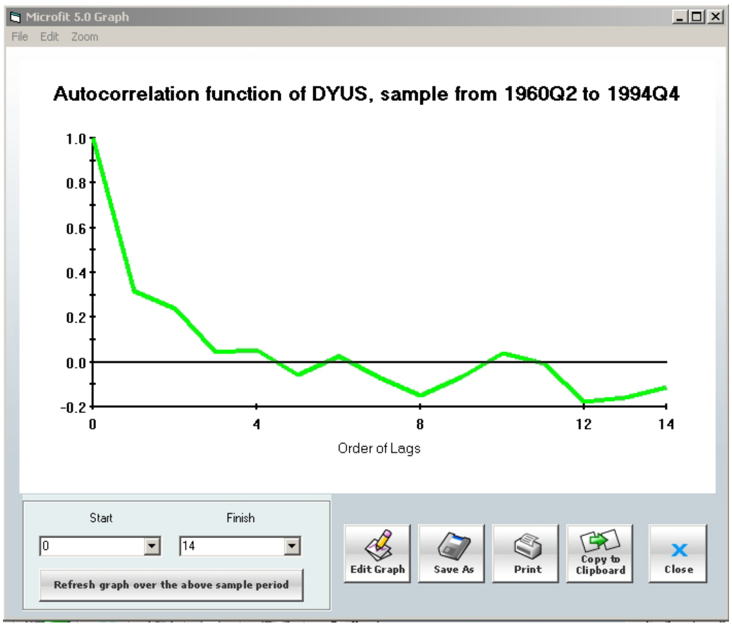


Figure 10.8: Autocorrelation function of $DYUS$ (sample from 1960Q1 to 1994Q4)

Then clear the editor type

SPECTRUM $DYUS$

You should see three different estimates of the standardized spectral density function of $DYUS$ on the screen. These estimates, and their asymptotic standard errors, are based on Bartlett, Tukey, and Parzen windows (see Section 21.3 for the details of the algorithms and the relevant references to the literature). The window size is set to the default value of $2\sqrt{n}$, where n is the number of observations. In the present application, $n = 139$, and the window size is equal to 24 (to override the default value for the window size see Section 4.4.27). The estimates of the spectral density are scaled and standardized using the unconditional variance of $DYUS$, and if evaluated at zero frequency provide a consistent estimate of Cochrane (1988) measure of persistence. Click to save these estimates in a result file, or to print. If you click you will be presented with four screens. The first three show the plots of the alternative estimates of the spectral density function (under Bartlett, Tukey and Parzen windows) and their standard error bands. For the purpose of comparing the different windows, the fourth screen displays all three estimates of the spectral density function on one graph (See Figure 10.9). Notice that the spectrum peaks at frequency 0.26, suggesting a cycle with periodicity equal to 24 quarters or 6 years.

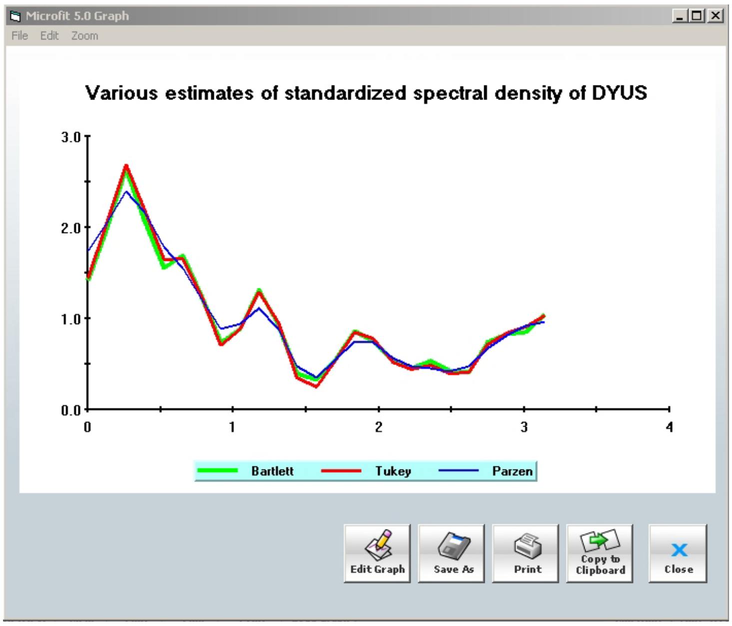


Figure 10.9: Various estimates of standardized spectral density of *DYUS* (sample from 1960Q1 to 1994Q4)

10.12 Lesson 10.12: Constructing a geometrically declining distributed lag variable: using the SIM command

Suppose you are interested in constructing a geometrically declining distributed lag function of the UK inflation rate, stored in the special *Microfit* file UKCON.FIT. Let Π_t be the inflation rate, and denote its geometric distributed lag function by Π_t^e. Then

$$\Pi_t^e = (1-\lambda)\sum_{i=0}^{\infty} \lambda^i \Pi_{t-i-1}, \qquad for \quad t = 1960Q1, ..., 1995Q1 \tag{10.1}$$

with $\lambda = 0.8$, and $\Pi_{1960Q1}^e = \Pi_{1960Q1}$. First notice that (10.1) can also be written recursively as

$$\Pi_t^e = \lambda \Pi_{t-1}^e + (1-\lambda)\Pi_{t-1}, \qquad \text{for} \quad t = 1960Q1, ..., 1995Q1$$

or

$$\Pi_t^e - \Pi_{t-1}^e = (1-\lambda)(\Pi_{t-1} - \Pi_{t-1}^e), \qquad \text{for} \quad t = 1960Q1, ..., 1995Q1$$

The last equation is immediately recognizable as the first-order adaptive expectations model.

To compute Π_t^e, for $t = 1960Q1, ..., 1995Q1$, load the special *Microfit* file, UKCON.FIT, and when presented with the Command and Data Transformation box, type

BATCH $UKCON$;
SAMPLE $1960Q1$ $1995Q1$; $PIE = PI$;
SAMPLE $1960Q2$ $1995Q1$;
SIM $PIE = 0.8 * PIE(-1) + 0.2 * PI(-1)$;
SAMPLE $1960Q1$ $1995Q1$ RUN

The variable PIE (Π_t^e) will now be created in your workspace, and you should see it added to the list of your existing variables. For a graphical presentation of the relationship between the inflation rate (PI), and the adaptively formed inflation expectations (PIE), type

PLOT PI PIE RUN

You should see the plot of PI and PIE against time on the screen (Figure 10.10). It can be clearly seen from this graph that the adaptive expectations tend to underestimate the actual rate of inflation when inflation is accelerating, and overestimate it when inflation is decelerating. A proof of this phenomenon can be found in Pesaran (1987a), pp. 18-19.

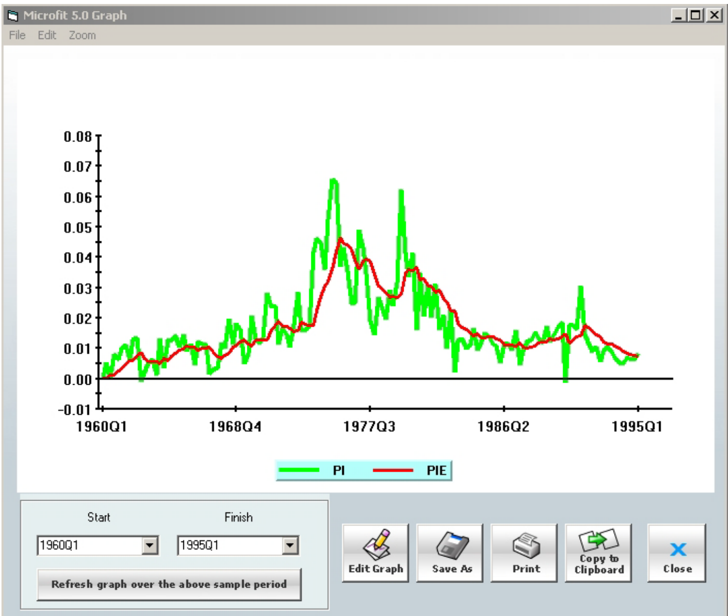


Figure 10.10: Actual and (adaptive) expected inflation in the UK (with adaptive coefficient=0.80)

10.13 Lesson 10.13: Computation of OLS estimators using formulae and commands

In this lesson we show how the function **SUM**, described in Section 4.3.29, can be used to compute *OLS* estimators of the coefficients of a simple regression equation from first principles. This type of application of *Microfit* is particularly useful for undergraduate courses in statistics and econometrics in which students need to be shown the details of the various steps involved in the computations.

Suppose you are interested in computing the *OLS* estimates of the regression of C (the real consumption expenditure) on Y (the real disposable income) using quarterly UK observations over the period 1960(1)-1994(4)

$$C_t = \alpha + \beta Y_t + u_t, \quad t = 1, 2, ..., n \tag{10.2}$$

where u_t is the error term. The *OLS* estimators of the coefficients α and β in (10.2) are given by

$$\begin{aligned} \hat{\beta} &= \sum_{t=1}^{n} \left(Y_t - \bar{Y}\right)\left(C_t - \bar{C}\right) \Big/ \sum_{t=1}^{n} \left(Y_t - \bar{Y}\right)^2 \\ \hat{\alpha} &= \bar{C} - \hat{\beta}\bar{Y} \end{aligned}$$

where $\bar{C}$ and $\bar{Y}$ are the arithmetic means of C and Y, respectively.

To carry out the necessary computations, read in the special *Microfit* file UKCON.FIT, and type the following instructions in the Commands and Data Transformations box

SAMPLE $1960Q1$ $1994Q4$;
$n =$ **SUM**(1);
$CBAR =$ **SUM**$(C)/n; CD = C - CBAR$;
$YBAR =$ **SUM**$(Y)/N; YD = Y - YBAR$;
$BHAT =$ **SUM**$(YD * CD)/$**SUM**$(YD\hat{}2)$;
$AHAT = CBAR - BHAT * YBAR$ RUN

The variables $AHAT$ and $BHAT$ will now contain the *OLS* estimates of α and β, respectively. You can list these estimates using the **LIST** command.

You can also use the **SUM** function to compute other statistics, such as the estimates of the standard error of the *OLS* estimates, the squared multiple correlation coefficient (R^2), the adjusted squared multiple correlation coefficient ($\overline{R}^2$), and the Durbin-Watson statistic. (The formulae for these statistics can be found in Section 21.6.1). The BATCH file OLS.BAT in the tutorial directory contains the necessary instructions for carrying out these computations. It is reproduced here in Table 10.5. To run this BATCH file first ensure that the variables C and Y are in your workspace (click Variables), then type

BATCH *OLS* RUN

If the operations are successful, you should see the following additional variables in the workspace

$AHAT$	OLS estimate of α
$BHAT$	OLS estimates of β
$SEAHAT$	estimate of the standard error of $ahat\,(\hat{\alpha})$
$SEBHAT$	estimate of the standard error of $bhat(\hat{\beta})$
$ZIGMA$	$\hat{\sigma}$, the standard error of regression
RSQ	R^2, the square of the multiple correlation coefficient
$RBARSQ$	$\bar{R}^2$, the adjusted R^2
DW	Durbin-Watson statistic
E	OLS residuals

Table 10.5: The BATCH file OLS.BAT

```
$ Content of the file OLS.BAT on the Tutorial Directory.
$ This is an example of a batch file for the direct
$ computation of the OLS regression, C = a + b Y + u,
$ estimated over the sub-period 1960(1)-1994(4), using the
$ special Microfit file, UKCON.FIT.  This file contains
$ quarterly observations on C (consumption) and Y (income),
$ over the period 1948(1)-1995(1).
$
$
sample 60q1 94q4
$ Setting the sample size (n)
n=sum(1)
$ Computing sample means of C and Y and storing the results in
$ cbar and ybar
cbar=sum(c)/n
ybar=sum(y)/n
$ Computing deviations of C and Y from their sample means
cd=c-cbar
yd=y-ybar
$ Computing varinaces of Y and C in yvar and cvar
yvar=sum(yd^2)/(n-1)
cvar=sum(cd^2)/(n-1)
$ Computing OLS estimates of the coefficients a and b
bhat=sum(cd*yd)/sum(yd^2)
ahat=cbar-bhat*ybar
$ Computing OLS residuals (e)
e=c-ahat-bhat*y
$ Computing R-Squared, R-BAR-Squared, standard error of
$ the regression, and the Durbin-Watson statistic
rsq=1-sum(e^2)/sum(cd^2)
zigsq=sum(e^2)/(n-2)
zigma=sqrt(zigsq)
rbarsq=1-(zigsq/cvar)
seahat=zigma*sqrt(sum(y^2))/sqrt(n*sum(yd^2))
sebhat=zigma/sqrt(sum(yd^2))
dummy=1
sample 60q2 94q4
dummy=0
sample 60q1 94q4
e1sq=sum(dummy*(e^2))
sample 60q2 94q4
dw=sum((e-e(-1))^2)/(sum(e^2)+e1sq)
$ Re-setting the sample back to its full range
sample 55q1 95q1
$ End of the BATCH file
```

You can now use either the **LIST** or the **COR** commands to list/print the various estimators/statistics computed by the BATCH file OLS.BAT. If you type

COR $C\ Y\ CBAR\ YBAR\ CVAR\ YVAR\ AHAT\ SEAHAT$
$BHAT\ SEBHAT\ ZIGMA\ RSQ\ RBARSQ\ DW\ N$

you should obtain the results in Table 10.6.

Table 10.6: OLS regression results using the BATCH file OLS.BAT

Sample period	:1960Q2 to 1994Q4					
Variable(s) :	C	Y	CBAR	YBAR	CVAR	YVAR
Maximum :	90399.0	100831.0	61143.5	67763.6	2.43E+08	3.05E+08
Minimum :	39059.0	42059.0	61143.5	67763.6	2.43E+08	3.05E+08
Mean :	61301.6	67954.1	61143.5	67763.6	2.43E+08	3.05E+08
Std. Deviation :	15546.1	17400.8	.1095E-9	.1606E-9	.2991E-6	.9571E-6
Skewness :	.46381	.40399	1.00000	-1.00000	1.00000	1.0000
Kurtosis - 3 :	-1.0111	-.99184	-2.0000	-2.0000	-2.0000	-2.0000
Coef of Variation:	.25360	.25607	.0000	.0000	.0000	.0000

Sample period	:1960Q2 to 1994Q4					
Variable(s) :	AHAT	SEAHAT	BHAT	SEBHAT	ZIGMA	RSQ
Maximum :	954.1058	513.2711	.88823	.0073359	1512.2	.99067
Minimum :	954.1058	513.2711	.88823	.0073359	1512.2	.99067
Mean :	954.1058	513.2711	.88823	.0073359	1512.2	.99067
Std. Deviation :	.0000	.0000	.0000	0.00	.0000	.0000
Skewness :	*NONE*	*NONE*	*NONE*	*NONE*	*NONE*	*NONE*
Kurtosis - 3 :	*NONE*	*NONE*	*NONE*	*NONE*	*NONE*	*NONE*
Coef of Variation:	.0000	.0000	.0000	0.00	.0000	.0000

Sample period	:1960Q2 to 1994Q4		
Variable(s) :	RBARSQ	DW	N
Maximum :	.99061	.44194	140.0000
Minimum :	.99061	.44194	140.0000
Mean :	.99061	.44194	140.0000
Std. Deviation :	.0000	.0000	0.00
Skewness :	*NONE*	*NONE*	*NONE*
Kurtosis - 3 :	*NONE*	*NONE*	*NONE*
Coef of Variation:	.0000	.0000	0.00

As we shall see in Chapter 11 (see Lesson 11.1), the same results (and more) can be readily computed by using the OLS option in the Linear Regression Menu.

10.14 Lesson 10.14: Construction of indices of effective exchange rates and foreign prices

In this lesson we provide an example of how a BATCH file can be used to compute the indices of the effective exchange rate (EER) and foreign prices (PF) for a given country (which we denote by 'j') with respect to its main trading partners.

Denote the effective exchange rate index of the jth country by E_{jt}. Then

$$E_{jt} = \sum_{i=1}^{N} w_{ji} \left(\frac{E_{jit} * 100}{E_{ji,85}} \right)$$

where w_{ji} is the share of country jth trade with the ith country, so that $\sum_{i=1}^{N} w_{ji} = 1$, and E_{jit} is the market rate of exchange of the jth currency in terms of the ith currency, computed as

$$E_{jit} = \left[\frac{j\text{th country national currency}}{\text{US dollar}}\right] \times \left[\frac{\text{US dollar}}{i\text{th country national currency}}\right]$$

$E_{ji,85}$ is the average value of the E_{jit} variable over the quarters in 1985,

$$E_{ji,85} = \tfrac{1}{4} \sum_{t=85q1}^{85q4} E_{jit}$$

Let PF_j be the jth country foreign price index, defined as the weighted average of the wholesale price indices of the main trading partners of the jth country,

$$PF_{jt} = \sum_{i=1}^{N} w_{ji} P_{it}$$

Where P_{it} is the wholesale price index (WPI) of the ith country.

The BATCH file G7EXCH.BAT contains the instructions for computing the variables EER and PF for the UK. But it can be readily modified to compute these variables for any other G7 country (see below). Table 10.7 reproduces this BATCH file. The data needed to run the BATCH file are stored in the file G7EXCH.FIT.

This file contains the variables E_i and P_i, with $i = 1, 2, ..., 10$, where

$E_1 =$ Japan market rate (Yen versus US$)
$E_2 =$ Germany market rate (DM versus US$)
$E_3 =$ France market rate (FF versus US$)
$E_4 =$ UK market rate (UK£ versus US$)
$E_5 =$ Italy market rate (Lira versus US$)
$E_6 =$ Canada market rate (Can$ versus US$)
$E_7 =$ The Netherlands market rate (NGuil versus US$)
$E_8 =$ Switzerland market rate (SF versus US$)
$E_9 =$ Belgium market rate (BF versus US$)
$E_{10} =$ Austria market rate (AS versus US$)

Table 10.7: Content of The BATCH file G7EXCH.BAT

```
$ Content of the file G7EXCH.BAT on the Tutorial Directory.
$ The following batch file is an example of how to compute the
$ effective exchange rate index (EER) and the foreign price
$ index (PF) needed to test the PPP hypothesis for a given
$ country j versus its main trading partners.
$
$ Setting the sample period
SAMPLE 72Q1 92Q3
$
$ Defining the Pound Sterling/Dollar rate as the domestic
$ currency j
EJ= E4
$ Specification of UK's major trading partner currencies
$
EJ0=E4
EJ1=EJ/E1
EJ2=EJ/E2
EJ3=EJ/E3
EJ4=EJ/E4
EJ5=EJ/E5
EJ6=EJ/E6
EJ7=EJ/E7
EJ8=EJ/E8
EJ9=EJ/E9
EJ10=EJ/E10
$
$ Constructing the dummy variable, D85, equal to zero except
$ for the four quarters in 1985 where it is set equal to one.
$
D85=0
SAMPLE 85Q1 85Q4
D85=1
SAMPLE 72Q1 92Q3
$
$ Computing the currency weights in the base year, 1985
D0= SUM(EJ0*D85)/4
D1= SUM(EJ1*D85)/4
D2= SUM(EJ2*D85)/4
D3= SUM(EJ3*D85)/4
D4= SUM(EJ4*D85)/4
D5= SUM(EJ5*D85)/4
D6= SUM(EJ6*D85)/4
D7= SUM(EJ7*D85)/4
D8= SUM(EJ8*D85)/4
D9= SUM(EJ9*D85)/4
D10=SUM(EJ10*D85)/4
$
$ Exchange rate indices with 1985=100
EJ0IND=(EJ0*100)/D0
EJ1IND=(EJ1*100)/D1
EJ2IND=(EJ2*100)/D2
EJ3IND=(EJ3*100)/D3
EJ4IND=(EJ4*100)/D4

EJ5IND=(EJ5*100)/D5
EJ6IND=(EJ6*100)/D6
EJ7IND=(EJ7*100)/D7
EJ8IND=(EJ8*100)/D8
EJ9IND=(EJ9*100)/D9
EJ10IND=(EJ10*100)/D10

$
$ Setting the values of the trading weights ( for the UK)
$
WJ0=0.2281
WJ1=0.0530
WJ2=0.2276
WJ3=0.1552
WJ4=0.0
WJ5=0.0814
WJ6=0.0349
WJ7=0.1422
WJ8=0.0
WJ9=0.0776
WJ10=0.0
$
$ Computing the EER index
$
EER=WJ0*EJ0IND+WJ1*EJ1IND+WJ2*EJ2IND+WJ3*EJ3IND+&
WJ4*EJ4IND+WJ5*EJ5IND+WJ6*EJ6IND+WJ7*EJ7IND+WJ8*EJ8IND+&
WJ9*EJ9IND+WJ10*EJ10IND
$
$ Computing the PF index
$
PF=WJ0*P0+WJ1*P1+WJ2*P2+WJ3*P3+WJ4*P4+WJ5*P5+WJ6*P6+WJ7*P7+&
WJ8*P8+WJ9*P9+WJ10*P10
$
$ Giving titles to the variables
$
ENTITLE EER PF
UK effective exchange rate index
UK foreign price index
$
$ Deleting unnecessary variables.
$
DELETE EJ EJ0 EJ1 EJ2 EJ3 EJ4 EJ5 EJ6 EJ7 EJ8 EJ9 EJ10 D85 &
D0 D1 D2 D3 D4 D5 D6 D7 D8 D9 D10 &
EJ0IND EJ1IND EJ2IND EJ3IND EJ4IND EJ5IND EJ6IND EJ7IND EJ8IND &
EJ9IND EJ10IND WJ0 WJ1 WJ2 WJ3 WJ4 WJ5 WJ6 WJ7 WJ8 WJ9 WJ10
$
$ End of batch file.
```

and

$P_0 =$ USA WPI
$P_1 =$ Japan WPI
$P_2 =$ Germany WPI
$P_3 =$ France CPI(1972-1979), WPI(1980-1992)
$P_4 =$ UK WPI
$P_5 =$ Italy WPI
$P_6 =$ Canada WPI
$P_7 =$ The Netherlands WPI
$P_8 =$ Switzerland WPI
$P_9 =$ Belgium WPI
$P_{10} =$ Austria WPI

The chosen domestic country, j, in the BATCH file G7EXCH.BAT is the UK, so that $j = 4$, $P_j = P_4, E_j = E_4$. The eight main trading partners are taken to be the USA, Japan, Germany, France, Italy, Canada, the Netherlands and Belgium. Note, however, that the BATCH file G7EXCH.BAT can be easily modified to compute the EER and PF indices for the other G7 countries. You simply need to edit it so that EJ is set to the currency of your choice, and the main trading weights in the construction of the indices are adjusted appropriately. The relevant weights for the G7 countries are given in Table 10.8.

Table 10.8: Trading weights of the G7 countries

	USA	Japan	Germany	France	UK	Italy	Canada
Main trading partners							
USA	-	0.7232	0.1513	0.1355	0.2281	0.1634	0.8546
Japan	0.2996	-	0.0493	0.0345	0.0530	0.0262	0.0613
Germany	0.0948	0.0778	-	0.2654	0.2276	0.3024	0.0201
France	0.0504	0.0256	0.1951	-	0.1552	0.2432	0.0111
UK	0.0841	0.0488	0.1422	0.1376	-	0.1084	0.0286
Italy	0.0473	0.0171	0.1362	0.1762	0.0814	-	0.0098
Canada	0.3601	0.0726	0.0156	0.0151	0.0349	0.0164	-
Netherlands	0.0366	0.0197	0.1803	0.0927	0.1422	0.0767	0.0820
Belgium	0.0261	0.0151	0.1301	0.1431	0.0776	0.0632	0.0061

To compute the two indices for the UK, load the file G7EXCH.FIT into *Microfit*. Check the definitions of the variables E_i, P_i, by clicking the 

button. Run the BATCH file G7EXCH.BAT by typing

BATCH *G7EXCH* 

in the Commands and Data Transformations box. If the operations are successful, you should see the following additional variables in the workspace

EER UK effective exchange rate index.

PF UK foreign price index.

To see a time-plot of these indices type

PLOT EER PF 

The screen should now appear as shown in Figure 10.11.

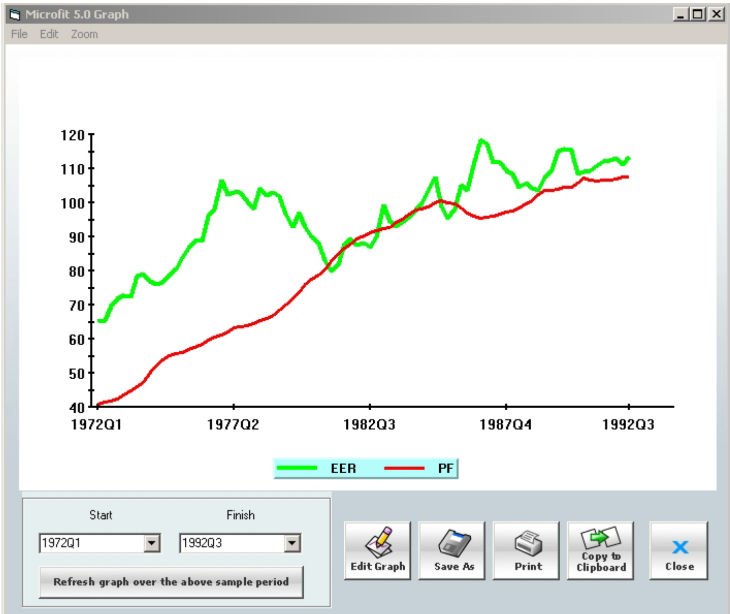


Figure 10.11: Effective and weighted foreign price indices for the UK (1985=100)

10.15 Lesson 10.15: Non-parametric density estimation of futures returns

In this lesson we demonstrate how to apply the **NONPARM** command (4.4.18) to estimate the density function of asset returns. In particular, consider daily data on returns of the equity futures index Nikkei (NK). The special *Microfit* file FUTURESDATA.FIT contains daily data on futures returns on a number of currencies, bonds and equity indices, and covers the period from 31-*Dec*-93 to 01-*Jan*-07.[1] To avoid lengthy computations in the present application we will only use data from 2005 to 2007 (a total of 522 observations).

Go to the Commands and Data Transformations box, and type

SAMPLE 31-*Dec*-04 01-*Jan*-07; **NONPARM** 1 2 3 4 NK 0

Microfit carries out the necessary computations and presents you an output window with the list of kernel density estimates for variable NK using Guassian and Epanechnikov kernels, and Silverman and least squares cross-validation as band widths. The list of observations on the variable NK, as well as the list of points at which the nonparametric functions are evaluated are also provided.

Since the vector of evaluation data points is not included in the command line, these are automatically supplied by the program. Also note that the use of the least squares cross-validation band width requires the evaluation of the kernel function at n^2 data points, and in applications where n is relatively large (for example, larger than 1000), this could take considerable amount of time (see Section 21.2 for further details).

If you now close the output window, you will be presented with the Kernel Density Estimation Menu, where you can display, plot or save your kernel density estimates and evaluation data points. Select option 2, then choose to inspect the plot of Gaussian kernel with least squares cross-validation band width against evaluation points. The graph is displayed in Figure 10.12.

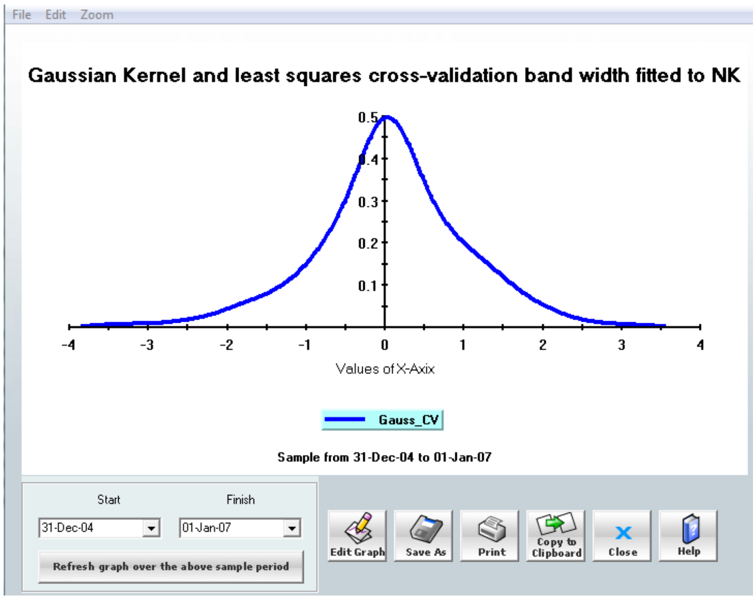


Figure 10.12: Gaussian kernel and least squares cross-valication band width for the variable NK

[1] See Section 20.1 for further details concerning this data set.

Now close the graph window and in the Kernel Density Estimation Menu click on option 3. Again, choose to inspect the plot of Gaussian kernel with least squares cross-validation band width. Results are displayed in Figure 10.13. Notice that the estimated density function for NK has a more acute peak and thicker tails than the normal distribution (i.e., there is evidence of excess kurtosis). As an exercise, use the command **COR** in the Commands and Data Transformations box to compute the kurtosis statistic and check that $b_2 - 3 > 0$ (see also Section 21.1).

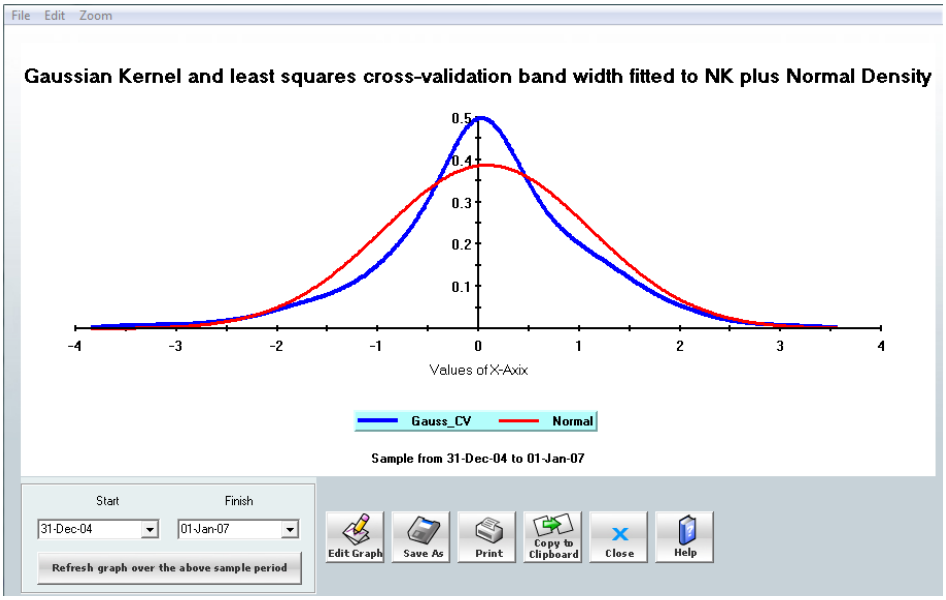


Figure 10.13: Gaussian kernel and least squares cross-valication band width for the variable NK, plus the normal density function

Finally, use option 4 from the Kernel Density Estimation Menu to plot the Gaussian kernel with least squares cross-validation band width for the variable NK, together with the Student's t-distribution. You will be asked to specify the degrees of freedom for the t-distribution. Type in, for example, 5, and click the button. The graph, displayed in Figure 10.14, indicates that the t-distribution captures the excess kurtosis of the distribution better than the normal distribution.

In this application the use of Epanechnikov kernel with least squares cross-validated band width produces a very uneven estimated density. Special care must be exercised in the choice of the kernel and the band width procedure.

Warning: In the case of large data sets avoid using the cross validation procedure. You can do this by using options 1 and 3 after the **NONPARM** command. See (4.4.18) for further detail.

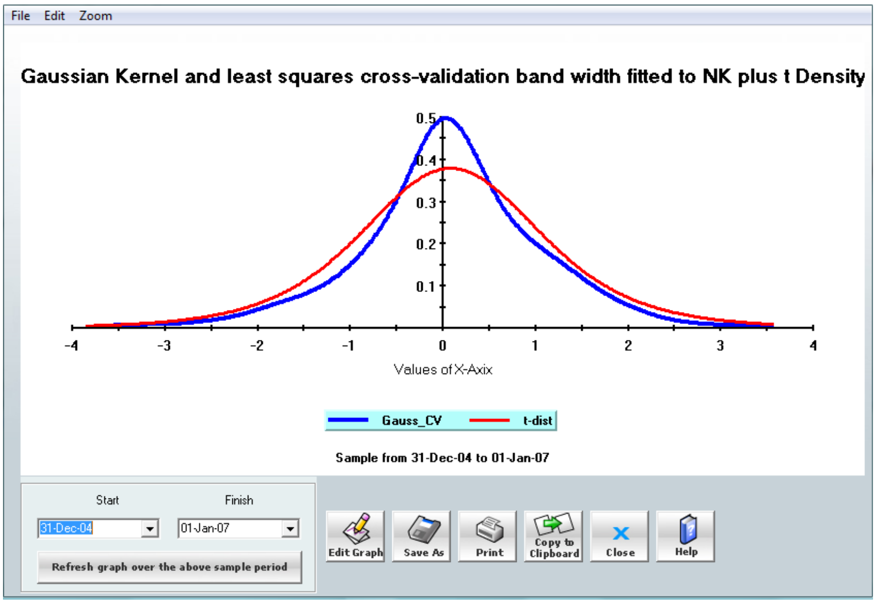


Figure 10.14: Gaussian kernel and least squares cross-valication band width for the variable NK, plus the t-density function

10.16 Lesson 10.16: Principal components analysis of US macroeconomic time series

In this lesson we perform a principal components analysis (PCA, 4.4.19 and 22.12) to summarize the empirical content of a large number of time series for the US economy. Data are available in the special *Microfit* file MSW90.FIT, from the Tutorial directory. This file contains a well known data set on monthly data on a large number of variables. The data set covers the period 1959 to 2002 (for a total of 528 months) on 90 variables describing different aspects of the US economy, such as income/output, employment, and construction/inventories. We refer to Marcellino, Stock, and Watson (2006) for further details. Due to the presence of some missing values, in this application we only consider data over the years 1967-2000. Load this file into *Microfit*, and use the  button to inspect the variables and their descriptions.[2] Once the data set is successfully entered, move to the Process window and use the  button to select from the Tutorial directory the EQU file PCATRANS.EQU. This file contains some instructions for creating an intercept, and for stardardizing the set of variables that will enter in the principal components analysis, so that they have unit (sample) variance. Clear the Commands and Data Transformations box, and retrieve from the Tutorial directory the LST or EQU files PCAPROD.LST or PCAPROD.EQU, which contains the list of 22 (standardized) variables on income/output

[2] Not all variables have a description in the data set. See Marcellino, Stock, and Watson (2006) for a detailed description of all the variables included.

to be included in the principal components analysis. Notice that the **PCA** command can be applied to a maximum of 102 variables. By clicking , *Microfit* carries out the necessary computations and presents you an output window with the list of non-zero eigenvalues, cumulative and percent cumulative eigenvalues, and eigenvectors associated with the selected set of variables. We refer to Section 22.12 for the relevant formula and for further information on principal components analysis. The output screen is partly reproduced in Table 10.9. Notice that, since variables have been standardized, the sum of non-zero eigenvalues associated to the **S** matrix is approximately equal to 22, the number of variables included in the principal components analysis. Further, we observe that there are only 2 eigenvalues larger than 1. This implies that there exist only two components, or factors, that explain at least as much as the equivalent of one original variable. Also, the eigenvalues indicate that these two factors account for about 97 per cent of the total variance, thus providing a reasonable summary of the data.

Table 10.9: Principal components analysis of US macroeconomic time-series

```
                       Principal Components Analysis
**************************************************************************
 Estimation period from 1967M1 to 2000M12, 408 observations.

 List of 22 variables included in the principal components analysis:
 IPS10ST         IPS11ST         IPS299ST        IPS12ST         IPS13ST
 IPS18ST         IPS25ST         IPIST           IPS32ST         IPS34ST
 IPS38ST         IPS43ST         IPDST           IPNST           IPMINST
 IPUTST          MSMQST          MSDQST          MSNQST          WTQST
 WTDQST          WTNQST

 The above variables have been filtered by the following variable:
 INPT

 List of 22 non-zero eigenvalues in descending order:
 20.1930    1.1789    .41763   .060608   .040962   .018841   .014624
 .0067359  .0053443  .0040499  .0022267  .0016353  .5906E-3  .3614E-3
 .2890E-3  .1138E-3  .4745E-4  .2044E-4  .1484E-4  .3920E-5  .2508E-5
 .1481E-5

 Cumulative Eigenvalues:
 20.1930   21.3720   21.7896   21.8502   21.8912   21.9100   21.9246
 21.9314   21.9367   21.9408   21.9430   21.9446   21.9452   21.9456
 21.9459   21.9460   21.9460   21.9461   21.9461   21.9461   21.9461
 21.9461

 Percent Cumulative Eigenvalues:
 92.0120   97.3840   99.2870   99.5632   99.7498   99.8357   99.9023
 99.9330   99.9574   99.9758   99.9860   99.9934   99.9961   99.9978
 99.9991   99.9996   99.9998   99.9999  100.0000  100.0000  100.0000
 100.0000
```

Now close the output window, and choose Option 2 to plot eigenvalues and cumulative eigenvalues from principal components analysis. These are reproduced in Figure 10.15 and 10.16, and display the eigenvalues and cumulative eigenvalues on the vertical axis and the principal component number on the horizontal axis.

Notice that these graphs show a sudden decrease in eigenvalues in the first two principal components, and relatively low contributions after the second principal component. This agrees with our preceding conclusion that two principal components provide a good summary of the data under consideration. We now save the first two principal components in a FIT file, and produce their plot against time. To this end, close the graph window, and in the

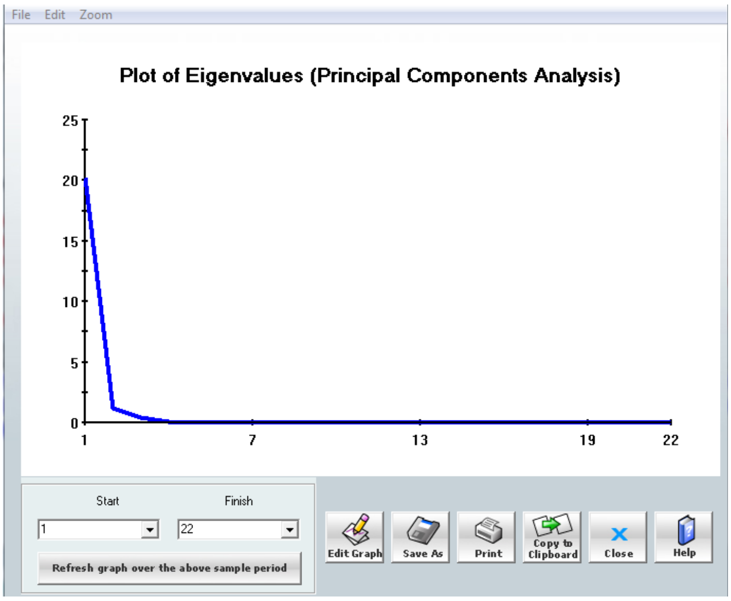


Figure 10.15: Plot of eigenvalues from principal components analysis

Post Estimation Menu (Principal Components Analysis) use options 3 and 4 to save the first two principal components in the file MSW90(2PC).FIT. Click  to return to the Process window and use the option 'Add a Special Microfit File to Workspace' from the file menu to add the file MSW90(2PC).FIT to the current data set. Then clear the Commands and Data Transformations box and type

SAMPLE $1967M1$ $2000M12$;
PLOT PC_1 PC_2 

The plot is reported in Figure 10.17. Notice that PC_1 summarizes a general trend of the variables included in the principal components analysis.

10.17 Lesson 10.17: Canonical correlation analysis of bond and equity futures

In this lesson we use canonical correlation analysis (CCA, 4.4.7 and 22.13) to explore the relationship between returns on bond and equity futures. Daily observations on futures returns on a number of currencies, bonds and equity indices over the period from 31-*Dec*-93 to 01-*Jan*-07 are stored in the special *Microfit* file FUTURESDATA.FIT.[3] In this lesson we use data over the period from 31-*Dec*-95 to 01-*Jan*-07 on four government bond futures: US ten year Treasury Note, ten year government bonds issued by Germany, UK and Japan,

[3] See Section 20.1 for further details regarding this data set.

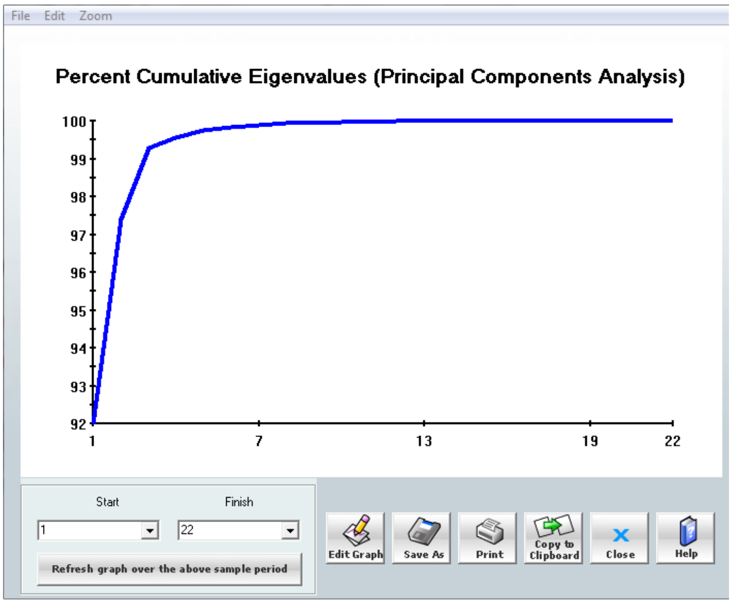


Figure 10.16: Plot of percent cumulative eigenvalues from principal components analysis

denoted by BU, BG, BE, and BJ, respectively; and five equity index futures in US, UK, Germany, France and Japan, namely S&P 500, FTSE, DAX, CAC and Nikkei, denoted by SP, $FTSE$, DAX, CAC, and, NK, respectively. Load the FUTURESDATA.FIT file into *Microfit*, go to the Commands and Data Transformations box and type

SAMPLE 31-*Dec*-95 01-*Jan*-07;
CCA BU BG BE BJ & SP $FTSE$ DAX CAC NK & C

where C denotes an intercept, included in the data set. *Microfit* starts the computation, and when finished, presents an output screen with the list of non-zero squared canonical correlations, the eigenvectors, the statistic for testing the independence between the two sets of variables and the canonical variates. Due to the length of the output, Table 10.10 only shows part of it. Notice that we have only four canonical variates, which is equal to the number of variables in the smaller of the two data sets. The chi-squared statistic (equal to 277.1414) is large and highly significant, indicating that there exists a significant degree of correlation between bond and equity futures returns. It is worth noting that the first canonical variate explains over 63 per cent of the canonical correlation between the two sets of variables.

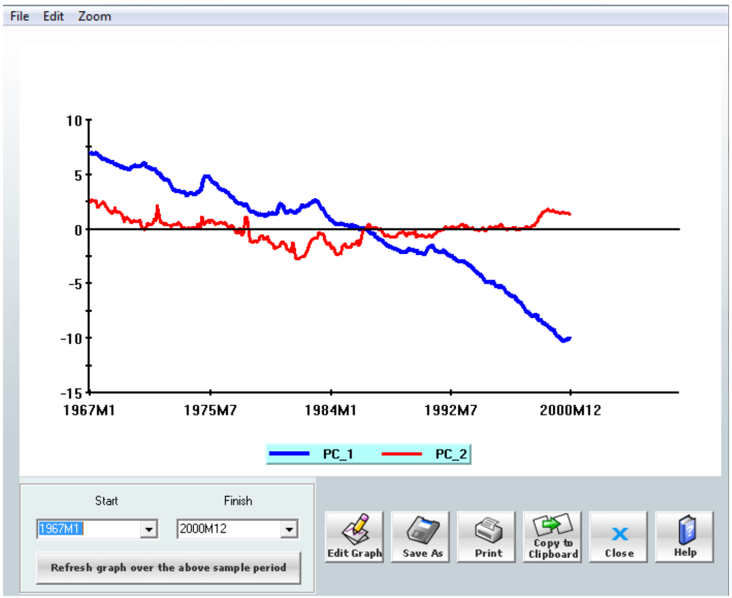


Figure 10.17: Plot of the first two principal components

Table 10.10: Canonical correlation analysis of bonds and equity futures

```
                       Canonical Correlation Analysis
******************************************************************************
 Estimation period from 29-Dec-95 to 01-Jan-07, 2872 observations.

 List of 4 Y-variables included in the canonical correlation analysis:
 BU              BG              BE              BJ
 List of 5 X-variables included in the canonical correlation analysis:
 SP              FTSE            DAX             CAC             NK

 The above variables have been filtered by the following variable:
 C

 List of 4 non-zero squared canonical correlations in descending order:
 .061168   .027111   .0065116  .0017067

 Cumulative Squared Canonical Correlations:
 .061168   .088279   .094791   .096498

 Test statistic for testing the independence of Y and X variables distributed
 as chi-squared with (4-1)*(5-1)=12 degrees of freedom =  277.1414[.000]

 Percent Cumulative Variances:
 63.3882   91.4835   98.2314  100.0000

 The number of chosen canonical variates is 4.
******************************************************************************
 List of Y-eigenvectors associated with non-zero canonical correlations:

               BU           BG           BE           BJ
 CCY_1       .023297      .014986     -.0048772     .053449
 CCY_2       .044427     -.016182      .011457     -.040457
 CCY_3       .025294     -.063072      .0046130     .022176
 CCY_4       .033206      .050800     -.10156     -.0048740

 List of X-eigenvectors associated with non-zero canonical correlations:

               SP           FTSE         DAX          CAC          NK
 CCX_1      -.0014205    -.0025342     .0086860    -.5923E-3     .0099787
 CCX_2       .5471E-3     .0050039     .013752     -.0090362    -.0082044
 CCX_3      -.0098473    -.0073288    -.0053886     .022001     -.0039208
 CCX_4      -.0065311     .026433     -.0056999    -.0088073     .4228E-3
******************************************************************************
```

10.18 Exercises in data processing

10.18.1 Exercise 10.1

Combine the two Special *Microfit* files UKSTOCK.FIT and USSTOCK.FIT and compute the rates of change of consumer prices (say $USPI$ and $UKPI$) in the two countries. Compare the histograms, estimated autocorrelation functions and spectrums of the two inflation rates. Comment on their differences and similarities.

10.18.2 Exercise 10.2

Load the file USCON.FIT into *Microfit* and retrieve the file USCON.EQU into the Command and Data Transformation box at the data processing stage. Process the content of the editor, and then plot the scatter of the rate of change of real non-durable consumption on the rate of change of real disposable income. Using the function **RATE(•)**, compute the average growth of UK real disposable income over the four sub-periods 1960(1)-1969(4), 1970(1)-1979(4), 1980(1)-1989(4), and 1990(1)-1994(4), and comment on your results. Repeat these calculations by computing the quarterly rate of change of real disposable income as first differences of the logarithm of the real disposable income. Are your conclusions affected by the method used to compute the average growth rates?

10.18.3 Exercise 10.3

Read in the file UKCON.FIT into *Microfit* and compute the Pesaran-Timmermann non-parametric statistic for testing the degree of association between the rates of change of consumption expenditure and real disposable income. Compare the results of this test with that based on the correlation coefficient between these variables.

10.18.4 Exercise 10.4

Use the special *Microfit* file G7EXCH.FIT and the associated BATCH file G7EXCH.BAT to construct the indices of effective exchange rates and foreign prices for Germany and France. The weights to be used in the construction of these indices are shown in Table 10.8.

Chapter 11

Lessons in Linear Regression Analysis

The lessons in this chapter are concerned with estimation, hypothesis testing, and prediction problems in the context of linear regression models. They use a variety of time-series and cross-sectional observations to show how the options in *Microfit* can be used to test for residual serial correlations, heteroscedasticity, non-normal errors, structural change, and prediction failure, how to carry out estimation of models with serially correlated errors, compute recursive and rolling regressions, test linear and non-linear restrictions on the regression coefficients, and detect when multicollinearity is likely to be a problem.

11.1 Lesson 11.1: OLS estimation of simple regression models

When you have finished your data transformations you can estimate, test, or forecast using a variety of estimation methods. You will need to specify your regression equation, the period over which you wish your regression to be estimated, and, in the case of linear regression, the number of observations you would like to set aside for predictive failure/structural stability tests.

In this lesson we shall consider two applications: first we estimate the simple regression equation (10.2) already estimated in Lesson 10.13 by running a BATCH file containing formulae and commands. Later we estimate a more complicated regression. Here we show how the computations can be carried out more simply using the *OLS* option. The relevant data are in the special *Microfit* file UKCON.FIT (see Lessons 10.1 and 10.2). Load this file (using the 'Open File' option in the the File Menu), and in the Commands and Data Transformations box create an intercept term by typing

$$INPT = 1$$

Click the Univariate Menu button on the main menu bar, choose the Linear Regression Menu and make sure option 1 Ordinary Least Squares is selected. Type the specifications of the regression equation in the Commands and Data Transformations box:

$$C \quad INPT \quad Y$$

Now enter the sample period

$$1960Q1 \quad 1994Q4$$

into the Start and End fields. Click RUN, and you will be presented with the OLS results reproduced in Table 11.1. Compare these estimates with those in Table 10.6.

Table 11.1: OLS estimates of a simple linear consumption function

```
                      Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is C
 140 observations used for estimation from 1960Q1 to 1994Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                     954.1058          513.2711              1.8589[.065]
 Y                          .88823           .0073359           121.0796[.000]
*******************************************************************************
 R-Squared                    .99067   R-Bar-Squared                     .99061
 S.E. of Regression           1512.2   F-Stat.    F(1,138)    14660.3[.000]
 Mean of Dependent Variable  61143.5   S.D. of Dependent Variable       15602.7
 Residual Sum of Squares    3.15E+08   Equation Log-likelihood          -1222.6
 Akaike Info. Criterion      -1224.6   Schwarz Bayesian Criterion       -1227.6
 DW-statistic                 .44194
*******************************************************************************


                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version       *          F Version          *
*******************************************************************************
*                     *                         *                             *
* A:Serial Correlation*CHSQ(4)  =  90.9952[.000]*F(4,134)      =  62.2049[.000]*
*                     *                         *                             *
* B:Functional Form   *CHSQ(1)  =   4.6340[.031]*F(1,137)      =   4.6899[.032]*
*                     *                         *                             *
* C:Normality         *CHSQ(2)  =   3.6022[.165]*       Not applicable        *
*                     *                         *                             *
* D:Heteroscedasticity*CHSQ(1)  =  25.5183[.000]*F(1,138)      =  30.7606[.000]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

Consider now the estimation of a slightly more complicated consumption function involving lagged values, namely the $ARDL(1,1)$ specification in logarithms[1]

$$\log c_t = \beta_1 + \beta_2 \log c_{t-1} + \beta_3 \log y_t + \beta_4 \log y_{t-1} + u_t \tag{11.1}$$

[1] In most applications the log-linear specification performs better than the linear specification. The coefficients of the log-linear specification, being elasticities, and hence scale-invariant, are also much easier to interpret. For a formal test of the linear versus the log-linear specification and vice versa, see Lesson 11.9.

For empirical analysis it is often more appropriate to consider an 'error correction' form of (11.1) given by

$$\Delta \log c_t = \alpha_1 + \alpha_2 \Delta \log y_t + \alpha_3 \log c_{t-1} + \alpha_4 \log y_{t-1} + u_t \tag{11.2}$$

where $\Delta \log c_t = \log c_t - \log c_{t-1}$, $\Delta \log y_t = \log y_t - \log y_{t-1}$, $\alpha_1 = \beta_1$, $\alpha_2 = \beta_3$, $\alpha_3 = -(1-\beta_2)$ and $\alpha_4 = \beta_4 + \beta_3$. To run the regression (11.2) first return to the Process window (click , choose in the next menu, then click) to generate the following variables:

$$LC = \mathbf{LOG}(C); \quad LY = \mathbf{LOG}(Y); \quad INPT = 1;$$
$$DLC = LC - LC(-1); \quad DLY = LY - LY(-1)$$

Alternatively, you can either retrieve the equation file UKCON.EQU into the Commands and Data Transformations box, or run the BATCH file UKCON.BAT. Once the above variables have been generated, choose Linear Regression Menu from the Univariate Menu and choose option 1 Ordinary Least Squares for the specification of the regression equation. Type the dependent variable, DLC, followed by the regressors

$$DLC \quad INPT \quad DLY \quad LC(-1) \quad LY(-1)$$

Choose the start and end dates 1955(1) and 1992(4) from the drop-down lists. Click . You can save the variables list for future use in a file using the button. Since the observations 1993(1)-1994(4) are not used up in the estimation, you will now be asked to specify the number of observations to be used in the predictive failure/structural stability tests. Type in 8 to choose all the eight remaining observations (notice that the observation 1995(1) for y_t is missing) and click . The results given in Table 11.2 should now appear on the screen. The diagnostic statistics that follow the estimation results suggest statistically significant evidence of residual serial correlation and non-normal errors.

Table 11.2: Error correction form of the $ARDL(1,1)$ model of consumption and income in the UK

```
                     Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is DLC
 151 observations used for estimation from 1955Q2 to 1992Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                      .044746            .040844              1.0955[.275]
 DLY                        .27680            .062296              4.4433[.000]
 LC(-1)                   -.072844            .042586             -1.7105[.089]
 LY(-1)                    .068540            .040599              1.6882[.093]
*******************************************************************************
 R-Squared                     .11929   R-Bar-Squared                   .10131
 S.E. of Regression           .011589   F-Stat.    F(3,147)     6.6366[.000]
 Mean of Dependent Variable  .0061324   S.D. of Dependent Variable     .012225
 Residual Sum of Squares      .019743   Equation Log-likelihood       460.8790
 Akaike Info. Criterion      456.8790   Schwarz Bayesian Criterion    450.8445
 DW-statistic                  2.3424
*******************************************************************************


                              Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version        *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(4)  =  13.5294[.009]*F(4,143)     =   3.5184[.009]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =  .017570[.895]*F(1,146)     =  .016990[.896]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =  62.5209[.000]*       Not applicable        *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =   1.1656[.280]*F(1,149)     =   1.1591[.283]*
*                     *                          *                            *
* E:Predictive Failure*CHSQ(8)  =   .84989[1.00]*F(8,147)     =   .10624[.999]*
*                     *                          *                            *
* F:Chow Test         *CHSQ(4)  =   .27225[.992]*F(4,151)     =  .068063[.991]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
   E:A test of adequacy of predictions (Chow's second test)
   F:Test of stability of the regression coefficients
```

To leave the OLS result screen click Close. You will now be presented with the Post Regression Menu (see Section 6.20), giving a number of options to further analyse your regression results. For example, suppose you wish to test the hypothesis that in (11.2) $\alpha_3 = \alpha_4 = 0$. Choose option 2 in this menu and then option 5 in the Hypothesis Testing Menu (see Section 6.23) that follows, and after clearing the content of the box editor if necessary (by clicking the Clear button), type

$$LC(-1) \quad LY(-1)$$ RUN

The results in Table 11.3 should now appear on the screen.

Table 11.3: Statistical significance of the level variables in the $ARDL(1,1)$ model of income and consumption in the UK

```
                      Variable Deletion Test (OLS case)
*******************************************************************************
 Dependent variable is DLC
 List of the variables deleted from the regression:
 LC(-1)          LY(-1)
 151 observations used for estimation from 1955Q2 to 1992Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                     .0045088           .0010256            4.3962[.000]
 DLY                       .23859            .058154             4.1026[.000]
*******************************************************************************
 Joint test of zero restrictions on the coefficients of deleted variables:
 Lagrange Multiplier Statistic     CHSQ(2)=   2.9891[.224]
 Likelihood Ratio Statistic        CHSQ(2)=   3.0191[.221]
 F Statistic                       F(2,147)=   1.4844[.230]
*******************************************************************************
```

The various statistics for testing the joint restrictions $\alpha_3 = \alpha_4 = 0$ are given at the lower end of Table 11.3. For example, the likelihood ratio (LR) statistic is 3.0191. Notice that the critical value of this test depends on whether or not $\log y_t$ is integrated. See Pesaran, Shin, and Smith (2000), and Lesson 16.5 for further details. However, in the present application, the value of the LR statistic is small enough for us to safely conclude that the hypothesis that $\alpha_3 = \alpha_4 = 0$ cannot be rejected. Therefore, the $ARDL(1,1)$ specification in (11.1) does not provide a stable long relationship between real disposable income and consumption in the UK.

To see a plot of the actual and fitted values choose option 3 in the Post Regression Menu, and when presented with the Display/Save Residuals and Fitted Values Menu (see Section 6.21) click OK. You should see Figure 11.1 on the screen.

You can save this figure in a variety of formats by using Save As (see Section 5.2 for further details). Figure 11.1 clearly shows that none of the sharp falls in the consumption expenditure are explained by the simple $ARDL(1,1)$ model in (11.1).

You can also compute static forecasts of $\Delta \log c_t$ over the period 1993(1)-1994(4). Click Close to leave Figure 11.1, then click OK, and choose option 8 in the Post Regression Menu. You will be asked to select the forecast interval, by entering the initial and the final forecast period. Click OK to obtain forecasts of $\Delta \log c_t$ together with a number of summary statistics. If you then press the Close button, you will be presented with a plot of actual and forecast values of $\Delta \log c_t$. You can also obtain this graph by choosing option 9 in the Post Regression Menu (see Figure 11.2).

Note that the forecasts generated in the present application are 'static' in the sense that for every quarter in the period 1993(1)-1994(4), actual values of $\log c_{t-1}$ are used in forecasting $\log c_t$. (See Section 21.26.1 for further details).

Note that in the above example, although the estimation period is specified as 1955(1)-1992(4), because of the missing initial values for the lagged variables $\log y_{t-1}$ and $\log c_{t-1}$, the program automatically adjusts the sample period to take account of these missing observations and selects 1955(2)-1992(4) as the estimation period.

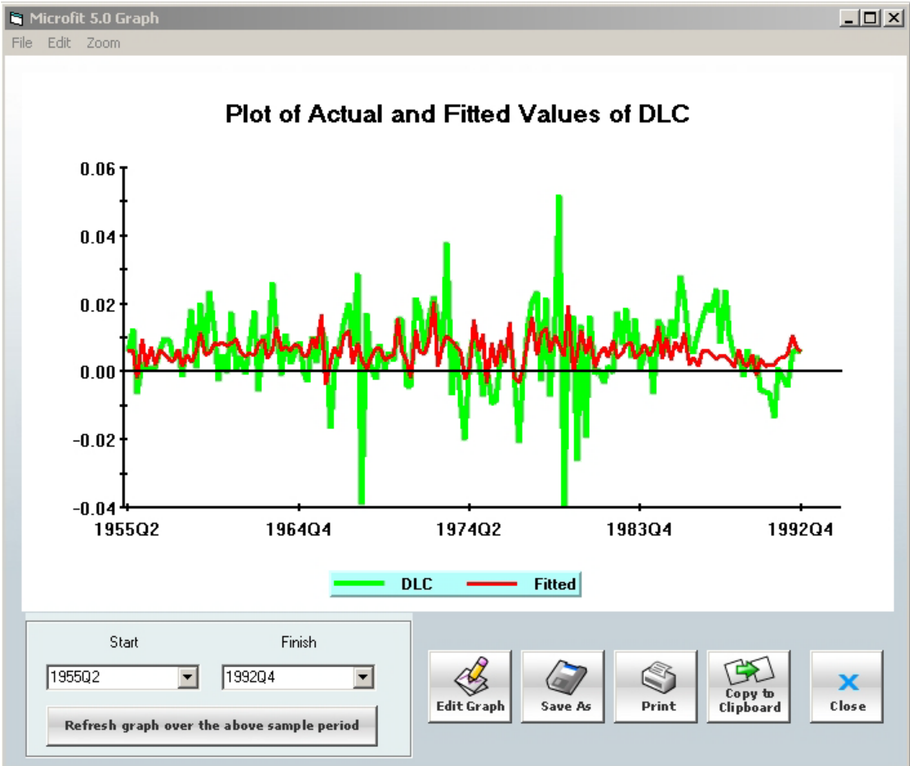


Figure 11.1: Plot of actual and fitted values of $\Delta \log c_t$

11.2 Lesson 11.2: Two alternative methods of testing linear restrictions

This lesson describes two different methods of testing the hypothesis of constant returns to scale in the context of a Cobb-Douglas (CD) production function.

Consider the CD production function

$$Y_t = AK_t^{\alpha} L_t^{\beta} e^{u_t}, \quad t = 1, 2, ..., n \tag{11.3}$$

where Y_t = Output, K_t = Capital Stock, L_t = Employment.

The unknown parameters A, α and β are fixed, and u_ts are serially uncorrelated disturbances with zero means and a constant variance. We also assume that u_ts are distributed independently of K_t and L_t. Notice that for simplicity of exposition we have not allowed for technical progress in (11.3). The constant returns to scale hypothesis postulates that proportionate changes in inputs (K_t and L_t) result in the same proportionate change in output. For example, doubling K_t and L_t should, under the constant returns to scale hypothesis, lead also to the doubling of Y_t. This imposes the following parametric restriction on (11.3):

$$H_0 : \quad \alpha + \beta = 1$$

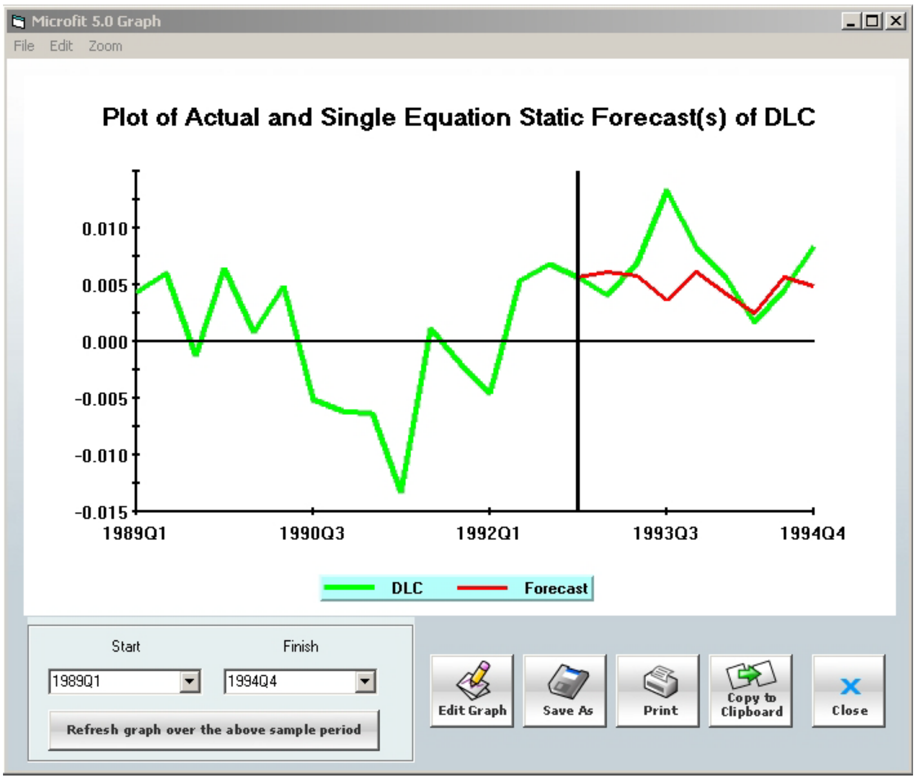


Figure 11.2: Plot of actual and single-equation static forecast(s)

which we consider as the null hypothesis and derive an appropriate test of it against the two-sided alternative:

$$H_1: \quad \alpha + \beta \neq 1$$

In order to implement the test of H_0 against H_1 we first take logarithms of both sides of (11.3), which yields the log-linear specification

$$LY_t = a + \alpha LK_t + \beta LL_t + u_t \tag{11.4}$$

where

$$LY_t = \log(Y_t), \quad LK_t = \log(K_t), \quad LL_t = \log(L_t)$$

and $a = \log(A)$.

It is now possible to obtain estimates of α and β by running OLS regressions of LY_t on LK_t and LL_t (for $t = 1, 2, ..., n$), including an intercept in the regression. Denote the OLS estimates of α and β by $\hat{\alpha}$ and $\widehat{\beta}$, and define a new parameter, δ, as

$$\delta = \alpha + \beta - 1 \tag{11.5}$$

The hypothesis $\alpha + \beta = 1$ against $\alpha + \beta \neq 1$ can now be written equivalently as

$$\begin{array}{ll} H_0: & \delta = 0 \\ H_1: & \delta \neq 0 \end{array}$$

We now consider two alternative methods of testing $\delta = 0$: a *direct method* and a *regression method*.

Direct method of testing $\delta = 0$

This method directly focuses on the OLS estimates of δ, namely $\widehat{\delta} = \hat{\alpha} + \widehat{\beta} - 1$, and examines whether this estimate is significantly different from zero. For this we need an estimate of the variance of $\widehat{\delta}$. We have

$$V(\widehat{\delta}) = V(\hat{\alpha}) + V(\widehat{\beta}) + 2\,\mathrm{Cov}\left(\hat{\alpha}, \widehat{\beta}\right)$$

where $V(\cdot)$ and $\mathrm{Cov}(\cdot)$ stand for the variance and the covariance operators, respectively. The OLS estimator of $V(\widehat{\delta})$ is given by

$$\widehat{V}(\widehat{\delta}) = \widehat{V}(\hat{\alpha}) + \widehat{V}(\widehat{\beta}) + 2\widehat{\mathrm{Cov}}(\hat{\alpha}, \widehat{\beta})$$

where $\hat{}$ denotes the estimate. The relevant test-statistic for testing $\delta = 0$ against $\delta \neq 0$ is now given by

$$t_{\widehat{\delta}} = \frac{\widehat{\delta}}{\sqrt{\widehat{V}(\widehat{\delta})}} = \frac{\hat{\alpha} + \widehat{\beta} - 1}{\sqrt{\widehat{V}(\hat{\alpha}) + \widehat{V}(\widehat{\beta}) + 2\widehat{\mathrm{Cov}}(\hat{\alpha}, \widehat{\beta})}} \tag{11.6}$$

and under $\delta = 0$, has a t-distribution with $n - 3$ degrees of freedom.

The regression method

This method starts with (11.4) and replaces β (or α) in terms of δ and α (or β). Using (11.5) we have

$$\beta = \delta - \alpha + 1$$

Substituting this in (11.4) for β now yields

$$LY_t - LL_t = a + \alpha(LK_t - LL_t) + \delta LL_t + u_t \tag{11.7}$$

or

$$Z_t = a + \alpha W_t + \delta LL_t + u_t \tag{11.8}$$

where $Z_t = \log(Y_t/L_t) = LY_t - LL_t$ and $W_t = \log(K_t/L_t) = LK_t - LL_t$. A test of $\delta = 0$ can now be carried out by first regressing Z_t on W_t and LL_t (including an intercept term), and then carrying out the usual t-test on the coefficient of LL_t in (11.8). The t-ratio of δ in (11.7) will be identical to $t_{\widehat{\delta}}$ defined by (11.6).

We now apply the two methods discussed above to the historical data on Y, K, and L used originally by Cobb and Douglas (1928). The relevant data are stored in the special

Microfit file CD.FIT, and covers the period 1899-1922. Read this file (using the Open File option in the File Menu), and in the Commands and Data Transformations box type

$$\begin{aligned} LY &= \mathbf{LOG}(Y); \quad LL = \mathbf{LOG}(L); \quad INPT = 1; \\ Z &= \mathbf{LOG}(Y/L); \quad W = \mathbf{LOG}(K/L) \end{aligned}$$

to generate the variables LY, LL, Z and W defined above. Then move to the Univariate Menu on the main menu bar and choose the Ordinary Least Squares option from the Linear Regression Menu. Type

$$LY \quad INPT \quad LK \quad LL$$

and click 

. You should see the OLS estimates on the screen (see Table 11.4).

Table 11.4: Estimates of the log-linear Cobb-Douglas production function

```
                        Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is LY
 24 observations used for estimation from 1899 to 1922
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                      -.17731              .43429           -.40827[.687]
 LK                         .23305             .063530            3.6684[.001]
 LL                         .80728              .14508            5.5645[.000]
*******************************************************************************
 R-Squared                     .95742   R-Bar-Squared                   .95337
 S.E. of Regression           .058138   F-Stat.    F(2,21)    236.1219[.000]
 Mean of Dependent Variable    5.0773   S.D. of Dependent Variable      .26923
 Residual Sum of Squares      .070982   Equation Log-likelihood        35.8261
 Akaike Info. Criterion       32.8261   Schwarz Bayesian Criterion     31.0590
 DW-statistic                  1.5235
*******************************************************************************


                                Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version          *
*******************************************************************************
*                     *                          *                              *
* A:Serial Correlation*CHSQ(1)  =    .35950[.549]*F(1,20)      =    .30414[.587]*
*                     *                          *                              *
* B:Functional Form   *CHSQ(1)  =    2.1448[.143]*F(1,20)      =    1.9627[.177]*
*                     *                          *                              *
* C:Normality         *CHSQ(2)  =    1.3613[.506]*       Not applicable         *
*                     *                          *                              *
* D:Heteroscedasticity*CHSQ(1)  =    2.5774[.108]*F(1,22)      =    2.6469[.118]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

Click 

to move to the Post Regression Menu and choose option 4 and then option 1 in the Standard, White and Newey-West Adjusted Variance Menu. The following estimates of the variance covariance matrix of $\left(\hat{\alpha}, \widehat{\beta}\right)'$ should appear on the screen:

$$\begin{bmatrix} \widehat{V}(\hat{\alpha}) & \widehat{\text{Cov}}\left(\hat{\alpha}, \widehat{\beta}\right) \\ \widehat{\text{Cov}}\left(\hat{\alpha}, \widehat{\beta}\right) & \widehat{V}(\widehat{\beta}) \end{bmatrix} = \begin{bmatrix} 0.004036 & -0.0083831 \\ -0.0083831 & 0.021047 \end{bmatrix}$$

Using the above *OLS* estimates of α and β given in Table 11.4 (namely $\hat{\alpha} = 0.23305$ and $\widehat{\beta} = 0.80728$) and the above results in (11.6) gives

$$t_{\widehat{\delta}} = \frac{0.23305 + 0.80728 - 1}{\sqrt{0.004036 + 0.021047 - 2(0.0083831)}} = 0.442 \tag{11.9}$$

Comparing $t_{\widehat{\delta}} = 0.442$ and the 5 per cent critical value of the t-distribution with $T-3 = 24-3 = 21$ degrees of freedom (which is equal to 2.080), it is clear that since $t_{\widehat{\delta}} = 0.442 < 2.080$, then the hypothesis $\delta = 0$ or $\alpha + \beta = 1$ cannot be rejected at the 5 per cent level.

To implement the regression approach you need to return to the Commands and Data Transformations box to edit the regression equation. Click 

to clear the box editor and then type

$$Z \quad INPT \quad W \quad LL$$ 

You should see the results given in Table 11.5 on your screen.

Table 11.5: Log-linear estimates of the Cobb-Douglas production function in per capita terms

```
                        Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is Z
 24 observations used for estimation from 1899 to 1922
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                      -.17731             .43429            -.40827[.687]
 W                          .23305            .063530             3.6684[.001]
 LL                        .040332            .091197             .44225[.663]
*******************************************************************************
 R-Squared                     .63674   R-Bar-Squared                   .60215
 S.E. of Regression           .058138   F-Stat.    F(2,21)      18.4052[.000]
 Mean of Dependent Variable    .11461   S.D. of Dependent Variable     .092173
 Residual Sum of Squares      .070982   Equation Log-likelihood        35.8261
 Akaike Info. Criterion       32.8261   Schwarz Bayesian Criterion     31.0590
 DW-statistic                  1.5235
*******************************************************************************


                                 Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version       *          F Version          *
*******************************************************************************
*                     *                         *                             *
* A:Serial Correlation*CHSQ(1)  =    .35950[.549]*F(1,20)      =    .30414[.587]*
*                     *                         *                             *
* B:Functional Form   *CHSQ(1)  = .2608E-5[.999]*F(1,20)      = .2174E-5[.999]*
*                     *                         *                             *
* C:Normality         *CHSQ(2)  =    1.3613[.506]*       Not applicable        *
*                     *                         *                             *
* D:Heteroscedasticity*CHSQ(1)  =    8.8809[.003]*F(1,22)      =   12.9227[.002]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

The t-ratio of the coefficient of the LL variable in this regression is equal to 0.442 which is identical to $t_{\hat{\delta}}$ as computed in (11.9).

It is worth noting that the above estimates of α and β, which have played a historically important role in the literature, are very 'fragile', in the sense that they are highly sensitive to the sample period chosen in estimating them. For example, estimating the model (given in (11.4)) over the period 1899-1920 (dropping the observations for the last two years) yields $\hat{\alpha} = 0.0807(0.1099)$ and $\widehat{\beta} = 1.0935(0.2241)$! The figures in brackets are standard errors.

11.3 Lesson 11.3: Estimation of long-run effects and mean lags

In this lesson we show how option 5 in the Post Regression Menu (see Section 6.20) can be used to estimate long-run effects, mean lags, and other functions of the underlying parameters of a regression model, together with their standard errors.

As an example, consider the following $ARDL(1,1)$ model relating capital expenditure in the US manufacturing sector (Y_t) to capital appropriations (X_t)

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + u_t \tag{11.10}$$

Assuming that $|\beta_1| < 1$, we have

$$Y_t = \frac{\beta_0}{1-\beta_1 L} + \left(\frac{\beta_2 + \beta_3 L}{1-\beta_1 L}\right) X_t + \left(\frac{1}{1-\beta_1 L}\right) u_t$$

or

$$Y_t = a_0 + \theta(L) X_t + (1-\beta_1 L)^{-1} u_t$$

where L is the lag-operator such that $LY_t = Y_{t-1}$, and $\theta(L)$ is the distributed lag function operating on X_t. The long-run response of Y_t to a unit change in X_t is given by

$$LR = \theta(1) = \frac{\beta_2 + \beta_3}{1-\beta_1} \tag{11.11}$$

The mean lag of response of Y_t to a unit change in X_t is defined by

$$ML = \frac{1}{\theta(1)} \sum_{i=1}^{\infty} i\theta_i = \theta'(1)/\theta(1)$$

where $\theta'(1)$ denotes the first derivative of $\theta(L)$ with respect to L, evaluated at $L = 1$. It is now easily seen that[2]

$$ML = \frac{\theta'(1)}{\theta(1)} = \frac{\beta_1\beta_2 + \beta_3}{(1-\beta_1)(\beta_2+\beta_3)} \tag{11.12}$$

Suppose now you wish to compute the estimates of LR and ML and their standard errors using observations in the special *Microfit* file ALMON.FIT. This file contains quarterly observations on Y_t and X_t over the period 1953(1)-1967(4), which is an extended version of the data originally analysed by Almon (1965).

Choose option 1 in the Single Equation Estimation Menu (see Section 6.4) and type

$$Y \;\; INPT \; Y(-1) \;\; X \;\; X(-1)$$

You should now see the OLS results on the screen. Click 

to move to the Post Regression Menu and choose option 5 in this menu. You will be presented with a box editor. Type the two functional relations (11.11) and (11.12) in the following manner:

$$LR = (A3 + A4)/(1 - A2);$$
$$ML = (A2 * A3 + A4)/((1 - A2) * (A3 + A4))$$

---

[2]For more details see Dhrymes (1971) or Greene (2002), Chapter 19. Note that the concept of mean lag is meaningful if all the lag coefficients, θ_i, have the same signs.

Notice that *Microfit* assigns the coefficients $A1$, $A2$, $A3$, and $A4$ to the regressors $INPT$, $Y(-1), X$, and $X(-1)$, respectively.

The results in Table 11.6 should now appear on the screen.

Table 11.6: Estimates of the long-run coefficient and mean lag for the relationship between capital expenditures and capital appropriations in US manufacturing

```
                    Analysis of Function(s) of Parameter(s)
*******************************************************************************
 Based on  OLS regression of Y on:
 INPT            Y(-1)           X                X(-1)
 59 observations used for estimation from 1953Q2 to 1967Q4
*******************************************************************************
 Coefficients A1 to A4 are assigned to the above regressors respectively.
 List of specified functional relationship(s):
 LR=(A3+A4)/(1-A2);ML=(A2*A3+A4)/((1-A2)*(A3+A4))
*******************************************************************************
 Function                   Estimate       Standard Error         T-Ratio[Prob]
 LR                          1.0383            .055739            18.6287[.000]
 ML                          4.7030             .52036             9.0381[.000]
*******************************************************************************

        Estimated Variance Matrix of the Function(s) of the Parameters

*******************************************************************************
                LR         ML
 LR         .0031068   .017405

 ML          .017405    .27077

*******************************************************************************
```

According to these results, the hypothesis of a unit long-run coefficient on X cannot be rejected. The mean lag is also estimated with a reasonable degree of accuracy, and suggests a mean lag of 4.7 quarters between changes in capital appropriations and capital expenditures in US manufacturing. Table 11.6 also reports an estimate of the covariance matrix of the functions of parameters LR and ML.

11.4 Lesson 11.4: The multicollinearity problem

Multicollinearity is commonly attributed to situations in which there is a high degree of intercorrelation among the explanatory variables in a multivariate regression equation. Multicollinearity is particularly prevalent in the case of time-series data where there often exists the same common trend in two or more regressors in the regression equation. As a simple example consider the model

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \tag{11.13}$$

and assume for simplicity that (x_{1t}, x_{2t}) have a bivariate distribution with the correlation coefficient, ρ, that is, $\rho = Cov\,(x_{1t}, x_{2t}) / \{V(x_{1t})V(x_{2t})\}^{\frac{1}{2}}$. It is clear that as ρ approaches unity, *separate* estimation of the slope coefficients β_1 and β_2 becomes more and more problematic. Multicollinearity will be a problem if the coefficients of x_{1t} and x_{2t} are jointly statistically significant but neither are statistically significant when tested individually. Expressed differently, multicollinearity will be a problem when the hypotheses $\beta_1 = 0$ and $\beta_2 = 0$ cannot be rejected when tested separately, while the hypothesis $\beta_1 = \beta_2 = 0$, is rejected when tested jointly. This clearly happens when x_{1t} (or x_{2t}) is an exact linear function of x_{2t} (or x_{1t}). In this case, $x_{2t} = \gamma x_{1t}$ and (11.13) reduces to the simple regression equation

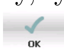
$$y_t = \alpha + (\beta_1 + \beta_2\gamma)x_{1t} + u_t \tag{11.14}$$

and it is only possible to estimate $\beta_1 + \gamma\beta_2$. Neither β_1 nor β_2 can be estimated (or tested) separately. This is the case of 'perfectly multicollinearity' and arises out of faulty specification of the regression equation. One such example is when four seasonal dummies are included in a quarterly regression model that already contains an intercept term.

The multicollinearity problem is also closely related to the problem of low power when separately testing hypotheses involving the regression coefficients. It is worth noting that no matter how large the correlation coefficient between x_{1t} and x_{2t}, so long as it is not exactly equal to ±1, a test of $\beta_1 = 0$ (or $\beta_2 = 0$) will have the correct size, assuming that all the other classical normal assumptions are satisfied. The high degree of correlation between x_{1t} and x_{2t} causes the power of the test to be low, and as a result we may end up not rejecting the null hypothesis that $\beta_1 = 0$ even if it is false.[3]

To demonstrate the multicollinearity problem and its relation to the problem of low power, consider the following (simulated) model

$$\begin{array}{l} x_1 \ \sim \ N(0,1) \\ x_2 = x_1 + 0.15v \\ v \ \sim \ N(0,1) \\ y = \alpha + \beta_1 x_1 + \beta_2 x_2 + u \\ u \ \sim \ N(0,1) \end{array}$$

with $\alpha = \beta_1 = \beta_2 = 1$, and where x_1, v and u are generated as independent standardized normal variates using respectively the 'seed' of 123, 321 and 4321 in the normal random generator (see the function **NORMAL** in Section 4.3.14). To generate x_1, x_2 and y choose Input Data from the Keyboard from the File Menu. In the New data set dialogue choose undated frequency, type 500 for the number of observations, and 0 for the number of variables and then click 

. Type the following formulae in the box editor that appears on the screen to generate the variables Y, $X1$ and $X2$,each having 500 observations:[4]

SAMPLE 1 500;
$X1 =$ **NORMAL**(123); $V =$ **NORMAL**(321);
$U =$ **NORMAL**(4321);
$X2 = X1 + 0.15 * V;$ $Y = 1 + X1 + X2 + U;$ $INPT = 1$ 

[3] The power of a test is defined as the probability of rejecting the null hypothesis when it is false.

[4] Alternatively, you can retrieve the equation file MULTI.EQU into the Commands and Data Transformations box and then click 

to process.

Then move to the Single Equation Estimation Menu (the Univariate Menu on the main menu bar), choose option 1 and run the OLS regression of Y on $INPT$, $X1$ and $X2$ using only the first 50 observations. You should see the results in Table 11.7 on the screen.

Table 11.7: An example of a multicollinear regression based on simulated data

```
                         Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is Y
 50 observations used for estimation from 1 to 50
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                        .90469             .12994            6.9625[.000]
 X1                          1.0950             1.0403            1.0526[.298]
 X2                          .87191             1.0200            .85483[.397]
*******************************************************************************
 R-Squared                     .84982   R-Bar-Squared                   .84343
 S.E. of Regression            .88903   F-Stat.    F(2,47)    132.9788[.000]
 Mean of Dependent Variable    1.4024   S.D. of Dependent Variable      2.2468
 Residual Sum of Squares      37.1474   Equation Log-likelihood       -63.5187
 Akaike Info. Criterion      -66.5187   Schwarz Bayesian Criterion    -69.3868
 DW-statistic                  2.0705
*******************************************************************************


                                 Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version        *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(1)  =    .37159[.542]*F(1,46)     =    .34442[.560]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =  .043743[.834]*F(1,46)      =  .040279[.842]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =    .21521[.898]*       Not applicable       *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =    .85284[.356]*F(1,48)     =    .83293[.366]*
*                     *                          *                            *
* E:Predictive Failure*CHSQ(450)= 528.4668[.006]*F(450,47)    =    1.1744[.253]*
*                     *                          *                            *
* F:Chow Test         *CHSQ(3)  =    .44144[.932]*F(3,494)    =    .14715[.932]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
   E:A test of adequacy of predictions (Chow's second test)
   F:Test of stability of the regression coefficients
```

The value of F statistics $F(2,47)$ for testing the joint hypothesis $H_0^J : \beta_1 = \beta_2 = 0$, against $H_1^J : \beta_1 \neq 0$,and/or $\beta_2 \neq 0$ is equal to 132.9788, which is well above the 95 per cent critical value of the F-distribution with 2 and 47 degrees of freedom, and strongly rejects the joint hypothesis that $\beta_1 = \beta_2 = 0$. The t-statistics for the separate induced tests of $H_0^I : \beta_1 = 0$, against $H_1^I : \beta_1 \neq 0$, and of $H_0^{II} : \beta_2 = 0$, against $H_1^{II} : \beta_2 \neq 0$, are 1.0526 and 0.8548, respectively. Neither are statistically significant and do not lead to the rejection of $\beta_1 = 0$ and $\beta_2 = 0$ when these restrictions are considered separately. The joint hypothesis that $\beta_1$and β_2 are both equal to zero is strongly rejected, but neither of the hypotheses that β_1 and β_2 are separately equal to zero can be rejected. This is clearly a multicollinearity problem. The sample correlation coefficient of x_1 and x_2 computed using the first 50 observations is equal to 0.99316, which is apparently too high given the sample size and the fit of the underlying equation, for the β_1 and β_2 coefficients to be estimated separately with any degree of precision. In short, the separate induced tests lack the necessary power to allow rejection of $\beta_1 = 0$ and/or $\beta_2 = 0$, separately.

The relationship between the F statistic used to test the hypothesis $\beta_1 = \beta_2 = 0$ jointly, and the t-statistics used to test $\beta_1 = 0$ and $\beta_2 = 0$ separately, can also be obtained theoretically, and is given by

$$F = \frac{t_1^2 + t_2^2 + 2\hat{\rho} t_1 t_2}{2(1 - \hat{\rho}^2)} \tag{11.15}$$

where $\hat{\rho}$ is the sample correlation coefficient between x_{1t} and x_{2t}.[5] This relationship clearly shows that even for small values of t_1 and t_2 it is possible to obtain quite large values of F so long as $\hat{\rho}$ happens to be close enough to 1.

In the case of regression models with more than two regressors the detection of the multicollinearity problem becomes even more complicated. For example, when there are three coefficients, namely testing them separately: $\beta_1 = 0$, $\beta_2 = 0$, $\beta_3 = 0$, in pairs: $\beta_1 = \beta_2 = 0$, $\beta_2 = \beta_3 = 0$, $\beta_1 = \beta_3 = 0$, and jointly: $\beta_1 = \beta_2 = \beta_3 = 0$. Only in the case where the results of separate induced tests, the 'pairs' tests and the joint test are free from contradictions can we be confident that the multicollinearity is not a problem.

There are a number of measures in the literature that purport to detect and measure the seriousness of the multicollinearity problem. These measures include the 'condition number' defined as the square root of the largest to the smallest eigenvalue of the matrix $\mathbf{X}'\mathbf{X}$, and the variance-inflation factor, defined as $\left(1 - R_i^2\right)$ for the β_i coefficient where R_i^2 is the squared multiple correlation coefficient of the regression of x_i on the other regressors in the regression equation. Both these measures only examine the intercorrelation between the regressors, and at best present a partial picture of the multicollinearity problem, and can often 'lead' to misleading conclusions.

To illustrate the main source of the multicollinearity problem in the present application, return to the simulation exercise, and use all the 500 observations (instead of the first 50 observations) in computing the regression of y on x_1 and x_2. The results are

$$\begin{array}{lllll} y_t = & \underset{(0.0428)}{0.9307} & + \underset{(0.28343)}{1.1045}\, x_{1t} & + \underset{(0.27981)}{0.93138}\, x_{2t} & + \hat{u}_t \qquad t = 1, 2, ..., 500 \end{array}$$

$$R^2 = 0.8333, \qquad \hat{\sigma} = 0.95664, \qquad F_{2,497} = 1242.3$$

As compared with the estimates based on the first 50 observations (see Table 11.7), these estimates have much smaller standard errors, and using the 95 per cent significance level we arrive at the same conclusion whether we test $\beta_1 = 0$ and $\beta_2 = 0$ separately or jointly. Yet, the sample correlation coefficient between x_{1t} and x_{2t} estimated over the 500 observations is equal to 0.9895, which is only marginally smaller than the estimate obtained for the first 50 observations. By increasing the sample size from 50 to 500 we have increased the precision with which β_1 and β_2 are estimated and the power of testing $\beta_1 = 0$ and $\beta_2 = 0$, both separately and jointly.

[5] In the simulation exercise we obtained $t_1 = 1.0526$, $t_2 = 0.8548$ and $\hat{\rho} = 0.99316$. Using these estimates in (11.15) yields $F = 132.9791$, which is only slightly different from the F statistic reported in Table 11.7. The difference between the two values is due to the rounding of errors.

The above illustration also points to the fact that the main cause of the multicollinearity problem is a lack of adequate observations (or information), and hence the imprecision with which the parameters of interest are estimated. Assuming that the regression model under consideration is correctly specified, the appropriate solution to the problem is to increase the information on the basis of which the regression is estimated. The new information could be either in the form of additional observations on y, x_1 and x_2, or it could be some *a priori* information concerning the parameters. The latter fits well with the Bayesian approach, but is difficult to accommodate within the classical framework. There are also other approaches suggested in the literature such as the ridge regression and the principle component regression, to deal with the multicollinearity problem. A review of these approaches can be found in Judge, Griffiths, Hill, Lütkepohl, and Lee (1985).

11.5 Lesson 11.5: Testing common factor restrictions

Consider the following $ARDL(1,1,1)$ model relating logarithm of real consumption expenditures ($\log c_t$) to the logarithm of the real disposable income ($\log y_t$) and the rate of inflation (Π_t) in the UK:

$$(1-\beta_1 L)\log c_t = \beta_0 + (\beta_2 + \beta_3 L)\log y_t + (\beta_4 + \beta_5 L)\Pi_t + u_t \tag{11.16}$$

where L represents the backward lag operator. The idea of testing for common factor restrictions was originally proposed by Sargan (1964). The test explores the possibility of simplifying the dynamics of (11.16) by testing the hypothesis that the lag polynomials operating on $\log c_t$, $\log y_t$, and Π_t have the same factor in common. The procedure can also be viewed as a method of testing the dynamics in the deterministic part of the regression model against the dynamics in the stochastic part (see Hendry, Pagan, and Sargan (1984), Section 2.6). In the case of the present example, the common factor restrictions are[6]

$$\left.\begin{array}{l} \beta_1\beta_2 + \beta_3 = 0 \\ \beta_1\beta_4 + \beta_5 = 0 \end{array}\right\} \tag{11.17}$$

A test of these restrictions can be readily carried out using *Microfit*. Here we assume that (11.16) is to be estimated by the OLS method, but the procedure outlined below is equally applicable if (11.16) is estimated by the IV method.

We use quarterly observations in the special *Microfit* file UKCON.FIT to carry out the test. First read the UKCON.FIT and make sure that the variables $LC = \log c_t$, $LY = \log y_t$, $P_t = CNOM/C$ and $PI = \log(P_t/P_{t-1})$ are on the workspace. To generate these variables go to the Data Processing Stage and retrieve the file UKCON.EQU, or equivalently run the BATCH file UKCON.BAT. (See also Lessons 10.1 and 10.2).

Choose option 1 in the Single Equation Estimation Menu (the Univariate Menu on the main menu bar: Section 6.4) and type

$$LC \quad INPT \quad LC(-1) \quad LY \quad LY(-1) \quad PI \quad PI(-1)$$

RUN

[6] To derive the restrictions in (11.17) note that for the lag polynomials $1-\beta_1 L$ and $\beta_4 + \beta_5 L$ to have the same factor in common, it is necessary that β_1^{-1}, the root of $1-\beta_1 L = 0$, should also be a root of $\beta_2 + \beta_3 L = 0$ and $\beta_4 + \beta_5 L = 0$.

You should see the OLS regression results on the screen. Move to the Post Regression Menu and choose option 2 (see Section 6.20). This takes you to the Hypothesis Testing Menu (see Section 6.23). Now choose option 7 in this menu to carry out a Wald test of the common factor restrictions in (11.17). In the box editor that appears on the screen type

$$A2 * A3 + A4 = 0; \quad A2 * A5 + A6 = 0$$

You should now see the test results shown in Table 11.8.

Table 11.8: Testing for common factor restrictions

```
                 Wald test of restriction(s) imposed on parameters
*******************************************************************************
 Based on  OLS regression of LC on:
 INPT            LC(-1)          LY              LY(-1)          PI
 PI(-1)
 158 observations used for estimation from 1955Q3 to 1994Q4
*******************************************************************************
 Coefficients A1 to A6 are assigned to the above regressors respectively.
 List of restriction(s) for the Wald test:
 A2*A3+A4=0;  A2*A5+A6=0
*******************************************************************************
 Wald Statistic                  CHSQ(2)=  22.0612[.000]
*******************************************************************************
```

The Wald statistic for testing the two non-linear restrictions in (11.17) is equal to 22.06, which implies a strong rejection of the common factor restrictions. It is, however, important to note that the Wald statistic is sensitive to the way the non-linear restrictions are specified. See, for example, Gregory and Veall (1985) and Gregory and Veall (1987). See also Exercise 11.3.

11.6 Lesson 11.6: Estimation of regression models with serially correlated errors

Suppose now you wish to estimate the saving equation

$$s_t = \alpha_0 + \alpha_1 s_{t-1} + \alpha_2 \Delta \log y_t + \alpha_3 \left(\Pi_t - \Pi_t^e\right) + u_t \tag{11.18}$$

using UK quarterly observations in the special *Microfit* file UKCON.FIT

s_t	saving rate (the variable S on the workspace)
$\Delta \log y_t$	the rate of change of real disposable income (DLY)
Π_t	Actual rate of inflation (PI)
Π_t^e	Adaptive expectations of Π_t as computed in Lesson 10.12

subject to the $AR(1)$ error specification by the Cochrane-Orcutt method

$$u_t = \rho_1 u_{t-1} + \epsilon_t \tag{11.19}$$

Equation (11.18) is a modified version of the saving function estimated by Deaton (1977).[7]

To carry out the computations for the above estimation problem first go through the steps described in Lesson 10.12 to generate the variable PIE with $\lambda = 0.8$ in your workspace. Alternatively, read UKCON.FIT and then retrieve the file UKCON.EQU into the Commands and Data Transformations box. If you now click the variables S, DLY, PI and $INPT$ will be created on the workspace. Click to clear the editor and then type

$$LAMBDA = 0.8; \quad \textbf{BATCH} \quad PIE$$

to generate the inflation expectations variable PIE. Then create the unanticipated inflation variable $\Pi_t - \Pi_t^e$ by typing

$$DPIE = PI - PIE$$

Now move to the Single Equation Estimation Menu (the Univariate Menu), and choose option 4 from the Linear Regression Menu (see Section 6.9). Type

$$S \quad INPT \quad S(-1) \quad DLY \quad DPIE$$

and choose the start and end dates 1960(1) and 1994(4). Click and, when prompted, type 1 and press . You will now be presented with a menu for initializing the estimation process (see Section 6.13.1). Choose option 3 to see the plot of the concentrated log-likelihood function, showing the log-likelihood profile for different values of ρ, in the range $[-0.99, 0.99]$ (see Figure 11.3).

As you can see, the log-likelihood function is bimodal for a positive and a negative value of ρ_1. The global maximum of the log-likelihood is achieved for $\rho_1 < 0$. Bimodal log-likelihood functions frequently arise in estimation of models with lagged dependent variables subject to a serially correlated error process, particularly in cases where the regressors show a relatively low degree of variability. The bimodal problem is sure to arise if apart from the lagged values of the dependent there are no other regressors in the regression equation.

To compute the ML estimates click to return to the menu for initialization of the unknown parameter ρ_1. Choose option 2 and type -0.2 as the initial estimate for ρ_1 and click . The results in Table 11.9 should now appear on the screen. The iterative algorithm has converged to the correct estimate of ρ_1 (i.e. $\hat{\rho}_1 = -0.22838$) and refers to the global maximum of the log-likelihood function given by $LL(\hat{\rho}_1 = -0.22838) = 445.3720$. Notice also that the estimation results are reasonably robust to the choice of the initial estimates chosen for ρ_1, so long as negative or small positive values are chosen. However, if the iterations are started from $\rho_1^{(0)} = 0.5$ or higher, the results in Table 11.10 will be obtained. The iterative process has now converged to $\hat{\rho}_1 = 0.81487$ with the maximized value for the log-likelihood function given by $LL(\hat{\rho}_1 = 0.81487) = 444.3055$, which is a local maximum. (Recall from Table 11.9 that $LL(\hat{\rho}_1 = -0.22838) = 445.3720$). This example clearly shows the importance of experimenting with different initial values when estimating regression models (particularly

[7] However, note that the saving function estimated by Deaton (1977) assumes that the inflation expectations Π_t^e are time invariant.

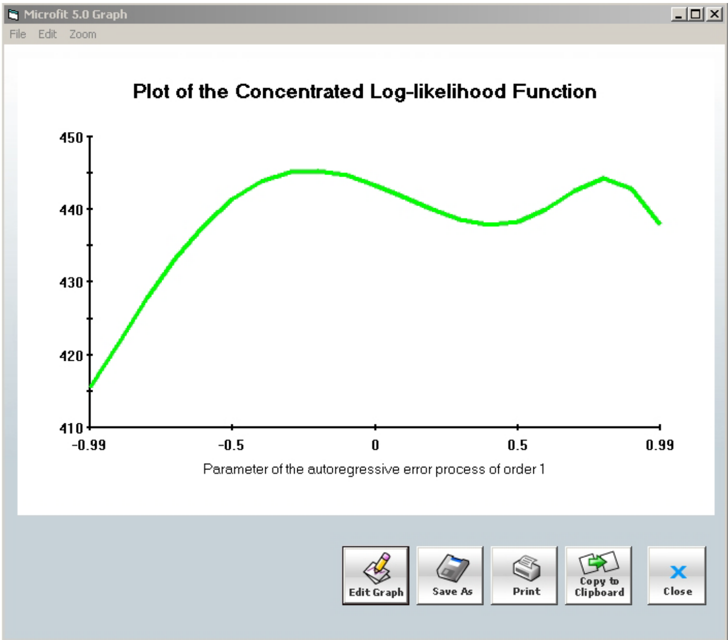


Figure 11.3: Log-likelihood profile for different values of ρ_1

when they contain lagged dependent variables) with serially correlated errors. Suppose now you wish to estimate equation (11.18) subject to the following $AR(4)$ error process with zero restrictions on two of its coefficients:

$$u_t = \rho_1 u_{t-1} + \rho_4 u_{t-4} + \epsilon_t$$

Return to the Single Equation Estimation window via the Backtracking Menu and choose option 5 in the Liner Regression estimation Menu, run the same calculation, and when prompted, type

1 4 

Choose option 1 to use the initial estimate supplied by the program. The results in Table 11.11 should now appear on the screen.

Table 11.9: Cochrane-Orcutt estimates of a UK saving function

```
               Cochrane-Orcutt Method AR(1) converged after 3 iterations
*******************************************************************************
 Dependent variable is S
 140 observations used for estimation from 1960Q1 to 1994Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                    -.0032323          .0041204            -.78448[.434]
 S(-1)                      .99250           .040347            24.5989[.000]
 DLY                        .66156           .060082            11.0111[.000]
 DPIE                       .31032           .093382             3.3231[.001]
*******************************************************************************
 R-Squared                      .76673   R-Bar-Squared                   .75977
 S.E. of Regression            .010004   F-Stat.    F(4,134)    110.1102[.000]
 Mean of Dependent Variable    .096441   S.D. of Dependent Variable     .020696
 Residual Sum of Squares       .013412   Equation Log-likelihood       445.3720
 Akaike Info. Criterion       440.3720   Schwarz Bayesian Criterion    433.0179
 DW-statistic                   1.9615
*******************************************************************************

             Parameters of the Autoregressive Error Specification
*******************************************************************************
 U=    -.22838*U(-1)+E
   (   -2.5135)[.013]
 T-ratio(s) based on asymptotic standard errors in brackets
*******************************************************************************
```

Table 11.10: An example in which the Cochrane-Orcutt method has converged to a local maximum

```
               Cochrane-Orcutt Method AR(1) converged after 7 iterations
*******************************************************************************
 Dependent variable is S
 140 observations used for estimation from 1960Q1 to 1994Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                     .075353           .0098576             7.6441[.000]
 S(-1)                     .19990            .084385             2.3689[.019]
 DLY                       .55758            .052907            10.5388[.000]
 DPIE                      .45522             .10271             4.4322[.000]
*******************************************************************************
 R-Squared                      .76312   R-Bar-Squared                   .75605
 S.E. of Regression            .010081   F-Stat.    F(4,134)    107.9234[.000]
 Mean of Dependent Variable    .096441   S.D. of Dependent Variable     .020696
 Residual Sum of Squares       .013619   Equation Log-likelihood       444.3055
 Akaike Info. Criterion       439.3055   Schwarz Bayesian Criterion    431.9514
 DW-statistic                   2.2421
*******************************************************************************

             Parameters of the Autoregressive Error Specification
*******************************************************************************
 U=     .81487*U(-1)+E
   (   16.1214)[.000]
 T-ratio(s) based on asymptotic standard errors in brackets
*******************************************************************************
```

Table 11.11: ML estimates of a saving equation with restricted $AR(4)$ error process maximum likelihood estimation: fixed initial values of disturbances

```
          Maximum Likelihood Estimation:Fixed Initial Values of Disturbances
                Error TERM : Restricted AR(4) converged after 7 iterations
*******************************************************************************
 Dependent variable is S
 140 observations used for estimation from 1960Q1 to 1994Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                    -.0048180            .0042275            -1.1397[.256]
 S(-1)                     1.0080             .041256             24.4335[.000]
 DLY                       .65590             .061110             10.7330[.000]
 DPIE                      .29880             .093110              3.2091[.002]
*******************************************************************************
 R-Squared                      .76567   R-Bar-Squared                   .75666
 S.E. of Regression            .010094   F-Stat.    F(5,130)     84.9560[.000]
 Mean of Dependent Variable    .096441   S.D. of Dependent Variable     .020696
 Residual Sum of Squares       .013247   Equation Log-likelihood       435.1181
 Akaike Info. Criterion       429.1181   Schwarz Bayesian Criterion    420.2932
 DW-statistic                   1.9598
*******************************************************************************

              Parameters of the Autoregressive Error Specification
*******************************************************************************
 U=    -.25410*U(-1)+   -.014586*U(-4)+E
   (   -2.7908)[.006]  (    -.16654)[.868]
 T-ratio(s) based on asymptotic standard errors in brackets
*******************************************************************************
```

These estimates seem to be quite robust to the choice of the initial values for ρ_1 and ρ_4. For example, starting the iterations with $\rho_1^{(0)} = 0.8$ and $\rho_2^{(0)} = 0.0$ yields the same results as in Table 11.11.

11.7 Lesson 11.7: Estimation of a 'surprise' consumption function: an example of two-step estimation

A simple version of the life cycle rational expectations theory of consumption predicts that changes in real consumption expenditures (or their logarithms) are only affected by innovations in real disposable income. Muellbauer (1983), building on the seminal work of Hall (1978), has estimated the following 'surprise' aggregate consumption function for the UK

$$\Delta \log c_t = a_0 + a_1(\log y_t - \widehat{\log y_t}) + u_t \qquad (11.20)$$

where c_t = real consumption expenditures, y_t = real disposable income, and $\widehat{\log y_t}$ is the predictor of $\log y_t$ based on information at time $t-1$.

In this lesson we use quarterly observations on c_t and y_t in the file UKCON.FIT to estimate (11.20). This is an example of the two-step estimation method for rational expectations discussed in Pagan (1984) and Pesaran (1987b) and Pesaran (1991).

In the first step the predicted values of $\log y_t$ are obtained by running the OLS regression of $\log y_t$ on its past values, and possibly lagged values of other variables. In what follows we estimate a second-order autoregressive process, $AR(2)$, in $\log y_t$. In the second step (11.20) is estimated by running the OLS regression of $\Delta \log c_t$ on a constant term and the residuals obtained from the regression in the first step.

To carry out the computations in the first step, read in the special *Microfit* file UKCON.FIT, generate the variables $LC = \log c_t$ and $LY = \log y_t$, and an intercept term, and then choose option 1 in the Single Equation Estimation Menu (Univariate Menu: see Section 6.4), make sure the OLS option is selected, and type

$$LY \quad INPT \quad LY(-1) \quad LY(-2)$$ 

When the table appears, click 

, and from the Post Regression Menu choose option 3. You should now see the Display/Save Residuals and Fitted Values Menu on the screen. Choose option 6, and when prompted type

$$DRLY \quad \text{Unanticipated change in } \log(Y)$$ 

The variable $DRLY \left(= \log y_t - \widehat{\log} y_t\right)$ is saved in the workspace, and you can now carry out the computations in the second step.

Return to the Single Equation Estimation window (making sure the OLS option from the Linear Regression Menu is selected), click 

to clear the box editor and then type

$$DLC \quad INPT \quad DRLY$$ 

The results in Table 11.12 should now appear on the screen.

Table 11.12: 'Surprise' consumption function for the UK

```
                    Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is DLC
 158 observations used for estimation from 1955Q3 to 1994Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                     .0061568           .8913E-3             6.9076[.000]
 DRLY                       .27395            .056753             4.8269[.000]
*******************************************************************************
 R-Squared                     .12995   R-Bar-Squared                   .12437
 S.E. of Regression           .011203   F-Stat.    F(1,156)    23.2993[.000]
 Mean of Dependent Variable  .0061568   S.D. of Dependent Variable     .011973
 Residual Sum of Squares      .019581   Equation Log-likelihood       486.4768
 Akaike Info. Criterion      484.4768   Schwarz Bayesian Criterion    481.4142
 DW-statistic                  2.4947
*******************************************************************************


                            Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version        *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(4)  =   13.5373[.009]*F(4,152)      =   3.5609[.008]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =    .19194[.661]*F(1,155)      =   .18852[.665]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =   99.9521[.000]*       Not applicable       *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =    1.1184[.290]*F(1,156)      =   1.1121[.293]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

The t-ratio of the coefficient of $DRLY$ is 4.8269, which is much higher than the critical value of the t distribution with $158 - 2 = 156$ degrees of freedom; thus suggesting that innovations in income growth have significant impact on consumption growth.[8]

Repeat the above exercise using the USCON.FIT file. Also try additional regressors, such as $\Delta \log y_{t-1}$ and $\Delta \log c_{t-1}$ in (11.20). Are your results sensitive to the order of the AR process chosen to compute $\widehat{\log y_t}$?

[8]Notice that

$$\begin{aligned} \log y_t - \widehat{\log y_t} &= (\log y_t - \log y_{t-1}) - \left(\widehat{\log y_t} - \log y_{t-1}\right) \\ &= \Delta \log y_t - \Delta\widehat{\log y_t}. \end{aligned}$$

11.8 Lesson 11.8: An example of non-nested hypothesis testing

Suppose you are faced with the following models:

$$\begin{aligned} M_1 &: \Delta \log c_t = \alpha_0 + \alpha_1 \log y_t + \alpha_2 \log c_{t-1} + \alpha_3 \log y_{t-1} + \alpha_4 \Pi_t + u_{t1} \\ M_2 &: \Delta \log c_t = \beta_0 + \beta_1 (\log y_t - \widehat{\log y_t}) + \beta_2 \Pi_t + u_{t2} \end{aligned}$$

Model M_1 is an inflation-augmented version of the error correction model (11.1) in Lesson 11.1. The inflation rate, Π_t, is measured as the change in the logarithm of the implicit price deflator of consumption.

Model M_2 is the inflation-augmented 'surprise' consumption function, and is estimated in Lesson 11.7. First read the special *Microfit* file UKCON.FIT and make sure that the following variables are in the list:

DLC	$\Delta \log c_t$
DLY	$\Delta \log y_t$
PI	$\log(p_t/p_{t-1})$
P	Implicit price Deflator of Consumption Expenditure
$DRLY$	$\log y_t - \widehat{\log y_t}$
	$\log y_t - \hat{\rho}_0 - \hat{\rho}_1 \log y_{t-1} - \hat{\rho}_2 \log y_{t-2}$
$INPT$	Intercept term

If one or more of these variables are not in your workspace you need to consult the relevant lessons on how to generate them. (See Lessons 11.6 and 11.7).

Suppose now that you wish to test model M_1 against M_2 and vice versa. Choose option 1 in the Single Equation Estimation Menu (Univariate Menu: see Section 6.4), make sure the OLS option is selected, and type

$$DLC \quad INPT \quad DLY \quad LC(-1) \quad LY(-1) \quad PI$$ 

Click 

and then 

, and select option 2 in the Post Regression Menu. You should now see the Hypothesis Testing Menu (see Section 6.23) on the screen. Choose option 8, and when prompted, first click 

to clear the box editor and then type the regressors of model M_2:

$$INPT \; DRLY \quad PI$$ 

The results in Table 11.13 should now appear on the screen.

Table 11.13: Non-nested statistics for testing *ARDL* and 'surprise' consumption functions

```
                 Alternative Tests for Non-Nested Regression Models
*******************************************************************************
 Dependent variable is DLC 158 observations used from 1955Q3 to 1994Q4
 Regressors for model M1:
 INPT            DLY             LC(-1)          LY(-1)          PI
 Regressors for model M2:
 INPT            DRLY            PI
*******************************************************************************
 Test Statistic                  M1 against M2                   M2 against M1
     N-Test                      -3.6717[.000]                   -6.1894[.000]
    NT-Test                      -3.4159[.001]                   -2.2548[.024]
     W-Test                      -3.3767[.001]                   -2.2163[.027]
     J-Test                       3.3888[.001]                    2.7843[.005]
    JA-Test                       3.3888[.001]                   -2.7045[.007]
  Encompassing        F(1,152)   11.4840[.001]        F(3,152)    6.3517[.000]
*******************************************************************************
 Model M1:    DW    2.1543 ;R-Bar-Squared    .21629 ;Log-likelihood   497.2737
 Model M2:    DW    2.4300 ;R-Bar-Squared    .19057 ;Log-likelihood   493.6973
 Model M1+M2: DW    2.2193 ;R-Bar-Squared    .26655 ;Log-likelihood   503.0276
 Akaike's Information Criterion of M1 versus M2=   1.5764 favours M1
 Schwarz's Bayesian Criterion of M1 versus M2=  -1.4862 favours M2
*******************************************************************************
```

All the non-nested tests suggest that both models should be rejected. There is also a conflict between the two model selection criteria, with the Akaike information criterion favouring M_2, and the Schwarz Bayesian criterion favouring M_1. The test results point to another model, possibly a combination of models M_1 and M_2, as providing a more satisfactory specification.

11.9 Lesson 11.9: Testing linear versus log-linear models

Suppose you are interested in testing the following linear form of the inflation augmented $ARDL(1,1)$ model:

$$M_1 : c_t = \alpha_0 + \alpha_1 c_{t-1} + \alpha_2 y_t + \alpha_3 y_{t-1} + \alpha_4 \Pi_t + u_{1t}$$

against its log-linear form

$$M_2 : \log c_t = \beta_0 + \beta_1 \log c_{t-1} + \beta_2 \log y_t + \beta_3 \log y_{t-1} + \beta_4 \Pi_t + u_{2t}$$

where

c_t	Real Non-durable Consumption Expenditure in the US
y_t	Real Disposable Income in the US
π_t	Inflation Rate

First read the special *Microfit* file USCON.FIT, and generate the necessary variables for running the above regressions, (for example by using the 

button to retrieve the file USCON.EQU into the Commands and Data Transformation box). Choose option 1 in the

Single Equation Estimation Menu (Univariate Menu, see Section 6.4), make sure the OLS option is selected, and type

$$C \quad INPT \quad C(-1) \quad Y \quad Y(-1) \quad PI$$

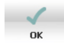
Click 

, then 

, and select option 2. You should now see the Hypothesis Testing Menu (see Section 6.23) on the screen. Choose option 9 and, when prompted, type the regressors of model M_2, namely

$$INPT \quad LC(-1) \quad LY \quad LY(-1) \quad PI$$

You will now be asked to specify the nature of the transformation of the dependent variable in model M_1. Choose the linear option 1.

A similar menu concerning the nature of the transformation of the dependent variable in model M_2 now appears on the screen. Choose option 2. You will be prompted to specify the number of replications (R) to be used in the computations of the Cox statistic by simulation (see Section 21.9 and option 9 in Section 6.23). For most applications, values of R in the range 100-250 will be adequate. Enter 100, and press 

for computations to start. Once the computations are completed, the results in Table 11.14 should appear on the screen. This table gives the parameter estimates under both models. The estimates of the parameters of M_1 computed under M_1 are the OLS estimates ($\hat{\alpha}$), while the estimates of the parameters of M_1 computed under M_2 are the pseudo-true estimators ($\hat{\alpha}_* = \hat{\alpha}_*(\hat{\beta})$). If model M_1 is correctly specified, one would expect $\hat{\alpha}$ and $\hat{\alpha}_*$ to be near to one another. The same also applies to the estimates of the parameters of model M_2 (β).

Table 11.14: Testing linear versus log-linear consumption functions

```
                          Non-Nested Tests by Simulation
*******************************************************************************
 Dependent variable in model M1 is C
 Dependent variable in model M2 is LOG(C)
 136 observations used from 1960Q2 to 1994Q1. Number of replications 100
*******************************************************************************
   Estimates of parameters of M1              Estimates of parameters of M2
                 Under M1  Under M2                          Under M2  Under M1
 INPT             20.1367   24.7609           INPT             .14429   .098439
 C(-1)             .93128    .90987           LC(-1)           .89781    .94227
 Y                .092935   .098593           LY               .29526    .27626
 Y(-1)           -.076891  -.077510           LY(-1)          -.22532   -.23860
 PI             -160.9665 -156.5597           PI              -.23880   -.23041
 Standard Error    4.6528    4.8476           Standard Error  .0057948  .0061028
 Adjusted Log-L -399.5709 -404.4709           Adjusted Log-L -399.0233 -405.2882
*******************************************************************************

                   Non-Nested Test Statistics and Choice Criteria
*******************************************************************************
 Test Statistic                  M1 against M2                    M2 against M1
     S-Test                      -2.5592[.010]                    -1.8879[.059]
    PE-Test                       2.3021[.021]                    -.26402[.792]
    BM-Test                       2.0809[.037]                    -.50743[.612]
    DL-Test                       2.0006[.045]                     1.8004[.072]
 Sargan's Likelihood Criterion for M1 versus M2= -.54764       favours M2
 Vuong's  Likelihood Criterion for M1 versus M2= -1.7599[.078] favours M2
*******************************************************************************
 S-Test is the SC_c test proposed by Pesaran and Pesaran (1995) and is
 the simple version of the simulated Cox test statistic.
 PE-Test is the PE test due to MacKinnon, White and Davidson.
 BM-Test is due to Bera and McAleer.
 DL-Test is the double-length regression test statistic due to Davidson
 and MacKinnon.
```

The bottom part of Table 11.14 gives a number of different statistics for testing the linear versus the log-linear model and vice versa. This table also gives the Sargan (1964) and Vuong (1989) likelihood function criteria for the choice between the two models. For other details and references to the literature, see Section 21.9.

In the present application all the tests reject the linear model against the log-linear model, and none reject the log-linear model against the linear one at the 5 per cent significance level, although the simulated Cox and the double-length tests also suggest rejection of the log-linear model at the 10 per cent significance level. Increasing the number of replications to 500 does not alter this conclusion. The two choice criteria also favour the log-linear specification over the linear specification.

11.10 Lesson 11.10: Testing for exogeneity: computation of the Wu-Hausman statistic

In this lesson we show how the variable addition test option in the Hypothesis Testing Menu (see Section 6.23) can be used to compute the Wu (1973) T_2 statistic for testing the independence (or more precisely the lack of correlation) of the regressors, $\log y_t$ and Π_t, and the disturbance term, u_t, in the following regression equation estimated on UK data:[9]

$$\log c_t = \alpha_0 + \alpha_1 \log c_{t-1} + \alpha_2 \log y_t + \alpha_3 \log y_{t-1} + \alpha_4 \Pi_t + u_t \tag{11.21}$$

We assume that you have read in the file UKCON.FIT and that the variables $\log c_t$, $\log y_t$, and Π_t are in the variable list (these variables can be generated by running the batch file UKCON.BAT). We also assume that the variables, $\log y_{t-1}$, $\log y_{t-2}$, $\log c_{t-1}$, $\log c_{t-2}$, Π_{t-1}, Π_{t-2}, can be used as instruments for this test.[10]

Computation of the Wu-Hausman T_2 statistic can be carried out in the following manner.

1. Run OLS regressions of LY ($\log y_t$) and PI (Π_t) on the variables $INPT$, $LY(-1)$, $LY(-2)$, $LC(-1)$, $LC(-2)$, $PI(-1)$, and $PI(-2)$, over the period 1960(1)-1994(4), and save the residuals (using option 6 in the Display/Save Residuals and Fitted Values Menu), in the variables RLY and RPI, respectively (see Lesson 11.7 on how to do this). More specifically, choose option 1 in the Single Equation Estimation Menu (see Section 6.4), make sure the OLS option is selected, choose the estimation period 1960(1) to 1994(4), then enter

   $$LY \;\; INPT \;\; LY\{1-2\} \quad LC\{1-2\} \quad PI\{1-2\}$$ RUN

   When the table appears, click Close. Choose option 3 from the Post Regression Menu and option 6 from the Display/Save Residuals and Fitted Values Menu. When prompted, enter

   RLY Residuals from LY regression

   Press OK, and then choose option 0 to return to the Single Equation Estimation window. Replace LY by PI in the screen editor box. Click RUN and when the table appears click Close. Choose option 3 from the Post Regression Menu and Option 6 from the Display/Save Residuals and Fitted Values. When prompted, enter the following string:

   RPI Residuals from PI regression

   Click OK to move to the Post Regression Menu. RUN Choose option 0 to return to the Single Equation Estimation window.

2. Make sure that the variables RLY and RPI are correctly saved in your workspace. Then choose option 1 in the Single Equation Estimation Menu (the Univariate Menu on the menu bar) and make sure the OLS option is selected. Click Clear to clear the

[9] Wu's T_2 statistic is also known as the Wu-Hausman statistic. For details see Wu (1973), Hausman (1978), Nakamura and Nakamura (1981), and Pesaran and Smith (1990).

[10] The Wu-Hausman test is also asymptotically equivalent to testing the statistical significance of the difference between the OLS and the Two-Stage Least Squares estimates of the regression coefficients in (11.21). It is also advisable to carry out Sargan's general mis-specification test given in the result table in the case of IV regressions.

box editor, and ensure that the start and end dates are set to 1960(1) and 1994(4). Then type

$$LC \quad INPT \quad LC(-1) \quad LY\{0-1\} \quad PI$$

click the 

button and proceed. When the table appears, click 

. In the Post Regression Menu select option 2 and then choose option 6 from the Hypothesis Testing Menu. In the box editor enter

$$RLY \quad RPI$$


Table 11.15: The Wu-Hausman statistic for testing the exogeneity of LY and PI

```
                    Variable Addition Test (OLS case)
*******************************************************************************
 Dependent variable is LC
 List of the variables added to the regression:
 RLY             RPI
 140 observations used for estimation from 1960Q1 to 1994Q4
*******************************************************************************
 Regressor                Coefficient        Standard Error          T-Ratio[Prob]
 INPT                        .017173             .039516              .43458[.665]
 LC(-1)                       1.3181              .15256              8.6400[.000]
 LY                          -1.2561              .48159             -2.6083[.010]
 LY(-1)                       .94106              .34011              2.7669[.006]
 PI                          -.26192             .094060             -2.7846[.006]
 RLY                          1.5019              .48568              3.0923[.002]
 RPI                         -.21711              .15261             -1.4227[.157]
*******************************************************************************
 Joint test of zero restrictions on the coefficients of additional variables:
 Lagrange Multiplier Statistic     CHSQ(2)=  11.0589[.004]
 Likelihood Ratio Statistic        CHSQ(2)=  11.5201[.003]
 F Statistic                       F(2,133)=   5.7035[.004]
*******************************************************************************
```

The Wu-Hausman statistic (Wu's T_2 statistic) is equal to the value of the F-statistic in Table 11.15, which is computed as 5.70, and under the null hypothesis (of exogeneity) is distributed approximately as an F with 2 and 133 degrees of freedom. The exogeneity test can also be based on the Lagrange multiplier, or the likelihood ratio statistic reported in the above table. All three tests are asymptotically equivalent, and in the case of the present application, reject the null hypothesis that the income and inflation variables are exogenous in the inflation augmented $ARDL(1,1)$, in (11.21). However, as can be seen from Table 11.15, the t-ratio of the inflation variable, RPI, is -1.4227, and suggests that the hypothesis that the inflation rate is exogenous cannot be rejected.

Also, the rejection of the exogeneity of $\log y_t$ in (11.21) crucially depends on the exclusion of $\log y_{t-2}$ from the $ARDL$ specification. As an exercise, include $\log y_{t-2}$ among the regressors of (11.21) and try the above exogeneity test again.

11.11 Lesson 11.11: Recursive prediction of US monthly excess returns

The literature on the predictability of returns or excess returns on common stocks is quite extensive. It has been shown that a substantial part of variations in excess returns at different time intervals is predictable. See, for example, Campbell (1987), Fama and French (1989), Pesaran (1991), Breen, Glosten, and Jagannathan (1989), Glosten, Jagannathan, and Runkle (1993) and Pesaran and Timmermann (1994). In this lesson we replicate some of the excess return regressions reported in Pesaran and Timmermann (1994) at monthly frequencies, and show how to use such regressions to generate recursive predictions of excess returns on Standard and Poor 500 (SP500) portfolio using only *ex ante* dated variables.

The special *Microfit* file PTMONTH.FIT contains monthly observations on a number of financial and macroeconomic variables over the period 1948(1) to 1992(12) for the US economy (notice, however, that there are missing observations for most of the variables during the 1948-1951 period). Read this file and run the batch file PTMONTH.BAT on it in the Commands and Data Transformations box. The following variables should now be in the variables list:

$DI11$	$i1 - i1(-1)$
$DIP12$	$\log(ip12/ip12(-12))$
$ERSP$	$nrsp - ((1 + (i1(-1)/100)\hat{}(1/12)) + 1$
$INPT$	1
$PI12$	$\log(ppi12/ppi12(-12))$
YSP	$divsp/psp$

where

$DIVSP$	Twelve-month average of dividends on SP500
$I1$	One-month t-bill rate (Fama-Bliss)
IP	Index of industrial production
$IP12$	$\mathbf{MAV}(ip, 12)$
$NRSP$	$(psp - psp(-1) + divsp)\,/psp(-1)$
PPI	Producer price index
$PPI12$	$\mathbf{MAV}(ppi, 12)$
PSP	SP500 price index (end of month)

$ERSP$ is the excess return on SP500 defined as the difference between the nominal return on SP500 ($NRSP$) minus the lagged one month Treasury Bill (TB) rate converted from an annual rate to a monthly rate (allowing for compounding).

$DI11$	Change in the one-month T-bill rate of the index of industrial production
$DIP12$	Rate of change of twelve-month moving average of the index of industrial production
$PI12$	Rate of change of the twelve-month moving average of the producer price index
YSP	Dividend yield defined as the ratio of dividends to share prices

Notice that in computing the twelve-month moving averages of the industrial production and producer price indices we have made use of the 'moving-average' function $\mathbf{MAV}(\cdot,\cdot)$ described in Section 4.3.11.

The excess return regressions reported in Pesaran and Timmermann (1994) were initially estimated over the period 1954(1)-1990(12), but were later extended to include the two years 1991 and 1992 (see Tables III and PIII in Pesaran and Timmermann (1994)). Here we consider estimating the following monthly excess return regression over the whole period 1954(1)-1992(12):

$$\begin{aligned} ERSP_t &= \beta_0 + \beta_1 YSP_{t-1} + \beta_2 PI12_{t-2} + \beta_3 DI11_{t-1} \\ &\quad + \beta_{4t} DIP12_{t-2} + u_t \end{aligned} \tag{11.22}$$

Under the joint hypothesis of risk neutrality and market efficiency it should not be possible to predict the excess returns, $ERSP_t$, using publicly available information. It is therefore important that observations on the regressors in (11.22) are available publicly at time $t-1$, when $ERSP_t$ is being forecast. Such information is readily available for the interest rates, share prices and dividends, but not for the production and the producers' price indices. Observations on these latter variables are released by the US government with a delay. In view of this the variables $PI12$ and $DIP12$ are included in the excess return regression with a lag of two months.

To replicate the OLS results in table PIII in Pesaran and Timmermann (1994), p.61, choose option 1 in the Single Equation Estimation Menu, and when prompted type[11]

$$ERSP \quad INPT \quad YSP(-1) \quad PI12(-2) \quad DI11(-1) \quad DIP12(-2)$$

The OLS results in Table 11.16 should appear on the screen. Check that these estimates are identical with those reported by Pesaran and Timmermann (1994), p. 61. To estimate (11.22) recursively choose option 2 in the Single Equation Estimation Menu (Section 6.4). The following variable list should be in the box editor:

$$ERSP \quad INPT \quad YSP(-1) \quad PI12(-2) \quad DI11(-1) \quad DIP12(-2)$$

Set the number of observations used for updating estimation equal to 1, and click RUN. The program now carries out the necessary computations and presents you with the Recursive OLS Regression Results Menu (see Section 6.14.1). You can use option 1 in this menu to plot the recursive estimates. For example, if you choose to see the recursive estimates of the coefficient of the dividend yield variable, $YSP(-1)$, the plot in Figure 11.4 will appear on the screen.

To save recursive predictions of excess return choose option 8, and when prompted type

$ERHAT$ Recursive Predictions of Excess Return on SP500

[11] The relevant variable list is saved in the special *Microfit* file PTMONTH.LST, and can be retrieved using the Retrieve button.

Click , and you will now be asked to supply the variable name for the standard errors of the recursive predictions, themselves computed recursively. Type[12] $ERSE$ standard errors of the recursive predictions of excess return on SP500, click and choose option 0 to return to the box editor. You should now see the variables $ERHAT$ and $ERSE$ in the variable list (click to to display). To estimate the extent to which the recursive predictions of excess return ($ERHAT$) and the actual excess returns ($ERSP$) are correlated, you can either compute simple correlation coefficients between these variables, or use the function **PTTEST**$(\cdot,\cdot)$ to compute the Pesaran and Timmermann predictive failure test statistic. To avoid uncertain initial estimates we suggest computing these statistics over the period 1960(1)-1992(12). You need return to the Process window (click), clear the editor and type

SAMPLE $60m1$ $92m12$; **COR** $ERSP$ $ERHAT$;
$STAT =$ **PTTEST**$(ERSP, ERHAT)$; **LIST** $STAT$

The sample correlation between $ERSP$ and $ERHAT$ is 0.2066, and the coefficients of variations computed for these variables suggest that actual returns are 3 times more variable than recursively predicted returns. The Pesaran-Timmermann test statistic is 2.8308, which is well above 1.67, the 95 per cent critical value for a one-sided test. There is clearly significant evidence that monthly excess returns are predictable using *ex ante* dated variables.

[12] Note that the descriptions that follow $ERHAT$ and $ERSE$ are optional.

Table 11.16: Regression of excess returns on Standard and Poor 500 portfolio

```
                        Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is ERSP
 468 observations used for estimation from 1954M1 to 1992M12
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                     -.024013           .0096634            -2.4850[.013]
 YSP(-1)                   14.2719             3.3204             4.2983[.000]
 PI12(-2)                  -.27856            .063504            -4.3865[.000]
 DI11(-1)                -.0068798           .0024969            -2.7554[.006]
 DIP12(-2)                 -.15856            .040177            -3.9465[.000]
*******************************************************************************
 R-Squared                      .086960   R-Bar-Squared                  .079072
 S.E. of Regression             .040716   F-Stat.    F(4,463)    11.0243[.000]
 Mean of Dependent Variable   .0059055   S.D. of Dependent Variable     .042428
 Residual Sum of Squares         .76758   Equation Log-likelihood       836.5758
 Akaike Info. Criterion        831.5758   Schwarz Bayesian Criterion    821.2046
 DW-statistic                    1.9900
*******************************************************************************

                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *         F Version          *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(12) =  10.2137[.597]*F(12,451)    =   .83852[.611]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =   1.4695[.225]*F(1,462)     =   1.4552[.228]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =  70.2152[.000]*       Not applicable        *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =   .20274[.653]*F(1,466)     =   .20196[.653]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

11.12 Lesson 11.12: Rolling regressions and the Lucas critique

This lesson will illustrate the use of rolling regression to examine parameter variation. In an influential paper Lucas (1976) argued that estimated econometric parameters are unlikely to be stable, since as policy regimes change, people will change how they form their expectations, and this will change the estimated decisions rules. The issue is discussed in more detail in Alogoskoufis and Smith (1991b). Consider the simple expectations augmented Phillips Curve

$$\Delta w_t = f(u_t) + \beta E(\Delta p_t \mid \Omega_{t-1}) \tag{11.23}$$

where w_t is the logarithm of money wages, u_t the unemployment rate, p_t the logarithm of a general price index, and Ω_{t-1} the information set at $t-1$. We would expect $\beta = 1$, if workers lacked money illusion. Now suppose the evolution of inflation could be described by a first order autoregression, with time-varying parameters

$$E(\Delta p_t \mid \Omega_{t-1}) = \pi_t(1-\rho_t) + \rho_t \Delta p_{t-1} \tag{11.24}$$

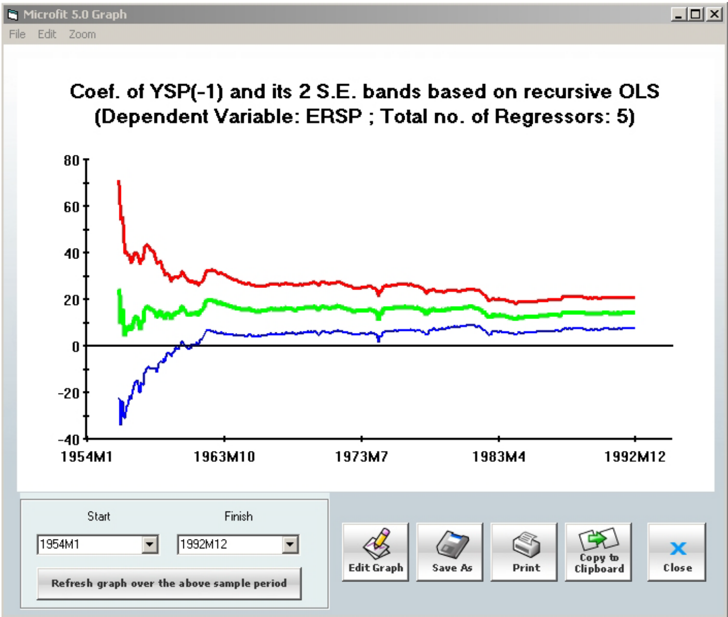


Figure 11.4: Recursive estimates of the coefficients of the dividend yield variable in equation (11.22)

where π_t is the steady-state rate of inflation and ρ_t measures the persistence of inflation. Alogoskoufis and Smith (1991b) estimate the above equations using UK data over the period 1855-1987. Over this 130-year period, with its varying policy regimes - the Gold Standard, the two World Wars, Bretton Woods, and so on - we would not expect either steady state inflation or its persistence to be constant. Substituting, the expectations equation (11.24) into the inflation-augmented Phillips Curve (11.23) gives

$$\Delta w_t = f(u_t) + \beta\pi_t(1-\rho_t) + \beta\rho_t\Delta p_{t-1} \tag{11.25}$$

The coefficient on lagged inflation in the above Phillips Curve should move in line with the coefficient of lagged inflation in the inflation expectations equation (11.24). The two coefficients must be equal when $\beta = 1$. To determine whether this is the case, we need to obtain time-varying parameter estimates of the coefficients in the two equations and compare their movements. We do this using the Rolling Regression Option in *Microfit*.

Read the special *Microfit* file PHILLIPS.FIT, which contains annual observations over the period 1855-1987 on the following variables

E	Log Employment
N	Log Labour Force
P	Log Consumer|Prices
W	Log Earnings
Y	Log Real GDP

In the Commands and Data Transformations box type

$$\begin{aligned} C &= 1; DW = W - W(-1); DP = P - P(-1); \\ U &= N - E; DU = U - U(-1) \end{aligned}$$

to create a constant term in C, the rate of change of money wages in DW, the rate of price inflation in DP, the rate of unemployment in U, and the change in the rate of unemployment in DU. Click the Univariate Menu option, choose Rolling Linear Regression Menu and then Rolling Least Squares. In the box editor that appears on the screen specify the equation as

$$DW \quad C \quad DU \quad U \quad DP(-1)$$

This allows both the level and change in unemployment to influence the rate of growth of wages. Set to 1 the number of observations used for updating estimation and click . Select the whole sample, choose the rolling least squares option, and in the window size field type 25. Whereas recursive regression extends the sample by one each time it re-estimates, rolling regression keeps the sample size the same, here at 25 years. Alogoskoufis and Smith (1991b) keep the sample size fixed because they think that information about the persistence of inflation before World War I is probably not informative about the persistence of inflation after World War II. The program will now estimate the equation over all sub-samples of 25 years. When it has stopped calculating, choose to plot the rolling coefficients and standard errors. Choose $DP(-1)$ to see the plot in Figure 11.5. You will see that the estimated coefficients of $DP(-1)$ were not significantly different from zero on pre-World War I samples, shot upwards after 1914, tended to decline thereafter and started rising towards the end of the period. You can examine the other coefficients as well. Since we wish to compare the coefficient on $DP(-1)$ with that from another equation, click Close to leave the graph, and choose option 3 in the Rolling Regression Results Menu to save the rolling coefficients.

Once again choose the regressor $DP(-1)$ and in response to the prompt name it $PCCOEF$ (Phillips Curve Coefficient). Backtrack and edit the regression

$$DP \quad C \quad DP(-1)$$

Repeat the earlier process and save the OLS rolling coefficients on $DP(-1)$ for this autoregression as $ARCOEF$. Backtrack, clear the Commands and Data Transformations box, then type

PLOT $PCCOEF \quad ARCOEF$

You should see Figure 11.6 on the screen. As the theory suggests, they show quite similar movements, though the match is less good before World War I. The correlation between the coefficients is 0.69.

11.13 Exercises in linear regression analysis

11.13.1 Exercise 11.1

Use the data in the file UKCON.FIT to estimate the equation (11.16) by the IV method. Using 1, $\log y_{t-1}$, $\log y_{t-2}$, $\log c_{t-1}$, $\log c_{t-2}$, Π_{t-1} and Π_{t-2} as instruments, compare these

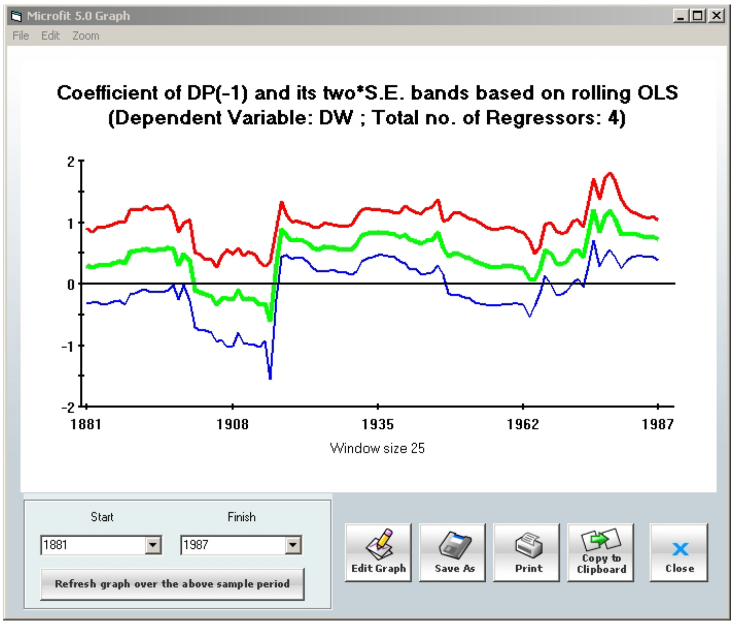


Figure 11.5: Rolling *OLS* estimates of the coefficient of the lagged inflation in the inflation equation

estimates with the corresponding *OLS* estimates. What is the interpretation of Sargan's mis-specification test statistic given in the IV regression result table.

11.13.2 Exercise 11.2

Carry out the non-nested tests in Lesson 11.9 using the UK quarterly consumption data in UKCON.FIT.

11.13.3 Exercise 11.3

Check the sensitivity of the Wald test statistics to the way the non-linear restrictions in Lesson 11.5 are specified by computing the relevant statistics for testing the restrictions $\beta_1 + (\beta_3/\beta_2) = 0$ and $\beta_1 + (\beta_5/\beta_4) = 0$. Notice that when it is known *a priori* that $\beta_2 \neq 0$ and $\beta_4 \neq 0$ the above restrictions are algebraically equivalent to those in Lesson 11.5 (see equation (11.17)).

11.13.4 Exercise 11.4

Carry out the test of the common factor restrictions in Lesson 11.5 on the US data using the file USCON.FIT.

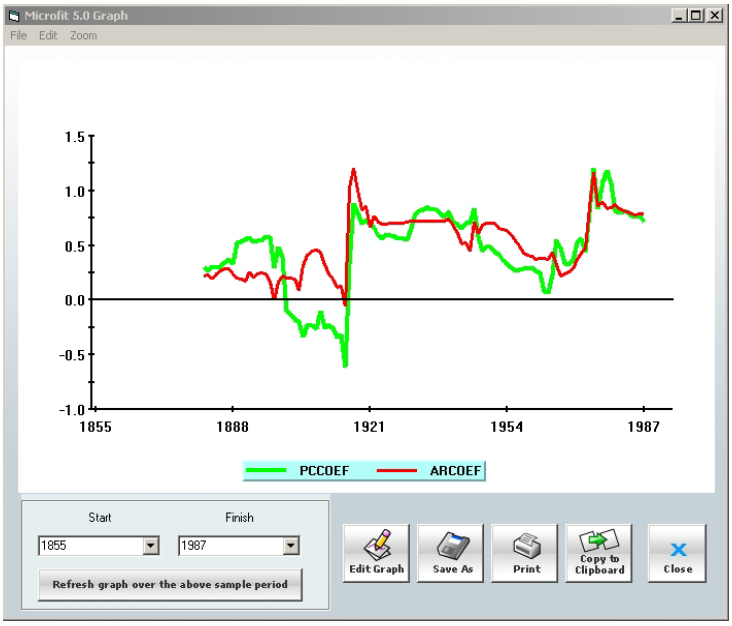


Figure 11.6: Rolling *OLS* coefficients of the lagged inflation rate in the Phillips and inflation equations

11.13.5 Exercise 11.5

Repeat the computations in Lesson 11.11 using the returns on the Dow Jones portfolio instead of the SP500 portfolio.

11.13.6 Exercise 11.6

Carry out the analysis in Lesson 11.12 using a larger window size (say 30), and discuss the robustness of your conclusions in the choice of the observation window.

Chapter 12

Lessons in Univariate Time-Series Analysis

The lessons in this chapter show you how to identify, estimate, and calculate dynamic forecasts using univariate $ARMA(p,q)$, and $ARIMA(p,d,q)$ models. Univariate $ARMA$ and $ARIMA$ processes have been used extensively in the time series literature and are particularly useful for short-term (one-step ahead) forecasting. Following the seminal work of Box and Jenkins (1970), a number of dedicated computer packages for the estimation of $ARIMA$ models have been developed. In *Microfit*, the MA option in the Linear Regression Menu can be readily used to provide estimates of univariate $ARIMA$ models containing up to 12 unknown parameters in their MA part.

The $ARIMA\,(p,d,q)$ model for the variable x is given by

$$y_t = f(t) + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} \tag{12.1}$$

where $y_t = \Delta^d x_t = (1-L)^d x_t$, and $f(t)$ is the deterministic trend in y_t (if any). In most economic applications either $d = 0$ and $f(t) = \alpha + \delta t$, or $d = 1$ and $f(t) = \mu$. The first step in univariate analysis concerns the selection of the orders p, d, and q. In Box-Jenkins' terminology, this is called the identification of the univariate model. The selection of these orders are often carried out in two stages: in the first stage, d, the order of integration of the process is determined using the Augmented Dickey-Fuller (ADF) tests.[1] Once the order of integration of the process is established, the orders of the $ARMA$ process, p, and q, are then selected either by plotting the correlogram of a time-series and comparing it with the theoretical correlogram of a specific time-series model (Box-Jenkins approach) or by specifying a relatively high order $ARMA$ model, as the most general one, and then using Likelihood Ratio tests or one of the popular model selection criteria such as the AIC, or the SBC to select a more parsimonious model (see Section 21.7). The ADF test statistics and plots of the correlogram can be obtained automatically by using commands **ADF** and **COR** in the Process window: see Sections 4.4.2 and 4.4.8. Values of AIC and SBC can

[1] A process is said to be integrated of order d if it must be differenced d times before it is rendered stationary. An integrated process of order d is often denoted by $I(d)$.

be computed by estimating $ARMA$ models of different orders using the MA Option in the Linear Regression Estimation Menu: see Section 6.5. The model with the highest value for the information criteria is selected. Notice that the use of different information criteria can lead to different models. The use of SBC tends to select a more parsimonious model as compared to the use of AIC. See Section 21.7 for more details and relevant references to the literature.

12.1 Lesson 12.1: Using the ADF command to test for unit roots

In this lesson we shall consider the problem of testing for a unit root in the US real GNP using the command **ADF** in the Process window.

The file GDP95.FIT contains quarterly observations over the period 1960(1) to 1995(1) on $USGNP$ (the US Gross National Product at 1987 prices, source: Citibase), and quarterly observations over the period 1955(1) to 1995(1) on $UKGDP$ (the UK Gross Domestic Product at 1990 prices, source: CSO95 Macroeconomic Variables). See also Lessons 10.8 and 10.9.

Use the Open File option in the File Menu on the main menu bar to load the file GDP95.FIT. In the Process window create the log of output, its first difference, constant and time trend by typing in the box editor

$$YUS = \mathbf{LOG}(USGNP);\ \ DYUS = YUS - YUS(-1);$$
$$INPT = 1;\ T = \mathbf{CSUM}(1)$$

Before applying the $ARMA$ methodology to the US output series, it is important to check if it is difference or trend stationary. If it is (trend) stationary, we use the $ARMA$ model for YUS plus a deterministic trend, while if it is difference-stationary of degree d (or integrated of order d, $I(d)$) we should use the autoregressive integrated moving average ($ARIMA$) model for YUS.[2]

Consider the univariate $AR(1)$ process

$$y_t = \alpha + (1-\phi)\delta t + \phi y_{t-1} + \epsilon_t,\ t = 1, ..., n \tag{12.2}$$

where ϵ_t is $iid(0, \sigma^2)$. If $|\phi| < 1$, $\{y_t\}$ is trend stationary, while if $\phi = 1$, $\{y_t\}$ is difference stationary with a non-zero drift, α. Attempts to distinguish the difference stationary process from the trend stationary series has generally taken the form of a (one-sided) test of the null hypothesis of a unit root against the alternative of stationarity; that is,

$$H_0 : \phi = 1 \text{ against } H_1 : \phi < 1$$

It is important to note that when using the t-statistic for testing $\phi = 1$, we should use the critical values of the non-standard Dickey-Fuller unit root distribution rather than the

[2] See Box and Jenkins (1970).

standard normal distribution.[3] In the more general case where the disturbances, $\epsilon_t, t = 1, 2, ...n$ are serially correlated, then we should use the augmented Dickey-Fuller (ADF) unit roots test statistic,[4] which is proposed to accommodate error autocorrelation by adding lagged differences of y_t:[5]

$$y_t = \alpha + (1 - \phi)\delta t + \phi y_{t-1} + \sum_{i=1}^{p+1} \phi_i y_{t-i} + \epsilon_t,\ t = 1, 2, ..., n \tag{12.3}$$

which can also be rewritten as

$$\Delta y_t = \alpha + \rho\delta t - \rho y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \epsilon_t,\ t = 1, 2, ..., n \tag{12.4}$$

where the null is now $H_0 : \rho = 1 - \phi = 0$. When using the ADF tests and interpreting the results, the following points are worth bearing in mind:

- Although ADF has good power characteristics as compared to other unit roots tests in the literature, it is nevertheless not very powerful in finite samples for alternatives $H_1 : \phi = \phi_0 < 1$, when ϕ_0 is near unity.

- There is size-power trade-off depending on the order of augmentation used in dealing with the problem of residual serial correlation.[6] Therefore, it is often crucial that an appropriate value is chosen for p, the order of augmentation of the test.[7]

To test for the unit root in log of $USGNP$, clear the Commands and Data Transformations box and type

ADF $YUS(5)$

---

[3]Critical values have been generated by Monte Carlo simulation for the three cases of (i) $\alpha = 0$, $\delta = 0$, (ii) $\alpha \neq 0$, $\delta = 0$, and (iii) $\alpha \neq 0$, $\delta \neq 0$. See Fuller (1996), Dickey and Fuller (1979) and Hamilton (1994) Chapter 17. For a comprehensive survey of the unit root literature see Stock (1994).

[4]There are other unit root test statistics such as the semi-parametric approach of Phillips and Perron (1988), and the tests of the stationarity hypothesis proposed in Kwiatkowski, Phillips, Schmidt, and Shin (1992). The Phillips-Perron test can be computed in *Microfit* as the ratio of the OLS estimate of ρ in the simple DF regression $\Delta y_t = \alpha + \rho\delta t - \rho y_{t-1} + \epsilon_t$, to its Newey-West standard error obtained using options 4 to 6 in the Standard, White and Newey-West Adjusted Variance Menu (see Section 6.22). The critical values for the Phillips-Perron test are the same as those for the Dickey-Fuller test, and depend on whether the DF regression contains an intercept term or a time trend. See the previous footnote. For an application see Lesson 16.1.

[5]Said and Dickey (1984) show that if the order p is suitably chosen, the ADF test statistics have the same asymptotic distribution as the simple DF statistic under IID errors. If ϵ_t follows the $AR(p)$ process, the number of lagged differences in the ADF regression must be at least as large as p. If ϵ_t has an MA component, the order p must be allowed to increase with the sample size though at a slower rate (at the rate of $n^{1/3}$).

[6]In the case of autocorrelated disturbances, the size distortion of the uncorrected DF test is quite considerable. The $ADF(p)$ test performs better, with a sufficiently high value for p. This happens, however, at the expense of power. ADF tests with p large relative to the sample size have almost no power.

[7]Since we do not know the true order of p, the two-step procedure might be used whereby the model selection criteria such as the Akaike information criterion (AIC) or the Schwarz Bayesian criterion (SBC) is used to select the order of the ADF regression, and the test is then performed.

where we have chosen $p = 5$ as the upper bound for the order of augmentation. Click on the 

button, and, when prompted, check the 'Simulate critical values' checkbox to obtain simulated critical values for the ADF test. Click 

to accept the default number of replications, number of observations to be used in your simulations, and significance level of the test. Then press 

. The results in Table 12.1 should appear on the screen.

Table 12.1: *ADF* unit root tests for variable *YUS*

```
                           ADF tests for variable YUS
          The Dickey-Fuller regressions include no intercept and no trend
*******************************************************************************
 135 observations used in the estimation of all ADF regressions.
 Sample period from 1961Q3 to 1995Q1
*******************************************************************************
        Test Statistic    CV          LL          AIC          SBC          HQC
 DF           9.3906    -1.9795    444.3374     443.3374     441.8847     442.7470
 ADF(1)       5.1652    -2.0251    451.8280     449.8280     446.9228     448.6474
 ADF(2)       3.9752    -2.0255    453.6384     450.6384     446.2805     448.8674
 ADF(3)       3.9958    -1.9949    453.9237     449.9237     444.1132     447.5625
 ADF(4)       3.6174    -1.9984    454.0324     449.0324     441.7693     446.0809
 ADF(5)       3.7266    -1.9466    454.4760     448.4760     439.7602     444.9341
*******************************************************************************
 CV  = 95% simulated critical value using 135 obs. and 1000 replications.
 LL  = Maximized log-likelihood        AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion      HQC = Hannan-Quinn Criterion
                           ADF tests for variable YUS
         The Dickey-Fuller regressions include an intercept but not a trend
*******************************************************************************
 135 observations used in the estimation of all ADF regressions.
 Sample period from 1961Q3 to 1995Q1
*******************************************************************************
        Test Statistic    CV          LL          AIC          SBC          HQC
 DF          -2.4732    -2.9102    448.1691     446.1691     443.2638     444.9885
 ADF(1)      -1.8788    -2.8820    454.0377     451.0377     446.6798     449.2667
 ADF(2)      -1.6595    -2.9898    455.3623     451.3623     445.5517     449.0010
 ADF(3)      -1.7318    -2.9740    455.8193     450.8193     443.5561     447.8678
 ADF(4)      -1.6948    -2.9781    455.8539     449.8539     441.1380     446.3120
 ADF(5)      -1.7756    -2.9162    456.4991     449.4991     439.3306     445.3669
*******************************************************************************
 95% published asymptotic critical value corresponding to ADF(0) =  -2.8828
 CV  = 95% simulated critical value using 135 obs. and 1000 replications.
 LL  = Maximized log-likelihood        AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion      HQC = Hannan-Quinn Criterion

                           ADF tests for variable YUS
        The Dickey-Fuller regressions include an intercept and a linear trend
*******************************************************************************
 135 observations used in the estimation of all ADF regressions.
 Sample period from 1961Q3 to 1995Q1
*******************************************************************************
        Test Statistic    CV          LL          AIC          SBC          HQC
 DF          -2.9715    -3.3908    451.6814     448.6814     444.3235     446.9105
 ADF(1)      -3.0616    -3.4022    458.0269     454.0269     448.2164     451.6657
 ADF(2)      -3.2291    -3.3483    459.9455     454.9455     447.6823     451.9939
 ADF(3)      -3.1084    -3.3884    460.0603     454.0603     445.3444     450.5184
 ADF(4)      -3.1760    -3.4063    460.3419     453.3419     443.1734     449.2097
 ADF(5)      -3.0175    -3.3947    460.5375     452.5375     440.9164     447.8150
*******************************************************************************
 95% published asymptotic critical value corresponding to ADF(0) =  -3.4435
 CV  = 95% simulated critical value using 135 obs. and 1000 replications.
 LL  = Maximized log-likelihood        AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion      HQC = Hannan-Quinn Criterion
```

This table consists of three panels: the upper panel provides results on the *ADF* statistics for models with no intercept and no trends; the middle panel gives the *ADF* statistics for models with an intercept term but no time trends; and the bottom part gives the *ADF* statistics for models with an intercept and a linear deterministic trend. Since $USGNP$ is trended, only the bottom panel of Table 12.1 is relevant. Before carrying out the unit roots tests, the order of the ADF regression defined by (12.4) needs to be selected. The values reported in Table 12.1 for the different model selection criteria suggest that the correct order is likely to be between 1 and 2, with the SBC selecting the lower order. The ADF statistics for $p = 1$ and 2 are equal to -3.0616 and -3.2291, respectively, which are both (in absolute value) below their (asymptotic) 95 per cent critical value given at the foot of the table (-3.4435). The ADF statistics for $p = 1$ and 2 are also below (in absolute value) the 95 per cent of the simulated critical values, given by the column headed CV. The same is also true if we consider higher values of p. It is therefore not possible to reject the null of a unit root ($H_0 : \phi = 1$) in the log of $USGNP$ at the 5 per cent significance level.

Similar conclusions can be reached using other unit roots tests provided by *Microfit*, such as the GLS-ADF, WS-ADF, or the PP-ADF test statistics (see Sections 4.4.3 to 4.4.5 and 4.4.10). For example, to apply the **GLS_ADF** procedure to the log of $USGNP$ series you need to enter the following command:

ADF_GLS $USGNP$ RUN

As another example, suppose you wish to test for a unit root in $DYUS$ (the growth of $USGNP$). Given the quarterly nature of the series it is advisable to use a maximum lag of at least 5. To this end you need to enter

ADF $DYUS(5)$ RUN

Since the output growth is not trended, the relevant ADF statistics are given in the second part of Table 12.2. The model selection criteria suggest selecting either $p = 0$ or 1. However, irrespective of the order of the augmentation chosen for the ADF test the absolute values of the ADF statistics are all well above the 95 per cent critical value of the test given at the foot of the table (-2.8830) and above the corresponding simulated critical values, and hence the hypothesis that growth rate of $USGNP$ has a unit root is firmly rejected.

Table 12.2: ADF unit root tests for the variable $DYUS$

ADF tests for variable DYUS

The Dickey-Fuller regressions include no intercept and no trend

134 observations used in the estimation of all ADF regressions.
Sample period from 1961Q4 to 1995Q1

	Test Statistic	CV	LL	AIC	SBC	HQC
DF	-5.9272	-1.9652	435.9639	434.9639	433.5150	434.3751
ADF(1)	-3.8875	-2.0114	442.4476	440.4476	437.5498	439.2700
ADF(2)	-3.4819	-2.0093	442.6603	439.6603	435.3136	437.8939
ADF(3)	-2.8621	-2.0583	444.0466	440.0466	434.2510	437.6915
ADF(4)	-2.7113	-2.0085	444.0674	439.0674	431.8228	436.1234
ADF(5)	-2.2270	-1.9853	445.5868	439.5868	430.8933	436.0541

CV = 95% simulated critical value using 134 obs. and 1000 replications.
LL = Maximized log-likelihood AIC = Akaike Information Criterion
SBC = Schwarz Bayesian Criterion HQC = Hannan-Quinn Criterion

ADF tests for variable DYUS

The Dickey-Fuller regressions include an intercept but not a trend

134 observations used in the estimation of all ADF regressions.
Sample period from 1961Q4 to 1995Q1

	Test Statistic	CV	LL	AIC	SBC	HQC
DF	-8.3346	-2.8228	448.5896	446.5896	443.6918	445.4120
ADF(1)	-5.7556	-2.8694	450.2928	447.2928	442.9460	445.5264
ADF(2)	-5.4562	-2.8888	450.5486	446.5486	440.7529	444.1934
ADF(3)	-4.7231	-2.8879	450.6403	445.6403	438.3957	442.6964
ADF(4)	-4.7182	-2.9785	451.0983	445.0983	436.4047	441.5655
ADF(5)	-4.0889	-2.9692	451.3262	444.3262	434.1838	440.2047

95% published asymptotic critical value corresponding to ADF(0) = -2.8830
CV = 95% simulated critical value using 134 obs. and 1000 replications.
LL = Maximized log-likelihood AIC = Akaike Information Criterion
SBC = Schwarz Bayesian Criterion HQC = Hannan-Quinn Criterion

ADF tests for variable DYUS

The Dickey-Fuller regressions include an intercept and a linear trend

134 observations used in the estimation of all ADF regressions.
Sample period from 1961Q4 to 1995Q1

	Test Statistic	CV	LL	AIC	SBC	HQC
DF	-8.4806	-3.4006	449.5871	446.5871	442.2403	444.8207
ADF(1)	-5.8735	-3.3768	450.9750	446.9750	441.1794	444.6199
ADF(2)	-5.6057	-3.3910	451.3615	446.3615	439.1169	443.4175
ADF(3)	-4.8797	-3.3672	451.4015	445.4015	436.7080	441.8687
ADF(4)	-4.9100	-3.4278	452.0155	445.0155	434.8730	440.8939
ADF(5)	-4.2808	-3.4232	452.1493	444.1493	432.5579	439.4389

95% published asymptotic critical value corresponding to ADF(0) = -3.4437
CV = 95% simulated critical value using 134 obs. and 1000 replications.
LL = Maximized log-likelihood AIC = Akaike Information Criterion
SBC = Schwarz Bayesian Criterion HQC = Hannan-Quinn Criterion

The model selection criteria in the second part of Table 12.2 also suggest that the process of US output growth can be approximated by a low-order AR process, with the AIC selecting an $AR(1)$ process and the SBC selecting an $AR(0)$ process for the US output growth. Estimating an $AR(1)$ process for US output growth over the period 1960(3)-1995(1) gives

$$\begin{array}{rccccl} DYUS_t = & \underset{(0.000927)}{0.00498} & + & \underset{(0.0806)}{0.31881} & DYUS_{t-1} + \hat{u}_t \end{array}$$

thus suggesting that $\log(USGNP_t)$ is an $I(1)$ process with a non-zero drift, estimated as $0.00498/(1-0.31881) = 0.00732(0.0011)$, representing the quarterly average rate of growth of the US economy. The figure in parentheses is the asymptotic standard error of the average growth rate and can be obtained using option 5 in the Post Regression Menu, once the OLS estimation of the $AR(1)$ process in $DYUS$ is completed.

12.2 Lesson 12.2: Spectral analysis of US output growth

The spectral approach enables us to investigate the properties of time-series in the frequency domain. Let $\{x_t,\ t = -\infty, ..., \infty\}$ be a univariate covariance stationary process with mean $E(x_t) = \mu$ and the kth autocovariance function

$$E(x_t - \mu)(x_{t-k} - \mu) = \gamma_k = \gamma_{-k}, \quad k = 0, 1, 2, ...$$

The main goal of the spectral analysis is to determine the importance of cycles of different frequencies in accounting for the behaviour of x_t.[8] Assuming that autocovariances are absolutely summable ($\sum_{k=0}^{\infty} \gamma_k$ is finite), the population spectrum can be written as[9]

$$f(\omega) = \frac{1}{\pi}\left\{\gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k \cos(\omega k)\right\}, \qquad 0 \leq \omega < \pi$$

- If x_t is a white noise process ($\gamma_0 = \sigma^2$ and $\gamma_k = 0$ for $k \neq 0$), then $f(\omega)$ is flat at σ^2/π for all $\omega \in [0, \pi]$.
- If x_t is a stationary $AR(1)$ process, $x_t = \mu + \phi x_{t-1} + \epsilon_t$ with $|\phi| < 1$ and ϵ_t being a white noise process, then $f(\omega)$ is monotonically decreasing in ω for $\phi > 0$, and a monotonically increasing function of ω for $\phi < 0$.
- If x_t is a stationary $MA(1)$ process, $x_t = \mu + \epsilon_t + \theta\epsilon_{t-1}$ with $|\theta| < 1$ and ϵ_t being a white noise process, then $f(\omega)$ is monotonically decreasing (increasing) in ω for $\theta > 0$ (for $\theta < 0$).

The sample spectral density function (or the sample periodogram) can be estimated by

$$\hat{f}(\omega) = \frac{1}{\pi}\left\{\hat{\gamma}_0 + 2\sum_{k=1}^{\infty}\hat{\gamma}_k \cos(\omega k)\right\}, 0 \leq \omega < \pi$$

[8] Any covariance stationary process has both a time-domain and a frequency-domain representation, and any feature of the data that can be described by one representation can equally be described by the other. For an introductory account of the spectral analysis see Chatfield (2003). A more advanced treatment can be found in Priestley (1981).

[9] Given the population spectrum, $f(\omega)$, then the j-th autocovariance of the covariance stationary process x_t can be written as

$$\int_{-\pi}^{\pi} f(\omega)\exp(i\omega j)d\omega = \gamma_j,\ j = 0, 1, ...$$

For example, in the special case when $j = 0$, the variance of x_t can be obtained by

$$\int_{-\pi}^{\pi} f(\omega)d\omega = \gamma_0.$$

In general, $\int_{-\omega_1}^{\omega_1} f(\omega)d\omega = 2\int_0^{\omega_1} f(\omega)d\omega$ represents the portion of the variance of x_t that is attributable to the periodic random components with frequency less than or equal to ω_1.

where $\hat{\gamma}_k$ is the sample autocovariances obtained by

$$\hat{\gamma}_k = n^{-1} \sum_{t=k+1}^{\infty} (x_t - \bar{x})(x_{t-k} - \bar{x}), \text{ for } k = 0, 1, ..., n-1$$

and $\bar{x}$ is the sample mean. However, one serious limitation of the use of the sample periodogram is that the estimator $\hat{f}(\omega)$ is not consistent; that is it is not becoming more accurate even as the sample size n increases. This is because in estimating it we have made use of as many parameter estimates ($\hat{\gamma}_k$, for $k = 0, 1, ..., n-1$) as we had observations (x_1, ..., x_n). Alternatively, the population spectrum can be estimated non-parametrically by use of *kernel* estimates given by

$$\hat{f}(\omega_j) = \sum_{i=-m}^{m} \lambda(\omega_{j+i}, \omega_j) \hat{f}(\omega_{j+i})$$

where $\omega_j = j\pi/m$, m is a *bandwidth* parameter[10] indicating how many frequencies $\{\omega_j, \omega_{j\pm1}, ..., \omega_{j\pm m}\}$ are used in estimating the population spectrum, and the *kernel* $\lambda(\omega_{j+i}, \omega_j)$ indicates how much weight each frequency is to be given, where $\sum_{i=-m}^{m} \lambda(\omega_{j+i}, \omega_j) = 1$. Specification of *kernel* $\lambda(\omega_{j+i}, \omega_j)$ can equivalently be described in terms of a weighting sequence $\{\lambda_j,\ j = 1,\ ...,\ m\}$, so that

$$\hat{f}(\omega_j) = \frac{1}{\pi} \left\{ \hat{\gamma}_0 + 2 \sum_{k=1}^{m} \lambda_k \hat{\gamma}_k \cos(\omega_j k) \right\} \tag{12.5}$$

Microfit computes a scaled and standardized version of $\hat{f}(\omega_j)$ by multiplying it by $\pi/\hat{\gamma}_0$, and gives

$$\hat{f}_*(\omega_j) = 1 + 2 \sum_{k=1}^{m} \lambda_k (\hat{\gamma}_k/\hat{\gamma}_0) \cos(\omega_j k) \tag{12.6}$$

and their estimated standard errors using Bartlett, Tukey and Parzen lag windows at the frequencies $\omega_j = j\pi/m,\ j = 0, 1, 2, ..., m$ (see Section 21.3). Each of these frequencies are associated with the *period* $= 2\pi/\omega_j = \frac{2m}{j},\ j = 0, 1, 2, ..., m$. For example, at zero frequency the periodicity of the series is infinity. The value of the standardized spectrum at zero frequency refers to the long-run properties of the series. The higher this value, the more persistent are the effects of deviations of x_t from its trend. The spectrum of a unit root process at zero frequency is unbounded.

In this lesson we carry out the spectrum analysis of the detrended and first-differenced log of $USGNP$. The unit-root analysis of YUS (the logarithm of $USGNP$) in the previous lesson suggests that we should first-difference the output series before carrying out any spectral analysis of it. Spectral analysis of trended or non-stationary processes can be very

[10] One important problem is the choice of the bandwidth parameter, m. One practical guide is to plot an estimate of the spectrum using several different bandwidths and rely on subjective judgement to choose the bandwidth that produces the most plausible estimate. Another possibility often recommended in practice is to set $m = 2\sqrt{n}$. This is the default value chosen by *Microfit*. For more formal statistical procedures see, for example, Andrews (1991), and Andrews and Monahan (1992).

misleading. To illustrate this we first consider the spectral analysis of output series, when it is de-trended using a simple regression of YUS on a linear trend.

Load the file GDP95.FIT, create the variables YUS and so on, as described in Lesson 12.1). In the box editor, type

$$YUS \quad INPT \quad T$$ 

which is a regression of YUS on a constant term ($INPT$) and a time trend (T). When the results table appears click 

, and in the Post Regression Menu choose option 3 (List/Plot/Save residuals), and then option 6. When prompted, type the name of the residual and its description as

$RESYUS$ Residual obtained from the regression of YUS on a linear trend

Then, click 

as many times as required to move back to the Commands and Data Transformations box, go to the Process window, clear it, and type

SPECTRUM $RESYUS$ 

Since we have not specified any window size after the variable, $RESYUS$, the program uses the default value of $2\sqrt{n} = 2\sqrt{144} = 24$ for the window (bandwidth) size (notice that observations on US output are available only over the period 1960(1)-1995(1), and hence $n = 144$). You will be presented with three alternative estimates of the $SPECTRUM$ followed by their plots. To save the graphs, click the 

button, and select an appropriate option (see Section 5.2 on details of how to print/save/retrieve graphs). The plots of the estimated standardized spectral density function of $RESYUS$ using the Bartlett, Tukey and Parzen lag windows are shown in Figure 12.1.

One prominent feature of this plot is that the contribution to the sample variance of the lowest-frequency component is much larger than the contributions of other frequencies (for example at business cycle frequencies). This is due to the non-stationary nature of the detrended series, $RESYUS$. In general, if the process is non-stationary (for example if it has a unit root), then its spectral density becomes dominated by the value of the spectrum at the zero frequency, and drops dramatically immediately thereafter, thus hiding possible peaks at higher (business cycle) frequencies.

To avoid this problem one could either try other detrending procedures such as the Hodrick-Prescott method described in Sections 4.3.7 and 21.4 and Lesson 10.8, or apply the **SPECTRUM** command to $DYUS$, the first-difference of output series (in logarithms). Here we do the latter.

Move to the Commands and Data Transformations box, clear it, and type

SPECTRUM $DYUS(24)$ 

The results in Table 12.3 should now appear on the screen, followed by plots of the $SPECTRUM$ for three different windows: see Figure 12.2.

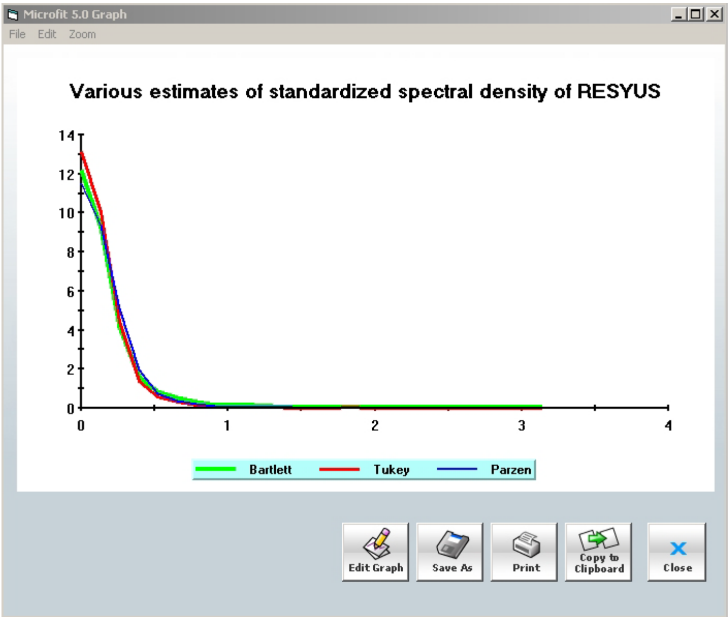


Figure 12.1: Alternative estimates of the spectrum of deviations of US output (in logs) from a linear trend

Table 12.3: Spectral density functions of $DYUS$ for the period 1960(1) to 1995(1)

Standardized spectral density functions of DYUS, sample 1960Q2 to 1995Q1
Estimated asymptotic standard errors in brackets

Frequency	Period	Bartlett	Tukey	Parzen
0.00	*NONE*	1.4186	1.4387	1.7345
		(.67822)	(.72957)	(.74570)
.13090	48.0000	1.9744	2.0362	2.0490
		(.66747)	(.73012)	(.62288)
.26180	24.0000	2.6527	2.6935	2.4019
		(.89677)	(.96580)	(.73017)
.39270	16.0000	2.0804	2.1879	2.1714
		(.70332)	(.78452)	(.66009)
.52360	12.0000	1.5628	1.6517	1.7797
		(.52834)	(.59224)	(.54104)
.65450	9.6000	1.6946	1.6545	1.5514
		(.57286)	(.59325)	(.47163)
.78540	8.0000	1.2296	1.2257	1.2044
		(.41569)	(.43951)	(.36614)
.91630	6.8571	.73965	.70097	.88047
		(.25005)	(.25135)	(.26766)
1.0472	6.0000	.88126	.89154	.94651
		(.29792)	(.31968)	(.28774)
1.1781	5.3333	1.3180	1.2947	1.1098
		(.44557)	(.46424)	(.33737)
1.3090	4.8000	.93593	.95044	.88559
		(.31640)	(.34080)	(.26922)
1.4399	4.3636	.39230	.35453	.47872
		(.13262)	(.12712)	(.14553)
1.5708	4.0000	.32515	.24582	.35209
		(.10992)	(.088145)	(.10704)
1.7017	3.6923	.54794	.53500	.54209
		(.18524)	(.19183)	(.16479)
1.8326	3.4286	.86221	.85046	.75330
		(.29148)	(.30495)	(.22900)
1.9635	3.2000	.76574	.78221	.73309
		(.25887)	(.28048)	(.22286)
2.0944	3.0000	.53634	.52077	.56553
		(.18131)	(.18673)	(.17192)
2.2253	2.8235	.45729	.44015	.47264
		(.15459)	(.15782)	(.14368)
2.3562	2.6667	.53392	.48736	.45624
		(.18050)	(.17475)	(.13870)
2.4871	2.5263	.43472	.39481	.42347
		(.14696)	(.14157)	(.12873)
2.6180	2.4000	.41271	.40593	.47895
		(.13952)	(.14556)	(.14560)
2.7489	2.2857	.74051	.71057	.67058
		(.25034)	(.25479)	(.20386)
2.8798	2.1818	.83469	.84157	.82310
		(.28218)	(.30176)	(.25022)
3.0107	2.0870	.85445	.90842	.91736
		(.28886)	(.32573)	(.27888)
3.1416	2.0000	1.0467	1.0316	.97073
		(.50040)	(.52313)	(.41733)

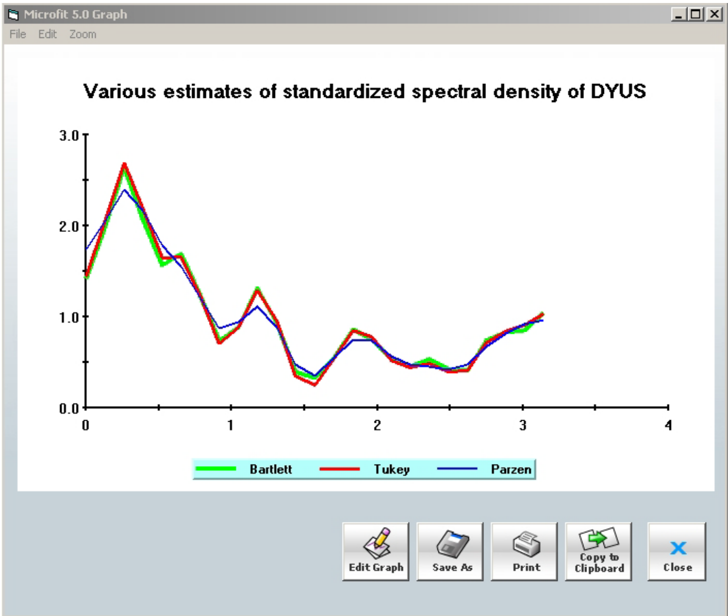


Figure 12.2: Alternative estimates of the spectrum of US output growth

In the case of output growths, we find that the spectral density function of $DYUS$, as compared to that of $RESYUS$, is relatively flat, but decreases over the entire frequency. This may imply that $DYUS$ is a stationary process with a positive autocorrelation coefficient.

The value of the scaled spectrum at zero frequency measures the long-run variance of $DYUS$, and is proposed by Cochrane (1988) as a measure of persistence of shocks to real output. In the case of the US output growth it is estimated to be 1.4186 (0.6782), 1.4387 (0.7296), and 1.7345 (0.7457), for the Bartlett, Tukey and Parzen windows, respectively. See the first row of Table 12.3, and Lesson 12.4.

If the process is over-differenced, then its spectral density at zero frequency becomes zero. For example, consider the spectrum of the second-difference of YUS. Clear the Commands and Data Transformations box and type

$$DDYUS = DYUS - DYUS(-1);\ \textbf{SPECTRUM}\ DDYUS(20)$$

RUN

The plots of the estimated standardized spectral density function of $DDYUS$ are given in Figure 12.3, from which we find that $DDYUS$ is in fact over-differenced, since its spectral density at zero frequency is very close to zero.

The spectrum analysis in general supports our finding in Lesson 12.1 that YUS is an $I(1)$ process.

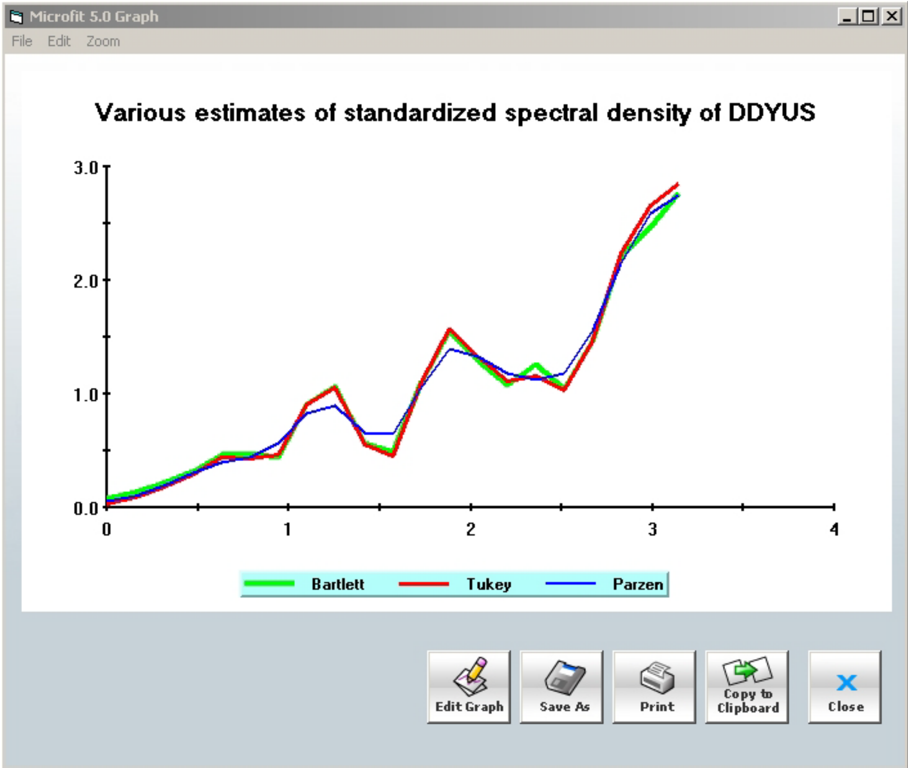


Figure 12.3: Alternative estimates of the spectrum of changes in US output growth

12.3 Lesson 12.3: Using an ARMA model for forecasting US output growth

In the previous two lessons we found that YUS (the logarithm of US GNP) is an integrated process of order 1 but $DYUS$ (the growth of US GNP) is a persistent stationary process. In this lesson we consider the problem of selecting the orders p and q for an $ARIMA(p, 1, q)$ model of YUS, or the $ARMA(p, q)$ model of $DYUS$:

$$\phi(L)DYUS_t = \mu + \theta(L)\epsilon_t \tag{12.7}$$

where $\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$ and $\theta(L) = 1 + \sum_{i=1}^{q} \theta_i L^i$. We assume that (12.7) satisfies the necessary stability and invertibility conditions.[11] To select the orders of p and q we first set them equal to a maximum value of 3. Since we lose 4 observations at the start of the sample when p and q take the maximum value of 3, we must estimate all the 16 $ARMA(p, q)$, $p, q = 0, 1, 2, 3$ models over the same sample period, namely 1961(1)-1993(4) (132 observations). We are keeping the 5 observations over the period 1994(1)-1995(1) for forecasting. We show how the $ARMA(1, 1)$ model is estimated and then report the results

[11] That is, the roots of $1 - \phi_1 z - \cdots - \phi_p z^p = 0$ and $1 + \theta_1 z + \cdots + \theta_q z^q = 0$ lie outside the unit circle.

for the other models. Choose option 1 Linear Regression Menu from the Single Equation Menu (the Univariate Menu), selecting option 7 MA errors. In the editor, type

$$DYUS \quad INPT \quad DYUS(-1)$$

which implies that we specify the AR order to be 1, and then specify sample period as

$$1961Q1 \quad 1993Q4$$

Click 

and, when prompted, type 1, which means that we specify the MA order to be 1. Next, we see that there are two choices of initial estimates. When there is a convergence failure, use option 3 to look at the plot of the concentrated log-likelihood function (which shows that the maximum is found when the $MA(1)$ parameter is around -0.20), and decide the initial estimate, accordingly. For example, using 0.0 as the initial estimate, we obtain the estimation results for the $ARMA(1,1)$ specification in Table 12.4. The values of the Akaike information criterion (AIC) and Schwarz Bayesian criterion (SBC) for this model are 437.92 and 433.60, respectively.

Table 12.4: An $ARMA(1,1)$ model for US output growth

```
                     Exact Maximum Likelihood Estimation Method
                Error TERM : MA(1) converged after 5 iterations
*******************************************************************************
 Dependent variable is DYUS
 132 observations used for estimation from 1961Q1 to 1993Q4
*******************************************************************************
 Regressor                Coefficient       Standard Error         T-Ratio[Prob]
 INPT                       .0036509           .0014541             2.5108[.013]
 DYUS(-1)                     .51107             .18049             2.8317[.005]
*******************************************************************************
 R-Squared                        .10502   R-Bar-Squared                 .091142
 S.E. of Regression             .0086714   F-Stat.     F(2,129)      7.5684[.001]
 Mean of Dependent Variable     .0073650   S.D. of Dependent Variable   .0090958
 Residual Sum of Squares        .0096999   Equation Log-likelihood      440.9237
 Akaike Info. Criterion         437.9237   Schwarz Bayesian Criterion   433.5995
 DW-statistic                     2.0294
*******************************************************************************

            Parameters of the Moving Average Error Specification
*******************************************************************************
 U=E+    -.21465*E(-1)
     (   -1.1184)[.265]
 T-ratio(s) based on asymptotic standard errors in brackets
*******************************************************************************
```

Repeat this procedure to estimate the other $ARMA(p,q)$ models for, $p,q = 0,1,2,3$. Notice that for $ARMA(p)$ models (i.e., when q is set to zero) you should use the OLS option 1 from the Univariate Menu. Then comparing the values of the AIC and\or the SBC, select the model specification with the highest value. Use as initial values 0.0, 0.0 (for the MA lag1 and 2, respectively) for the $ARMA(2,2)$, and 0.0, 0.1 for the $ARMA(3,2)$ while use the initial values provided by the program for all remaining combinations of p and q. When using AIC, we have

$p\backslash q$	0	1	2	3
0	432.59	436.47	438.87	438.02
1	438.35	437.92	438.08	437.10
2	438.76	438.21	437.22	436.25
3	438.13	437.29	436.30	435.35

and therefore, the $ARMA(0, 2)$ specification for $DYUS$ is selected. When using SBC, we have

$p\backslash q$	0	1	2	3
0	431.15	433.59	434.55	432.25
1	435.46	433.60	432.32	429.89
2	434.43	432.44	430.02	427.60
3	432.37	430.08	427.65	425.26

and $ARMA(1, 0)$ is selected. In what follows we base the forecasts on the model selected by SBC, namely the $ARMA(1, 0)$ or simply the $AR(1)$ specification.

To compute the forecasts for the growth rate of the US GNP for the period 1994(1)-1995(1), now set the sample period to

$$1961Q4 \ 1993Q4$$

and use the $AR(1)$ specification for $DYUS$ to estimate the parameters. In the Post Regression Menu, choose option 8 (Forecast) and then press the  button to compute dynamic forecasts of the US output growth over the period 1994(1)-1995(1). The results in Table 12.5 should appear on the screen.

Table 12.5: Forecasts of US output growth based on an $AR(1)$ model

```
                         Single Equation Dynamic Forecasts
*******************************************************************************
 Based on  OLS regression of DYUS on:
 INPT             DYUS(-1)
 132 observations used for estimation from 1961Q1 to 1993Q4
*******************************************************************************
 Observation          Actual          Prediction          Error        S.D. of Error
  1994Q1             .0083959          .0093314        -.9355E-3         .0087240
  1994Q2             .0090418          .0080202         .0010216         .0091424
  1994Q3             .0092593          .0076139         .0016454         .0091837
  1994Q4              .010412          .0074880         .0029242         .0091888
  1995Q1             .0077799          .0074490         .3309E-3         .0091897
*******************************************************************************

              Summary statistics for single equation dynamic forecasts
*******************************************************************************
 Based on 5 observations from 1994Q1 to 1995Q1
 Mean Prediction Errors       .9973E-3    Mean Sum Abs Pred Errors     .0013715
 Sum Squares Pred Errors      .2657E-5    Root Mean Sumsq Pred Errors  .0016301
 Predictive failure test    F(5,130)   =  .027714[1.00]
 Structural stability test F(2,133)    =  .050736[.951]
*******************************************************************************
```

The forecasts are very close to the actual values for the first two quarters, and then settle down to around 0.0074, which is only slightly above the average quarterly rate of growth of US real GNP. Option 9 also provides a plot of actual and forecast values, which are presented in Figure 12.4.

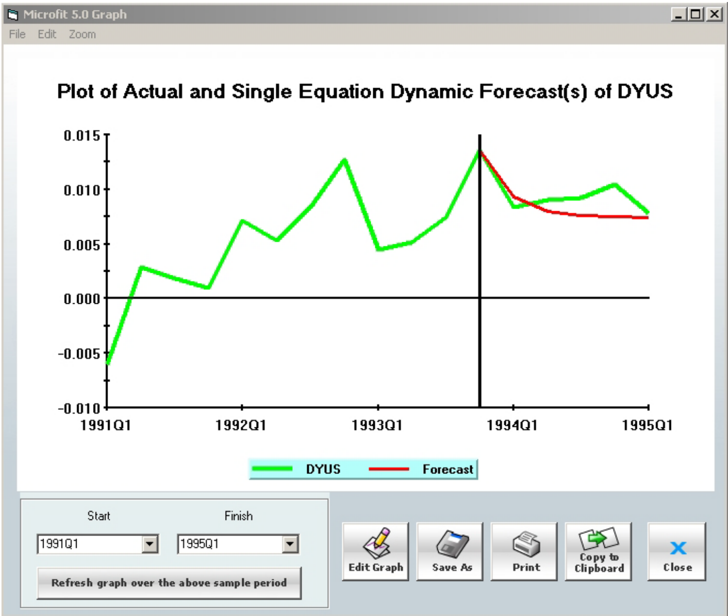


Figure 12.4: Dynamic forecasts of US output growth over the period 1994(1)-1995(1)

12.4 Lesson 12.4: Alternative measures of persistence of shocks to US real GNP

One of the important features of unit root processes lies in the fact that the effect of shocks on these series (or random deviations from their trend) do not die out. In the case of random walk models the long-run impact of the shocks is unity. But for more general $I(1)$ processes this long-run impact could be more or less than unity. The most satisfactory overall measure of persistence is the value of the spectral density of the first-differences of the series evaluated at zero frequency, and then appropriately scaled. In the case of the $ARIMA(p, 1, q)$ process:

$$\Delta x_t = \mu + \phi_1 \Delta x_{t-1} + \cdots + \phi_p \Delta x_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} \tag{12.8}$$

where $\epsilon_t \sim IID(0, \sigma^2)$, the spectral density of Δx_t at zero frequency, is given by

$$f_{\Delta x}(0) = \frac{\sigma^2}{\pi} \left(\frac{1 + \theta_1 + \theta_2 + \cdots + \theta_q}{1 - \phi_1 - \phi_2 - \cdots - \phi_q} \right)^2$$

The measure proposed by Campbell and Mankiw (1987) is given by

$$P_{cm} = \left\{ \frac{\pi f_{\Delta x}(0)}{\sigma^2} \right\}^{\frac{1}{2}} = \frac{1 + \theta_1 + \theta_2 + \cdots + \theta_q}{1 - \phi_1 - \phi_2 - \cdots - \phi_q} \tag{12.9}$$

Cochrane (1988) suggests scaling $f_{\Delta x}(0)$ by the unconditional variance of Δx_t, namely

$$P_c = \left\{ \frac{\pi f_{\Delta x}(0)}{V(\Delta x_t)} \right\}^{\frac{1}{2}} = \left\{ \frac{\sigma^2}{V(\Delta x_t)} \right\}^{\frac{1}{2}} P_{cm} \tag{12.10}$$

Notice, however, that $\sigma^2/V(\Delta x_t) = 1 - R^2$, where R^2 is the squared multiple correlation coefficient of the $ARIMA$ model (12.8). Hence

$$P_c = \left(1 - R^2\right)^{\frac{1}{2}} P_{cm} \tag{12.11}$$

In practice $f_{\Delta x}(0)$ can be estimated by using either the non-parametric approach of the spectral analysis (see Lesson 12.2), or by first estimating the $ARIMA$ model (12.8) and then computing P_{cm} by replacing the unknown parameters $\theta_1, \theta_2, ..., \theta_q$, $\phi_1, \phi_2, ..., \phi_q$, in (12.9) by their ML estimates. Cochrane (1988) favours the former, while Campbell and Mankiw (1987) employ the latter approach. The two estimates can differ a great deal in practice. This is primarily because the estimates obtained using the $ARIMA$ specification are conditional on the orders p and q being *correctly* selected. But the estimates based on the spectral density estimation are less rigidly tied up to a given parametric model, and hence are often much less precisely estimated. In the case of both estimates it is important that their standard errors are also computed.

Here we estimate the two persistent measures, P_{cm}, and P_c, for the US real GNP both using the spectral density and $ARIMA$ modelling approaches. We use the data in the special *Microfit* file GDP95.FIT to estimate the following $ARIMA(0, 1, 2)$ process for y_t (the logarithm of the US real GNP) over the period 1960(1)-1995(1):[12,13]

$$\Delta y_t = \mu + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \epsilon_t$$

Load the file GDP95.FIT into *Microfit*. In the Commands and Data Transformations window create the variables

$$INPT = 1; \quad DYUS = \mathbf{LOG}(USGNP/USGNP(-1))$$

Then choose option 1 in the Single Equation Estimation Menu (Univariate Menu), and select option 7 MA Errors. Enter the start and end dates of 1960Q1 and 1995Q1, then type

$$DYUS \quad INPT$$

[12] Time series observations on the US real GNP are analyzed extensively in the recent literature on the measurement of the persistence of shocks to the US aggregate output. See, for example, Campbell and Mankiw (1987), Stock and Watson (1988), Evans (1989), and Pesaran, Pierce, and Lee (1993). The present data set extends the data used in these studies.

[13] The choice of the ARIMA(0,1,2) for y_t is based on the Akaike information criterion. See Lesson 12.3.

When prompted type 1, then 2 and click the RUN button and then OK, to obtain the results in Table 12.6.

Table 12.6: A $MA(2)$ model for US output growth

```
                     Exact Maximum Likelihood Estimation Method
                Error TERM : MA(2) converged after 5 iterations
*******************************************************************************
 Dependent variable is DYUS
 140 observations used for estimation from 1960Q2 to 1995Q1
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                     .0071916          .0010445              6.8851[.000]
*******************************************************************************
 R-Squared                      .12172   R-Bar-Squared                    .10889
 S.E. of Regression           .0084593   F-Stat.    F(2,137)      9.4930[.000]
 Mean of Dependent Variable   .0072140   S.D. of Dependent Variable    .0089613
 Residual Sum of Squares      .0098037   Equation Log-likelihood       470.9812
 Akaike Info. Criterion       467.9812   Schwarz Bayesian Criterion    463.5687
 DW-statistic                   1.9677
*******************************************************************************

             Parameters of the Moving Average Error Specification
*******************************************************************************
 U=E+     .27280*E(-1)+     .20932*E(-2)
     (    3.2601)[.001] (    2.8332)[.005]
 T-ratio(s) based on asymptotic standard errors in brackets
*******************************************************************************
```

Both of the moving average coefficients, $\hat{\theta}_1 = 0.2728\,(3.26)$ and $\hat{\theta}_2 = 0.2093\,(2.83)$ are statistically significant. The figures in () are asymptotic t-ratios. The Campbell and Mankiw (1987) measure of persistence, P_{cm}, for the $ARIMA(0,1,2)$ specification is given by

$$P_{cm} = 1 + \theta_1 + \theta_2$$

An estimate of P_{cm} together with its (asymptotic) standard error can be readily computed by first choosing option 5 in the Post Regression Menu, and typing

$$PCM = (1 + B1 + B2)$$ RUN

to obtain the results shown in Table 12.7.

Table 12.7: Estimates of Campbell and Mankiw persistence measure for US output growth based on an $MA(2)$ specification

```
                   Analysis of Function(s) of Parameter(s)
*******************************************************************************
 Based on stochastic initial value(s)  MA(2) regression of DYUS on:
 INPT
 140 observations used for estimation from 1960Q2 to 1995Q1
*******************************************************************************
 Coefficients A1 to A1 are assigned to the above regressors respectively.
 Coeffs. B1 to B2 are assigned to MA parameters respectively.
 List of specified functional relationship(s):
 PCM=(1+B1+B2)
*******************************************************************************
 Function                  Estimate       Standard Error         T-Ratio[Prob]
 PCM                        1.4821             .11730            12.6348[.000]
*******************************************************************************

       Estimated Variance Matrix of the Function(s) of the Parameters

*******************************************************************************
                PCM
 PCM          .013760

*******************************************************************************
```

This yields $\hat{P}_{cm} = 1.4821(0.1173)$, where the bracketed figure is the asymptotic standard error of $\hat{P}_{cm}$. This estimate is smaller than the one obtained by Campbell and Mankiw (1987), and it is still significantly larger than unity, which is the persistence measure for a pure random walk model. The t-statistic for this latter test is computed as

$$t_{P_{cm}=1} = \left(\hat{P}_{cm} - 1\right) / \sqrt{\widehat{V}(\hat{P}_{cm})} = \frac{1.4821 - 1.0}{0.1173} = 4.11$$

Consider now the measure of persistence proposed by Cochrane (1988) in (12.10). The non-parametric estimate of this measure is given by the standardized spectral density function of $DYUS$ at zero frequency. To obtain this estimate move to the Commands and Data Transformation box in the Process window and type

SPECTRUM $DYUS$ RUN

You should see the estimates of P_c^2 using the Bartlett, Tukey and Parzen windows in the first row of the result table that appears on the screen, namely 1.4186(0.6782), 1.4387(0.7296) and 1.7345(0.7457), respectively. To make these estimates comparable to the estimates for the Campbell and Mankiw measure, P_{cm}, we need to divide their square root by $\left(1 - R^2\right)^{\frac{1}{2}}$, where R^2 is estimated to be 0.12 in the present application. Therefore, the point estimates of the persistence measures are rather similar, although not surprisingly the non-parametric estimates based on the spectrum are much less precisely estimated than the estimate based on the parametric approach.

12.5 Lesson 12.5: Non-stationarity and structural breaks in real GDP

Suppose you are interested in testing for a unit root in a time-series whilst at the same time allowing for the presence of *known* structural break(s). As an example consider log real GDP in the UK for the years from 1942 to 1987, and suppose that there has been a single mean shift in UK real GDP in 1973, the time of the first oil price shock. The file PHILLIPS.FIT in the tutorial directory contains annual aggregate UK data for the years 1855-1987 (inclusive) on the following five variables[14]

E	Logarithm of employment
N	Logarithm of labour force
Y	Logarithm of real GDP
P	Logarithm of a price index
W	Logarithm of a nominal wage index

Read this file and in the Commands and Data Transformations box in the Process window create an intercept, a time trend and a dummy variable for the year 1973 by typing

$$\textbf{SAMPLE}\quad 1942\quad 1987; C = 1;\quad T = \textbf{CSUM}(1); DU = 0;$$
$$\textbf{SAMPLE}\quad 1973\quad 1973;\quad DU = 1;$$

Now clear the Commands and Data Transformations box and type

$$\textbf{SAMPLE}\quad 1942\quad 1987;\quad \textbf{ADF}\quad Y$$

Click on the  button, and, when prompted, check the 'Simulate critical values' checkbox to obtain simulated critical values for the ADF test. Click  and then  . The results are reported in Table 12.8. The ADF statistics are above the 95 per cent simulated critical values in the three cases, no intercept no trend, intercept but no trend, intercept and trend. Hence, it is not possible to reject the unit root hypothesis in the variable Y at the 5 per cent significance level.

[14] It should also be remembered that there are likely to be substantial measurement errors in such historical data.

Table 12.8: *ADF* unit roots tests on the variable Y

```
                          ADF tests for variable Y
      The Dickey-Fuller regressions include no intercept and no trend
******************************************************************************
 46 observations used in the estimation of all ADF regressions.
 Sample period from 1942 to 1987
******************************************************************************
         Test Statistic     CV         LL         AIC         SBC        HQC
 DF          4.9957      -1.9632    102.9432    101.9432    101.0288   101.6007
 ADF(1)      2.9330      -1.9059    104.9258    102.9258    101.0971   102.2408
******************************************************************************
 CV  = 95% simulated critical value using 46 obs. and 1000 replications.
 LL  = Maximized log-likelihood       AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion     HQC = Hannan-Quinn Criterion

                          ADF tests for variable Y
     The Dickey-Fuller regressions include an intercept but not a trend
******************************************************************************
 46 observations used in the estimation of all ADF regressions.
 Sample period from 1942 to 1987
******************************************************************************
         Test Statistic     CV         LL         AIC         SBC        HQC
 DF          1.0955      -2.9184    103.4285    101.4285     99.5998   100.7434
 ADF(1)      .76187      -2.8937    105.1654    102.1654     99.4224   101.1379
******************************************************************************
 95% published asymptotic critical value corresponding to ADF(0) =  -2.9256

 CV  = 95% simulated critical value using 46 obs. and 1000 replications.
 LL  = Maximized log-likelihood       AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion     HQC = Hannan-Quinn Criterion

                          ADF tests for variable Y
   The Dickey-Fuller regressions include an intercept and a linear trend
******************************************************************************
 46 observations used in the estimation of all ADF regressions.
 Sample period from 1942 to 1987
******************************************************************************
         Test Statistic     CV         LL         AIC         SBC        HQC
 DF         -3.0976      -3.5072    108.8904    105.8904    103.1474   104.8629
 ADF(1)     -3.4928      -3.5443    111.7552    107.7552    104.0979   106.3851
******************************************************************************
 95% published asymptotic critical value corresponding to ADF(0) =  -3.5088

 CV  = 95% simulated critical value using 46 obs. and 1000 replications.
 LL  = Maximized log-likelihood       AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion     HQC = Hannan-Quinn Criterion
```

Now repeat the same exercise allowing for a one-off break in the level of series Y. Specifically, clear the Commands and Data Transformations box and type

SAMPLE 1942 1987; **ADF** Y & C T DU

The use of & in conjunction with the command **ADF** allows running an *ADF* test that allows for the effects of deterministics or known exogenous variables that are listed after the & sign. See Section 4.4.5.

Notice that when **ADF** is used with &, *Microfit* only reports the *ADF* test for the case of no intercept and no trend. Since we believe that the series Y is trended, we have added C and T to the list of variables after &. The output is reported in Table 12.9. The *ADF* statistics are below the 95 per cent simulated critical values, for $p = 0, 1$. Therefore, once controlled for a single (one off) break due to the quadrupling of oil prices in 1973, the presence of a unit root the Y process is marginally rejected at the 5 per cent significance level. This outcome, however, is not robust and depends on the initial values (the War years) and the order of the ADF test. To see this issue the command

SAMPLE 1942 1987; **ADF** $Y(4)$ & C T DU

Table 12.9: ADF unit roots tests on the variable Y after controlling for DU

```
                        ADF tests for variable Y
   The Dickey-Fuller test allowing for specified deterministic variables
*******************************************************************************
 44 observations used in the estimation of all ADF regressions.
 Sample period from 1944 to 1987

 List of deterministic variables added to the ADF regression:

 C               T               DU
*******************************************************************************
        Test Statistic    CV          LL          AIC         SBC         HQC
 DF         -3.8406     -3.5173    103.1422    102.1422    101.2501    101.8113
 ADF(1)     -4.0275     -3.5767    104.7177    102.7177    100.9335    102.0560
*******************************************************************************
 CV  = 95% simulated critical value using 44 obs. and 1000 replications.
 LL  = Maximized log-likelihood      AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion    HQC = Hannan-Quinn Criterion
```

12.6 Lesson 12.6: Unit roots in US nominal wages and the stock market crash

In this lesson we investigate the non-stationarity of the logarithm of nominal wages in the US, allowing for a single structural break in 1929, the year of the stock market crash. Following Perron (1989), we estimate three equations to test for the unit root hypothesis. Specifically, we consider

$$y_t = a_0 + a_1 t + a_2 DC_t + \phi y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \varepsilon_t, \tag{12.12}$$

$$y_t = a_0 + a_1 t + a_2 DT_t + \phi y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \varepsilon_t, \tag{12.13}$$

$$y_t = a_0 + a_1 t + a_2 DC_t + a_3 DT_t + \phi y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \varepsilon_t, \tag{12.14}$$

where

$$\begin{aligned} DC_t &= 0 \ \text{ for } t \leq 1929 \\ DC_t &= 1 \ \text{ for } t > 1929 \\ DT_t &= t \cdot DC_t \end{aligned}$$

In this lesson we set $p = 4$. The above regressions allow for the possible existence of three kinds of structural breaks: model (12.12) allows for a break in the level (intercept) of the series; equation (12.13) allows for a break in the rate of growth (slope); and equation (12.14) allows for breaks in both, namely possible breaks in both the level and in the rate of growth of the series.

Load the *Excel* file NPLUS.XLS from the tutorial directory, and go to the Process window. To create the logarithm of NW, an intercept, a trend and the variables DC and DT, in the Commands and Data Transformations box type

SAMPLE 1930 1970 ; $C = 1$; $T =$ **CSUM**(1); $DC = 0$;
SAMPLE 1930 1970; $DC = 1$;
SAMPLE 1909 1970; $DT = DC * T$; $Y =$ **LOG**(NW); RUN

In the following, we apply the ADF-WS unit root test described in Section 4.4.5 to Y. We first consider equation (12.12), and test for a unit roots after allowing for a break in the intercept only. Clear the Commands and Data Transformations box and type

ADF_WS $Y(4)$ & C DC T

Click on the Simulation of Critical Values for Unit Root Tests (Off) button, and, when prompted, check the 'Simulate critical values' checkbox to obtain simulated critical values for the ADF test. Click Close and then RUN. The results in Table 12.10 should appear on the screen. The ADF statistics are above the 95 per cent simulated critical values, given by the column headed CV. It is, therefore, not possible to reject the null of a unit root ($H_0 : \phi = 1$) in the variable Y at the 5 per cent significance level, despite allowing for a break in the intercept.

Table 12.10: WS-ADF unit roots tests on the variable Y after controlling for a break in the intercept

```
                    ADF-WS tests for variable Y
       The Dickey-Fuller test allowing for specified deterministic variables
******************************************************************************
 57 observations used in the estimation of all ADF regressions.
 Sample period from 1914 to 1970

 List of deterministic variables added to the ADF regression:

 C               DC              T
******************************************************************************
        Test Statistic    CV          LL          AIC         SBC         HQC
 DF         -3.2656     -4.0688     57.1816     56.1816     55.1601     55.7846
 ADF(1)     -3.9468     -4.0712     59.4737     57.4737     55.4306     56.6797
 ADF(2)     -3.4744     -3.9773     59.4802     56.4802     53.4156     55.2892
 ADF(3)     -3.2481     -3.9975     59.5056     55.5056     51.4195     53.9176
 ADF(4)     -3.1151     -3.9660     59.5640     54.5640     49.4564     52.5790
******************************************************************************
 CV  = 95% simulated critical value using 57 obs. and 1000 replications.
 LL  = Maximized log-likelihood        AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion      HQC = Hannan-Quinn Criterion
```

Now clear the Commands and Data Transformations box and type

ADF_WS $Y(4)$ & C T DT RUN

The above instruction allows testing for $H_0 : \phi = 1$ in model (12.13). The results are reported in Table 12.11. As before, it is not possible to reject the null of a unit root in the log of NW at the 5 per cent significance level, if we only allow for a break in the rate of growth of the series.

Table 12.11: *WS-ADF* unit roots tests on the variable Y after controlling for a break in the trend

```
                         ADF-WS tests for variable Y
    The Dickey-Fuller test allowing for specified deterministic variables
********************************************************************************
 57 observations used in the estimation of all ADF regressions.
 Sample period from 1914 to 1970

 List of deterministic variables added to the ADF regression:

 C                T                DT
********************************************************************************
        Test Statistic   CV          LL          AIC         SBC         HQC
 DF        -2.4067      -3.9412     58.0822     57.0822     56.0607     56.6852
 ADF(1)    -2.9127      -3.8597     59.6981     57.6981     55.6551     56.9041
 ADF(2)    -2.4829      -3.7571     59.9446     56.9446     53.8800     55.7536
 ADF(3)    -2.2894      -3.7719     59.9767     55.9767     51.8906     54.3887
 ADF(4)    -2.1662      -3.6922     59.9819     54.9819     49.8742     52.9969
********************************************************************************
 CV  = 95% simulated critical value using 57 obs. and 1000 replications.
 LL  = Maximized log-likelihood      AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion    HQC = Hannan-Quinn Criterion
```

We now allow for a break both in the level and in the rate of growth, by issuing the command

$$\textbf{ADF\_WS} \quad Y(4) \quad \& \quad C \quad DC \quad T \quad DT$$

The test results are displayed in Table 12.12, and show that even if we allow for a simultaneous breaks in the intercept and the trend of Y in 1929, the series Y is still non-stationary for $p = 1, 2, 3, 4$ since the values of the *ADF-WS* statistic are below the simulated critical values. As an exercise, try replicating this lesson using the **ADF_MAX** and **ADF** commands. Similar results should be obtained.

Table 12.12: *WS-ADF* unit root tests on the variable Y after controlling for a break in the intercept and trend

```
                         ADF-WS tests for variable Y
    The Dickey-Fuller test allowing for specified deterministic variables
********************************************************************************
 57 observations used in the estimation of all ADF regressions.
 Sample period from 1914 to 1970

 List of deterministic variables added to the ADF regression:

 C                DC               T                DT
********************************************************************************
        Test Statistic   CV          LL          AIC         SBC         HQC
 DF        -3.3708      -4.4764     56.1012     55.1012     54.0797     54.7042
 ADF(1)    -4.0414      -4.4563     58.2738     56.2738     54.2308     55.4798
 ADF(2)    -3.5706      -4.2919     58.2763     55.2763     52.2117     54.0853
 ADF(3)    -3.3343      -4.3405     58.3016     54.3016     50.2155     52.7136
 ADF(4)    -3.1956      -4.2746     58.3628     53.3628     48.2552     51.3778
********************************************************************************
 CV  = 95% simulated critical value using 57 obs. and 1000 replications.
 LL  = Maximized log-likelihood      AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion    HQC = Hannan-Quinn Criterion
```

12.7 Exercises in univariate time-series analysis

12.7.1 Exercise 12.1

Detrend the logarithm of US real GNP using the Hodrick-Prescott filter, and then investigate the cyclical properties of the detrended series by using the **SPECTRUM** command. Show that the most likely periodic cycle in this series is 24 quarters. How robust is this conclusion to the choice of λ, the smoothing coefficient in the Hodrick-Prescott filter? Carry out this analysis on the UK GDP and compare the results.

12.7.2 Exercise 12.2

Estimate the spectral density function of the UK output growth evaluated at zero frequency, using both parametric and non-parametric approaches. The relevant data are given in the file GDP95.FIT.

12.7.3 Exercise 12.3

Estimate a suitable $ARIMA(p, d, q)$ model for UK consumer prices over the period 1965-1990 and use it to forecast the inflation rate over the period 1991-1993. Compare the forecasting performance of the model with a 'naive' forecast based on a random walk model, possibly with a drift.

Chapter 13

Lessons in Non-Linear Estimation

In this chapter we show how the non-linear option in the Single Equation Estimation Menu can be used to estimate simple non-linear models such as the Cobb-Douglas production function, the Phillips curve, Almon distributed lag functions, and parameters of the Euler equation that arise in inter-temporal optimization models. For the relevant estimation menus see Section 6.16, and for the underlying econometric and computational methods see Section 21.21.

13.1 Lesson 13.1: Non-linear estimation of Cobb-Douglas production function

The non-linear estimation option in *Microfit* provides a powerful tool for the estimation of non-linear equations and/or the estimation of linear equations subject to linear or non-linear parametric restrictions. Suppose you are interested in estimating the following non-linear form of the Cobb-Douglas production function

$$Y_t = AK_t^{\alpha} L_t^{1-\alpha} + u_t \tag{13.1}$$

Read into *Microfit* the file CD.FIT which contains the annual observations on US Output (Y), Capital Stock (K) and Labour Input (L), over the period 1899-1922 originally analysed by Cobb and Douglas (1928). Choose option 4 in the Single Equation Estimation Menu (see Section 6.4), and in the Commands and Data Transformation box type (see Section 6.16.1)

$$Y = A0 * (K\ \hat{}\ A1) * (L\,\hat{}\,(1 - A1)) \qquad \text{RUN}$$

You will now be prompted to specify the initial estimates for the parameters $A0$ and $A1$ (see Section 6.16.2). The initial choice of $A0$ and $A1$ is often critical for the convergence of the iterative process. For example, the iterative process is unlikely to converge if the iterations are started with very small values of $A0$ and $A1$. For example, starting the iterations with

Table 13.1: Non-linear estimates of the Cobb-Douglas production function

```
                    Non-Linear Least Squares Estimation
                 The estimation method converged after 2 iterations
*******************************************************************************
 Non-linear regression formula:
 Y=A0*(K^A1)*(L^(1-A1))
 24 observations used for estimation from 1899 to 1922
*******************************************************************************
 Parameter                  Estimate        Standard Error         T-Ratio[Prob]
 A0                          1.0182             .028085           36.2552[.000]
 A1                          .24962             .048076            5.1921[.000]
*******************************************************************************
 R-Squared                      .94175   R-Bar-Squared                     .93910
 S.E. of Regression            10.7970   F-Stat.    F(1,22)     355.6956[.000]
 Mean of Dependent Variable   165.9167   S.D. of Dependent Variable      43.7532
 Residual Sum of Squares        2564.6   Equation Log-likelihood       -90.1128
 Akaike Info. Criterion       -92.1128   Schwarz Bayesian Criterion    -93.2909
 DW-statistic                   1.5256
*******************************************************************************


                              Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *          F Version          *
*******************************************************************************
*                     *                          *                             *
* A:Serial Correlation*CHSQ(1)  =    .14424[.704]*F(1,21)      =    .12697[.725]*
*                     *                          *                             *
* B:Functional Form   *CHSQ(1)  =    .55738[.455]*F(1,21)      =    .49930[.488]*
*                     *                          *                             *
* C:Normality         *CHSQ(2)  =    3.9850[.136]*       Not applicable        *
*                     *                          *                             *
* D:Heteroscedasticity*CHSQ(1)  =    4.1286[.042]*F(1,22)      =    4.5708[.044]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

$A0 = A1 = 0$ will result in the message

> *An invalid operation has been carried out or a very large number has been calculated.*
>
> *Either initial values of parameters are not appropriate or variables need to be scaled! Try again.*

There are no general rules concerning the choice of the initial estimates for the unknown parameters, but in most applications a preliminary linear regression based on a Taylor series expansion of the non-linear equation can be very helpful in providing a reasonable set of initial estimates for the unknown parameters. In the case of the present example, one can obtain initial estimates for these parameters by running the constrained linear regression of $\log Y_t - \log L_t$ on an intercept term and $(\log K_t - \log L_t)$. This yields the initial estimates of 0.0145 and 0.2541 for the parameters $\log A0$ and $A1$, respectively. Therefore, using $1.0147\,(= exp(0.0145))$ and 0.25 as initial estimates for $A0$ and $A1$ we obtain the results shown in Table 13.1.

It is generally advisable to carry out the iterations from different initial estimates to guard against the possibility of local optima. For the present application, we retried the

Table 13.2: Non-linear estimates of the Cobb-Douglas production function

```
                 Non-Linear Two-Stage Least Squares Estimation
               The estimation method converged after 3 iterations
*******************************************************************************
 Non-linear regression formula:
 Y=A0*(K^A1)*(L^(1-A1))
 List of instruments:
 INPT           LK(-1)          LK(-2)          LL(-1)          LL(-2)
 22 observations used for estimation from 1901 to 1922
*******************************************************************************
 Parameter                 Estimate       Standard Error          T-Ratio[Prob]
 A0                         1.0114            .033088             30.5683[.000]
 A1                         .26253            .056247              4.6674[.000]
*******************************************************************************
 R-Squared                      .92682  R-Bar-Squared                    .92316
 GR-Squared                     .81828  GR-Bar-Squared                   .80920
 S.E. of Regression            11.2670  F-Stat.    F(1,20)    253.2879[.000]
 Mean of Dependent Variable   171.8636  S.D. of Dependent Variable      40.6451
 Residual Sum of Squares        2538.9  Value of IV Minimand             1454.7
 DW-statistic                   1.5378  Sargan's    CHSQ(3)    11.4596[.009]
*******************************************************************************


                              Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version          *          F Version         *
*******************************************************************************
*                     *                            *                            *
* A:Serial Correlation*CHSQ(1)  =   .17498[.676]*       Not applicable          *
*                     *                            *                            *
* B:Functional Form   *CHSQ(1)  =   6.7514[.009]*       Not applicable          *
*                     *                            *                            *
* C:Normality         *CHSQ(2)  =   1.2579[.533]*       Not applicable          *
*                     *                            *                            *
* D:Heteroscedasticity*CHSQ(1)  =   5.5735[.018]*       Not applicable          *
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

computations using the initial values $(1, 0.8)$, $(3, 0.75)$, $(5, 0.01)$ and $(10, 0.95)$, and arrived at exactly the same estimates as above.

Suppose now that you are interested in estimating (13.1) by the non-linear 2-stage least squares (or $NLS - IV$) method using $\{1,\ \log K_{t-1},\ \log K_{t-2},\ \log L_{t-1},$ and $\log L_{t-2}\}$ as instruments. First specify the non-linear equation using option 4 in the Single Equation Estimation Menu (Univariate Menu), and selecting the Non-linear 2-stage Least Squares option. Use the NLS estimates of $A0$ and $A1$, namely 1.02 and 0.25 as initial estimates and, when prompted, type

$$INPT \quad LK(-1) \quad LK(-2) \quad LL(-1) \quad LL(-2)$$

RUN

You should see the results in Table 13.2 on the screen. In this example the NLS estimates in Table 13.1 and the NLS-IV estimates in Table 13.2 are very similar.

13.2 Lesson 13.2: Estimation of Euler equations by the NLS-IV method

The non-linear instrumental variable option in *Microfit* can also be used to estimate the parameters of Euler equations obtained as first-order conditions of a representative agent's utility maximization problem under uncertainty. Typically, a Euler equation takes the form of

$$E\left\{H(\mathbf{x}_t, \boldsymbol{\gamma})\,|\boldsymbol{\Omega}_{t-1}\right\} = 0, \tag{13.2}$$

where $E(\cdot\,|\boldsymbol{\Omega}_{t-1})$ denotes conditional expectation, $\boldsymbol{\gamma}$ is a $k \times 1$ vector of unknown parameters to be estimated, $\mathbf{x}_t$ is a vector of observable variables, and $\boldsymbol{\Omega}_{t-1}$ is the information known to the agent (but not necessarily to the econometrician) at time $t-1$ (see, for example, Hansen and Singleton (1983)). As an example, consider the optimization problem of a representative consumer with a constant relative risk averse utility function who faces a consumption/investment decision in an uncertain environment. Assuming that only investment in stocks is being considered, the Euler equation for this optimization problem will be given by

$$u_t = \beta(\mathbf{x}_{1t})^{\sigma}\mathbf{x}_{2t} - 1 \tag{13.3}$$

where $E(u_t\,|\boldsymbol{\Omega}_{t-1}) = 0$, β is the discount factor, σ is the constant relative risk aversion parameter, $\mathbf{x}_{1t} = (c_{t-1}/c_t)$, and $\mathbf{x}_{2t}$ is the one-period real return on stocks (see, for example, Grossman and Shiller (1981), Hansen and Singleton (1983), and Pesaran (1991)). On the assumption that the parameters of (13.2) are identified, the *NLS-IV* option in *Microfit* can be used to obtain consistent estimates of β and σ by defining u_t to be a vector with all its elements set equal to zero, using $E\left(E\left(u_t\,|\Omega_{t-1}\right)\right) = 0$ and then running the non-linear regression of u_t on $\mathbf{x}_{1t}$ and $\mathbf{x}_{2t}$ as in (13.3). To implement this procedure proceed in the following manner:

1. First read the file HS.FIT into the workspace. This file covers monthly observations over the period 1959(3)-1978(12) on the following variables

$X1$	the ratio of consumption in time period $t-1$ to consumption in time period t
$X2$	the one-period real return on stocks

   This is the corrected version of the data set used by Hansen and Singleton (1983).

2. In the Commands and Data Transformations box generate the variables

   $$INPT = 1; \quad U = 0$$

3. From the Single Equation Estimation Menu (Univariate Menu) choose option 4, and when prompted type

   $$U = b * (X1\ \hat{}\, s) * X2 - 1$$ RUN

   You will now be asked to give initial estimates for the unknown parameters b and s. These correspond to the discount factor, β, and the risk aversion coefficient, σ. Try the initial values of 0.8 and 1 for these two parameters, respectively, and click OK. You need to list at least 2 instruments. You can try different lagged values of $X1$ and $X2$, and the unit vector (namely $INPT$ in the workspace) as your instruments. If you choose the variables

   $$INPT \quad X1(-1) \quad X2(-1)$$

Table 13.3: Euler equation estimates of the Hansen-Singleton consumption-based asset pricing model

```
                  Non-Linear Two-Stage Least Squares Estimation
                    The estimation method converged after 4 iterations
*******************************************************************************
 Non-linear regression formula:
 U=b*(X1^s)*X2-1
 List of instruments:
 INPT            X1(-1)          X2(-1)
 237 observations used for estimation from 1959M4 to 1978M12
*******************************************************************************
 Parameter                  Estimate       Standard Error        T-Ratio[Prob]
 B                           .99895          .0049466           201.9465[.000]
 S                           .86474           2.0461              .42263[.673]
*******************************************************************************
 R-Squared                      *NONE*   R-Bar-Squared                   *NONE*
 GR-Squared                     *NONE*   GR-Bar-Squared                  *NONE*
 S.E. of Regression            .041545   F-Stat.    F(1,235)           *NONE*
 Mean of Dependent Variable       0.00   S.D. of Dependent Variable        0.00
 Residual Sum of Squares        .40561   Value of IV Minimand          .0027813
 DW-statistic                   1.8293   Sargan's    CHSQ(1)    1.6114[.204]
*******************************************************************************


                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version         *
*******************************************************************************
*                     *                          *                             *
* A:Serial Correlation*CHSQ(12) =  14.5944[.264]*       Not applicable         *
*                     *                          *                             *
* B:Functional Form   *CHSQ(1)  = .0090165[.924]*       Not applicable         *
*                     *                          *                             *
* C:Normality         *CHSQ(2)  =  10.6268[.005]*       Not applicable         *
*                     *                          *                             *
* D:Heteroscedasticity*CHSQ(1)  =   7.2347[.007]*       Not applicable         *
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

as instruments, the results given in Table 13.3 should appear on the screen. The results seem to be robust to the choice of the initial parameter estimates, and yield the estimate 0.865 (2.046) for the risk aversion coefficient, which is very poorly estimated. In contrast the discount factor, β, is estimated much more precisely.

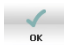
Notes

1. The diagnostic and other summary statistics that follow the $NLS\text{-}IV$ results for the estimation of the Euler equation should be treated with caution. Given the way the non-linear regression (13.3) is set up, the functional form and the heteroscedasticity test statistics are not appropriate and should be ignored.

2. In the context of non-linear rational expectations models the disturbances u_t need not be homoscedastic and are likely to be serially correlated if the observation horizon exceeds the decision horizon of the economic agent (see, for example, Chapter 7 in Pesaran (1987b)). In these circumstances the appropriate standard errors for the parameter estimates are given by White's or Newey and West's adjusted estimates. To

Table 13.4: Euler equation estimates of the Hansen-Singleton consumption-based asset pricing model

```
                    Non-Linear Two-Stage Least Squares Estimation
                       The estimation method converged after 4 iterations
  Based on Newey-West adjusted S.E.'s Bartlett weights, truncation lag = 12
*******************************************************************************
 Non-linear regression formula:
 U=b*(X1^s)*X2-1
 List of instruments:
 INPT            X1(-1)          X2(-1)
 237 observations used for estimation from 1959M4 to 1978M12
*******************************************************************************
 Parameter                  Estimate       Standard Error         T-Ratio[Prob]
 B                           .99895          .0041580           240.2460[.000]
 S                           .86474           1.6366              .52839[.598]
*******************************************************************************
```

compute these standard errors first choose option 4 in the Post Regression Menu (see Section 6.20), and then choose option 4 (Newey-West adjusted with Bartlett weights) in the Standard, White and Newey-West Adjusted Variance Menu that follows and, when asked to specify the truncation point (or horizon) for the Bartlett window type 12 and press . The Newey-West adjusted variance matrix of ($\hat{\beta}$ and $\hat{\sigma}$) should appear on the screen. Click  followed by  to obtain the estimation results in Table 13.4 giving the GMM estimates with Newey-West adjusted standard error using the Bartlett weights.

1. The results in Table 13.4 are only marginally different from the estimates based on the unadjusted standard errors given in Table 13.3.

2. The *NLS-IV* procedure applied to Euler equations also provides a simple method of implementing the Generalized Method of Moments (GMM) due to Hansen (1982). The possible effects of serial correlation and heteroscedasticity in u_ts on the standard errors can be readily dealt with using option 4 in the Post Regression Menu (see note 2 above).

3. Sargan's general mis-specification test statistic described in Section 21.10.3 can also be readily computed for the present application, and is given by $0.0027813/0.041545^2 = 1.61$, which should be compared with the critical value of a chi-squared variate with one degree of freedom (the difference between the number of instruments and the number of unknown parameters).

13.3 Lesson 13.3: Estimation of Almon distributed lag models

Suppose you are interested in estimating the following polynomial distributed lag model:

$$Y_t = \sigma + \sum_{i=0}^{m} w_i X_{t-i} + u_t \tag{13.4}$$

where the weights w_i are determined by polynomials of order r

$$w_i = b_0 + b_1 i + b_2 i^2 + \cdots + b_r i^r \tag{13.5}$$

for $i = 0, 1, ..., m$. The above model is also known as the Almon distributed lag model, $ALMON(m, r)$, and in the case where $r < m$ it imposes $m - r$ restrictions on the lag coefficient w_i (see for example Greene (2002), Chapter 19), and the original paper by Almon (1965)). Here we show how to estimate such a model using *Microfit*.

There are two different ways of estimating the polynomial distributed lag model (13.5).

Method A, using the BATCH command One possibility would be to construct the following weighted averages:

$$Z_{tj} = \sum_{i=0}^{m} i^j X_{t-i}, \quad j = 0, 1, ..., r \tag{13.6}$$

and regress Y_t on an intercept term and the variables Z_{t0}, Z_{t1}, ..., Z_{tr}. This would then yield the estimates of a, b_0, b_1,...,b_r. The construction of the Zs can be carried out in the Commands and Data Transformations in the Process windows, preferably using the **BATCH** command on a previously prepared batch file.

Method B, using the non-linear option Alternatively, one can estimate the Almon distributed lag model directly using the non-linear least squares option. You simply need to type the formula for the distributed lag model (13.4) in the screen editor box for the non-linear estimation option, substituting (13.5) for the weights w_i.

As an example, consider the estimation of a polynomial distributed lag model with $m = 8$ and $r = 3$, ALMON(8,3), between appropriations (X), and capital expenditures (Y) for the US manufacturing sector. The relevant data are in the special *Microfit* file ALMON.FIT, and contain observations on Y and X over the period 1953(1)-1967(4). This is an extended version of the data analysed originally by Almon (1965). The special *Microfit* file ALMON.FIT, and two other related files, namely ALMON83.BAT and ALMON83.EQU, should all be in the tutorial directory.

Table 13.5: Estimation of the Almon distributed lag model using constructed variables (method A)

```
                          Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is Y
 52 observations used for estimation from 1955Q1 to 1967Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                       62.3626              57.4400           1.0857[.283]
 Z0                         .051777              .030495           1.6979[.096]
 Z1                          .11843              .051512           2.2991[.026]
 Z2                        -.033676              .016538          -2.0363[.047]
 Z3                        .0023995             .0013856           1.7318[.090]
*******************************************************************************
 R-Squared                      .98924   R-Bar-Squared                   .98832
 S.E. of Regression           124.8792   F-Stat.    F(4,47)      1080.1[.000]
 Mean of Dependent Variable     3253.9   S.D. of Dependent Variable      1155.6
 Residual Sum of Squares      732955.9   Equation Log-likelihood      -322.1783
 Akaike Info. Criterion      -327.1783   Schwarz Bayesian Criterion   -332.0564
 DW-statistic                   .45710
*******************************************************************************


                                   Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version         *
*******************************************************************************
*                     *                          *                             *
* A:Serial Correlation*CHSQ(4)  =  32.8237[.000]*F(4,43)      =  18.4005[.000]*
*                     *                          *                             *
* B:Functional Form   *CHSQ(1)  =  .031974[.858]*F(1,46)      =  .028303[.867]*
*                     *                          *                             *
* C:Normality         *CHSQ(2)  =   8.7992[.012]*       Not applicable         *
*                     *                          *                             *
* D:Heteroscedasticity*CHSQ(1)  =  .011755[.914]*F(1,50)      =  .011305[.916]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

Method A First load the file ALMON.FIT into *Microfit*. To implement method A, in the Commands and Data Transformations in the Process window run the batch file ALMON83.BAT by typing

$$\textbf{BATCH} \quad ALMON83$$

This creates the variables $Z0$, $Z1$, $Z2$, and $Z3$ defined by (13.6). Now run a linear regression of Y on an intercept term and the variable $Z0$, $Z1$, $Z2$, and $Z3$. To do this you need to select option 1 from the Single Equation Estimation Menu (Univariate Menu) choosing the OLS option. Type

$$Y \quad INPT \quad Z0 \quad Z1 \quad Z2 \quad Z3$$

The results in Table 13.5 should now appear on the screen.

The estimates of b_0, b_1, b_2 and b_3 are given by the coefficients of $Z0$, $Z1$, $Z2$ and $Z3$ respectively. Notice, however, that the very low value obtained for the DW static suggests the possibility of a serious dynamic mis-specification.

Method B The same results can be obtained using the non-linear option. Choose option 4 in the Single Equation Estimation Menu (Univariate Menu), selecting the Nonlinear

Table 13.6: Estimation of the Almon distributed lag model using constructed variables (method B)

```
                         Non-Linear Least Squares Estimation
                         The estimation method converged after 2 iterations
*******************************************************************************
 Non-linear regression formula:
 Y=a+b0*x+(b0+b1+b2+b3)*x(-1)+(b0+2*b1+4*b2+8*b3)*x(-2)+(b0+3*b1+9*b2+27*b3)*x(
 -3)+(b0+4*b1+16*b2+64*b3)*x(-4)+(b0+5*b1+25*b2+125*b3)*x(-5)+(b0+6*b1+36*b2+21
 6*b3)*x(-6)+(b0+7*b1+49*b2+343*b3)*x(-7)+(b0+8*b1+64*b2+512*b3)*x(-8)
 52 observations used for estimation from 1955Q1 to 1967Q4
*******************************************************************************
 Parameter                  Estimate         Standard Error         T-Ratio[Prob]
 A                          62.3626             57.4400             1.0857[.283]
 B0                         .051777             .030495             1.6979[.096]
 B1                          .11843             .051512             2.2991[.026]
 B2                        -.033676             .016538            -2.0363[.047]
 B3                        .0023995            .0013856             1.7318[.090]
*******************************************************************************
 R-Squared                      .98924   R-Bar-Squared                   .98832
 S.E. of Regression           124.8792   F-Stat.    F(4,47)       1080.1[.000]
 Mean of Dependent Variable     3253.9   S.D. of Dependent Variable      1155.6
 Residual Sum of Squares      732955.9   Equation Log-likelihood      -322.1783
 Akaike Info. Criterion      -327.1783   Schwarz Bayesian Criterion   -332.0564
 DW-statistic                   .45710
*******************************************************************************


                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version        *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(4)  =  32.8237[.000]*F(4,43)      =  18.4005[.000]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =  .031975[.858]*F(1,46)      =  .028303[.867]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =   8.7992[.012]*        Not applicable       *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =  .011755[.914]*F(1,50)      =  .011305[.916]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
```

Least Square option. In the box editor, retrieve the equation file ALMON83.EQU using the  button, and then select the file. This gives the equation for the estimation of the Almon distributed lag model between X and Y with $m = 8$ and $r = 3$. Click  to accept the equation. You will now be prompted to specify the initial parameter estimates. Since the equation is inherently linear in the unknown parameters, click  to accept the default values of zeros for the initial estimates. The results of the non-linear estimation are reproduced in Table 13.6.

13.4 Lesson 13.4: Estimation of a non-linear Phillips curve

Phillips (1958) estimated his famous curve, a non-linear relationship between the rate of growth of wages and unemployment, graphically from pre-World War I UK data. In this lesson using econometric methods, we estimate the same curve on the same data.[1] The relevant data are in the file PHILLIPS.FIT (for more details see Lesson 12.5).

[1] The background to the estimation of the Phillips curve is discussed in Alogoskoufis and Smith (1991a).

Read this file, and in the Commands and Data Transformations box in the Process window create an intercept, C, the rate of growth of money wages, DW, and the (log) unemployment rate, U, and then graphically examine the pre-WWI Phillips Curve by typing

$$C = 1; DW = W - W(-1); U = N - E;$$

SAMPLE 1861 1913; **SCATTER** DW U

The scatter diagram associated with the original Phillips curve is shown in Figure 13.1.

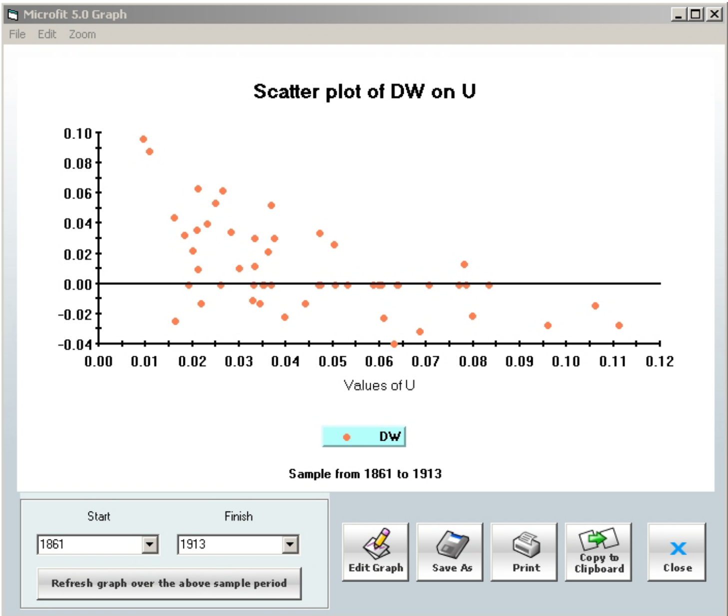


Figure 13.1: Scatter of changes in money wages against the unemployment rate 1861-1913

Although the fit is not very close, there is clearly a negative relationship between DW and U, with some evidence of non-linearity: at high rates of unemployment the effect on wage growth is much smaller than at low rates of unemployment. Using averages of data for 1861-1913 Phillips obtained estimates of a curve of the form

$$DW_t = a_1 + a_2 U_t^{a_3} + \xi_t \tag{13.7}$$

where ξ_t represents an error term. His estimate of a_3 was -1.4, close to a linear relationship between wage growth and the *reciprocal* of the unemployment rate.

To evaluate the importance of non-linearity in the Phillips curve we first estimate two linear relationships: one between the wage growth and the level of the unemployment variable, and the other between the wage growth and the reciprocal of the unemployment variable. We can use these estimates to obtain initial values for the non-linear estimation of (13.7)

Table 13.7: *OLS* estimates of a linear Phillips curve

```
                         Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is DW
 53 observations used for estimation from 1861 to 1913
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 C                          .041152           .0069416            5.9284[.000]
 U                          -.68564             .13556           -5.0580[.000]
*******************************************************************************
 R-Squared                       .33406   R-Bar-Squared                    .32100
 S.E. of Regression              .024063  F-Stat.    F(1,51)      25.5831[.000]
 Mean of Dependent Variable      .010278  S.D. of Dependent Variable     .029202
 Residual Sum of Squares         .029531  Equation Log-likelihood       123.3507
 Akaike Info. Criterion         121.3507  Schwarz Bayesian Criterion    119.3804
 DW-statistic                     1.5404
*******************************************************************************

                                Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version        *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(1)  =   2.6976[.101]*F(1,50)      =   2.6814[.108]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =   4.0615[.044]*F(1,50)      =   4.1496[.047]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =   .59581[.742]*       Not applicable        *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =  11.6186[.001]*F(1,51)      =  14.3193[.000]*
*                     *                          *                            *
* E:Predictive Failure*CHSQ(74) = 766.3090[.000]*F(74,51)     =  10.3555[.000]*
*                     *                          *                            *
* F:Chow Test         *CHSQ(2)  =  38.0303[.000]*F(2,123)     =  19.0151[.000]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
   E:A test of adequacy of predictions (Chow's second test)
   F:Test of stability of the regression coefficients
```

that we shall be carrying out subsequently. Close the scatter plot and click 

to clear the content of the Commands and Data Transformations box, and type

SAMPLE 1855 1987; $RU = 1/U$ 

to create the inverse of the unemployment rate in RU. Choose option 1 from the Single Equation Estimation Menu (Univariate Menu), making sure that the *OLS* option is selected. Specify the equation as $DW \quad C \quad U$, set the sample to 1861- 1913, and click 

. You should obtain the results in Table 13.7. The unemployment rate is statistically highly significant with a negative coefficient, but the failure of the functional form and heteroscedasticity tests suggests that there may be important non-linearities in the relationship. For a comparison of this linear specification with the non-linear ones to be estimated below also note that $\bar{R}^2$, AIC, and SBC for this regression are given by 0.321, 121.35, and 119.38, respectively.

Click 

to move to the Post Regression Menu, and then backtrack and edit the regression equation to

$$DW \quad C \quad RU$$

Table 13.8: OLS estimates of a non-linear Phillips curve

```
                         Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is DW
 53 observations used for estimation from 1861 to 1913
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 C                        -.020083           .0054899             -3.6582[.001]
 RU                       .9873E-3           .1500E-3              6.5831[.000]
*******************************************************************************
 R-Squared                    .45939   R-Bar-Squared                   .44879
 S.E. of Regression          .021681   F-Stat.    F(1,51)       43.3377[.000]
 Mean of Dependent Variable  .010278   S.D. of Dependent Variable     .029202
 Residual Sum of Squares     .023973   Equation Log-likelihood       128.8761
 Akaike Info. Criterion     126.8761   Schwarz Bayesian Criterion    124.9058
 DW-statistic                 1.7224
*******************************************************************************

                                  Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version        *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(1)  =   .96053[.327]*F(1,50)      =   .92289[.341]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =   .10424[.747]*F(1,50)      =  .098535[.755]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =   1.2645[.531]*       Not applicable        *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =   .14444[.704]*F(1,51)      =   .13937[.710]*
*                     *                          *                            *
* E:Predictive Failure*CHSQ(74) = 986.3012[.000]*F(74,51)     =  13.3284[.000]*
*                     *                          *                            *
* F:Chow Test         *CHSQ(2)  =  24.9302[.000]*F(2,123)     =  12.4651[.000]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
   E:A test of adequacy of predictions (Chow's second test)
   F:Test of stability of the regression coefficients
```

Estimating this equation by the OLS method over the same sample period (1861-1913) yields the results in Table 13.8.

Note that the functional form and heteroscedasticity tests are now acceptable and the equation is preferred on the basis of all three model selection criteria. The values of $\bar{R}^2$, AIC, and SBC for this specification are 0.449, 126.88, and 124.91, respectively.

Now backtrack to the Single Equation Estimation Menu and choose the Non-Linear Regression option 4, selecting the Non-Linear Least Squares option. Type

$$DW = A1 + A2 * U\hat{\ }A3$$ 

and set the sample from 1861 to 1913. Click . For the initial values of $A1$, $A2$, and ,$A3$ choose 0.1, 0.1, and -1, respectively. Click . You should obtain the results in Table 13.9.

The estimate of a_3 is very close to and not significantly different from -1, though a_2 is not significantly different from zero. Such mixed results do arise in non-linear estimation, and special care needs to be taken in interpreting them. Firstly, if $a_2 = 0$, then a_3 is not

Table 13.9: Phillips curve estimated by non-linear least squares

```
                    Non-Linear Least Squares Estimation
                    The estimation method converged after 6 iterations
*******************************************************************************
 Non-linear regression formula:
 DW=A1+A2*U^A3
 53 observations used for estimation from 1861 to 1913
*******************************************************************************
 Parameter                 Estimate       Standard Error          T-Ratio[Prob]
 A1                       -.020325           .016049             -1.2665[.211]
 A2                       .0010202          .0020655              .49391[.624]
 A3                        -.99311           .47045              -2.1110[.040]
*******************************************************************************
 R-Squared                     .45939   R-Bar-Squared                   .43777
 S.E. of Regression            .021896  F-Stat.    F(2,50)      21.2441[.000]
 Mean of Dependent Variable   .010278   S.D. of Dependent Variable     .029202
 Residual Sum of Squares       .023973  Equation Log-likelihood       128.8762
 Akaike Info. Criterion       125.8762  Schwarz Bayesian Criterion    122.9207
 DW-statistic                  1.7227
*******************************************************************************


                              Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *          F Version         *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(1)  =   .96172[.327]*F(1,49)       =   .90557[.346]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =   3.5802[.058]*F(1,49)       =   3.5498[.065]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =   1.2727[.529]*       Not applicable        *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =   .15415[.695]*F(1,51)       =   .14876[.701]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

identified. We would also expect a high covariance between the two estimates. For instance, given the negative relationship between DW and U, if a_3 is positive a_2 will be negative and vice versa. Comparing the results to the regression using RU above, where a_3 is set to -1, we see that the reciprocal formulation is preferred by all three model selection criteria, and has a much sharper estimate of a_2.

13.5 Lesson 13.5 Estimating a non-linear Phillips curve with serially correlated errors

The non-linear option is very flexible and can be used to estimate linear regression models subject to linear and/or non-linear parametric restrictions or to estimate non-linear models with serially correlated errors. Suppose that we wish to estimate the Phillips equation defined by (13.7) subject to a first-order autoregressive error process

$$DW_t = a_1 + a_2 U_t^{a_3} + \xi_t$$

where

$$\xi_t = \rho \xi_{t-1} + \epsilon_t, \qquad \epsilon_t \sim iid(0, \sigma^2)$$

We first note that the above relations can also be written as

$$DW_t = a_1 + a_2 U_t^{a_3} + \rho(DW_{t-1} - a_1 - a_2 U_{t-1}^{a_3}) + \epsilon_t$$

which is another, albeit more complicated, non-linear equation. To estimate this model using the Phillips' original data set first read the file PHILLIPS.FIT and follow the steps in Lesson 13.4, to create the rate of growth of money wages in DW and the logarithm of the rate of unemployment in U. Choose option 4 in the Single Equation Estimation Menu (Univariate Menu), selecting Non-linear Least Squares, and type

$$DW = A1 + A2 * U\hat{\ } A3 + A4 * (DW(-1) - A1 - A2 * U(-1)\hat{\ } A3)$$

Select 1861-1913 as the sample period and click 

. For the initial values of $A1$, $A2$, $A3$, and $A4$ choose 0.1,0.1,-1.0, and 0.0, respectively, and then click 

. You should see the results in Table 13.10 on the screen.

The estimate of ρ, denoted by $A4$ in Table 13.10, is not significantly different from zero, and is in line with the diagnostic test results obtained in Lesson 13.4.

13.6 Exercises in non-linear estimation

13.6.1 Exercise 13.1

Use the data in the special *Microfit* file CD.FIT to estimate the Cobb-Douglas production function:

$$Y_t = A K_t^{\alpha} L_t^{\beta} + u_t,$$

Use option 7 in the Hypothesis Testing Menu (see Section 6.23) to test the constant returns to scale restriction $\alpha + \beta = 1$. Compare your results with those in Table 13.1.

Table 13.10: Estimates of Phillips curve with $AR(1)$ serially correlated residuals

```
                     Non-Linear Least Squares Estimation
                The estimation method converged after 9 iterations
*******************************************************************************
 Non-linear regression formula:
 DW=A1+A2*U^A3+A4*(DW(-1)-A1-A2*U(-1)^A3)
 53 observations used for estimation from 1861 to 1913
*******************************************************************************
 Parameter                  Estimate       Standard Error        T-Ratio[Prob]
 A1                        -.018585           .015696          -1.1841[.242]
 A2                        .7996E-3          .0017130           .46680[.643]
 A3                         -1.0444            .51522          -2.0271[.048]
 A4                          .11842            .14609           .81062[.422]
*******************************************************************************
 R-Squared                      .46940   R-Bar-Squared                  .43691
 S.E. of Regression            .021913   F-Stat.    F(3,49)     14.4494[.000]
 Mean of Dependent Variable    .010278   S.D. of Dependent Variable    .029202
 Residual Sum of Squares       .023529   Equation Log-likelihood      129.3714
 Akaike Info. Criterion       125.3714   Schwarz Bayesian Criterion   121.4308
 DW-statistic                   1.9403
*******************************************************************************


                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version         *
*******************************************************************************
*                     *                          *                             *
* A:Serial Correlation*CHSQ(1)  =    .87106[.351]*F(1,48)       =    .80206[.375]*
*                     *                          *                             *
* B:Functional Form   *CHSQ(1)  = .4156E-4[.995]*F(1,48)       = .3764E-4[.995]*
*                     *                          *                             *
* C:Normality         *CHSQ(2)  =    1.9032[.386]*       Not applicable        *
*                     *                          *                             *
* D:Heteroscedasticity*CHSQ(1)  =    .16699[.683]*F(1,51)       =    .16119[.690]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

13.6.2 Exercise 13.2

Estimate the Euler equation (13.3) in Lesson 13.2 by the non-linear least squares method and compare your results with those in Table 13.3.

13.6.3 Exercise 13.3

Re-estimate the Phillips curve in Lesson 13.4 over the entire sample period and the sub-periods 1954-1969, 1970-1987, allowing for the possible effects of current and past changes in inflation and unemployment on changes in money wages. Compare your results with those obtained by Phillips (1958).

13.6.4 Exercise 13.4

Use the quarterly observations on US real GNP ($USGNP$) in the file GDP95.FIT to estimate the following non-linear autoregressive model, known as the threshold autoregressive (TAR) model

$$\begin{aligned} y_t &= a_{01} + a_{01} * I(y_{t-2} - b) + \\ &\quad a_{11} * y_{t-1} + a_{11} * y_{t-1} * I(y_{t-2} - b) \end{aligned} \tag{13.8}$$

where $y_t = \log(USGNP/USGNP(-1))$, and $I(y_{t-2} - b)$ represents the indicator (or sign) function such that it is equal to unity when $y_{t-2} > b$, and zero otherwise. You may find it easier to estimate the model for different values of the threshold parameter, b. Note that the sample mean of y_t is 0.0072, which corresponds to an annual average growth rate of 2.9 per cent.

For a discussion of this class of models see Tong (1990). For an application to US output see Potter (1995).

Chapter 14

Lessons in Probit and Logit Estimation

The lessons in this section demonstrate the Logit/Probit options described in Section 6.19.

14.1 Lesson 14.1: Modelling the choice of fertilizer use by Philippine farmers

The file PHIL.FIT contains observations on fertilizer use by 491 small farmers in the Philippines, together with 5 explanatory variables. The dependent variable to be explained is $FERUSE$, a binary variable equal to 1 if fertilizer is used and zero otherwise. The explanatory variables are

$CREDIT$	Amount of credit (per hectares) held by the farmer
$DMARKET$	Distance of the farm from the nearest market
$HOURMEET$	Number of hours the farmer spent with an agricultural 'expert'
$IRSTAT$	Binary variable equal to 1 if irrigation is used, 0 otherwise
$OWNER$	Binary variable equal to 1 if the farmer owns the land, 0 otherwise
ONE	Vector of 1s

The appropriate probability model for explaining the binary choice variable $FERUSE$ is defined by

$$\Pr(FERUSE_i = 1) = F(\boldsymbol{\beta}'\mathbf{x}_i), \quad i = 1, 2, ..., 491$$

where $\mathbf{x}_i$ is a 6×1 vector of the regressors for the ith farmer. The program allows you to compute ML estimates of β both when $F(\cdot)$ is the cumulative distribution function of the standard normal (the Probit model) and when it has the logistic form (the Logit model). See Sections 6.19 and 21.20 for further details.

Read the *Microfit* file $PHIL.FIT$ and choose Logit from option 7 (Logit and Probit models) in the Single Equation estimation Menu (Univariate Menu: see Section 6.4). List the dependent variable $FERUSE$ followed by the explanatory variables in the editor:

$$FERUSE \quad ONE \quad CREDIT \quad DMARKET$$
$$HOURMEET \quad IRSTAT \quad OWNER$$

For the estimation sample enter 1 and 450 into the start and end fields, thus keeping the remaining 41 observations for forecasting. Click RUN. The results in Table 14.1 should appear on the screen. Similar results are also obtained using the Probit option. See Table 14.2.

Table 14.1: Probability of fertilizer use by Philippine farmers by Logit maximum likelihood estimation

```
                    The estimation method converged after 6 iterations
*******************************************************************************
 Dependent variable is FERUSE
 450 observations used for estimation from 1 to 450
*******************************************************************************
 Regressor               Coefficient        Standard Error        T-Ratio[Prob]
 ONE                       -1.5111               .23152            -6.5270[.000]
 CREDIT                   .2720E-3             .1330E-3             2.0453[.041]
 DMARKET                 -.026518              .021826             -1.2149[.225]
 HOURMEET                 .033875              .016509              2.0519[.041]
 IRSTAT                    1.7645               .22330              7.9021[.000]
 OWNER                     .48739               .22475              2.1686[.031]
*******************************************************************************
 Factor for the calculation of marginal effects =   .24416
 Maximized value of the log-likelihood function =-252.5396
 Akaike Information Criterion =-258.5396
 Schwarz Bayesian Criterion =-270.8673
 Hannan-Quinn Criterion =-263.3984
 Mean of FERUSE =   .42889
 Mean of fitted FERUSE =   .50667
 Goodness of fit =   .71778
 Pesaran-Timmermann test statistic =   9.3863[.000]
 Pseudo-R-Squared =   .17833
*******************************************************************************
```

Table 14.2: Probability of fertilizer use by Philippine farmers by Probit maximum likelihood estimation

```
                 The estimation method converged after 6 iterations
*******************************************************************************
 Dependent variable is FERUSE
 450 observations used for estimation from 1 to 450
*******************************************************************************
 Regressor              Coefficient        Standard Error         T-Ratio[Prob]
 ONE                       -.91034              .13245            -6.8733[.000]
 CREDIT                  .1522E-3             .7652E-4             1.9893[.047]
 DMARKET                 -.014658              .012261            -1.1955[.233]
 HOURMEET                 .018227             .0089651             2.0332[.043]
 IRSTAT                    1.0770               .13225             8.1440[.000]
 OWNER                     .28423               .13323             2.1334[.033]
*******************************************************************************
 Factor for the calculation of marginal effects =   .39184
 Maximized value of the log-likelihood function =-252.8353
 Akaike Information Criterion =-258.8353
 Schwarz Bayesian Criterion =-271.1631
 Hannan-Quinn Criterion =-263.6942
 Mean of FERUSE =   .42889
 Mean of fitted FERUSE =   .50667
 Goodness of fit =   .71778
 Pesaran-Timmermann test statistic =   9.3863[.000]
 Pseudo-R-Squared =   .17737
*******************************************************************************
```

Although the maximized value of the log-likelihood function is slightly larger for the Logit model, the two models fit the data equally well. The goodness of fit measure, computed as the proportion of observations with correctly predicted values of $FERUSE$, and the associated Pesaran-Timmermann test statistic, are the same for both models. In what follows we focus on the Logit estimates.

The estimated coefficients have the expected signs, with the variables $CREDIT$, $HOURMEET$, $IRSTAT$ and $OWNER$ having a positive effect, and the $DMARKET$ (the distance of the farm from the nearest market) variable having a negative effect on the probability of fertilizer use.[1]

To estimate the marginal effect of a unit change in, say, the $CREDIT$ variable, computed at sample means on the probability of the fertilizer use, you must multiply the factor 0.24416 given in the second part of Table 14.1 by the coefficient of $CREDIT$, using $0.24416 \times 0.00272 = 0.00066$ (see Section 6.19.2). Similarly, the marginal effect of the $DMARKET$ variable on probability of fertilizer use (evaluated at sample means) is given by $0.24416 \times (-0.026518) = -0.0065$.

The standard errors reported in Tables 14.1 and 14.2 allow you to carry out tests on the individual coefficients in β. To implement joint linear/non-linear tests on these coefficients you need to choose option 5 (Wald test of linear/non-linear restrictions) in the Post Estimation Menu (Logit Model). Suppose you wish to test the joint hypothesis that coefficients of the $CREDIT$ and $DMARKET$ variables are zero. Type

$$A2 = 0; \quad A3 = 0$$ RUN

[1] Notice that the magnitudes of the coefficients reported for the Logit and the Probit models in Tables 14.1 and 14.2 respectively, are not comparable. To make them comparable the coefficients estimated under the Probit option must be multiplied by 1.814.

The result of the test is given in Table 14.3, and the value of the Wald statistic for this test is equal to 5.1871, which is below the critical value of the χ^2 distribution with two degrees of freedom at the 95 per cent level.

A plot of actual values and fitted probabilities for the Logit specification is shown in Figure 14.1.

Table 14.3: Testing joint restrictions on the parameters of the probability model of fertilizer use by Philippine farmers

```
                Wald test of restriction(s) imposed on parameters
*******************************************************************************
 Based on Logit regression of FERUSE on:
 ONE              CREDIT           DMARKET          HOURMEET         IRSTAT
 OWNER
 450 observations used for estimation from 1 to 450
*******************************************************************************
 Coefficients A1 to A6 are assigned to the above regressors respectively.
 List of restriction(s) for the Wald test:
 A2=0;  A3=0
*******************************************************************************
 Wald Statistic                   CHSQ(2)=    5.1871[.075]
*******************************************************************************
```

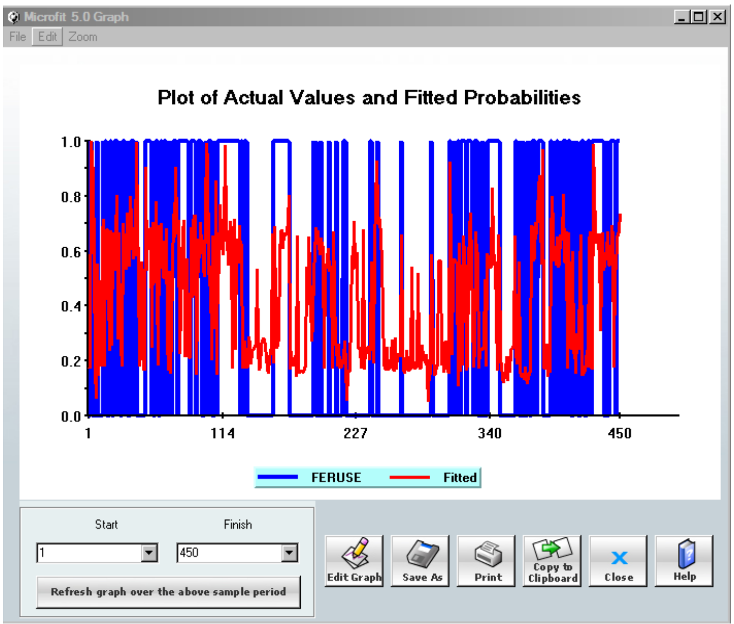


Figure 14.1: Actual values of $FERUSE$ and the fitted probabilities for the Logit specification

14.1.1 Forecasting with Probit/Logit models

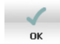
Forecasts of the probability of fertilizer use for the remaining 41 farmers in the sample can be computed using option 7 in the Logit Post Estimation Menu (see Section 6.19.4). When you choose this option you will be asked to choose the forecast sample. Click  to select all the remaining farmers in the sample. You will now be presented with the results in Table 14.4. The second part of this table gives a number of summary statistics, for the estimation and the prediction samples. As might be expected, the fitted values match the actual observations much better over the estimation sample as compared to the prediction sample. The Pesaran-Timmermann statistic is equal to -1.062 over the forecast sample, which is no longer statistically significant.

Table 14.4: Forecasting the probability of fertilizer use by Philippine farmers

```
                   Actual and Forecast Values of Regression
******************************************************************************
 Based on Logit regression of FERUSE on:
 ONE             CREDIT          DMARKET         HOURMEET        IRSTAT
 OWNER
 450 observations used for estimation from 1 to 450
******************************************************************************
 Observation             Actual     Predicted Probability    Predicted
  451                    1.0000           .67313              1.0000
  452                    1.0000           .79683              1.0000
  453                    1.0000           .92595              1.0000
  454                    1.0000           .17116                0.00
  455                      0.00           .27194                0.00
  456                    1.0000           .66279              1.0000
  457                    1.0000           .59973              1.0000
  458                    1.0000           .92810              1.0000
  459                    1.0000           .76485              1.0000
  460                      0.00           .57036              1.0000
  461                      0.00           .16929                0.00
  462                    1.0000           .63879              1.0000
  463                    1.0000           .27300                0.00
  464                    1.0000           .73589              1.0000
  465                    1.0000           .63212              1.0000
  466                      0.00           .59318              1.0000
  467                      0.00           .70850              1.0000
  468                    1.0000           .71838              1.0000
  469                    1.0000           .56605              1.0000
  470                      0.00           .55321              1.0000
  471                    1.0000           .67297              1.0000
  472                      0.00           .65805              1.0000
  473                    1.0000           .25842                0.00
  474                    1.0000           .17180                0.00
  475                    1.0000           .26025                0.00
  476                    1.0000           .28074                0.00
  477                    1.0000           .40079                0.00
  478                    1.0000           .34445                0.00
  479                    1.0000           .64146              1.0000
  480                    1.0000           .25012                0.00
  481                    1.0000           .63523              1.0000
  482                    1.0000           .25162                0.00
  483                    1.0000           .18879                0.00
  484                    1.0000           .17630                0.00
  485                    1.0000           .19336                0.00
  486                    1.0000           .22671                0.00
  487                    1.0000           .19940                0.00
  488                      0.00           .60793              1.0000
  489                    1.0000           .23203                0.00
  490                    1.0000           .19978                0.00
  491                      0.00           .18909                0.00
******************************************************************************
           Summary Statistics for Residuals and Prediction Errors
******************************************************************************
                        Estimation Period           Forecast Period
                          1 to 450              451 to 491
******************************************************************************
 Mean of FERUSE                           .42889             .78049
 Mean of predicted FERUSE                          .50667              .51220
 Goodness of fit                           .71778             .43902
 Pesaran-Timmermann Stat.     9.3863[.000]    -1.0624[.288]
******************************************************************************
```

14.2 Lesson 14.2: Fertilizer use model estimated over a sub-sample of farmers

Microfit readily allows you to estimate regression or Probit/Logit models over a sub-sample of observations selected according to a particular set of criteria. For example, suppose you wish to estimate the probability of fertilizer use only over the sample of farmers that use irrigation and own their own farms. Move to the Commands and Data Transformations in the Process window and generate the variable

$$X = 1 - \mathbf{SIGN}(IRSTAT) * \mathbf{SIGN}(OWNER)$$

It is clear that

$$\begin{aligned} X &= 0 \quad \text{if} \quad IRSTAT = 1 \quad \text{and} \quad OWNER = 1 \\ X &= 1 \quad \text{if either} \quad IRSTAT \quad \text{or} \quad OWNER \quad \text{is equal to zero} \end{aligned}$$

Now reorder the observations so that all farmers that own their farm and use irrigation are put at the top of the workspace. This can be done by issuing the **REORDER** command (see Section 4.4.21)

$$\mathbf{REORDER} \quad X; \quad N = \mathbf{SUM}(1 - X)$$

The number of farmers that own their farm and use irrigation is given by N (in the present example $N = 96$). The first 96 observations of the reordered data set represent the observations on farmers that own their land and use irrigation. To see that this is in fact the case use the **LIST** command to list the observations.[2]

To estimate the Logit model for this sub-sample of farmers, move to the Single Equation Estimation Menu, and type

$$FERUSE \quad ONE \quad CREDIT \quad DMARKET \quad HOURMEET$$

Enter 1 and 96 into the start and end fields, select the Logit model from option 7 in the Single Equation Estimation Menu, and click . The results in Table 14.5 should now appear on the screen. For this sub-sample only the credit variable is significant. But the Pesaran-Timmermann statistic is equal to 1.70 and does not reject the hypothesis that the fitted and actual values for this sub-sample are not related at the 5 per cent level of significance.

[2] Notice that you can use the **RESTORE** command to restore the original ordering of your observations. See Section 4.4.22.

Table 14.5: Probability of fertilizer use by a sub-sample of Philippine farmers owning their farms and using irrigation

```
                         Logit Maximum Likelihood Estimation

                    The estimation method converged after 6 iterations
******************************************************************************
 Dependent variable is FERUSE
 96 observations used for estimation from 1 to 96
******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 ONE                        .86021             .38301            2.2459[.027]
 CREDIT                   .0020980           .9248E-3            2.2686[.026]
 DMARKET                  -.077630            .066965           -1.1593[.249]
 HOURMEET                  .015895            .022009            .72222[.472]
******************************************************************************
 Factor for the calculation of marginal effects =    .15999
 Maximized value of the log-likelihood function = -49.8423
 Akaike Information Criterion = -53.8423
 Schwarz Bayesian Criterion = -58.9710
 Hannan-Quinn Criterion = -55.9154
 Mean of FERUSE =    .73958
 Mean of fitted FERUSE =    .98958
 Goodness of fit =    .75000
 Pesaran-Timmermann test statistic =    1.7030[.089]
 Pseudo-R-Squared =   .094685
******************************************************************************
```

14.3 Exercises in Logit/Probit estimation

14.3.1 Exercise 14.1:

Use the observations in the file PHIL.FIT to estimate a Probit model of the probability of fertilizer use for the sub-sample of farmers who reside within the two-miles radius of the market (see Lessons 14.1 and 14.2).

14.3.2 Exercise 14.2:

Read the special *Microfit* file PTMONTH.FIT and run the batch file on it to generate the variables $INPT$, $ERSP$, YSP, $DI11$, $DIP12$, and $PI12$. See Lesson 10.11 for more details. Use the **SIGN(•)** function to construct a dummy variable that takes the value of unity if $ERSP$ (the excess returns on $SP500$) is unity and zero otherwise. Then estimate Logit and Probit regressions of this $(1, 0)$ variable on $INPT$, $YSP(-1)$, $PI12(-2)$, $DI11(-1)$, and $DIP12(-2)$ over the period 1954(1)-1992(12). Compare these results with the OLS estimates in Table 11.16. Comment on the relative merits of the two estimation approaches.

Re-estimate the Probit/Logit regressions over the period 1954(1)-1992(11), and forecast the probability of a negative excess return in 1992(12).

Chapter 15

Lessons in VAR Modelling

Lessons in this chapter demonstrate some of the main features of the unrestricted VAR options in *Microfit.* The relevant menus and options for these lessons are described in Section 7.4, and the related econometric methods are briefly reviewed in 22.4.

The lessons are based on a trivariate VAR model in output growths of USA, Japan and Germany. These series have been already analysed in some detail in Canova (1995) Section 8, where he provides some empirical justification for modelling output growths rather than output levels. Canova's study, however, covers the period 1955(1)-1986(4), while the data that we will be using cover the period 1963(1)-1993(4).

The file G7GDP.FIT contains quarterly observations on GDP(GNP) at constant 1990 prices for the G7 countries, namely Canada, France, Germany, Italy, Japan, USA and the UK. This file also contains data on patents granted by the US Patent Office to all the G7 countries.[1]

15.1 Lesson 15.1: Selecting the order of the VAR

In this lesson we consider the problem of selecting the order of the trivariate VAR model in the output growths of USA, Japan and Germany.

The special *Microfit* file G7GDP.FIT contains quarterly observations on 83 different variables over the period 1963(1)-1993(4). Read this file into *Microfit.* Check that the following variables are in the workspace (click 

)

$DLYUSA$	US output growth
$DLYJAP$	Japan's output growth
$DLYGER$	Germany's output growth
$CONST$	Vector of ones

[1]The source of the output data is the International Financial Statistics, IMF. The output series for Japan and Germany refer to GNP at constant 1990 prices. The quarterly patent data has been compiled by Silvia Fabiani from primary sources (the file PATSIC supplied by the US Patent Office). We are grateful to her for providing us with this data set. For more details see Fabiani (1995), and the file G7READ.ME in the tutorial directory.

Open the System Estimation Menu (the Multivariate Menu on the main menu bar) and choose option 1 (See Section 7.3). In the Commands and Data Transformation box type

$$DLYUSA \quad DLYJAP \quad DLYGER \quad \& \quad CONST$$

This specifies an unrestricted VAR model in the output growths of USA, Japan and Germany, and includes a vector of intercepts in the model. For the estimation period type

$$1963Q1 \quad 1992Q4$$

and keep the quarterly observations in 1993 for forecasting purposes. You will now be asked to specify the order of the VAR. Since the aim is to select an 'optimal' order for the VAR, it is important that at this stage a high enough order is selected such that one is reasonably confident that the optimal order will not exceed it. In the case of the present application we recommend using 6 as the maximum order for the VAR. Therefore, enter 6 into the Order of the VAR field. Click 

. You will now be presented with the Unrestricted VAR Post Estimation Menu (see Section 7.4.1). Choose option 4 to move to the VAR Hypothesis Testing Menu (see Section 7.4.3). Option 1 in this menu presents you with the results in Table 15.1. All the seven $VAR(p)$, $p = 0, 1, 2, ..., 6$, models are estimated over the same sample period, namely 1964(3)-1992(4), and as to be expected the maximized values of the log-likelihood function given under the column headed LL increase with p. However, the Akaike and the Schwarz criteria select the orders 1 and 0, respectively. The log-likelihood ratio statistics (whether or not adjusted for small samples) reject order 0, but do not reject a VAR of order 1. In the light of these we choose the $VAR(1)$ model. Notice that it is quite usual for the SBC to select a lower order VAR as compared with the AIC.

Table 15.1: Selecting the order of a trivariate VAR model in output growths

Test Statistics and Choice Criteria for Selecting the Order of the VAR Model

Based on 114 observations from 1964Q3 to 1992Q4. Order of VAR = 6
List of variables included in the unrestricted VAR:
DLYUSA DLYJAP DLYGER
List of deterministic and/or exogenous variables:
CONST

Order	LL	AIC	SBC	LR test	Adjusted LR test
6	1128.4	1071.4	993.3935	------	------
5	1125.2	1077.2	1011.5	CHSQ(9)= 6.4277[.696]	5.3564[.802]
4	1120.4	1081.4	1028.1	CHSQ(18)= 15.8583[.602]	13.2152[.779]
3	1112.8	1082.8	1041.8	CHSQ(27)= 31.1322[.266]	25.9435[.522]
2	1108.1	1087.1	1058.4	CHSQ(36)= 40.5268[.277]	33.7724[.575]
1	1101.5	1089.5	1073.1	CHSQ(45)= 53.7352[.175]	44.7793[.481]
0	1084.2	1081.2	1077.1	CHSQ(54)= 88.2988[.002]	73.5824[.039]

AIC=Akaike Information Criterion SBC=Schwarz Bayesian Criterion

Having chosen the order of the VAR it is prudent to examine the residuals of individual equations for serial correlation. Click 

, backtrack, and estimate a $VAR(1)$ model over the period 1964(3)-1992(4). Then choose option 1 in the Unrestricted VAR Post Estimation Menu to inspect the results on individual equations in the VAR. Tables 15.2 to 15.4 show the regression results for the US, Japan and Germany, respectively. There is no evidence of residual serial correlation in the case of the US and Germany's output equations, but there is a statistically significant evidence of residual serial correlation in the case of Japan's output equation. There is also important evidence of departures from normality in the case of output equations for the USA and Japan. A closer examination of the residuals of these equations suggest considerable volatility during early 1970s as a result of the abandonment of the Bretton Wood system and the quadrupling increase in oil prices. Therefore, it is likely that the remaining serial correlation in the residuals of Japan's output equation may be due to these unusual events. Such a possibility can be handled by introducing a dummy variable for the oil shock in the VAR model (see Lesson 15.2).

Table 15.2: US output growth equation

```
           OLS estimation of a single equation in the Unrestricted VAR
*******************************************************************************
 Dependent variable is DLYUSA
 114 observations used for estimation from 1964Q3 to 1992Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 DLYUSA(-1)                 .28865             .095475            3.0233[.003]
 DLYJAP(-1)                .029389             .080403            .36552[.715]
 DLYGER(-1)                .072501             .085957            .84346[.401]
 CONST                    .0039865            .0014390            2.7704[.007]
*******************************************************************************
 R-Squared                     .10155   R-Bar-Squared                   .077044
 S.E. of Regression          .0092075   F-Stat.    F(3,110)     4.1442[.008]
 Mean of Dependent Variable  .0068049   S.D. of Dependent Variable     .0095841
 Residual Sum of Squares     .0093256   Equation Log-likelihood        374.6786
 Akaike Info. Criterion      370.6786   Schwarz Bayesian Criterion     365.2062
 DW-statistic                  2.0058   System Log-likelihood            1101.5
*******************************************************************************


                              Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version         *           F Version        *
*******************************************************************************
*                     *                           *                            *
* A:Serial Correlation*CHSQ(4)  =    2.5571[.634]*F(4,106)     =    .60806[.658]*
*                     *                           *                            *
* B:Functional Form   *CHSQ(1)  =    .28508[.593]*F(1,109)     =    .27326[.602]*
*                     *                           *                            *
* C:Normality         *CHSQ(2)  =    9.0063[.011]*       Not applicable        *
*                     *                           *                            *
* D:Heteroscedasticity*CHSQ(1)  =    2.1578[.142]*F(1,112)     =    2.1609[.144]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

Table 15.3: Japanese output growth equation

```
         OLS estimation of a single equation in the Unrestricted VAR
*******************************************************************************
 Dependent variable is DLYJAP
 114 observations used for estimation from 1964Q3 to 1992Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 DLYUSA(-1)               .065140             .10992              .59262[.555]
 DLYJAP(-1)                .21958            .092566              2.3721[.019]
 DLYGER(-1)                .23394            .098960              2.3640[.020]
 CONST                   .0081150           .0016567              4.8984[.000]
*******************************************************************************
 R-Squared                     .12935   R-Bar-Squared                   .10560
 S.E. of Regression            .010600  F-Stat.    F(3,110)      5.4473[.002]
 Mean of Dependent Variable    .013128  S.D. of Dependent Variable     .011209
 Residual Sum of Squares       .012361  Equation Log-likelihood       358.6189
 Akaike Info. Criterion       354.6189  Schwarz Bayesian Criterion    349.1465
 DW-statistic                   2.1004  System Log-likelihood           1101.5
*******************************************************************************

                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version         *           F Version       *
*******************************************************************************
*                     *                           *                           *
* A:Serial Correlation*CHSQ(4)  =  13.5014[.009]*F(4,106)     =   3.5601[.009]*
*                     *                           *                           *
* B:Functional Form   *CHSQ(1)  =   1.3036[.254]*F(1,109)     =   1.2608[.264]*
*                     *                           *                           *
* C:Normality         *CHSQ(2)  =  39.9038[.000]*       Not applicable        *
*                     *                           *                           *
* D:Heteroscedasticity*CHSQ(1)  =   2.8502[.091]*F(1,112)     =   2.8720[.093]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
```

Table 15.4: Germany's output growth equation

```
         OLS estimation of a single equation in the Unrestricted VAR
*******************************************************************************
 Dependent variable is DLYGER
 114 observations used for estimation from 1964Q3 to 1992Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 DLYUSA(-1)                .19307             .10327              1.8695[.064]
 DLYJAP(-1)                .23045            .086968              2.6499[.009]
 DLYGER(-1)              -.031787            .092975             -.34188[.733]
 CONST                   .0025928           .0015565              1.6658[.099]
*******************************************************************************
 R-Squared                     .10562   R-Bar-Squared                  .081232
 S.E. of Regression           .0099594  F-Stat.    F(3,110)      4.3303[.006]
 Mean of Dependent Variable   .0067337  S.D. of Dependent Variable     .010390
 Residual Sum of Squares       .010911  Equation Log-likelihood       365.7306
 Akaike Info. Criterion       361.7306  Schwarz Bayesian Criterion    356.2582
 DW-statistic                   2.0799  System Log-likelihood           1101.5
*******************************************************************************

                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version         *           F Version       *
*******************************************************************************
*                     *                           *                           *
* A:Serial Correlation*CHSQ(4)  =   7.1440[.128]*F(4,106)     =   1.7717[.140]*
*                     *                           *                           *
* B:Functional Form   *CHSQ(1)  =   .30999[.578]*F(1,109)     =   .29721[.587]*
*                     *                           *                           *
* C:Normality         *CHSQ(2)  =   .56351[.754]*       Not applicable        *
*                     *                           *                           *
* D:Heteroscedasticity*CHSQ(1)  =   .52835[.467]*F(1,112)     =   .52150[.472]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

15.2 Lesson 15.2: Testing for the presence of oil shock dummies in output equations

Consider the VAR model of output growths of US, Germany and Japan in Lesson 15.1, and suppose we are interested in testing the significance of an oil shock dummy variable, which takes the value of unity in the first quarter of 1974 and zeros elsewhere, in that model.

Load the file G7GDP.FIT, and in the Commands and Data Transformations box in the process window type

$D74 = 0;$ **SAMPLE** $74Q1 \quad 74Q4;$ $D74 = 1;$
SAMPLE $63Q1 \quad 93Q4$

You should now see the dummy variable $D74$ among the variables in the variable list (click to view it). Choose the unrestricted VAR option in the System Estimation Menu, and specify the following augmented VAR model:

$$DLYUSA \quad DLYJAP \quad DLYGER \quad \& \quad CONST \quad D74$$

For the sample period choose

$$1964Q1 \quad 1992Q4$$

For the order of the VAR enter 1. Click and move to the VAR Hypothesis Testing Menu (see Section 7.4.3). Select option 2 and type

$$D74$$

The results in Table 15.5 should now appear on the screen.

Table 15.5: Testing for the effect of the oil shock in the VAR model

```
     LR Test of Deletion of Deterministic/Exogenous Variables in the VAR
*******************************************************************************
 Based on  116 observations from 1964Q1 to 1992Q4. Order of VAR = 1
 List of variables included in the unrestricted VAR:
 DLYUSA          DLYJAP          DLYGER
 List of deterministic and/or exogenous variables:
 CONST           D74
 Maximized value of log-likelihood =    1124.0
*******************************************************************************
 List of variables included in the restricted VAR:
 DLYUSA          DLYJAP          DLYGER
 List of deterministic and/or exogenous variables:
 CONST
 Maximized value of log-likelihood =    1119.7
*******************************************************************************
 LR test of restrictions, CHSQ(3) =   8.6083[.035]
*******************************************************************************
```

As can be seen from this table, the log-likelihood ratio statistic for testing the deletion of the oil shock dummy from all three output equations is 8.61, which is statistically significant at the 3.5 per cent level.

To check the significance of the oil shock dummy in individual output equations you need to choose option 1 in the Unrestricted VAR Post Estimation Menu. The dummy variable is significant in Japan's output growth equation, marginally significant in the US output growth equation, and not statistically significant in Germany's equation. Also note that the inclusion of the dummy has reduced the significance of the residual serial correlation in Japan's output equation, but has not eliminated the problem. Therefore, there may be other factors (such as non-linear effects) that should be taken into account. Another possibility is to try a higher order for the VAR model; although this course is not recommended by the model selection criteria or the likelihood ratio test statistics.

15.3 Lesson 15.3: International transmission of output shocks

One important issue in the analysis of international business cycle is the extent to which output shocks are transmitted from one country to another. In this lesson we examine this issue by Granger non-causality tests applied to the trivariate VAR in the US, Japanese and German output growths.

Read file G7GDP.FIT into *Microfit* and go through the steps in Lesson 15.2, and specify the augmented $VAR(1)$ model

$$DLYUSA \quad DLYJAP \quad DLYGER \quad \& \quad CONST \quad D74$$

to be estimated over the period 1964(1)-1992(4). For the order of the VAR choose 1. Then choose option 3 in the VAR Hypothesis Testing Menu (see Section 7.4.3). You will be asked to list the sub-set of variables with respect to which you wish to carry out the block non-causality tests. Type

$$DLYJAP \quad DLYGER$$ 

to test for the non-causality of the Japanese and German output growths in the US output equation. You should now see the test results on the screen, shown in Table 15.6.

Table 15.6: Granger non-causality test of US output growth with respect to the output growths of Germany and Japan

```
                LR Test of Block Granger Non-Causality in the VAR
*******************************************************************************
 Based on  116 observations from 1964Q1 to 1992Q4. Order of VAR = 1
 List of variables included in the unrestricted VAR:
 DLYUSA          DLYJAP          DLYGER
 List of deterministic and/or exogenous variables:
 CONST           D74
 Maximized value of log-likelihood =     1124.0
*******************************************************************************
 List of variable(s) assumed to be non-causal under the null hypothesis:
 DLYJAP          DLYGER
 Maximized value of log-likelihood =     1123.7
*******************************************************************************
 LR test of block non-causality, CHSQ(2) =    .50437[.777]
*******************************************************************************
 The above statistic is for testing the null hypothesis that the coefficients
 of the lagged values of:
 DLYJAP          DLYGER
 in the block of equations explaining the variable(s):
 DLYUSA
 are zero. The maximum order of the lag(s) is 1.
*******************************************************************************
```

The log-likelihood ratio statistic for this test is equal to 0.50, which is asymptotically distributed as a χ^2 variate with 2 degrees of freedom, and is clearly not significant statistically. Carrying out a similar exercise for the other output equations we obtain the LR statistic of 4.93 [0.085] when testing for the non-causality of the US and Germany's output growth in the Japanese output equation, and the LR statistic of 11.975 [0.003] for testing the non-causality of the US and Japanese output growth in the Germany's output equation. The figures in square brackets refer to rejection probabilities. Therefore, if there is any transmission of output shocks between these three countries, it seems that it goes from the US to other two countries rather than the reverse.

15.4 Lesson 15.4: Contemporaneous correlation of output shocks

Another aspect of the international transmission of output shocks is the extent to which shocks in different output equations are contemporaneously correlated. Again using the augmented $VAR(1)$ model of output growths of the US, Japan, and Germany, in this lesson we test the hypothesis that

$$H_0 : \sigma_{12} = \sigma_{13} = \sigma_{23} = 0$$

against the alternative that

$$H_1 : \sigma_{12} \neq 0, \quad \sigma_{13} \neq 0, \quad \sigma_{23} \neq 0$$

where σ_{ij}, denotes the contemporaneous covariance between the shocks in the output equations of countries i and j.

One possible method of testing the above hypothesis is to compute the log-likelihood ratio statistic

$$\mathfrak{LR}\,(H_0\,|H_1\,) = 2\,(LL_U - LL_R)$$

where LL_U and LL_R are the maximized values of the log-likelihood function under H_1 (the unrestricted model), and under H_0 (the restricted model), respectively. See also Section 22.2.2.

To compute LL_U for the present application, follow the steps in Lesson 15.2 and estimate the augmented $VAR(1)$ model in

$$\mathbf{x}_t = (DLYUSA, DLYJAP, DLYGER)$$

augmented with the variables $\mathbf{w}_t = (CONST, D74)$, over the period 1964(1)-1992(4). Choose option 1 in The Unrestricted VAR Post Estimation Menu, and press to see the regression results for the US output equation. The value of LL_U is given by the value of the 'system log-likelihood' shown in the bottom right hand-corner of the result table, namely

$$LL_U = 1124.0$$

To compute the restricted log-likelihood value, LL_R, we first note that under H_0

$$LL_R = LL_{US} + LL_{GER} + LL_{JAP}$$

where LL_{US}, LL_{GER} and LL_{JAP} are the single-equation log-likelihood values of the output equations for the US, Germany and Japan, respectively. They are readily computed using the OLS Option. For example, to compute LL_{US} choose option 1 in the Single Equation Estimation Menu (Univariate Menu: see Section 6.4). Click to clear the box editor and then type

$$DLYUSA \quad CONST \quad D74 \quad DLYUSA(-1) \quad DLYJAP(-1) \quad DLYGER(-1)$$

Enter the start and end dates $1964Q1$ and $1992Q4$, and click . You should see the OLS regression results on the screen. LL_{US} is given by the value of the 'Equation Log-likelihood' in this result table; that is, $LL_{US} = 381.95$. Now click , backtrack, change the dependent variable from $DLYUSA$ to $DLYGER$ and run the regression by the OLS to obtain $LL_{GER} = 373.31$. Similarly we have $LL_{JAP} = 367.02$. Therefore

$$\begin{aligned} LR\,(H_0 : H_1) &= 2(1124 - 381.95 - 373.31 - 367.02) \\ &= 3.44 \end{aligned}$$

which is asymptotically distributed as a χ^2 variate with 3 degrees of freedom. The 95 per cent critical value of the χ^2 distribution with 3 degrees of freedom is 7.81. Therefore, the null hypothesis that the shocks in different output equations are contemporaneously uncorrelated cannot be rejected.

The estimates of σ_{ij}, obtainable using option 2 in the Unrestricted VAR Post Estimation Menu, also corroborate this finding. These estimates are reproduced in Table 15.7 and show that the estimates $\hat{\sigma}_{12} = 0.1014 \times 10^{-4}$, $\hat{\sigma}_{13} = 0.6001 \times 10^{-5}$ and $\hat{\sigma}_{23} = 0.1282 \times 10^{-4}$ are about 1/10 of the estimated error variances, given by the diagonal elements of the 3×3 matrix in Table 15.7.

Table 15.7: Granger non-causality test of US output growth with respect to the output growths of Germany and Japan

	DLYUSA	DLYJAP	DLYGER
DLYUSA	.8447E-4	.1014E-4	.6001E-5
DLYJAP	.1014E-4	.1093E-3	.1282E-4
DLYGER	.6001E-5	.1282E-4	.9803E-4

15.5 Lesson 15.5: Forecasting output growths using the VAR

Here we use the augmented $VAR(1)$ model estimated in the previous lessons to compute multivariate, multi-step ahead forecasts of output growths.

Read the file G7GDP.FIT and follow the steps in Lesson 15.2 to estimate the $VAR(1)$ model of output growths of the US, Japan and Germany over the period 1964(1)-1992(4). Then choose option 5 in the Unrestricted VAR Post Estimation Menu, and when prompted press the  button to select 1993(4) as the final quarter of the forecast period. You will then be asked to choose the growth rate that you wish to forecast. Press  to choose the US output growth, and then select 'Level of $DLYUSA$' rather than 'Change in $DLYUSA$' to see the forecasts of the levels of output growth. You should now see the Multivariate Forecast Menu on the screen (See Section 7.4.4). Click  to see the forecasts and the forecast errors for the four quarters of 1993. These forecast results are reproduced in Table 15.8. As can be seen from the summary statistics, the size of the forecast errors and the in-sample residuals are very similar. A similar picture also emerges by plotting in-sample fitted values and out-of-sample forecasts (see Figure 15.1). It is, however, important to note that the US growth experience in 1993 may not have been a stringent enough test of the forecast performance of the VAR, as the US output growths have been positive in all the four quarters. A good test of forecast performance is to see whether the VAR model predicts the turning points of the output movements.

Similarly, forecasts of output growths for Japan and Germany can also be computed. For Japan the root mean sum of squares of the forecast errors over the 1993(1)-1993(4) period turned out to be 1.48 per cent, which is slightly higher than the value of 1.02 per cent obtained for the root mean sum of squares of residuals over the estimation period. It is also worth noting that the growth forecasts for Japan miss the two negative quarterly output growths that occurred in the second and fourth quarters of 1993.

A similar conclusion is also reached in the case of output growth forecasts for Germany.

Table 15.8: Multivariate dynamic forecasts for US output growth ($DLYUSA$)

```
           Multivariate dynamic forecasts for the level of DLYUSA
                  Unrestricted Vector Autoregressive Model
*******************************************************************************
 Based on  116 observations from 1964Q1 to 1992Q4. Order of VAR = 1
 List of variables included in the unrestricted VAR:
 DLYUSA          DLYJAP          DLYGER
 List of deterministic and/or exogenous variables:
 CONST           D74
*******************************************************************************
 Observation            Actual            Prediction             Error
  1993Q1               .0019499             .011379            -.0094288
  1993Q2               .0047025             .0084160           -.0037135
  1993Q3               .0070727             .0076889           -.6162E-3
  1993Q4                .016841             .0075384            .0093022
*******************************************************************************

            Summary Statistics for Residuals and Forecast Errors
*******************************************************************************
                     Estimation Period             Forecast Period
                     1964Q1 to 1992Q4             1993Q1 to 1993Q4
*******************************************************************************
 Mean                       -.0000                     -.0011141
 Mean Absolute              .0067166                    .0057652
 Mean Sum Squares           .8082E-4                    .4740E-4
 Root Mean Sum Squares      .0089902                    .0068848
*******************************************************************************
```

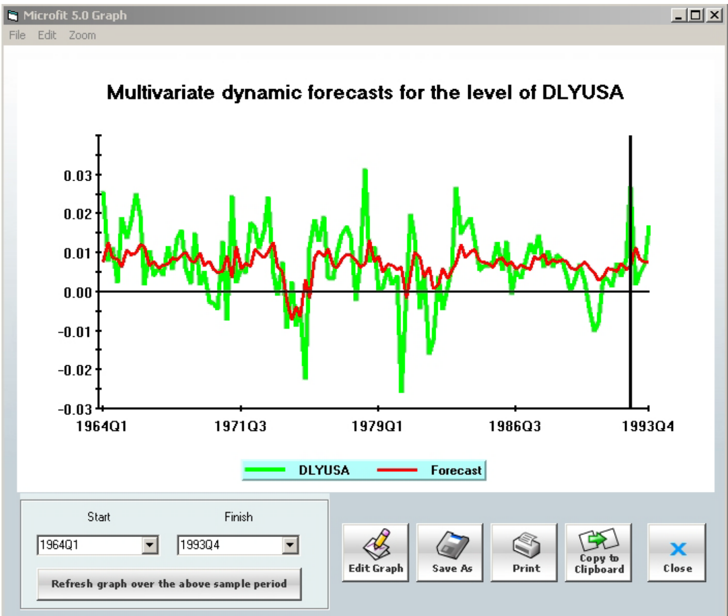


Figure 15.1: Multivariate dynamic forecasts of US output growth ($DLYUSA$)

15.6 Lesson 15.6: Impulse responses of the effects of output growth shocks

In this lesson we show how to use *Microfit* to compute/plot impulse responses (and forecast error variance decomposition) for the effect of unit shocks (equal to one standard error) to the US output growth equation on the growth of outputs in Germany and Japan within the trivariate VAR model analysed in the earlier lessons. Since in Lesson 15.4 we could not reject the hypothesis that the variance matrix of the errors in the VAR model is diagonal, we do not expect the orthogonalized and the generalized impulse responses (and the associated forecast error variance decompositions) to be very different. Also, in view of the results of Lesson 15.3, we estimate the model with the US output growth as the first variable in the VAR.

Read the file G7GDP.FIT and choose the Unrestricted VAR option in the System Estimation Menu and set up the VAR model in the box editor that appears on the screen by typing

$$DLYUSA \quad DLYGER \quad DLYJAP \quad \& \quad CONST \quad D74$$

Then select the sample period 1964(1)-1993(4), and a VAR of order 1. *Microfit* carries out the necessary computations and presents you with the Unrestricted VAR Post Estimation Menu (see Section 7.4.1). Choose option 3 to move to the Unrestricted VAR Dynamic Response Analysis Menu. (See Section 7.4.2). Initially choose option 1 to compute orthogonalized impulse responses, and then select the variable/equation $DLYUSA$ to shock. Since there is little persistence in output growth shocks it is advisable to select a short horizon for the impulse responses. We recommend choosing 25 when you are asked to specify the 'horizon for impulse responses'. The results are reproduced in Table 15.9.

Table 15.9: Orthogonalized impulse responses to one SE shock in the equation for US output growth ($DLYUSA$) using an unrestricted vector autoregressive model

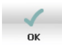
```
Orthogonalised Impulse Responses to one SE shock in the equation for DLYUSA
                    Unrestricted Vector Autoregressive Model
*******************************************************************************
Based on  120 observations from 1964Q1 to 1993Q4. Order of VAR = 1
List of variables included in the unrestricted VAR:
DLYUSA          DLYGER          DLYJAP
List of deterministic and/or exogenous variables:
CONST           D74
*******************************************************************************
```

Horizon	DLYUSA	DLYGER	DLYJAP
0	.0091174	.8103E-3	.9367E-3
1	.0022248	.0015059	.8355E-3
2	.6358E-3	.4971E-3	.6469E-3
3	.1960E-3	.2403E-3	.2981E-3
4	.6809E-4	.9687E-4	.1354E-3
5	.2525E-4	.4121E-4	.5773E-4
6	.9870E-5	.1704E-4	.2441E-4
7	.3960E-5	.7096E-5	.1020E-4
8	.1614E-5	.2941E-5	.4248E-5
9	.6632E-6	.1220E-5	.1764E-5
10	.2737E-6	.5059E-6	.7322E-6
11	.1132E-6	.2097E-6	.3037E-6
12	.4686E-7	.8693E-7	.1259E-6
13	.1941E-7	.3604E-7	.5220E-7
14	.8043E-8	.1494E-7	.2164E-7
15	.3333E-8	.6191E-8	.8969E-8
16	.1382E-8	.2566E-8	.3718E-8
17	.5726E-9	.1064E-8	.1541E-8
18	.2373E-9	.4409E-9	.6387E-9
19	.0000	.1827E-9	.2647E-9
20	.0000	.0000	.1097E-9
21	.0000	.0000	.0000
22	.0000	.0000	.0000
23	.0000	.0000	.0000
24	.0000	.0000	.0000
25	.0000	.0000	.0000

```
*******************************************************************************
```

As you can see, the effect of a unit shock to the US output growth has only a small impact on output growths of Germany and Japan, and this effect dies out very quickly with the forecast horizon.

To see the plot of impulse responses click 

to move to the Impulse Response Results Menu, and choose the graph option. You will be presented with a list of the jointly determined variables in the VAR. Select all three variables, $DLYUSA$, $DLYGER$ and $DLYJAP$, and click 

. Figure 15.2 should now appear on the screen.

As you can see, the impact of US output growth shocks on Germany and Japan is small on impact and generally tends to die out very quickly.

Consider now the generalized impulse responses of the effect of a unit shocks to $DLYUSA$, the US output growth. Since $DLYUSA$ is the first variable in the VAR, the orthogonalized and generalized impulse responses will be identical. To see that this is in fact the case choose option 2 in the Unrestricted VAR Dynamic Response Analysis Menu and select $DLYUSA$ to shock and 25 for the forecast horizon, and compare the results that appear on the screen with those in Table 15.9. This is not, however, the case if you choose to shock Japan's output growth equation. For example, in the case of the orthogonalized responses the effects of shocking Japan's output growth on the US and Germany's output growths on *impact* (at zero horizon) are zero, (by construction), but the corresponding generalized impulse responses are 0.08 and 0.12 per cent, respectively. However, due to the almost diagonal nature of the variance matrix of the shocks, the two impulse responses are not significantly different from one another in this particular application.

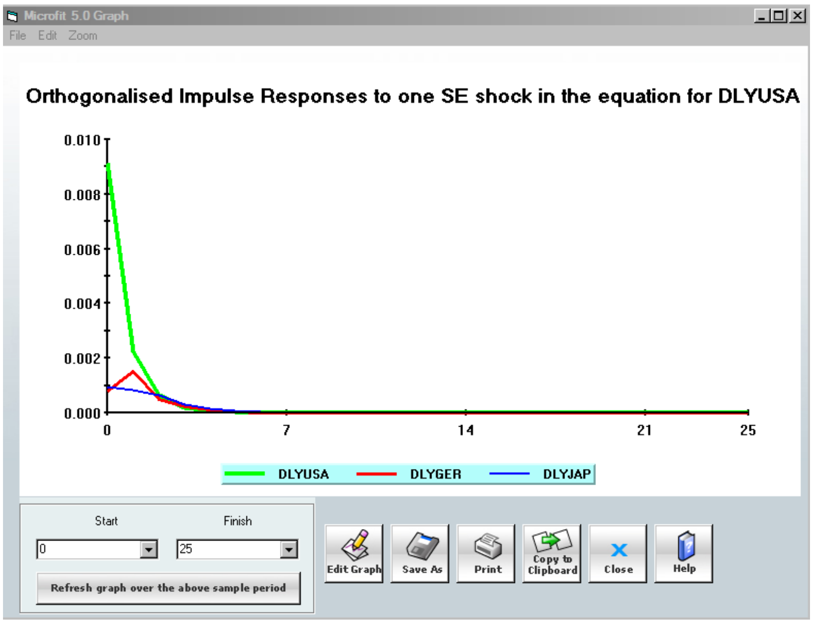


Figure 15.2: Orthogonalised impulse responses to one SE shock in the equation for $DLYUSA$

15.7 Exercises in VAR modelling

15.7.1 Exercise 15.1

Carry out Lessons 15.1 to 15.5 using a VAR model in output levels, and compare your findings with those obtained using the VAR model in growth rates.

15.7.2 Exercise 15.2

The file G7GDP.FIT (on the tutorial directory) also contains patent data for the G7 countries. Repeat the Lessons 15.1 to 15.5 using rates of change of patents granted by the US Patent Office to the US, Japan and Germany, instead of the output growth rates. Discuss the similarity and the differences between the two sets of results.

15.7.3 Exercise 15.3

Again using the file G7GDP.FIT, estimate an augmented VAR model in output growths and rates of change of patents granted in the case of the US, Japan and Germany.

- Select an optimal order for this 6-variable VAR model.
- Test the statistical significance of the $D74$ oil shock dummy variable in this combined VAR model.

- Test the Granger non-causality of output growths with respect to the rates of change of the number of patents granted and vice versa.

Chapter 16

Lessons in Cointegration Analysis

The lessons in this chapter are concerned with single equation and multiple equation approaches to cointegration analysis. They show how *Microfit* can be used to test for cointegration under a variety of circumstances, how to estimate long-run relations using both single equation approaches, such as the Phillips and Hansen (1990) Fully-Modified OLS estimators, and the Autoregressive-Distributed Lag ($ARDL$) estimators discussed in Pesaran and Shin (1999); and the system approaches such as the Full Maximum Likelihood procedure applied to the underlying VAR model. The lessons in this chapter also demonstrate *Microfit* capabilities in the areas of impulse response analysis and forecasting using cointegrating VAR models.

The literature on cointegrating VAR is vast and growing. Excellent accounts of the early developments can be found in Watson (1994), and Hamilton (1994) Chapter 19. The basic cointegrating VAR model underlying most of the lessons in this chapter is set out in Chapter 22, where further references to the literature can also be found. The $ARDL$ approach is described in Section 21.19.

These lessons deal with the problem of testing for cointegration, and cover both the residual-based methods proposed by Engle and Granger (1987), and a generalization of Johansen (1988) and Johansen (1991) Full Information Maximum Likelihood ($FIML$) approach. Two main procedures are currently used to test for cointegration. One is the residual-based ADF method proposed by Engle and Granger (1987), and the other is the Maximum Likelihood approach (Johansen (1988), Johansen (1991) and Johansen (1995)). There are also other procedures such as the common stochastic trends approach of Stock and Watson (1988), the auxiliary regression procedure of Park (1992), and variants of the residual based approach proposed by Phillips and Ouliaris (1990), extended by Hansen (1992). For a review of these tests see Watson (1994), and Hamilton (1994) Chapter 19. We shall also consider a new approach to testing for the existence of long-run relations when it is not known whether the underlying regressors are $I(1)$ or $I(0)$. This testing method is developed in Pesaran, Shin, and Smith (2001). For reasons that will become clear from Lesson 16.5, we shall refer to this as the *bounds test*.

With the exception of the $ARDL$ approach to cointegration analysis and the related bounds test, the cointegrating options in *Microfit* presume that the variables under consideration are first-difference stationary (or are integrated of order 1). On the problem of how to test for the order of integration of the variables, see the commands **ADF**, **ADF_GLS,** **ADF_MAX** and **ADF_W** (Section 4.4.2) and Lessons 12.1 and 12.2.

16.1 Lesson 16.1: Testing for cointegration when the cointegrating coefficients are known

It is often the case that economic theory suggests certain variables are cointegrated with a *known* cointegrating vector. Examples are the 'great ratios' and the 'purchasing power parity' (PPP) relations. In this lesson we use the extensive historical data analyzed by Alogoskoufis and Smith (1991b) to test the hypothesis that wages (W), prices (P), output (Y), and employment (E), all measured in logarithms, are cointegrated with coefficients $(+1, -1, -1, +1)$. Such a cointegrating relation implies that the logarithm of the share of wages in output has been mean-reverting to a constant value over time. Similarly, real wages, $WP = W - P$, and labour productivity, $YE = Y - E$, should cointegrate with coefficients $(1, -1)$. In this lesson we test this hypothesis using univariate procedures.

The file PHILLIPS.FIT contains UK data on the logarithms of employment, labour force, prices, wages and real GDP over the period 1855-1987 (for more details see Lesson 12.5). Load this file, and in the Commands and Data Transformations box in the Process window type

$$C = 1;\; WP = W - P;\; YE = Y - E;\; WPYE = WP - YE;$$
$$DWPYE = WPYE - WPYE(-1)$$

RUN

to create an intercept term ($C = 1$), the logarithm of the real wage, WP, the logarithm of labour productivity, YE, and the logarithm of the share of wages in output, $WPYE$.

In the case of most tests of cointegration, the hypothesis being tested is the null of 'non-cointegration'.[1] In the present application the relevant hypothesis is that $WPYE = W - P - Y + E$ is not a cointegrating relation, or equivalently that $WPYE$ contains a unit root. Any one of the unit roots tests can be used for this purpose. Initially, we use a standard ADF test; later we consider the application of the Phillips and Perron (1988) semi-parametric test procedure (see footnote 4 to Lesson 12.1). Click on the button Simulation of Critical Values for Unit Root Tests (Off), check the 'Simulate critical values' checkbox, and click Close to accept the default number of replications, significance level, and maximum number of observations used for simulating critical values.

Use the Clear button to clear the box editor and type

ADF $WPYE(4)$ RUN

The ADF test results, shown in Table 16.1, should appear on the screen.

[1] The exception is the test of the stationarity hypothesis proposed by Kwiatkowski, Phillips, Schmidt, and Shin (1992).

Table 16.1: Unit roots tests for variable $WPYE$

ADF tests for variable WPYE
The Dickey-Fuller regressions include no intercept and no trend

128 observations used in the estimation of all ADF regressions.
Sample period from 1860 to 1987

	Test Statistic	CV	LL	AIC	SBC	HQC
DF	.29586	-1.7939	265.8218	264.8218	263.3958	264.2424
ADF(1)	.19942	-1.8226	269.5075	267.5075	264.6555	266.3487
ADF(2)	.24707	-1.8176	271.8762	268.8762	264.5982	267.1380
ADF(3)	.29121	-1.8243	274.2355	270.2355	264.5314	267.9179
ADF(4)	.30074	-1.8303	274.6928	269.6928	262.5627	266.7958

CV = 95% simulated critical value using 128 obs. and 1000 replications.
LL = Maximized log-likelihood AIC = Akaike Information Criterion
SBC = Schwarz Bayesian Criterion HQC = Hannan-Quinn Criterion

ADF tests for variable WPYE
The Dickey-Fuller regressions include an intercept but not a trend

128 observations used in the estimation of all ADF regressions.
Sample period from 1860 to 1987

	Test Statistic	CV	LL	AIC	SBC	HQC
DF	-1.6854	-2.8402	267.2550	265.2550	262.4030	264.0962
ADF(1)	-2.4481	-2.8430	272.5111	269.5111	265.2331	267.7729
ADF(2)	-1.9485	-2.8655	273.8123	269.8123	264.1083	267.4947
ADF(3)	-1.5082	-2.8573	275.4135	270.4135	263.2834	267.5165
ADF(4)	-1.3525	-2.8812	275.6501	269.6501	261.0940	266.1738

95% published asymptotic critical value corresponding to ADF(0) = -2.8840

CV = 95% simulated critical value using 128 obs. and 1000 replications.
LL = Maximized log-likelihood AIC = Akaike Information Criterion
SBC = Schwarz Bayesian Criterion HQC = Hannan-Quinn Criterion

ADF tests for variable WPYE
The Dickey-Fuller regressions include an intercept and a linear trend

128 observations used in the estimation of all ADF regressions.
Sample period from 1860 to 1987

	Test Statistic	CV	LL	AIC	SBC	HQC
DF	-2.9562	-3.4368	270.7909	267.7909	263.5129	266.0527
ADF(1)	-3.8079	-3.4159	277.0338	273.0338	267.3298	270.7163
ADF(2)	-3.3022	-3.5046	277.9466	272.9466	265.8165	270.0496
ADF(3)	-2.8785	-3.3852	279.3377	273.3377	264.7816	269.8613
ADF(4)	-2.7328	-3.4043	279.5868	272.5868	262.6047	268.5310

95% published asymptotic critical value corresponding to ADF(0) = -3.4452

The Akaike information (AIC) and Schwarz Bayesian (SBC) criteria suggest selecting ADF regressions of order 3 and 1, respectively. But if all the $ADF(p)$ statistics reported in Table 16.1 are considered, the null of a unit root is only rejected by an $ADF(1)$ with trend. Since we would not expect the share to be trended, it is worth trying to determine what is happening. Plotting $WPYE$ indicates that the share seemed to have moved to a higher level and become somewhat more stable after World War II (see Figure 16.1). However, to formally allow for such shift in the mean of $WPYE$ requires a different set of critical values, and will be subject to the uncertainty associated with the point at which such shift in the mean share of wages may have occurred (see Perron (1989)).

A similar conclusion also follows from the Phillips and Perron (1988) test. To compute this test statistic, choose the linear regression option 1 from the Single Equation Estimation Menu (Univariate Menu), select the OLS option, and specify the following simple Dickey-

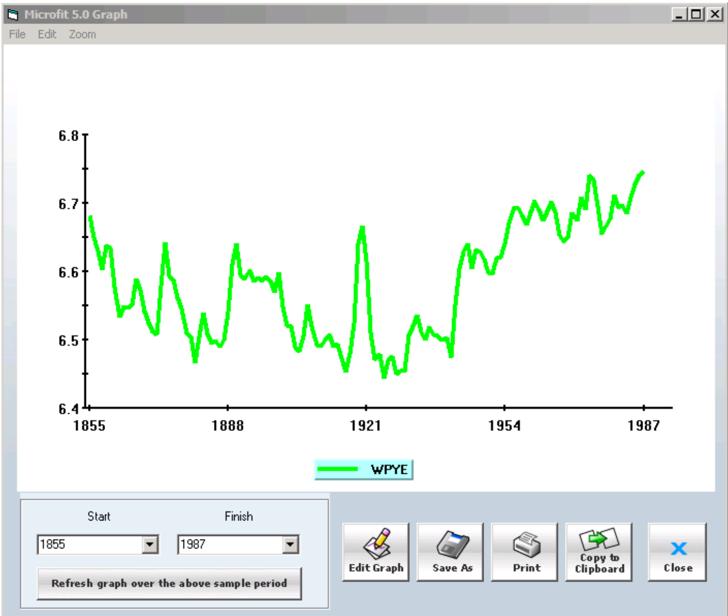


Figure 16.1: Logarithm of the share of wages in output in the UK

Fuller regression:

$$DWPYE \quad C \quad WPYE(-1)$$

where $DWPYE = WPYE - WPYE(-1)$. This regression is estimated over the whole sample period. The t-ratio of $WPYE(-1)$, namely -1.86, is the simple ADF statistic. To carry out the non-parametric correction to this statistic proposed by Phillips-Perron, choose option 4 (standard, White and Newey-West adjusted variance matrices) in the Post Regression Menu, choose the Newey-West adjusted variances with Bartlett weights, and when asked, use a window of 12 (say). Similar results are obtained with other window sizes in the range $(4, 20)$. Now click Close to leave the estimation results for the adjusted variance matrices, and choose option 1 (Display regression results for the adjusted covariance matrix). The t-ratio of $WPYE(-1)$ in this result screen (namely -2.368) is a Phillips-Perron-type statistic (see Table 16.2). The critical values for this test are the same as those for the ADF test. Once again the hypothesis that $WPYE$ contains a unit root cannot be rejected at the 95 per cent level. There is a clear evidence that the share of wages in output has changed over the 1855-1987 period. This is easily confirmed by adding a dummy variable, say $D46$, which takes the value of zero before World War II and unity thereafter, to the ADF regression. However, the inclusion of such a dummy variable alters the critical values for the unit roots tests. On this, see Perron (1989).

Table 16.2: Simple Dickey-Fuller regression with Newey-West adjusted standard errors

```
                        Ordinary Least Squares Estimation
  Based on Newey-West adjusted S.E.'s Bartlett weights, truncation lag = 12
*******************************************************************************
 Dependent variable is DWPYE
 132 observations used for estimation from 1856 to 1987
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 C                           .39937              .16752           2.3840[.019]
 WPYE(-1)                  -.060596             .025591          -2.3679[.019]
*******************************************************************************
```

16.2 Lesson 16.2: A residual-based approach to testing for cointegration

This lesson is concerned with the residual-based Augmented Dickey-Fuller (ADF) test of cointegration. Engle and Granger (1987) considered seven, asymptotically valid, residual-based test statistics for testing the null hypothesis of non-cointegration against the alternative of cointegration, but pointed out that in most applications the ADF test is preferable to the other six tests. More recently other residual-based type tests have also been proposed by Phillips and Ouliaris (1990), and extended by Hansen (1992) to allow for the possibility of a deterministic trend in the cointegrating relation. The ADF residual based tests and the Phillips-Ouliaris-Hansen tests are asymptotically equivalent and only differ in the way they deal with the problem of residual serial correlation in the simple ADF regression.

Here we show how the residual-based approach can be applied to test the stability of the share of wages in output analyzed in Lesson 16.1. Load the file PHILLIPS.FIT and ensure that the variables YE (log of output per man), WP (log of real wages), and the intercept term (C) are in the list of variables. You also need to satisfy yourself that the time series WP and YE are integrated processes of order 1.

Choose the linear regression option in the Single Equation Estimation Menu (Univariate Menu: see Section 6.4), selecting the OLS option. Then type

$$WP \quad C \quad YE$$ RUN

to set up the cointegrating equation (under the alternative hypothesis of cointegration).[2] This regression is estimated using the whole sample period, 1855-1987. Then move to the Hypothesis Testing Menu (see Section 6.23) and choose option 3 to compute ADF statistics of different orders (up to a maximum order of 12) based on the residuals of the cointegration regression. The $ADF(p)$ statistics are computed as the t-ratios of the estimated coefficient

[2] We could have equally considered the regression of YE on WP. Asymptotically the results should not depend whether the ADF test is applied to the residuals from the regression of WP on YE or the reverse regression. This may not, however, be the case in small samples. It is, therefore, important that the above computations are also carried out using the residuals from the regression of YE on WP. See Section 16.9.2.

of $R(-1)$ in the following OLS regressions:

$$DR = -\lambda R(-1) + \sum_{i=1}^{p} \delta_i R(-i) + error$$

for $p = 0, 1, .., P$, where R represents the residual of the OLS regression of WP on an intercept and YE, and $DR = R - R(-1)$. Now choose 4 for the maximum order of the ADF statistics, to obtain the results in Table 16.3 on the screen. The different model selection criteria in Table 16.3 favour a relatively low order for the ADF test, with the Akaike information criterion (AIC) selecting the order 2, and the Schwarz Bayesian criterion (SBC) selecting the order 1. For these orders the hypothesis of a unit root in the residuals is rejected. But the evidence is less convincing when higher order ADF statistics are considered. This is very similar to our finding in the previous lesson.

Table 16.3: Residual-based statistics for testing cointegration between real wages and labour productivity

```
                          Unit root tests for residuals
*******************************************************************************
 Based on  OLS regression of WP on:
 C               YE
 133 observations used for estimation from 1855 to 1987
*******************************************************************************
        Test Statistic       LL           AIC           SBC           HQC
 DF          -3.2833      263.5320      262.5320      261.1060      261.9526
 ADF(1)      -4.2689      270.3642      268.3642      265.5122      267.2054
 ADF(2)      -3.7199      271.0256      268.0256      263.7475      266.2874
 ADF(3)      -3.3064      271.8257      267.8257      262.1216      265.5081
 ADF(4)      -3.1371      272.0243      267.0243      259.8942      264.1273
*******************************************************************************
 95% critical value for the Dickey-Fuller statistic =  -3.3849
 LL  = Maximized log-likelihood      AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion    HQC = Hannan-Quinn Criterion
```

Consider now adding a dummy variable, say $D46$, which takes the value of zero before World War II, and unity thereafter, to the cointegrating relation. First return to the Process window and create the dummy variable $D46$ by typing in the Commands and Data Transformations box

$$D46 = 0;\ \textbf{SAMPLE}\ 1946\quad 1987;\ \ D46 = 1$$ 

Then choose the linear regression option in the Single Equation Estimation Menu (Univariate Menu), with the OLS option selected, add the variable $D46$ to the list of the regressors, and estimate the new regression equation over the whole sample period. If you now choose option 3 in the Hypothesis Testing Menu, with 4 as the maximum order for the ADF test, you should obtain the results reproduced in Table 16.4.

The residual-based ADF statistics are now much higher (in absolute value) than those given in Table 16.3, but due to the presence of the shift-dummy variable, $D46$, in the regression, the critical value given at the foot of Table 16.4 is not appropriate. The correct critical value for the test is also higher (in absolute value) than 3.3849. Overall, the tests in Lessons 16.1 and 16.2 suggest a reasonably stable share of wages in output, with an important shift in this share in the aftermath of World War II.

Notes:

1. The critical values for the *ADF* residual-based tests are computed using the response surface estimates given in MacKinnon (1991), and differ from the critical values reported when the **ADF** command is utilized in the Process window (see Section 4.4.2).

2. While it is possible to save the residuals and then apply the **ADF** command to the saved residuals at the Process window, it is important to note that in that case the reported critical values are not valid and can therefore result in misleading inferences. Therefore, it is important that residual-based *ADF* tests are carried out using option 3 in the Hypothesis Testing Menu (see Section 6.23).

Table 16.4: Residual-based statistics for testing cointegration between real wages and labour productivity allowing for a World War II dummy

```
                         Unit root tests for residuals
*******************************************************************************
 Based on  OLS regression of WP on:
 C               YE               D46
 133 observations used for estimation from 1855 to 1987
*******************************************************************************
        Test Statistic      LL           AIC           SBC           HQC
 DF          -4.1359      258.3337      257.3337      255.9077      256.7543
 ADF(1)      -5.3074      264.4931      262.4931      259.6410      261.3343
 ADF(2)      -5.0510      264.6720      261.6720      257.3940      259.9339
 ADF(3)      -4.3891      264.8753      260.8753      255.1713      258.5578
 ADF(4)      -4.0394      264.9369      259.9369      252.8068      257.0399
*******************************************************************************
 95% critical value for the Dickey-Fuller statistic =  -3.8090
 LL  = Maximized log-likelihood      AIC = Akaike Information Criterion
 SBC = Schwarz Bayesian Criterion    HQC = Hannan-Quinn Criterion
```

16.3 Lesson 16.3: Testing for cointegration: Johansen ML approach

The residual-based cointegration tests described in the previous lesson are inefficient and can lead to contradictory results, especially when there are more than two $I(1)$ variables under consideration. A more satisfactory approach would be to employ Johansen's ML procedure. This provides a unified framework for estimation and testing of cointegrating relations in the context of vector autoregressive (VAR) error correction models (see Section 7.5). Here we show how to use the cointegrating VAR options in *Microfit* to carry out $FIML$ cointegration tests.

In Lesson 13.3 we estimated a distributed lag relationship between capital expenditures (Y) and appropriations (X) for the US manufacturing sector employing the Almon (1965) polynomial distributed lag approach. However, we found that the estimated model suffered from a significant degree of residual serial correlation. In this lesson we re-examine the relationship between Y and X using cointegration techniques. The relevant data are in the special *Microfit* file ALMON.FIT, which contains observations on Y and X over the period 1953(1)-1967(4) (see Lesson 13.3 for more details). Read this file and plot the variables Y and X in the Process window. As can be seen from Figure 16.2, the two series are trended and generally move with one another, but by just looking at the graph it would not be possible to say whether they are cointegrated.

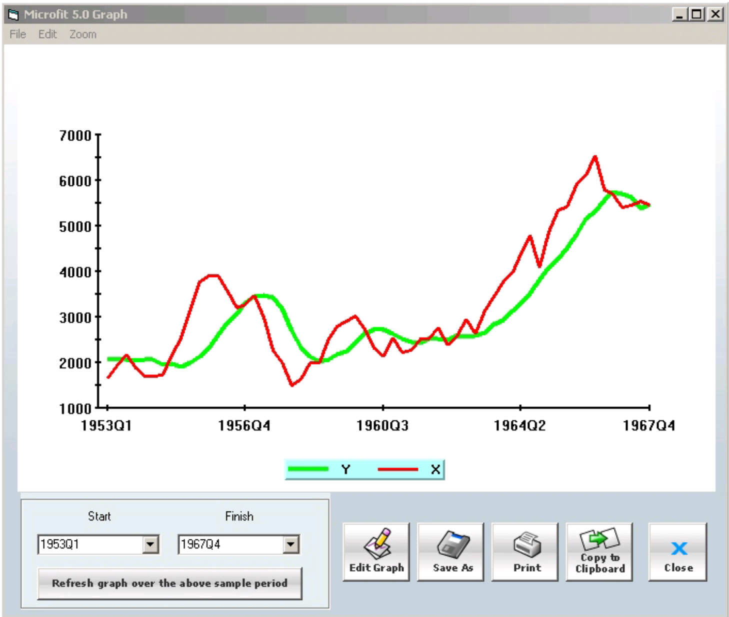


Figure 16.2: Capital expenditures (Y) and appropriations (X) in US manufacturing

Before using the cointegrating VAR options we need to ensure that the variables Y and X are in fact $I(1)$, ascertain the nature of the intercept/trend in the underlying VAR model, and choose the order for the VAR. Clear the Commands and Data Transformations box in the Process window and type

$$DY = Y - Y(-1); \quad DX = X - X(-1);$$

$$\textbf{ADF} \quad Y; \quad \textbf{ADF} \quad X; \quad \textbf{ADF} \quad DY; \quad \textbf{ADF} \quad DX$$

You should see the various ADF statistics needed for testing the unit root hypothesis on the screen. From these results it seems reasonable to conclude that Y and X are $I(1)$.

To select the order of the VAR in these variables, select the unrestricted VAR option from the system Estimation Menu (Multivariate Menu). In the box editor type

$$Y \quad X \quad \& \quad INPT$$

to specify a bivariate VAR model in Y and X, containing an intercept term ($INPT$) as its deterministic component.

Choose the whole sample and set the maximum order of the VAR to 4. In the Post Estimation Menu, choose the hypothesis testing option 4, and then choose testing/selecting the lag length, option 1 in the VAR Hypothesis Testing Menu. The results in Table 16.5 should now appear on the screen.

Table 16.5: Selecting the order of the VAR model in capital expenditures (Y) and appropriations (X) for US manufacturing

```
Test Statistics and Choice Criteria for Selecting the Order of the VAR Model
*****************************************************************************
Based on  56 observations from 1954Q1 to 1967Q4. Order of VAR = 4
List of variables included in the unrestricted VAR:
Y              X
List of deterministic and/or exogenous variables:
INPT
*****************************************************************************
Order    LL        AIC       SBC              LR test          Adjusted LR test
4  -720.3923 -738.3923 -756.6205              ------                ------
3  -720.8494 -734.8494 -749.0269  CHSQ(4)=    .91413[.923]     .76722[.943]
2  -725.2045 -735.2045 -745.3313  CHSQ(8)=   9.6244[.292]     8.0776[.426]
1  -738.3078 -744.3078 -750.3839  CHSQ(12)=  35.8310[.000]    30.0724[.003]
0  -916.7598 -918.7598 -920.7852  CHSQ(16)= 392.7350[.000]   329.6168[.000]
*****************************************************************************
AIC=Akaike Information Criterion     SBC=Schwarz Bayesian Criterion
```

The Schwarz Bayesian criterion (SBC) suggests a VAR of order 2, the Akaike information criterion (AIC) of order 3. Since we have a short time series (60 observations), we cannot take the risk of over-parameterization, and therefore choose 2 as the order of the VAR. In such situations it is, however, important to check the residuals of the individual equations in VAR for possible serial correlation. An inspection of the results suggests that this is not a problem in the present application.

Backtrack to the System Estimation window and from the System Estimation Menu (Multivariate Menu) choose the cointegrating VAR option, selecting the Restricted Intercepts option 2. In the box editor delete the intercept term ($INPT$), so that the content of the box editor is

$$Y \quad X$$

In the case of the cointegrating VAR option, specification of intercept and (linear) trend terms is done at a later stage. Use the whole period for estimation, and set the order of the VAR to 2. The choice of intercepts/trends is very important in testing for cointegration. In the present application, although the underlying variables are trended, they move together, and it seems unlikely that there will be a trend in the cointegrating relation between Y and X. The results in Table 16.6 should now appear on the screen. Both the maximum and trace eigenvalue statistics strongly reject the null hypothesis that there is no cointegration between Y and X, (namely that $r = 0$), but do not reject the hypothesis that there is one cointegrating relation between these variables (namely that $r = 1$).

Table 16.6: Testing for cointegration between capital expenditures (Y) and appropriations (X) in US manufacturing

```
      Cointegration with restricted intercepts and no trends in the VAR
   Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix
*******************************************************************************
 58 observations from 1953Q3 to 1967Q4. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 Y               X               Intercept
 List of eigenvalues in descending order:
.52365    .038066
*******************************************************************************
 Null    Alternative    Statistic     95% Critical Value  90% Critical Value
 r = 0      r = 1        43.0135          15.8700             13.8100
 r<= 1      r = 2         2.2510           9.1600              7.5300
*******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

      Cointegration with restricted intercepts and no trends in the VAR
          Cointegration LR Test Based on Trace of the Stochastic Matrix
*******************************************************************************
 58 observations from 1953Q3 to 1967Q4. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 Y               X               Intercept
 List of eigenvalues in descending order:
.52365    .038066
*******************************************************************************
 Null    Alternative    Statistic     95% Critical Value  90% Critical Value
 r = 0      r>= 1        45.2645          20.1800             17.8800
 r<= 1      r = 2         2.2510           9.1600              7.5300
*******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

      Cointegration with restricted intercepts and no trends in the VAR
Choice of the Number of Cointegrating Relations Using Model Selection Criteria
*******************************************************************************
 58 observations from 1953Q3 to 1967Q4. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 Y               X               Intercept
 List of eigenvalues in descending order:
.52365    .038066
*******************************************************************************
 Rank      Maximized LL        AIC              SBC              HQC
 r = 0       -772.7598       -776.7598        -780.8807        -778.3649
 r = 1       -751.2530       -759.2530        -767.4948        -762.4633
 r = 2       -750.1275       -760.1275        -770.4297        -764.1404
*******************************************************************************
 AIC = Akaike Information Criterion    SBC = Schwarz Bayesian Criterion
 HQC = Hannan-Quinn Criterion
```

A similar result also follows from the values of the various model selection criteria reported in the third panel of Table 16.6. This complete agreement between the three procedures for testing/selecting the number of cointegrating relations is very rare, however. In practice, these three methods often result in conflicting conclusions, and the decision concerning the choice of r, the number of cointegrating relations, must be made in view of other information, perhaps from economic theory.

As an example, consider the application of the cointegrating VAR option to the problem of testing the Purchasing Power Parity (PPP) hypothesis studied in Johansen and Juselius (1992) and Pesaran and Shin (1996). The relevant data are in the special *Microfit* file

PPP.FIT and contain the following variables.

P	Logarithm of UK wholesale prices
PF	Logarithm of foreign prices
E	Logarithm of the UK effective exchange rate
R	Domestic interest rate
RF	Foreign interest rate
DPO	Changes in real oil prices
$S1, S2, S3$	Quarterly Seasonal dummies
$INPT$	Intercept term

Load this file into *Microfit* and make sure that the above variables are on your workspace. Using the **ADF** command in the Process window, it is easily seen that we cannot reject the hypothesis that the variables P, PF, E, R, and RF are $I(1)$, and that DPO is $I(0)$. The second stage in the cointegration analysis is to decide on the order of the underlying VAR model and the nature of the intercepts/trends in the model. Using the Unrestricted VAR option in the System Estimation Menu, and choosing 4 as the maximum order for the following specification:

$$P \quad E \quad PF \quad R \quad RF \quad \& \quad INPT \quad S1 \quad S2 \quad S3 \quad DPO \quad DPO(-1)$$ 

the AIC and SBC criteria select order 3 and 1, respectively. Johansen and Juselius (1992) (JJ) select the order 2. Given the fact that the sample is relatively small (only 62 quarters) we follow JJ and select 2 for the order of the VAR.[3] As far as the specification of the intercept and trend in the VAR is concerned we also follow JJ, and assume that the underlying VAR model does not contain deterministic trends, but contains unrestricted intercepts. However, recall from the discussion in Sections 22.7 and 7.5 that such specifications will generate deterministic trends in the level of the variables (P, PF, E, R, and RF), when the long-run multiplier matrix is rank deficient, which will be the case in this application if one accepts the conclusion that variables P, PF, E, R, and RF are $I(1)$.

With the above considerations in mind choose the cointegrating VAR in the System Estimation Menu (Multivariate Menu), selecting the Unrestricted intercepts, no trends option 3, and type

$$P \quad E \quad PF \quad R \quad RF \quad \& \quad S1 \quad S2 \quad S3 \quad DPO \quad DPO(-1)$$

Select the whole period for estimation, choose 2 for the order of the VAR, and click . The results in Table 16.7 should now appear on the screen.

[3]However, an inspection of the regression results for the individual equations in the $VAR(2)$ model suggest important evidence of residual serial correlation for the RF equation. In fact this equation decisively fails all the four diagnostic tests (residual serial correlation, functional form, normality and heteroscedasticity) automatically supplied by *Microfit*.

Table 16.7: Testing for cointegration between prices, interest rates, and the exchange rate in the UK economy

```
    Cointegration with unrestricted intercepts and no trends in the VAR
  Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix
******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P               E               PF              R               RF
 List of unrestricted deterministic variables included in the VAR:
 S1              S2              S3              DPO             DPO(-1)
 List of eigenvalues in descending order:
.40680     .28525     .25423     .10228     .082827
******************************************************************************
 Null    Alternative    Statistic     95% Critical Value  90% Critical Value
 r = 0      r = 1        31.3337          33.6400              31.0200
 r<= 1      r = 2        20.1497          27.4200              24.9900
 r<= 2      r = 3        17.6006          21.1200              19.0200
 r<= 3      r = 4         6.4739          14.8800              12.9800
 r<= 4      r = 5         5.1875           8.0700               6.5000
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

    Cointegration with unrestricted intercepts and no trends in the VAR
         Cointegration LR Test Based on Trace of the Stochastic Matrix
******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P               E               PF              R               RF
 List of unrestricted deterministic variables included in the VAR:
 S1              S2              S3              DPO             DPO(-1)
 List of eigenvalues in descending order:
.40680     .28525     .25423     .10228     .082827
******************************************************************************
 Null    Alternative    Statistic     95% Critical Value  90% Critical Value
 r = 0      r>= 1        80.7454          70.4900              66.2300
 r<= 1      r>= 2        49.4117          48.8800              45.7000
 r<= 2      r>= 3        29.2620          31.5400              28.7800
 r<= 3      r>= 4        11.6614          17.8600              15.7500
 r<= 4      r = 5         5.1875           8.0700               6.5000
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).
    Cointegration with unrestricted intercepts and no trends in the VAR
Choice of the Number of Cointegrating Relations Using Model Selection Criteria
******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P               E               PF              R               RF
 List of unrestricted deterministic variables included in the VAR:
 S1              S2              S3              DPO             DPO(-1)
 List of eigenvalues in descending order:
.40680     .28525     .25423     .10228     .082827
******************************************************************************
 Rank      Maximized LL       AIC            SBC            HQC
 r = 0       900.3650       845.3650       787.7705       822.8367
 r = 1       916.0319       852.0319       785.0128       825.8170
 r = 2       926.1067       855.1067       780.7575       826.0247
 r = 3       934.9070       858.9070       779.3219       827.7769
 r = 4       938.1440       859.1440       776.4174       826.7850
 r = 5       940.7377       860.7377       776.9639       827.9692
******************************************************************************
 AIC = Akaike Information Criterion    SBC = Schwarz Bayesian Criterion
 HQC = Hannan-Quinn Criterion
```

The maximum eigenvalue and the trace statistics in this table are almost identical with those reported in Johansen and Juselius (1992) Table 2. However, there are important differences between the critical values used by *Microfit* and those reported in JJ and also by Osterwald-Lenum (1992). The reasons for these differences are explained in Pesaran, Shin, and Smith (2000). Irrespective of which sets of critical values are used there is a clear conflict between the test results based on the maximum eigenvalue statistics and the trace statistic. Taken literally, the maximum eigenvalue statistic does not reject $r = 0$ (no cointegration), while the trace statistic does not reject $r = 2$, at the 95 per cent significant level. Changing the significance level of the two tests to 90 per cent results in the maximum eigenvalue value statistic to select $r = 1$, and the trace statistic to select $r = 3$. Turning to the model selection criteria (given in panel three of the Table), we find that the AIC and SBC choose $r = 5$ and zero, the two opposite extremes; while the Hannan and Quinn criteria choose $r = 5$. The data in this application seems hopelessly uninformative on the choice of r. Turning to long-run economic theory based on arbitrage in the product and capital markets we would expect *two* cointegrating relations: the PPP relation

$$P - E - PF \sim I(0)$$

and the interest-rate arbitrage relation (which is a long-run implication of the uncovered interest parity hypothesis)

$$R - RF \sim I(0)$$

Combining this theoretical insight with the mixed test results (based on a very short sample) it seems reasonable to set $r = 2$.

16.4 Lesson 16.4: Testing for cointegration in models with $I(1)$ exogenous variables

In certain applications, particularly in the context of small open economies, it is reasonable to assume that one or more of the $I(1)$ variables in the cointegrating VAR model are 'long-run forcing' variables, in the sense that in the long-run they are not 'caused' by the other variables in the model. This does not, of course, rule out contemporaneous or short-run interactions between the $I(1)$ variables (for a more formal discussion see Section 22.7). For example, in the case of the five variable VAR model in (P, E, R, PF and RF) analyzed in Lesson 16.3, it seems *a priori* plausible to assume that in the long-run there are no feedbacks from UK prices, interest rates and exchange rates into foreign prices and interest rates; that is PF and RF are the 'long-run forcing' variables of the system. To test the purchasing power parity (PPP) hypothesis under these assumptions, first load the file PPP.FIT and choose option 2 in the System Estimation Menu (see Lesson 16.3 and Johansen and Juselius (1992) for further details) - selecting Unrestricted intercepts, no trends option 3. This is the intercept/trend specification chosen by Johansen and Juselius (1992) (JJ) which we also adopt in order to make our analysis comparable with theirs. To set up the VAR model with PF and RF as 'long-run forcing' variables, and changes in real oil prices (DPO), and seasonal dummies ($S1$, $S2$, and $S3$) as exogenous $I(0)$ variables, in the box editor you need to type

$$P \quad E \quad R; \quad RF \quad PF \quad \& \quad DPO \quad DPO(-1) \quad S1 \quad S2 \quad S3$$

The semicolon separates the $I(1)$ variables of the model into the set of jointly determined variables, P, E and R, and the long-run forcing variables PF and RF. Select the whole sample period, and for the order of the VAR choose 2 (as in JJ's analysis). Then click 

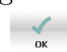
. You should now obtain the results in Table 16.8 on the screen.

Table 16.8: Testing for cointegration between prices, interest rates, and the exchange rate in the UK economy, treating foreign prices and interest rates as exogenous

```
     Cointegration with unrestricted intercepts and no trends in the VAR
   Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix
******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P               E               R               RF              PF
 List of I(1) exogenous variables included in the VAR:
 RF              PF
 List of unrestricted deterministic variables included in the VAR:
 DPO             DPO(-1)         S1              S2              S3
 List of eigenvalues in descending order:
.38501     .18433     .12159
******************************************************************************
 Null     Alternative    Statistic     95% Critical Value  90% Critical Value
 r = 0       r = 1        29.1690           27.7500             25.2100
 r<= 1       r = 2        12.2249           21.0700             18.7800
 r<= 2       r = 3         7.7788           14.3500             12.2700
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

     Cointegration with unrestricted intercepts and no trends in the VAR
          Cointegration LR Test Based on Trace of the Stochastic Matrix
******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P               E               R               RF              PF
 List of I(1) exogenous variables included in the VAR:
 RF              PF
 List of unrestricted deterministic variables included in the VAR:
 DPO             DPO(-1)         S1              S2              S3
 List of eigenvalues in descending order:
.38501     .18433     .12159
******************************************************************************
 Null     Alternative    Statistic     95% Critical Value  90% Critical Value
 r = 0       r>= 1        49.1728           46.4400             42.6700
 r<= 1       r>= 2        20.0037           28.4200             25.6300
 r<= 2       r = 3         7.7788           14.3500             12.2700
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).
     Cointegration with unrestricted intercepts and no trends in the VAR
Choice of the Number of Cointegrating Relations Using Model Selection Criteria
******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P               E               R               RF              PF
 List of I(1) exogenous variables included in the VAR:
 RF              PF
 List of unrestricted deterministic variables included in the VAR:
 DPO             DPO(-1)         S1              S2              S3
 List of eigenvalues in descending order:
.38501     .18433     .12159
******************************************************************************
 Rank      Maximized LL        AIC             SBC             HQC
 r = 0       521.5419        488.5419        453.9852        475.0249
 r = 1       536.1264        496.1264        454.2395        479.7421
 r = 2       542.2389        497.2389        450.1161        478.8066
 r = 3       546.1283        498.1283        447.8640        478.4672
******************************************************************************
 AIC = Akaike Information Criterion    SBC = Schwarz Bayesian Criterion
 HQC = Hannan-Quinn Criterion
```

Both the maximum eigenvalue and the trace statistic suggest $r = 1$. The hypothesis that $r = 0$ is rejected against $r = 1$, but the hypothesis that $r = 1$ cannot be rejected against $r = 2$, and so on. The Schwarz Bayesian (SBC) and the Hannan and Quinn (HQC) also favour $r = 1$, but the same is not true for the Akaike information criterion (AIC), which selects $r = 3$! In what follows we assume $r = 1$ and present the estimates of the cointegrating coefficients, normalized on the coefficient of P.

Click 

to move to the Cointegrating VAR Post Estimation Menu, choose option 2, and when prompted type 1 to specify $r = 1$. To obtain estimates of the cointegrating coefficients (together with their asymptotic standard errors), choose option 6 (Long Run Structural Modelling, IR Analysis and Forecasting), then choose option 4 and click 

.

You will be asked to specify exactly one restriction to identify the cointegrating relation. In the present application where $r = 1$, this can be achieved by normalizing on one of the coefficients. In the box editor type

$$A1 = 1$$

RUN

to normalize on the coefficient of P, the first variable in the cointegrating VAR. You should now see the estimates of the cointegrating coefficients and their asymptotic standard errors on the screen. See Table 16.9.

Table 16.9: ML estimates subject to exactly identifying restriction(s); estimates of restricted cointegrating relations (SEs in brackets)

```
             ML estimates subject to exactly identifying restriction(s)
        Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                           Converged after 1 iterations
       Cointegration with unrestricted intercepts and no trends in the VAR
*******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P                E                R                RF               PF
 List of I(1) exogenous variables included in the VAR:
 RF               PF
 List of unrestricted deterministic variables included in the VAR:
 DPO              DPO(-1)          S1               S2               S3
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                   Vector  1
 P                    1.0000
                  (    *NONE*)

 E                   -.94623
                  (    .33076)

 R                   -4.7556
                  (    1.4337)

 RF                  -.77615
                  (    .89249)

 PF                  -.90197
                  (    .19997)

*******************************************************************************
 LL subject to exactly identifying restrictions= 536.1264
*******************************************************************************
```

Notice that the maximized value of the log-likelihood function $LL(r = 1) = 536.1264$ which is identical to the value 'maximized LL' for $r = 1$ in the bottom panel of Table 16.8. Also only the coefficient of RF, estimated at $-0.77615(0.8925)$, with its asymptotic standard error in brackets, is not statistically significant. It is therefore reasonable to re-estimate the cointegrating relation imposing the over-identifying restriction $A4 = 0$, where $A4$ stands for the coefficient of RF. To impose this restriction, click Close to move to the IR Analysis and Forecasting Menu, choose option 0, and then click YES to return to the box editor for the specification of the coefficient restrictions. You should see the normalizing (or exactly identifying) restriction $A1 = 1$ on the screen. Add the restriction $A4 = 0$ to it. The box editor should now contain the two restrictions

$$A1 = 1; \quad A4 = 0$$

Click to estimate the cointegrating relations subject to these restrictions. You will be asked to specify/edit initial values for the cointegrating coefficients. Click to accept the initial values supplied by the program, and then press to choose the back substitution algorithm B (new to *Microfit 5.0*). If you then click , you are asked whether you wish to obtain bootstrapped critical values for the overidentifying restrictions on the long-run relationship. Click to accept the default number of replications, equal to 1,000, and significance levels. The results are shown in Table 16.10.

Table 16.10: ML estimates subject to over-identifying restriction(s); estimates of restricted cointegrating relations (SEs in brackets)

```
                               Converged after 1 iterations
           Cointegration with unrestricted intercepts and no trends in the VAR
*******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P                E               R               RF               PF
 List of I(1) exogenous variables included in the VAR:
 RF               PF
 List of unrestricted deterministic variables included in the VAR:
 DPO              DPO(-1)         S1              S2               S3
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                  Vector 1
 P                   1.0000
                  (   *NONE*)

 E                 -.79464
                  (   .26575)

 R                 -5.2257
                  (   1.4960)

 RF                   0.00
                  (   *NONE*)

 PF                -.98320
                  (   .16874)

*******************************************************************************
 LR Test of Restrictions          CHSQ(1)=    .74335[.389]
 95% Bootstrapped Critical Value =    7.3113
 90% Bootstrapped Critical Value =    5.3342
 Bootstrapped simulations based on 1000 SIMULATIONS.
 DF=Total no of restrictions(2) - no of just-identifying restrictions(1)
 LL subject to exactly identifying restrictions= 536.1264
 LL subject to over-identifying restrictions= 535.7547
*******************************************************************************
```

The log-likelihood ratio statistic for testing the restriction $A4 = 0$, is given by $CHSQ(1) = 0.74$ with a p-value of 0.389. Also, 0.74 is smaller that the bootstrapped critical values at the 5 per cent significance level. Hence, the log-likelihood ratio is not statistically significant, suggesting that the restriction $A4 = 0$ cannot be rejected.

The PPP hypothesis, however, postulates that the coefficients of E and PF should also be both equal to -1. To impose (and then test) this restriction, you can now click and then to move to the IR Analysis and Forecasting Menu (see Section 7.5.4), from where you can return to the box editor to specify the additional restrictions, $A2 = -1$ and $A5 = -1$. Again, if you close the output window and press the button, you should obtain the results in Table 16.11.

Table 16.11: ML estimates subject to over-identifying restriction(s); estimates of restricted cointegrating relations (SEs in brackets)

```
                           Converged after 9 iterations
    Cointegration with unrestricted intercepts and no trends in the VAR
*******************************************************************************
 60 observations from 1972Q3 to 1987Q2. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 P               E               R               RF              PF
 List of I(1) exogenous variables included in the VAR:
 RF              PF
 List of unrestricted deterministic variables included in the VAR:
 DPO             DPO(-1)         S1              S2              S3
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                  Vector  1
 P                   1.0000
                 (    *NONE*)

 E                  -1.0000
                 (    *NONE*)

 R                  -5.3198
                 (    .87522)

 RF                  -.0000
                 (    *NONE*)

 PF                 -1.0000
                 (    *NONE*)

*******************************************************************************
 LR Test of Restrictions          CHSQ(3)=    3.6052[.307]
 95% Bootstrapped Critical Value =  14.7041
 90% Bootstrapped Critical Value =  11.6359
 Bootstrapped simulations based on 1000 SIMULATIONS.
 DF=Total no of restrictions(4) - no of just-identifying restrictions(1)
 LL subject to exactly identifying restrictions= 536.1264
 LL subject to over-identifying restrictions= 534.3238
*******************************************************************************
```

The LR statistic for the three over-identifying restrictions ($A2 = -1$, $A4 = 0$, $A5 = -1$) is equal to 3.61, which is distributed asymptotically as a chi-squared variate with 3 degrees of freedom, and hence is not statistically significant. Also, notice that the LR statistic is below its 95 per cent bootstrapped critical value (equal to 14.70). Therefore, the restricted cointegrating relation is estimated as

$$P_t - E_t - PF_t - \underset{(0.8754)}{5.3198} \ R_t \sim I(0)$$

What is striking about this result is the significant positive long-run effect of the domestic interest rate on the PPP relation, defined by $P_t - E_t - PF_t$. This provides evidence against the validity of the PPP (in the case of the present data set), but also pinpoints for the departure from PPP, namely the effect of nominal interest rate.

Finally, to see the error correction form of the relations in the cointegrating VAR model you can choose option 7 in the IR Analysis and Forecasting Menu (see Section 7.5.4). The number of error-correction equations in the present application is 3, corresponding to the jointly determined variables of the model, namely P, E and R. The error correction equation for P is shown in Table 16.12.

Table 16.12: ECM for the variable P estimated by OLS based on cointegrating $VAR(2)$

```
 Dependent variable is dP
 60 observations used for estimation from 1972Q3 to 1987Q2
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 Intercept                   .17444              .055071           3.1676[.003]
 dP1                         .43577               .11238           3.8777[.000]
 dE1                        .040618              .034748           1.1689[.248]
 dR1                       -.070052               .11897          -.58880[.559]
 dRF1                      -.081188              .073639          -1.1025[.276]
 dPF1                        .12503               .11915           1.0494[.299]
 ecm1(-1)                  -.042265              .013610          -3.1053[.003]
 DPO                        .024890             .0077761           3.2009[.002]
 DPO(-1)                   .0012268              .012356          .099290[.921]
 S1                        .0056492             .0031539           1.7912[.080]
 S2                        .8278E-3             .0032346           .25593[.799]
 S3                        .5328E-3             .0032104           .16597[.869]
*******************************************************************************
 List of additional temporary variables created:
 dP = P-P(-1)

 dP1 = P(-1)-P(-2)
 dE1 = E(-1)-E(-2)
 dR1 = R(-1)-R(-2)
 dRF1 = RF(-1)-RF(-2)
 dPF1 = PF(-1)-PF(-2)
 ;ecm1 =    1.0000*P   -1.0000*E   -5.3198*R    -.0000*RF   -1.0000*PF
*******************************************************************************
 R-Squared                      .81889   R-Bar-Squared                    .77739
 S.E. of Regression           .0083111   F-Stat.    F(11,48)    19.7307[.000]
 Mean of Dependent Variable    .026359   S.D. of Dependent Variable      .017615
 Residual Sum of Squares      .0033156   Equation Log-likelihood        208.9679
 Akaike Info. Criterion       196.9679   Schwarz Bayesian Criterion     184.4018
 DW-statistic                   1.7786   System Log-likelihood          534.3238
*******************************************************************************


                             Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *          F Version         *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(4)  =    3.2948[.510]*F(4,44)      =   .63914[.637]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =    3.0856[.079]*F(1,47)      =   2.5481[.117]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  =    2.0421[.360]*       Not applicable       *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  = .4119E-3[.984]*F(1,58)      = .3982E-3[.984]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
 HQC = Hannan-Quinn Criterion
```

The coefficient of the error-correction term, $-0.0423(0.0136)$, has the correct sign and is statistically significant, but rather small, suggesting that it would take a long time for the equation to return to its equilibrium once it is shocked. The error correction terms in the exchange-rate and the domestic interest rate equations are also statistically significant, but only just. Notice that the positive coefficients obtained for the error correction terms in these two equations are correct. This is because E and R enter the error correction terms with negative coefficients.

16.5 Lesson 16.5: Long-run analysis of consumption, income and inflation: the ARDL approach

In this lesson we employ the testing and estimation procedure advanced in Pesaran, Shin, and Smith (1996) and Pesaran and Shin (1999) to examine the relationship between logarithm of non-durable consumption expenditures (LC), the logarithm of the real disposable income (LY) and the inflation rate (DP) in the US, using quarterly observations over the period 1960(1)-1994(1). The main advantage of this testing and estimation strategy (which we refer to as the $ARDL$ procedure) lies in the fact that it can be applied irrespective of whether the regressors are $I(0)$ or $I(1)$, and can avoid the pre-testing problems associated with the standard cointegration analysis which requires the classification of the variables into $I(1)$ and $I(0)$.

The $ARDL$ procedure involves two stages. At the first stage the existence of the long-run relation between the variables under investigation is tested by computing the F-statistic for testing the significance of the lagged levels of the variables in the error correction form of the underlying $ARDL$ model. However, the (asymptotic) distribution of this F-statistic is non-standard, irrespective of whether the regressors (in the present application LY and DP) are $I(0)$ or $I(1)$. Pesaran, Shin, and Smith (1996) have tabulated the appropriate critical values for different numbers of regressors (k), and determined whether the $ARDL$ model contains an intercept and/or trend. They give two sets of critical values: one set assuming that *all* the variables in the $ARDL$ model are $I(1)$, and another computed assuming all the variables are $I(0)$. For each application, this provides a band covering all the possible classifications of the variables into $I(0)$ and $I(1)$, or even fractionally integrated ones. If the computed F-statistic falls outside this band a conclusive decision can be made without needing to know whether the underlying variables are $I(0)$ or $I(1)$, or fractionally integrated. If the computed statistic falls within the critical value band the result of the inference is inconclusive and depends on whether the underlying variables are $I(0)$ or $I(1)$. It is at this stage in the analysis that the investigator may have to carry out unit roots tests on the variables.

The second stage in the analysis is to estimate the coefficients of the long-run relations and make inferences about their values using the $ARDL$ option (see Section 6.18.) Note that it is only appropriate to embark on this stage if you are satisfied that the long-run relationship between the variables to be estimated is not in fact spurious.

To apply the above approach to the US consumption data, first load the special *Microfit* file USCON.FIT and then run the transformations in the equation file USCON.EQU on this data in the Process window. You should have the following variables in the variable list

LC	Log of non-durable real consumption expenditures
DLC	$LC - LC(-1)$
LY	Log of real disposable income
DLY	$LY - LY(-1)$
PI	The rate of inflation
DPI	$PI - PI(-1)$
$INPT$	A vector of 1s

Since the observations are quarterly, for the maximum order of the lags in the $ARDL$ model we choose 4, and carry out the estimation over the period 1960(1)-1992(4), retaining the remaining five observations 1993(1)-1994(1) for predictions.

The error correction version of the $ARDL(5,5,5)$ model in the variables LC, LY and PI is given by

$$\begin{aligned} DLC_t &= a_0 + \sum_{i=1}^{4} b_i DLC_{t-i} + \sum_{i=1}^{4} d_i DLY_{t-i} + \sum_{i=1}^{4} e_i DPI_{t-i} \\ &\quad + \delta_1 LC_{t-1} + \delta_2 LY_{t-1} + \delta_3 PI_{t-1} + u_t \end{aligned} \tag{16.1}$$

Due to the high levels of cross-sectional and temporal aggregations involved, it is not possible to know *a priori* whether LY and PI are the 'long-run forcing' variables for aggregate consumption (LC), so we have excluded the current values of DLY and DPI from (16.1). We shall reconsider this issue once we have completed our *stability* tests, namely whether there exists a long-run relationship between LC, LY and PI.

The hypothesis that we will be testing is the null of 'non-existence of the long-run relationship' defined by

$$H_0 : \delta_1 = \delta_2 = \delta_3 = 0$$

against

$$H_1 : \delta_1 \neq 0, \quad \delta_2 \neq 0, \quad \delta_3 \neq 0$$

The relevant statistic is the familiar F-statistic for the joint significance of δ_1, δ_2 and δ_3. To compute this statistic choose option 1 in the Single Equation Estimation Menu (see Section 6.4). In the Commands and Data Transformations box type

$$DLC \quad INPT \quad DLC\{1-4\} \quad DLY\{1-4\} \quad DPI\{1-4\}$$

choose the sample 1960(1)-1992(4) for estimation, and click  to see the OLS results of the regression in first differences. This regression is of no direct interest. Click  and then  to move to the Hypothesis Testing Menu and choose option 6 (variable addition test). Now list the lagged values of the level variables by typing

$LC(-1) \quad LY(-1) \quad PI(-1)$ 

The F-statistic for testing the joint null hypothesis that the coefficients of these level variables are zero (namely there exists no long-run relationship between them) is given in the last row of the result table that appears on the screen. We denote it by $F(LC\,|LY,\ PI) = 5.43$. As we have already noted under $H_0 : \delta_1 = \delta_2 = \delta_3 = 0$, this statistic has a non-standard distribution irrespective of whether LC, LY, and PI are $I(0)$ or $I(1)$. The critical value bounds for this test are computed by Pesaran, Shin, and Smith (1996), and are reproduced as Tables B.1 and B.2 in Appendix B. Table B.2 gives the bounds for the W-statistic for three cases depending on whether the underlying regression contains an intercept or trend. Table B.1 gives the critical value bounds for the F-statistic version of the test. The relevant critical value bounds for the present application is given in the middle panel of Table B.1, and at the 95 per cent level is given by 3.219-4.378. Since $F(LC|LY,\ PI) = 5.43$ exceeds the upper bound of the critical value band, we can reject the null of no long-run relationship between LC, LY and PI irrespective of the order of their integration.

Consider now the significance of the lagged level variables in the error correction models explaining DLY_t and DPI_t. Backtrack to edit the regression equation, and change DLC (the dependent variable) to DLY and then follow the same steps as above to compute the F-statistic for the joint significance of $LC(-1)$, $LY(-1)$ and $PI(-1)$ in this new regression. You should obtain $F(LY|LC,\ PI) = 2.631$. Similarly for the PI equation, $F(PI|LC,\ LY) = 1.359$. Both these statistics fall well below the lower bound of the critical value band (which is 3.793-4.855), and hence the null hypothesis that the level variables do not enter significantly into the equations for DLY and DPI cannot be rejected. Once again this conclusion holds irrespective of whether the underlying variables are $I(0)$ and $I(1)$.

The above test results suggest that there exists a long-run relationship between LC, LY and PI, and the variables LY and PI can be treated as the 'long-run forcing' variables for the explanation of LC.

The estimation of the long-run coefficients and the associated error-correction model can now be accomplished using the $ARDL$ option in *Microfit* (see Section 6.18.) Move to the Single Equation Estimation Menu, and choose the $ARDL$ option (item 6 in this menu). Click  to clear the Commands and Data Transformations box, and type

$$LC \quad LY \quad PI \quad \& \quad INPT$$

Choose the sample 1960(1)-1992(4) for estimation, enter 4 for the maximum lag to be used in the model selection that follows, and click . *Microfit* estimates 125 regressions and presents you with the $ARDL$ Order Selection Menu (see Section 6.18.2). This gives a choice between different model selection criteria. The SBC and the AIC select the $ARDL(1,2,0)$ and $ARDL(2,2,3)$ specifications, respectively.

The estimates of the long-run coefficients based on these models are summarized in Table 16.13.

Table 16.13: Estimates of the long-run coefficients based on $ARDL$ models selected by AIC and SBC

	Model Selection Criteria	
Long-Run Coefficients	SBC-$ARDL$(1,2,0)	AIC-$ARDL$(2,2,3)
INPT	1.336	1.269
	(0.167)	(0.112)
LY	0.693	0.700
	(0.021)	(0.014)
PI	-2.595	-2.288
	(1.171)	(0.754)

The point estimates are very similar, but as to be expected the estimated standard errors obtained using the model selected by the AIC are considerably smaller given the much higher order $ARDL$ model selected by the AIC as compared to the SBC.

To obtain the estimates of the error correction model associated with these long-run estimates you need to choose option 3 in the Post $ARDL$ Model Selection Menu (see Section 6.18.3). The estimated error correction model selected using AIC is given in Table 16.14.

Table 16.14: A log-linear error correction model of US consumption

```
         Error Correction Representation for the Selected ARDL Model
          ARDL(2,2,3) selected based on Akaike Information Criterion
*******************************************************************************
 Dependent variable is dLC
 127 observations used for estimation from 1961Q2 to 1992Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 dLC1                       .11651            .085463             1.3633[.175]
 dLY                        .26694            .056676             4.7098[.000]
 dLY1                       .16621            .060825             2.7326[.007]
 dPI                       -.18461            .080063            -2.3058[.023]
 dPI1                       .18936            .087764             2.1576[.033]
 dPI2                       .25269            .078663             3.2123[.002]
 ecm(-1)                   -.12599            .036172            -3.4832[.001]
*******************************************************************************
 List of additional temporary variables created:
 dLC = LC-LC(-1)
 dLC1 = LC(-1)-LC(-2)
 dLY = LY-LY(-1)
 dLY1 = LY(-1)-LY(-2)
 dPI = PI-PI(-1)
 dPI1 = PI(-1)-PI(-2)
 dPI2 = PI(-2)-PI(-3)
 ecm = LC   -.70016*LY +   2.2877*PI   -1.2690*INPT
*******************************************************************************
 R-Squared                     .46234  R-Bar-Squared                    .42098
 S.E. of Regression          .0054148  F-Stat.    F(7,119)    14.3727[.000]
 Mean of Dependent Variable  .0055870  S.D. of Dependent Variable     .0071160
 Residual Sum of Squares     .0034305  Equation Log-likelihood        487.7664
 Akaike Info. Criterion      477.7664  Schwarz Bayesian Criterion     463.5455
 DW-statistic                  1.9835
*******************************************************************************
 R-Squared and R-Bar-Squared measures refer to the dependent variable
 dLC and in cases where the error correction model is highly
 restricted, these measures could become negative.
```

With the exception of the coefficient of DLC_{t-1}, all the other coefficients are statistically significant. The underlying $ARDL$ equation also passes all the diagnostic tests that are automatically computed by *Microfit*. The error correction coefficient, estimated at $-0.12599(0.036172)$ is statistically highly significant, has the correct sign and suggests a moderate speed of convergence to equilibrium. The larger the error correction coefficient (in absolute value) the faster will be the economy's return to its equilibrium, once shocked.

The F and W statistics for testing the existence of a long run level relationship between LC, LY, and PI, together with their critical value bounds at 90 and 95 per cent levels are also supplied at the bottom of the above result tables. These critical value bounds are close to the ones provide in Appendix B, but have the advantage that unlike the critical values in

Tables B.1 and B.2 they continue to be applicable even if shift dummy variables are included amongst the deterministic variables

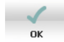
The above error correction model can also be used in forecasting the rate of change of consumption conditional on current and past changes in real disposable income and inflation. Proceed to the Post *ARDL* Model Selection Menu and choose option 4. Now click the  button twice and select to forecast change in *LC* over the period 1993(1)-1994(1). You should obtain the results in Table 16.15.

Table 16.15: Dynamic forecasts for the change in *LC*

```
                 Dynamic forecasts for the change in LC
                        Based on ARDL Regression
*****************************************************************************
 Based on  127 observations from 1961Q2 to 1992Q4.
 ARDL(2,2,3) selected using Akaike Information Criterion.
 Dependent variable in the ARDL model is LC included with a lag of 2.
 List of other regressors in the ARDL model:
 LY               LY(-1)           LY(-2)           PI               PI(-1)
 PI(-2)           PI(-3)           INPT
*****************************************************************************
 Observation              Actual             Prediction               Error
  1993Q1               -.0053759              .0034962              -.0088721
  1993Q2                .0065768              .0026170               .0039598
  1993Q3                .0090587              .0078788               .0011798
  1993Q4                .0066042              .0048395               .0017647
  1994Q1                 .010188              .0078575               .0023304
*****************************************************************************

             Summary Statistics for Residuals and Forecast Errors
*****************************************************************************
                     Estimation Period            Forecast Period
                     1961Q2 to 1992Q4             1993Q1 to 1994Q1
*****************************************************************************
 Mean                         -.3306E-8                   .7252E-4
 Mean Absolute                 .0040512                   .0036214
 Mean Sum Squares              .2701E-4                   .2087E-4
 Root Mean Sum Squares         .0051973                   .0045679
*****************************************************************************
```

The root mean squares of forecast errors of around 0.45 per cent per quarter compares favourably with the value of the same criterion computed over the estimation period. However, the model fails to forecast the extent of the fall in the non-durable consumption expenditures in the first quarter of 1993.

16.6 Lesson 16.6: Great ratios and long-run money demand in the US

In their paper, King, Plosser, Stock, and Watson (1991) (KPSW) examine long-run relations between (private) output (Y), consumption (C), investment (I), real money balances ($M - P$), the interest rate (R) and the rate of inflation ($DP = P - P(-1)$). With the exception of the interest rate all the other variables are in logarithms. Output, consumption, investment and money balances are measured on a *per capita* basis using civilian noninstitutional population as the deflator.[4] KPSW estimate two sets of cointegrated VAR models: a three-variable model containing the real variables, C, I and Y; and a six-variable model containing the real as well as the nominal variables. For the three-variable model they estimate restricted and unrestricted VARs over the 1949(1)-1988(4) period, but for the six-variable model they choose the shorter estimation period of 1954(1)-1988(4), to avoid dealing with the possible effects of the Korean War, and other rather special developments in the US economy on nominal variables. They also experimented with VARs of different orders, and settled on the order $p = 6$.

In this lesson we reconsider the cointegrating VAR models analysed by KPSW, and show how to analyze these models and their long-run properties using the cointegrating VAR options available in *Microfit.* The cointegrating VAR analysis involves a number of important steps, namely (i) ensuring that the jointly determined variables of the model are $I(1)$, (ii) deciding the order of the VAR model, (iii) identifying the nature of the deterministic variables such as intercepts and trends in the underlying VAR, (iv) resolving the identification problem of the long-run relations that arises when the number of the cointegrating relations is larger than unity, and (v) testing over-identifying restrictions on the long-run relations (if any). This analysis should also be supplemented by an examination of the short-run dynamic properties of the model, by considering the effect of variable-specific and system-wide shocks on the cointegrating (long-run relations) with the help of impulse response analysis and persistence profiles. See also Section 7.5.4.

With the above considerations in mind, load the file KPSW.FIT (containing quarterly observations over the period 1947(1)-1988(4) for the US economy), and ensure that the following variables are in the list:

C	Real *per capita* consumption ($GC82$, in logs)
DP	Inflation (rate of change of GNP deflator)
I	Investment *per capita* ($GIE82$, in logs)
MP	Real money balance $M2$ defined as log of $M2/P$, $P = GNP$ deflator
R	Interest rate ($FYGM3/100$)
Y	Real private output *per capita* ($GNP82 - GGE82$, in logs)

The first stage in the analysis is to ascertain the order of the integration of the variables. The simplest way to achieve this in *Microfit* is to use the **ADF** command, although as we have pointed out earlier the test results can be subject to a considerable margin of uncertainty.[5] To compute the ADF statistics for the six variables in the KPSW.FIT file, type in the box editor

ADF $C(6)$; **ADF** $I(6)$; **ADF** $MP(6)$;
ADF $Y(6)$; **ADF** $R(6)$; **ADF** $DP(6)$ RUN

[4] For further details of the variables and their sources see KPSW. We are grateful to Mark Watson for providing us with their data set.

[5] For more details on the ADF tests see Lesson 12.1.

The test results will appear on the screen. From the results for the ADF regression with both intercept and trend, we find that the null of a unit root is not rejected for C, MP, Y, and R, but may be rejected for the inflation rate, DP, and investment I. To check if any of these variables are $I(2)$, create their first-differences and apply the **ADF** command to them:

$$DC = C - C(-1);\ DI = I - I(-1);\ DMP = MP - MP(-1);$$
$$DY = Y - Y(-1);\ DR = R - R(-1);\ DDP = DP - DP(-1);$$
$$\textbf{ADF}\ DC(5);\ \textbf{ADF}\ DI(5);\ \textbf{ADF}\ DMP(5);$$
$$\textbf{ADF}\ DY(5);\ \textbf{ADF}\ DR(5);\ \textbf{ADF}\ DDP(5)$$

From the ADF statistics based on the regressions with an intercept term but no trend, we find that the null hypothesis that the first-differences of these variables have a unit root is strongly rejected in the case of all six variables. Hence, we conclude that C, MP, Y and R could be $I(1)$, but the evidence on DP and I is less certain, and they could be $I(0)$. In view of this uncertainty over the order of integration of DP, in what follows we shall focus on the remaining five $I(1)$ variables, namely C, I, MP, Y, and R. The analysis of the cointegrating VAR models when the orders of integration of the variables are unknown or are uncertain is beyond the scope of the present lesson. For later use we create a constant and a linear time trend by typing

$$INPT = 1;\ T = \textbf{CSUM}(1)$$

To decide on the order of the VAR, choose option 1 in the System Estimation Menu, and in the box editor that appears on the screen type

$$C \quad I \quad MP \quad Y \quad R \quad \& \quad INPT \quad T$$

where the intercept and time trend are included in each equation of the VAR system, since all variables seem to have a linear trend. Specify the sample period:

$$1954Q1 \quad 1988Q4$$

This is the same sample period as used by KPSW in their analysis of the six-variable VAR model. Set the maximum order of the VAR to 6 and click . Choose option 4 in the Unrestricted VAR Post Estimation Menu (see Section 7.4.1) and then option 1. The results in Table 16.16 should now appear on the screen.

Table 16.16: Selecting the order of the VAR model; King, Plosser, Stock and Watson (1991) data set

```
Test Statistics and Choice Criteria for Selecting the Order of the VAR Model
*******************************************************************************
Based on  140 observations from 1954Q1 to 1988Q4. Order of VAR = 6
List of variables included in the unrestricted VAR:
C               I               MP              Y               R
List of deterministic and/or exogenous variables:
INPT            T
*******************************************************************************
Order    LL        AIC       SBC             LR test           Adjusted LR test
6      2475.0    2315.0    2079.7              ------               ------
5      2449.4    2314.4    2115.9  CHSQ(25)=  51.1425[.002]   39.4528[.033]
4      2435.0    2325.0    2163.2  CHSQ(50)=  80.0707[.004]   61.7688[.123]
3      2422.1    2337.1    2212.0  CHSQ(75)= 105.9165[.011]   81.7070[.279]
2      2397.0    2337.0    2248.8  CHSQ(100)= 156.0384[.000]  120.3724[.081]
1      2327.5    2292.5    2241.0  CHSQ(125)= 294.9906[.000]  227.5642[.000]
0      1618.5    1608.5    1593.8  CHSQ(150)=    1712.9[.000]    1321.4[.000]
*******************************************************************************
AIC=Akaike Information Criterion     SBC=Schwarz Bayesian Criterion
```

On the basis of these results, the Akaike information criterion (AIC) selects order 3, and the Schwarz Bayesian criterion (SBC) selects the order 2. You can now inspect the estimates of the individual equations in the VAR for these orders. Since the values of AIC for the $VAR(2)$ and $VAR(3)$ specifications differ by a decimal point, and the sample size is small relative to the number of variables in the VAR, we choose the order 2 in the cointegrating VAR analysis that follows. Once again, the choice of the order of the VAR is subject to an important degree of uncertainty.

We now employ the cointegrating VAR option to test the null hypothesis that the five variables in the VAR are not cointegrated. If this hypothesis is rejected we can then move to the next stage and, on the assumption that there exists a cointegrating relation among the five variables, test the hypothesis that there are no more cointegrating relations among them, and so on. In this way we should obtain some idea as to the number of cointegrating relations that may exist among these variables. Given the complicated nature of this testing procedure, the overall size of the cointegration test is not known, even in large samples, and hence special care needs to be exercised when interpreting the test results.

Return to the System Estimation Menu, and this time choose option 2 from the System Estimation Menu (Multivariate Menu), selecting the Unrestricted intercepts, restricted trends option 4, since data are trended and we wish to avoid the possibility of quadratic trends in some of the variables.[6] Delete the intercept and the trend terms so that the Commands and Data Transformations box contains the following variables:

$$C \quad I \quad MP \quad Y \quad R$$

For the sample size type

$$1954Q1 \quad 1988Q4$$

[6]See Section 7.5 for the rationale behind the different options in this menu.

For the order of the VAR enter 2, then click . The test results in Table 16.17 should now appear on the screen.

Table 16.17: Cointegration properties of the King, Plosser, Stock and Watson (1991) model

```
 Cointegration with unrestricted intercepts and restricted trends in the VAR
   Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix
********************************************************************************
 140 observations from 1954Q1 to 1988Q4. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 C                I               MP               Y               R
 Trend
 List of eigenvalues in descending order:
.25832      .14277      .12331     .040600     .013589
********************************************************************************
 Null     Alternative      Statistic      95% Critical Value  90% Critical Value
 r = 0      r = 1           41.8376           37.8600              35.0400
 r<= 1      r = 2           21.5670           31.7900              29.1300
 r<= 2      r = 3           18.4237           25.4200              23.1000
 r<= 3      r = 4            5.8027           19.2200              17.1800
 r<= 4      r = 5            1.9155           12.3900              10.5500
********************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

 Cointegration with unrestricted intercepts and restricted trends in the VAR
          Cointegration LR Test Based on Trace of the Stochastic Matrix
********************************************************************************
 140 observations from 1954Q1 to 1988Q4. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 C                I               MP               Y               R
 Trend
 List of eigenvalues in descending order:
.25832      .14277      .12331     .040600     .013589
********************************************************************************
 Null     Alternative      Statistic      95% Critical Value  90% Critical Value
 r = 0      r>= 1           89.5464           87.1700              82.8800
 r<= 1      r>= 2           47.7088           63.0000              59.1600
 r<= 2      r>= 3           26.1418           42.3400              39.3400
 r<= 3      r>= 4            7.7182           25.7700              23.0800
 r<= 4      r = 5            1.9155           12.3900              10.5500
********************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

 Cointegration with unrestricted intercepts and restricted trends in the VAR
Choice of the Number of Cointegrating Relations Using Model Selection Criteria
********************************************************************************
 140 observations from 1954Q1 to 1988Q4. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 C                I               MP               Y               R
 Trend
 List of eigenvalues in descending order:
.25832      .14277      .12331     .040600     .013589
********************************************************************************
 Rank       Maximized LL         AIC              SBC              HQC
 r = 0         2352.2           2322.2           2278.1           2304.3
 r = 1         2373.1           2333.1           2274.3           2309.2
 r = 2         2383.9           2335.9           2265.3           2307.2
 r = 3         2393.1           2339.1           2259.7           2306.9
 r = 4         2396.0           2338.0           2252.7           2303.4
 r = 5         2397.0           2337.0           2248.8           2301.1
********************************************************************************
 AIC = Akaike Information Criterion     SBC = Schwarz Bayesian Criterion
 HQC = Hannan-Quinn Criterion
```

According to the maximum eigenvalue and the trace statistics in this Table, the null hypothesis of 'no cointegration' (namely $r = 0$) is rejected, but the null hypothesis that there exists one cointegrating relation (namely $r = 1$) cannot be rejected. So, we find only one statistically significant cointegrating relation among the five $I(1)$ variables. From economic theory, however, we expect three long-run relations among these variables, given by

$$z_1 = C - Y \tag{16.2}$$

$$z_2 = I - Y \tag{16.3}$$

$$z_3 = MP - c_4Y - c_5R \tag{16.4}$$

with $c_4 > 0$, and $c_5 < 0$. The z_1 and z_2 relations are known as the 'great ratios'. The balanced growth literature, and the more recent real business cycle literature both predict $C - Y$ and $I - Y$ as the two cointegrating relations. The relation defined by z_3 is the long-run money demand equation. The test results seem to be in conflict with economic theory. Notice, however, that the finite sample performance of Johansen's log-likelihood ratio statistics are not known, and also tends to be quite sensitive to the order of the VAR chosen. Hence, we will continue our analysis assuming that there are three cointegrating relations among C, I, MP, Y and R. Under this assumption we need three independent *a priori* restrictions on each of the three cointegrating relations to exactly identify them.[7]

In view of the theory-based restrictions implicit in the relations (16.2)-(16.4), we start the long-run structural analysis with the following just-identifying restrictions:

$$\begin{array}{lll} \beta_{11} = 1, & \beta_{12} = 0, & \beta_{13} = 0 \\ \beta_{21} = 0, & \beta_{22} = 1, & \beta_{23} = 0 \\ \beta_{31} = 0, & \beta_{32} = 0, & \beta_{33} = 1 \end{array}$$

where we have denoted the three cointegrating vector by

$$\begin{aligned} \boldsymbol{\beta}_1 &= (\beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}, \beta_{15}, \beta_{16})' \\ \boldsymbol{\beta}_2 &= (\beta_{21}, \beta_{22}, \beta_{23,} \beta_{24}, \beta_{25}, \beta_{26})' \\ \boldsymbol{\beta}_3' &= (\beta_{31}, \beta_{32}, \beta_{33}, \beta_{34}, \beta_{35}, \beta_{36})' \end{aligned}$$

Notice that there are six elements in each of these three vectors. The first five elements are the coefficients of the $I(1)$ variables C, I, MP, Y and R, respectively, and the last elements (β_{i6}, $i = 1, 2, 3$) refer to the time trend. Recall that under case 4, where the underlying VAR model contains linear deterministic trends with restricted coefficients, the trend term is automatically included as a part of the cointegrating relation. See Section 7.5.3. The above exactly identifying restrictions do not impose any testable restrictions on the cointegrating VAR model.

To estimate the VAR model subject to the above restrictions, click 

to leave the cointegration test results in Table 16.17, and move to the Cointegrating VAR Post Estimation Menu (see Section 7.5.3). Choose option 2 and set the number of cointegrating relations to 3. This returns to the Cointegrating VAR Post Estimation Menu to start your long-run structural analysis. Choose option 6 and in the Long-Run Structural Modelling Menu choose option 4. Type the exactly identifying restrictions in the box editor that appears on the screen:[8]

$$\begin{array}{lll} A1 = 1; & A2 = 0; & A3 = 0; \\ B1 = 0; & B2 = 1; & B3 = 0; \\ C1 = 0; & C2 = 0; & C3 = 1 \end{array}$$

---

[7] In principle, it seems more appropriate to select the order of the VAR and the number of the cointegrating relations simultaneously. This could be clearly achieved by means of the familiar model selction criteria, such as AIC or SBC. But as yet little is known about their small sample performance.

[8] Alternatively, you can retrieve the contents of the file CO3.EQU, in the tutorial directory, into the box editor by clicking the 

button.

The exactly identified ML estimates of the three cointegrating vectors $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_3$ should appear on the screen together with their asymptotic estimated standard errors in brackets. These results are shown in Table 16.18.

Table 16.18: An exactly identified structural long-run model for the US economy; King, Plosser, Stock and Watson (1991) data set

```
              ML estimates subject to exactly identifying restriction(s)
          Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                           Converged after 1 iterations
 Cointegration with unrestricted intercepts and restricted trends in the VAR
********************************************************************************
 140 observations from 1954Q1 to 1988Q4. Order of VAR = 2, chosen r =3.
 List of variables included in the cointegrating vector:
 C                I                MP               Y                R
 Trend
********************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                   Vector 1       Vector 2       Vector 3
 C                    1.0000          0.00           0.00
                   (    *NONE*)    (    *NONE*)    (    *NONE*)

 I                      0.00        1.0000           0.00
                   (    *NONE*)    (    *NONE*)    (    *NONE*)

 MP                     0.00          0.00         1.0000
                   (    *NONE*)    (    *NONE*)    (    *NONE*)

 Y                   -.88600       -1.2278        -1.4465
                   (   .095419)    (    .25405)    (    .14934)

 R                    .81756        1.2052         .61093
                   (    .27519)    (    .73451)    (    .40651)

 Trend             -.0010563      .1643E-3       .0013923
                   ( .4350E-3)    ( .0011584)    ( .6775E-3)

********************************************************************************
 LL subject to exactly identifying restrictions=   2393.1
********************************************************************************
```

The estimates in the first two columns (under vector 1 and vector 2) refer to the 'great ratios', and those under vector 3 refer to the money demand equation. Notice that the maximized value of the log-likelihood function given in this Table (namely 2393.1) is the same as the log-likelihood value reported in the third panel of Table 16.17 for $rank = 3$. Since we do not expect these long-run relations to include a linear trend we first test the following over-identifying restrictions:

$$\beta_{16} = \beta_{26} = \beta_{36} = 0$$

Click  to leave the screen with the exactly identified estimates, and then choose option 0 in the IR Analysis and Forecasting Menu (see Section 7.5.4). You will be asked whether you wish to test over-identifying restrictions on the cointegrating vectors (CVs). Click YES. The box editor appears on the screen containing the exactly identified restrictions. *Do not* delete these restrictions. Simply *add* the following over-identifying restrictions to them:

$$A6 = 0; \quad B6 = 0; \quad C6 = 0$$

You can add these restrictions anywhere in the box editor so long as all the restrictions are separated by semicolons. If you add these restrictions at the end of each row, your screen should look like Figure 16.3.

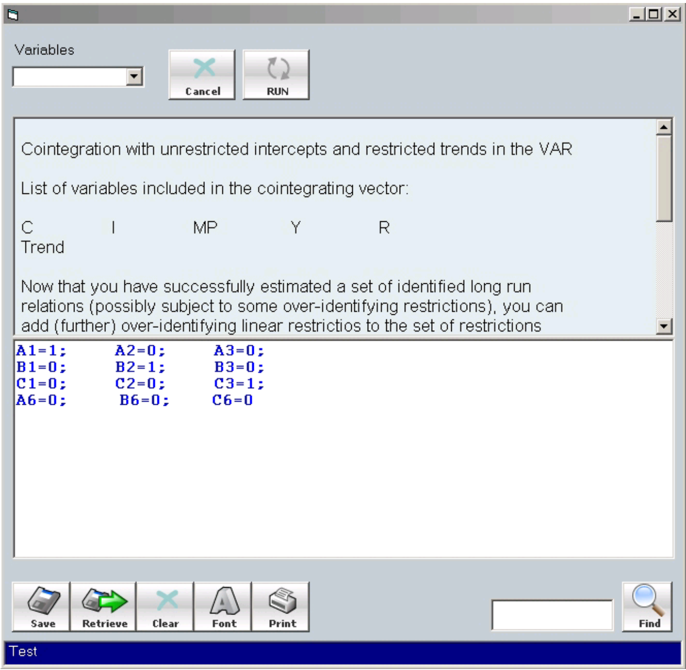


Figure 16.3: Testing for over-identifying restrictions

Now click  to process. Accept the initial estimates suggested by the program by clicking , and choose the 'Back substitution algorithm (B), new to *Microfit*'. Our experience suggests that in general this algorithm converges more often and faster than the back substitution algorithm (A) and the modified Newton-Rapson algorithm. The program now starts the computations and presents you with the ML estimates obtained subject to the over-identifying restrictions. Close the output window and click the button to accept the default number of replications and significance levels in the computation of critical values for the LR test statistic. The results are shown in Table 16.19.

Table 16.19: A structural long-run model for the US economy subject to three over-identifying restrictions; King, Plosser, Stock and Watson (1991) data set

```
              ML estimates subject to over identifying restriction(s)
        Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                          Converged after 23 iterations
 Cointegration with unrestricted intercepts and restricted trends in the VAR
*******************************************************************************
 140 observations from 1954Q1 to 1988Q4. Order of VAR = 2, chosen r =3.
 List of variables included in the cointegrating vector:
 C                I                MP               Y                R
 Trend
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                        Vector 1          Vector 2          Vector 3
 C                       1.0000            -.0000            -.0000
                       (    *NONE*)      (    *NONE*)      (    *NONE*)

 I                       -.0000            1.0000             .0000
                       (    *NONE*)      (    *NONE*)      (    *NONE*)

 MP                       .0000            -.0000            1.0000
                       (    *NONE*)      (    *NONE*)      (    *NONE*)

 Y                      -1.1521           -1.1881           -1.0999
                       (  .074588)       (   .10604)       (  .084131)

 R                        .99688           1.1903            .40384
                       (   .60665)       (   .84744)       (   .67691)

 Trend                    .0000             .0000            -.0000
                       (    *NONE*)      (    *NONE*)      (    *NONE*)

*******************************************************************************
 LR Test of Restrictions          CHSQ(3)=    6.4807[.090]
 95% Bootstrapped Critical Value =  14.7941
 90% Bootstrapped Critical Value =  12.4658
 Bootstrapped simulations based on 1000 SIMULATIONS.
 DF=Total no of restrictions(12) - no of just-identifying restrictions(9)
 LL subject to exactly identifying restrictions=   2393.1
 LL subject to over-identifying restrictions=   2389.9
*******************************************************************************
```

The LR statistic for testing the three over-identifying restrictions is computed to be 6.48 which is below the 95 per cent critical value of the χ^2 distribution with 3 degrees of freedom, and below its 95 per cent bootstrapped critical value (14.79). We therefore do not reject the hypothesis that there are no linear trends in the cointegrating relations, although there was a linear trend in the underlying VAR model. Hence the cointegrating relations are also 'co-trending'.

We can now consider imposing further over-identifying restrictions. In the context of the present model there are two sets of restrictions implied by the 'great ratios' (16.2) and (16.3): namely that the interest rates do not enter these relationships; and that Y enters with a coefficient of -1. To impose these restrictions choose option 0 in the IR Analysis and Forecasting Menu and say that you wish to test (further) over-identifying restrictions, and when presented with the box editor add the following four restrictions:

$$A4 = -1; \qquad A5 = 0;$$
$$B4 = -1; \qquad B5 = 0;$$

and then carry out the necessary computations by accepting all the defaults suggested by the program, and computing the boostrapped critical values for the LR test statistic. You should obtain the results in Table 16.20.

Table 16.20: A structural long-run model for the US economy subject to seven over-identifying restrictions; King, Plosser, Stock and Watson (1991) data set

```
           ML estimates subject to over identifying restriction(s)
      Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                       Converged after 14 iterations
 Cointegration with unrestricted intercepts and restricted trends in the VAR
*****************************************************************************
 140 observations from 1954Q1 to 1988Q4. Order of VAR = 2, chosen r =3.
 List of variables included in the cointegrating vector:
 C                I                MP               Y                R
 Trend
*****************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                   Vector  1        Vector  2        Vector  3
 C                   1.0000           -.0000            .0000
                   (   *NONE*)      (   *NONE*)      (   *NONE*)

 I                   -.0000           1.0000            .0000
                   (   *NONE*)      (   *NONE*)      (   *NONE*)

 MP                  -.0000            .0000           1.0000
                   (   *NONE*)      (   *NONE*)      (   *NONE*)

 Y                  -1.0000          -1.0000          -1.2252
                   (   *NONE*)      (   *NONE*)      (  .079324)

 R                   -.0000           -.0000           1.2622
                   (   *NONE*)      (   *NONE*)      (   .44862)

 Trend                .0000            .0000            .0000
                   (   *NONE*)      (   *NONE*)      (   *NONE*)

*****************************************************************************
 LR Test of Restrictions           CHSQ(7)=  17.3821[.015]
 95% Bootstrapped Critical Value =  28.2640
 90% Bootstrapped Critical Value =  25.4445
 Bootstrapped simulations based on 1000 SIMULATIONS.
 DF=Total no of restrictions(16) - no of just-identifying restrictions(9)
 LL subject to exactly identifying restrictions=    2393.1
 LL subject to over-identifying restrictions=   2384.4
*****************************************************************************
```

The LR statistic reported in this Table, if compared with the 95 per cent critical value of the χ^2 distribution with 7 degrees of freedom, rejects the over-identifying restrictions. However, notice that the LR statistic is below its 95 per cent bootstrapped critical value. Also, the estimates of the long-run income and interest rate elasticities in the money demand equation (vector 3) have the correct signs and are of orders of magnitudes that one expects. The long-run income elasticity is 1.225(0.079), and the long-run interest rate elasticity is $-1.26(0.45)$, although the latter is not very precisely estimated. The asymptotic standard errors are in brackets. In view of the tendency of the LR tests to over-reject the null hypothesis, and the strong theoretical underpinning of the three relations, we shall now adopt the cointegrating vectors in Table 16.20 and analyse the short-run dynamic properties of the model.

We shall first examine the effect of system-wide shocks on the cointegrating (or long-run) relations, by plotting the 'persistence profiles' of these relations. Choose option 4 in the IR Analysis and Forecasting Menu, and accept the default horizon of 50 quarters for the persistence profiles. You should see the estimates listed on the screen. To obtain a plot of these estimates click Close to move to the Impulse Response Results Menu, and choose option 2 to graph the profiles. Figure 16.4 shows the plots of the persistence profiles for all the three long-run relations.

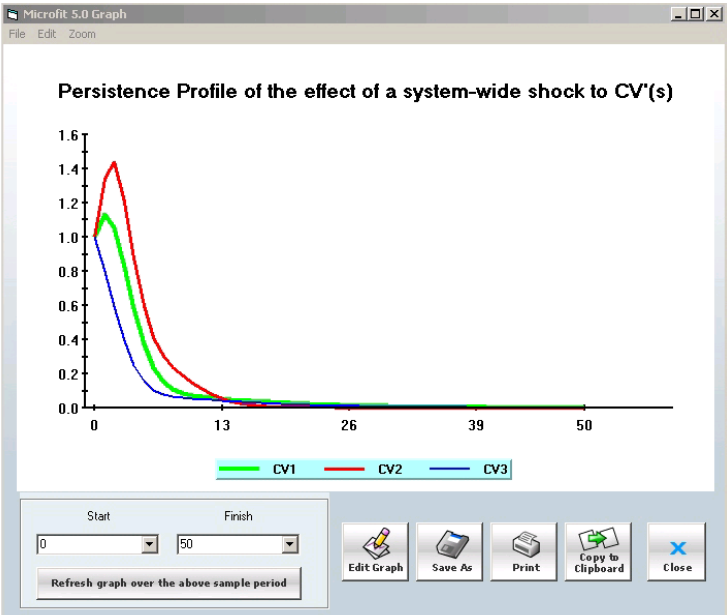


Figure 16.4: Persistence profile of the effect of a system-wide shock to CV(s)

These profiles clearly show that while all the three relations have a strong tendency to converge to their respective equilibria, the speed of convergence of the money demand equation (vector $CV3$) to its equilibrium is noticeably faster than those of the two great ratios, $CV1 = C - Y$ and $CV2 = I - Y$, with the consumption output ratio equilibrating faster than the investment-output ratio. These profiles also show a marked over-shooting effect for the consumption-output and investment-output ratios. Close the graph and select option 0. You are asked whether you wish to compute the boostrapped estimates of the confidence intervals of impulse responses. Click the Bootstrap button to accept the default number of replications and confidence level, or change to your desired values. *Microfit* starts the computations, and when finished presents you the list of the three cointegrating vectors, for which you can inspect the bootstrapped distribution. Choose for example $CV3$, click OK, and then close the output window to return to the 'Impulse Response Results Menu'. Select the graph option 2; and then choose the point estimate of $CV3$, its top 97.5 and lower 2.5 percentiles. You will be presented with Figure 16.5.

To see the effect of variable- (or equation-) specific shocks on the cointegrating vectors

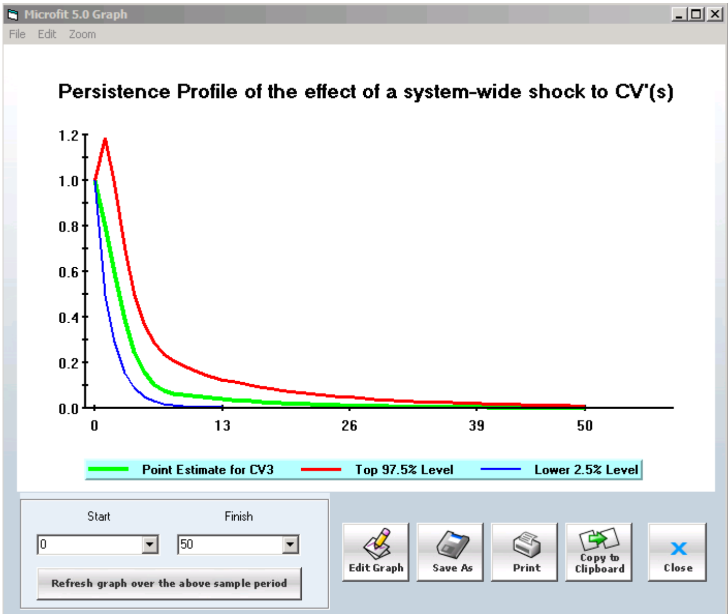


Figure 16.5: Persistence profile of the effect of a system-wide shock to $CV3$

you need to choose option 3 in the IR Analysis and Forecasting Menu. In this case you can either use the Orthogonalized or the Generalized Impulse Responses. If you choose the Orthogonalized IR option and choose to consider the effect of shocking the nominal interest on the cointegrating vectors you should obtain Figure 16.6.

Again, you can plot the mean, median, 97.5 and 2.5 percentiles of the bootstrapped distribution for the impulse response functions for each cointegrating vector.

Choosing the Generalized IR Option produces Figure 16.7.

In both cases the effect of the shocks on the cointegrating relations dies out, although the profiles have very different shapes. The two sets of profiles would have been, however, identical if we had decided to shock the first variable in the VAR, namely consumption (see Section 7.5.4).

Consider now the dynamic effects of a shock to the output equation on all the variables in the cointegrating VAR model. Choose option 1 in the IR Analysis and Forecasting Menu, select the Orthogonalized IR option, and then choose the Y equation to shock. The orthogonalized impulse responses should now appear on the screen. Notice that the impact effect of shocking Y on the variables that have been entered in the VAR before it (namely C, I and MP) are zero by construction. Also note that the impact effect of a unit shock in output (measured as one standard error) on the nominal interest rate is positive, but relatively small. To see the plots of these impulse responses, close the output window and then choose to plot the orthogonalized impulse responses for C, I, MP and Y. You should see Figure 16.8 on the screen.

A similar graph can also be obtained using the Generalized IR option. This option is

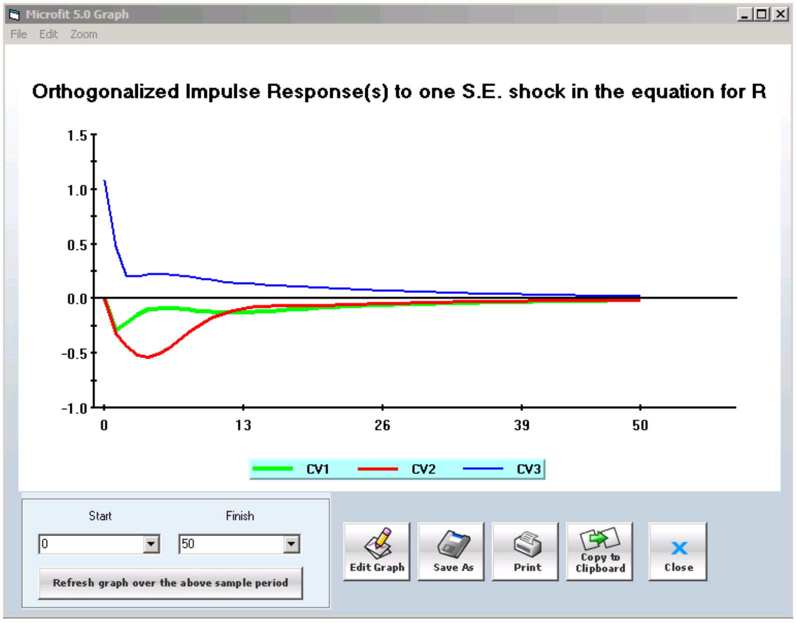


Figure 16.6: Orthogonalized impulse responses to one SE shock in the equation for R

more satisfactory, and unlike the orthogonalized IR option does not impose zero impact effects on C, I and MP when Y is shocked. For example, using the Generalized IR option the impact effect of a unit shock to output on investment is quite large, and is in fact slightly larger than the effect on output itself. The result is given in Figure 16.9, and suggests an important cyclical effect of an output shock on investment, with investment responding very strongly to an output shock and then declining very sharply.

Other options in the IR Analysis and Forecasting Menu can also be used to compute multivariate dynamic forecasts, based on the cointegrating VAR model subject to the over-identifying restrictions. See Lesson 16.7. You can also obtain error correction relations for all the variables in the VAR. Suppose you are interested in the error correction equation for consumption. Choose option 8 in this menu, and when asked for the choice of the variable, click 

to choose consumption. The results in Table 16.21 should now appear on the screen.

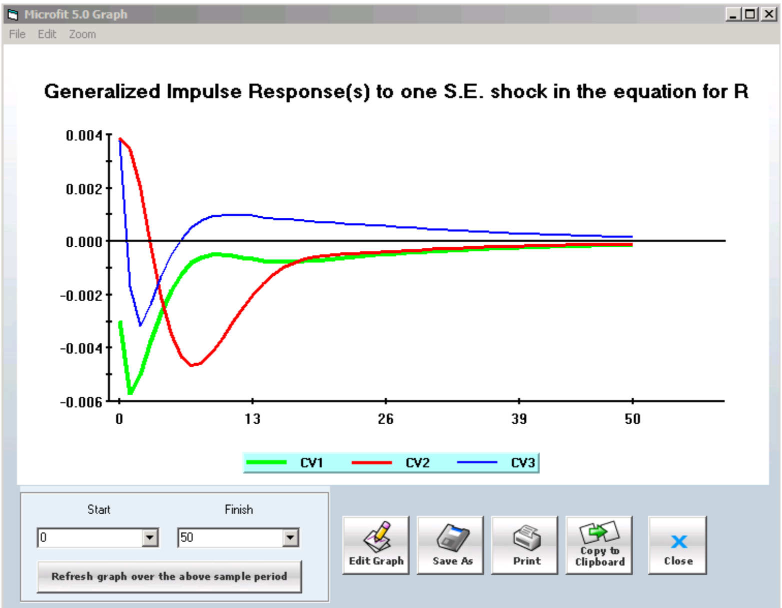


Figure 16.7: Generalized impulse responses to one SE shock in the equation for R

Table 16.21: Estimates of the error correction model for the variable C over the period 1954(1) to 1988(4)

```
      ECM for variable C estimated by OLS based on cointegrating VAR(2)
*******************************************************************************
 Dependent variable is dC
 140 observations used for estimation from 1954Q1 to 1988Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 Intercept                -.048594            .029784            -1.6315[.105]
 dC1                      -.032789             .10213            -.32104[.749]
 dI1                       .034235            .032717             1.0464[.297]
 dMP1                       .14640            .072354             2.0233[.045]
 dY1                        .12490            .069410             1.7995[.074]
 dR1                       -.23200            .073506            -3.1561[.002]
 ecm1(-1)                -.0038103            .030946            -.12313[.902]
 ecm2(-1)                -.0054789            .014163            -.38684[.700]
 ecm3(-1)                  .055074            .019651             2.8027[.006]
*******************************************************************************
 List of additional temporary variables created:
 dC = C-C(-1)

 dC1 = C(-1)-C(-2)
 dI1 = I(-1)-I(-2)
 dMP1 = MP(-1)-MP(-2)
 dY1 = Y(-1)-Y(-2)
 dR1 = R(-1)-R(-2)
 ;ecm1 =     1.0000*C     -.0000*I     -.0000*MP    -1.0000*Y +     .0000*R     -.000
 0*Trend;ecm2 =      .0000*C +    1.0000*I     -.0000*MP    -1.0000*Y     -.0000*R +
 .0000*Trend;ecm3 =     -.0000*C     -.0000*I +    1.0000*MP    -1.2252*Y +    1
 .2622*R +     .0000*Trend
*******************************************************************************
 R-Squared                       .27311   R-Bar-Squared                    .22872
 S.E. of Regression             .0063724   F-Stat.     F(8,131)     6.1526[.000]
 Mean of Dependent Variable     .0045623   S.D. of Dependent Variable    .0072560
 Residual Sum of Squares        .0053195   Equation Log-likelihood       513.8094
 Akaike Info. Criterion         504.8094   Schwarz Bayesian Criterion    491.5720
 DW-statistic                     2.1860   System Log-likelihood           2384.4
*******************************************************************************


                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version         *         F Version         *
*******************************************************************************
*                     *                           *                           *
* A:Serial Correlation*CHSQ(4)  =  11.3399[.023]*F(4,127)     =    2.7984[.029]*
*                     *                           *                           *
* B:Functional Form   *CHSQ(1)  =   2.0946[.148]*F(1,130)     =    1.9745[.162]*
*                     *                           *                           *
* C:Normality         *CHSQ(2)  =  16.7897[.000]*       Not applicable        *
*                     *                           *                           *
* D:Heteroscedasticity*CHSQ(1)  =   1.0879[.297]*F(1,138)     =    1.0808[.300]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

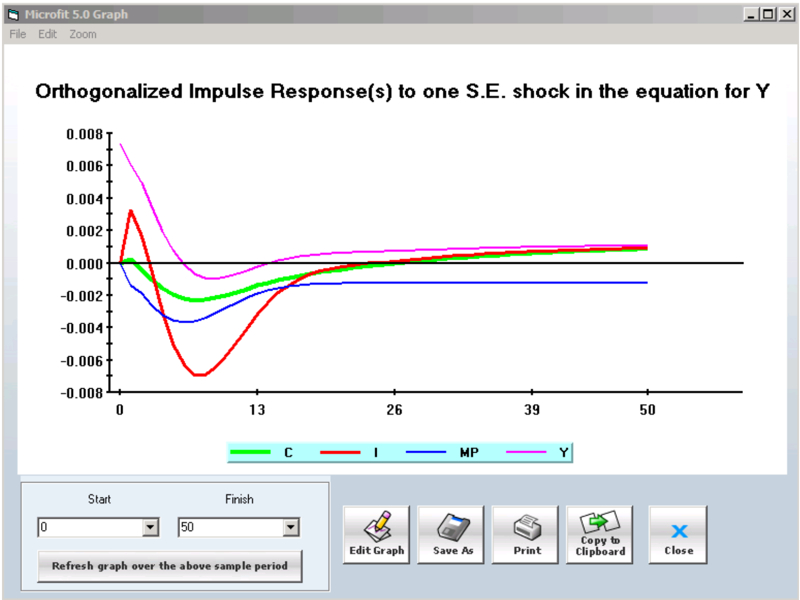


Figure 16.8: Orthogonalized impulse response(s) to one SE shock in the equation for Y

In the case of this equation only the error correction term associated with the long-run money demand equation has a significant impact on consumption growth. The equation also suffers from residual serial correlation, reflecting the fact that the order 2 chosen for the underlying VAR is not high enough to deal with the problem of the residual serial correlation. This is in line with the LR test statistics reported in Table 16.16.

16.7 Lesson 16.7: Application of the cointegrating VAR analysis to the UK term structure of interest rates

Let $R(k,t)$ be the rate of interest with a maturity of k periods as observed at the beginning of time t. The expectations theory of the term structure of interest rate postulates that

$$R(k,t) = \frac{1}{k}\left[E_t R(1,t) + E_t R(1,t+1) + ... + E_t R(1,t+k-1)\right] + L(k,t) \tag{16.5}$$

where $L(k,t)$ is the risk/liquidity premium. Notice that $E_t R(1,t) = R(1,t)$, which is known at the beginning of period t, and equation (16.5) simply states that the average expected rate of return of investing a certain sum of money in k successive time periods should be equal to the rate of return of investing this sum of money for k periods with the fixed rate of return $R(k,t)$, (after allowing for a risk/liquidity premium). Equation (16.5) can be rearranged to give

$$R(k,t) - R(1,t) = \frac{1}{k}\left[\sum_{i=1}^{k-1}\sum_{j=1}^{i} E_t\left\{\Delta R(1,t+j)\right\}\right] + L(k,t) \tag{16.6}$$

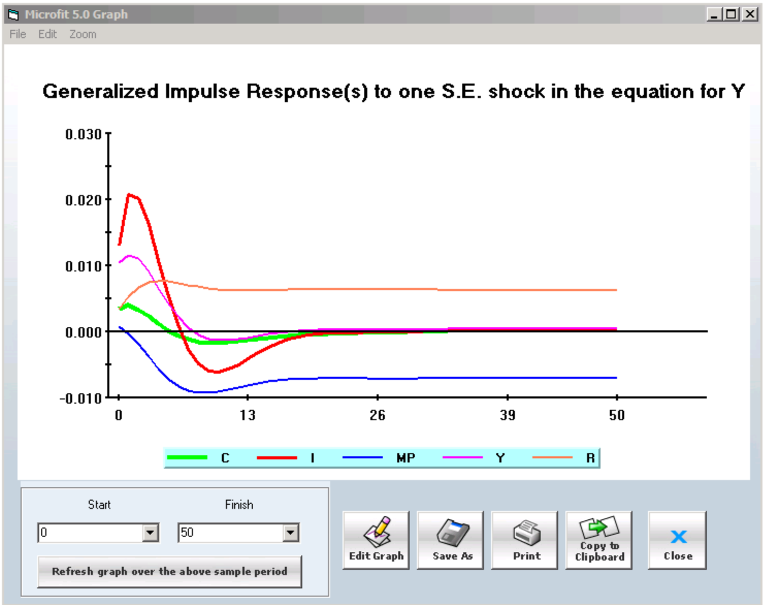


Figure 16.9: Generalized impulse response(s) to one SE shock in the equation for Y

Assuming that $R(k,t)$ is $I(1)$ and $L(k,t)$ is stationary, it can be seen that the right-hand side of (16.6) is stationary and hence the left-hand side of (16.6) which represents the spread between two interest rates $R(k,t)$ and $R(1,t)$, should also be stationary.

The above analysis shows that we should expect to find $(n-1)$ cointegrating relations between a set of n interest rates with different maturities. These can be represented by

$$R(k,t) = R(1,t) + a_k; \quad k = 2, n \tag{16.7}$$

Hall, Anderson, and Granger (1992) have applied the above model to analysis the US term structure of interest rates, and have found strong support for the existence of $(n-1)$ cointegrating vectors of the form (16.7) amongst a set of n interest rates with different maturities. We should point out that we expect a_k to be positive and rise with maturity.

In what follows we apply the cointegrating VAR techniques to the UK term structure, in the case of London Interbank Offer Rates (LIBOR) at different maturities (for further details see Pesaran and Wright 1995).

The special *Microfit* file TERMUK.FIT contains monthly observations on LIBORs with 1 month, 3 months, 6 months and 12 months maturities, which we denote by $R1$, $R3$, $R6$ and $R12$, respectively. This file also contains the UK effective exchange rate, EER.

In order to apply the cointegrating VAR technique to the above data set we need to make a number of decisions regarding:

1. The variables to be included as $I(1)$ endogenous (jointly determined) variables.

2. The variables to be included as $I(1)$ exogenous variables.

3. The variables to be included as additional $I(0)$ variables in the VAR.

4. The inclusion and nature (restricted or unrestricted) of the intercept and/or time trend in the VAR.

5. The selection of the order of the VAR.

We obviously should include $R1$, $R3$, $R6$ and $R12$, as endogenous $I(1)$ variables. As an exercise you should convince yourself that all these four variables are $I(1)$ before including them in the cointegrating VAR analysis. In this application there are no $I(1)$ exogenous variables.

The additional $I(0)$ variables included in the VAR allow for the short-run movements in the $I(1)$ variables which moves them away from their long-run equilibrium. We propose to include the lag of the percentage change of EER as well as three dummy variables which take the value of 1 in 1984(8), 1985(2) and 1992(10), and zeros elsewhere. The inclusion of the last dummy variable is intended to capture the effect of UK exit from the ERM, and the first two dummy variables are included because on each occasion the interest rates were raised by 2 per cent by the Bank of England, and could be regarded as outliers for our purposes (namely identification of the long-run relations). Since interest rates are not trended, we should not include a trend in the VAR, and given the relationships in (16.7), we expect each cointegrating vector to include an intercept term and moreover the intercept term, should be restricted.

In order to determine the order of the VAR we first run an unrestricted VAR of a relatively high order (12, for example). We also include an intercept term, $\Delta \log EER(-1)$ and the above three dummy variables as additional $I(0)$ deterministic/exogenous variables.

We therefore, load the file $TERMUK.FIT$ into *Microfit*, click Retrieve, and retrieve the file $TERMUK.EQU$ which contains the necessary transformations. Click RUN to create the above three dummy variables. The following variables should now be on your workspace

$$\begin{array}{lllllll} R1 & R3 & R6 & R12 & EER & INPT & DLEER \\ D84M8 & D85M2 & D92M10 & & & & \end{array}$$

Now move to the System Estimation Menu, choose option 1 (the unrestricted VAR) and type the following in the Commands and Data Transformations box:

$$\begin{array}{lllllll} R1 & R3 & R6 & R12 & \& & INPT & DLEER(-1) \\ D84M8 & D85M2 & D92M10 & & & & \end{array}$$

Choose all available observations and select the maximum order of the VAR to be 12. In order to decide whether it is justifiable to consider a VAR model with an order less than 12, choose option 4 in the VAR Post Estimation Menu, move to the Unrestricted VAR Hypothesis Testing Menu, and then choose option 1. You will be presented with Table 16.22.

Table 16.22: Selecting the order of the VAR for the analysis of the UK term structure of interest rates

```
Test Statistics and Choice Criteria for Selecting the Order of the VAR Model
*******************************************************************************
 Based on  165 observations from 1981M1 to 1994M9. Order of VAR = 12
 List of variables included in the unrestricted VAR:
 R1              R3              R6              R12
 List of deterministic and/or exogenous variables:
 INPT            DLEER(-1)       D84M8           D85M2           D92M10
*******************************************************************************
 Order    LL         AIC       SBC             LR test          Adjusted LR test
 12   223.7027   11.7027 -317.5275              ------               ------
 11   213.2903   17.2903 -287.0924  CHSQ(16)=  20.8248[.185]   14.1357[.589]
 10   201.6576   21.6576 -257.8775  CHSQ(32)=  44.0902[.076]   29.9279[.572]
 9   197.4085   33.4085 -221.2791  CHSQ(48)=  52.5885[.301]   35.6964[.905]
 8   189.1103   41.1103 -188.7296  CHSQ(64)=  69.1847[.307]   46.9618[.946]
 7   179.1601   47.1601 -157.8323  CHSQ(80)=  89.0853[.228]   60.4700[.949]
 6   155.9621   39.9621 -140.1827  CHSQ(96)= 135.4811[.005]   91.9629[.598]
 5   145.1763   45.1763 -110.1210  CHSQ(112)= 157.0529[.003]  106.6056[.626]
 4   134.5754   50.5754  -79.8743  CHSQ(128)= 178.2546[.002]  120.9970[.657]
 3   121.0112   53.0112  -52.5909  CHSQ(144)= 205.3830[.001]  139.4115[.592]
 2   110.4686   58.4686  -22.2859  CHSQ(160)= 226.4681[.000]  153.7238[.625]
 1    93.9668   57.9668    2.0598  CHSQ(176)= 259.4719[.000]  176.1264[.483]
 0  -312.1241 -332.1241 -363.1836  CHSQ(192)=    1071.7[.000]  727.4255[.000]
*******************************************************************************
 AIC=Akaike Information Criterion     SBC=Schwarz Bayesian Criterion
```

We can see that according to the AIC the order of the VAR should be 2, whilst the SBC indicates choosing 1 as the order of the VAR. Since we have a reasonable number of observations we choose the order of VAR to be 2, bearing in mind that choosing the higher order is less damaging than choosing the lower order, when the sample size is reasonably large. Also we notice that if we choose option 2 in the VAR Hypothesis Testing Menu to test for the significance of deletion of $DLEER(-1)$ in the VAR, and type

$$DLEER(-1)$$ RUN

in the box editor, we will be presented with Table 16.23 with a $\chi^2(4) = 23.0018$, which is highly significant, justifying the inclusion of $DLEER(-1)$ in each equation of the VAR.

Table 16.23: Testing for deletion of $DLBER(-1)$ from the VAR

```
     LR Test of Deletion of Deterministic/Exogenous Variables in the VAR
******************************************************************************
 Based on  165 observations from 1981M1 to 1994M9. Order of VAR = 12
 List of variables included in the unrestricted VAR:
 R1              R3              R6              R12
 List of deterministic and/or exogenous variables:
 INPT            DLEER(-1)       D84M8           D85M2           D92M10
 Maximized value of log-likelihood =  223.7027
******************************************************************************
 List of variables included in the restricted VAR:
 R1              R3              R6              R12
 List of deterministic and/or exogenous variables:
 INPT            D84M8           D85M2           D92M10
 Maximized value of log-likelihood =  212.2018
******************************************************************************
 LR test of restrictions, CHSQ(4) =  23.0018[.000]
******************************************************************************
```

We are now in a position to estimate a cointegrating VAR and we should choose option 2 in the System Estimation Menu (Multivariate Menu), selecting the Restricted Intercept, no trends option 2, and type

$$R1 \quad R3 \quad R6 \quad R12 \quad \& \quad DLEER(-1) \quad D84M8 \quad D85M2 \quad D92M10$$

in the box editor. Remember not to include $INPT$ amongst the above.

Choose all available observations, choose the order of VAR to be 2 and you will now be presented with Table 16.24, and as we expect we find support for the existence of 3 cointegrating vectors amongst these 4 interest rates.

Table 16.24: Sequence of log-likelihood ratio statistics for testing the rank of the long-run multiplier matrix

```
      Cointegration with restricted intercepts and no trends in the VAR
   Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix
********************************************************************************
 175 observations from 1980M3 to 1994M9. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 R1            R3              R6              R12             Intercept
 List of unrestricted deterministic variables included in the VAR:
 Intercept     D84M8           D85M2           D92M10
 List of eigenvalues in descending order:
.39641     .31188     .11064     .022479
********************************************************************************
 Null     Alternative     Statistic     95% Critical Value  90% Critical Value
 r = 0      r = 1          88.3497          28.2700              25.8000
 r<= 1      r = 2          65.4148          22.0400              19.8600
 r<= 2      r = 3          20.5200          15.8700              13.8100
 r<= 3      r = 4           3.9787           9.1600               7.5300
********************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

      Cointegration with restricted intercepts and no trends in the VAR
          Cointegration LR Test Based on Trace of the Stochastic Matrix
********************************************************************************
 175 observations from 1980M3 to 1994M9. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 R1            R3              R6              R12             Intercept
 List of unrestricted deterministic variables included in the VAR:
 Intercept     D84M8           D85M2           D92M10
 List of eigenvalues in descending order:
.39641     .31188     .11064     .022479
********************************************************************************
 Null     Alternative     Statistic     95% Critical Value  90% Critical Value
 r = 0      r>= 1          178.2633         53.4800              49.9500
 r<= 1      r>= 2           89.9136         34.8700              31.9300
 r<= 2      r>= 3           24.4987         20.1800              17.8800
 r<= 3      r = 4            3.9787          9.1600               7.5300
********************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

      Cointegration with restricted intercepts and no trends in the VAR
Choice of the Number of Cointegrating Relations Using Model Selection Criteria
********************************************************************************
 175 observations from 1980M3 to 1994M9. Order of VAR = 2, chosen r =1.
 List of variables included in the cointegrating vector:
 R1            R3              R6              R12             Intercept
 List of unrestricted deterministic variables included in the VAR:
 Intercept     D84M8           D85M2           D92M10
 List of eigenvalues in descending order:
.39641     .31188     .11064     .022479
********************************************************************************
 Rank       Maximized LL       AIC            SBC             HQC
 r = 0        -19.2767       -51.2767       -101.9133        -71.8163
 r = 1         24.8982       -15.1018        -78.3975        -40.7764
 r = 2         57.6056        11.6056        -61.1845        -17.9201
 r = 3         67.8656        17.8656        -61.2541        -14.2276
```

In order to identify and test relationships of the form (16.7), choose option 2 in the Cointegrating VAR Post Estimation Menu, then choose the number of cointegrating vectors to be three. Then choose option 6 to move to the Long-Run Structural Modelling Menu. In this menu choose Option 4 for carrying out an LR test of imposing general restrictions on the cointegrating vectors. You will be asked to specify exactly three identifying restrictions for each of the three cointegrating vectors. In the box editor you should type the following:

$$\begin{array}{lll} A2=-1; & A3=0; & A4=0; \\ B2=0; & B3=-1; & B4=0; \\ C2=0; & C3=0; & C4=-1; \end{array}$$

You should see Table 16.25.

Table 16.25: The exact identification of the cointegrating vectors for the UK term structure of interest rates

```
            ML estimates subject to exactly identifying restriction(s)
        Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                           Converged after 1 iterations
        Cointegration with restricted intercepts and no trends in the VAR
******************************************************************************
 175 observations from 1980M3 to 1994M9. Order of VAR = 2, chosen r =3.
 List of variables included in the cointegrating vector:
 R1              R3              R6              R12             Intercept
 List of unrestricted deterministic variables included in the VAR:
 Intercept       D84M8           D85M2           D92M10
******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                   Vector 1       Vector 2       Vector 3
 R1                   .98321         .93871         .86499
                   (  .011437)    (  .028626)    (  .050086)

 R3                  -1.0000           0.00           0.00
                   (   *NONE*)    (   *NONE*)    (   *NONE*)

 R6                     0.00        -1.0000           0.00
                   (   *NONE*)    (   *NONE*)    (   *NONE*)

 R12                    0.00           0.00        -1.0000
                   (   *NONE*)    (   *NONE*)    (   *NONE*)

 Intercept            .29809         .85018         1.7407
                   (   .13362)    (   .33287)    (   .58088)

******************************************************************************
 LL subject to exactly identifying restrictions=  67.8656
******************************************************************************
```

In order to impose and test the three over-identifying restrictions by restricting the coefficient of $R1$ to be unity in each of the three cointegrating vectors choose option 0 of the IR Analysis and Forecasting Menu and answer YES to the question which follows. You should add the restrictions one by one and gradually. Therefore, first add the restriction $A1 = 1$ to the existing set of restrictions and carry out the test. Then repeat the procedure for each of the extra restrictions $B1 = 1$ and $C1 = 1$. At each stage you should choose the initial values suggested by *Microfit*, choose the default Back substitution (B) algorithm, and compute the bootstrapped critical values of the LR statistic using the default number of replications and significance levels. The final result is presented in Table 16.26, which shows the LR statistic for testing the three over-identifying restrictions given by $\chi^2(3) = 8.738$.

Table 16.26: Test of over-identifying restrictions for the UK term structure of interest rates

```
            ML estimates subject to over identifying restriction(s)
       Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                          Converged after 9 iterations
       Cointegration with restricted intercepts and no trends in the VAR
*******************************************************************************
 175 observations from 1980M3 to 1994M9. Order of VAR = 2, chosen r =3.
 List of variables included in the cointegrating vector:
 R1                R3                R6                R12               Intercept
 List of unrestricted deterministic variables included in the VAR:
 Intercept         D84M8             D85M2             D92M10
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                   Vector 1          Vector 2          Vector 3
 R1                   1.0000            1.0000            1.0000
                   (   *NONE*)       (   *NONE*)       (   *NONE*)

 R3                  -1.0000             .0000            -.0000
                   (   *NONE*)       (   *NONE*)       (   *NONE*)

 R6                   -.0000           -1.0000             .0000
                   (   *NONE*)       (   *NONE*)       (   *NONE*)

 R12                   .0000            -.0000           -1.0000
                   (   *NONE*)       (   *NONE*)       (   *NONE*)

 Intercept            .13974            .26894            .45695
                   (  .058156)       (   .16579)       (   .32002)

*******************************************************************************
 LR Test of Restrictions           CHSQ(3)=    8.7382[.033]
 95% Bootstrapped Critical Value =  10.0094
 90% Bootstrapped Critical Value =   7.9149
 Bootstrapped simulations based on 1000 SIMULATIONS.
 DF=Total no of restrictions(12) - no of just-identifying restrictions(9)
 LL subject to exactly identifying restrictions=  67.8656
 LL subject to over-identifying restrictions=  63.4965
*******************************************************************************
```

If we compare the LR statistic with the 95 per cent critical value of a χ^2 with 3 degrees of freedom, theory restrictions are rejected. However, notice that the LR statistic is below the 95 per cent bootstrapped critical value. Hence, in view of the over-rejection tendency of the LR statistic, we do not reject the restrictions suggested by the economic theory.

It is also possible to estimate the generalized impulse responses of all interest rates to a unit shock in equation for $R1$ (say). To do this choose option 1 in the IR Analysis and Forecasting Menu for these restricted CVs. Select option 2 to choose the Generalized Impulse Response, and choose the horizon for the impulse responses to be 50 months. Then choose $R1$ to be shocked. You can inspect the results that follow, or more instructively you can choose option 2 in the following Impulse Response Results Menu to obtain a graphic display of these impulse responses. Choose all 4 variables to be included in the graph by checking all boxes, and then click 

. You will be presented with Figure 16.10.

As you can see, all interest rates converge to the same level after the effect of the shock dies away. You can also compute forecasts from this cointegrating VAR model using these restricted CVs. Choose option 6 in the IR Analysis and Forecasting Menu (Restricted CVs) and click OK to forecast all the remaining data periods. Choose to view the forecasts for the change in $R1$. You will be presented with the results in Table 16.27.

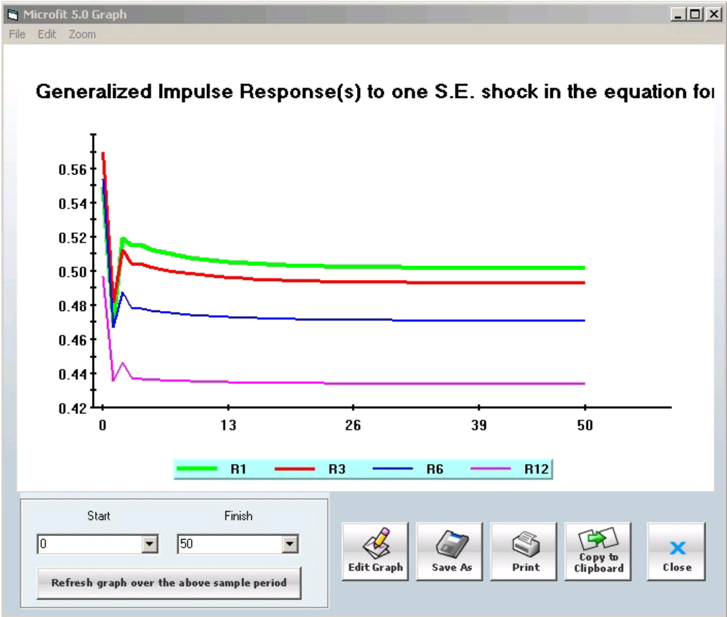


Figure 16.10: Generalized impulse response(s) to one SE shock in the equation for $R1$

Table 16.27: Multivariate dynamic forecasts of the change in the one-month LIBOR

Multivariate dynamic forecasts for the change in R1
Cointegration with restricted intercepts and no trends in the VAR

175 observations from 1980M3 to 1994M9. Order of VAR = 2, chosen r =3.
List of variables included in the cointegrating vector:

R1	R3	R6	R12	Intercept

List of unrestricted deterministic variables included in the VAR:

Intercept	D84M8	D85M2	D92M10

Observation	Actual	Prediction	Error
1994M10	.44000	.44561	-.0056118
1994M11	.29500	.088424	.20658

Summary Statistics for Residuals and Forecast Errors

	Estimation Period 1980M3 to 1994M9	Forecast Period 1994M10 to 1994M11
Mean	-.0056592	.10048
Mean Absolute	.37963	.10609
Mean Sum Squares	.28092	.021353
Root Mean Sum Squares	.53002	.14613

16.8 Lesson 16.8: Canonical correlations and cointegration analysis

In this lesson we perform a canonical correlation analysis to explore the relationship between two sets of variables. For technical and other details see and 22.13.

We use data from the special *Microfit* file ALMON.FIT on capital expenditures (Y) and appropriations (X) for the US manufacturing sector. The relevant data are in the special *Microfit* file ALMON.FIT, which contains observations on Y and X over the period 1953(1)-1967(4). See Lesson 13.3 for more details.

Our intent is to study the relationship between Y and X expressed in first differences and their lagged values, and observe the similarity between canonical correlation and cointegration analysis in terms of computations. See Lesson 16.3 for a detailed analysis on the non-stationarity and cointegration properties of these variables.

Read the file ALMON.FIT into *Microfit,* go to the Process window and in the Commands and Data Transformations box type

$$
\begin{aligned}
DY &= Y - Y(-1); \quad DX = X - X(-1)\,; \\
TREND &= \mathbf{CSUM}(1); \quad LY = Y(-1); \quad LX = X(-1)\,;
\end{aligned}
$$

These instructions allow you to create the first differences and first-order lags of the variables Y and X, as well as a trend variable. Now clear the Commands and Data Transformations box and type

$$\mathbf{CCA} \quad DY \quad DX \quad \& \quad LY \quad LX$$

Results from the output screen on non-zero squared canonical correlations, the eigenvectors and the statistic for testing the independence of Y and X are provided in Table 16.28. The "chi-squared statistic" (equal to 44.32) is large, although, given the non-stationarity of Y and X (see Lesson 13.3on this), this statistic is not distributed as chi-squared and hence the standard statistical tests cannot be applied.

Table 16.28: Canonical correlation analysis between DY, DX and LY, LX

```
                         Canonical Correlation Analysis
*******************************************************************************
 Estimation period from 1953Q2 to 1967Q4, 59 observations.

 List of 2 Y-variables included in the canonical correlation analysis:
 DY              DX
 List of 2 X-variables included in the canonical correlation analysis:
 LY              LX

 List of 2 non-zero squared canonical correlations in descending order:
 .74782  .0033638

 Cumulative Squared Canonical Correlations:
 .74782    .75119

 Test statistic for testing the independence of Y and X variables distributed
 as chi-squred with (2-1)*(2-1)=1 degrees of freedom =   44.3200[.000]

 Percent Cumulative Variances:
 99.5522  100.0000

 The number of chosen canonical variates is 2.
*******************************************************************************
 List of Y-eigenvectors associated with non-zero squared canonical correlations

             DY          DX
 CCY_1     .7431E-3   -.8746E-5
 CCY_2    -.1258E-3    .3679E-3

 List of X-eigenvectors associated with non-zero squared canonical correlations

             LY          LX
 CCX_1     .1818E-3   -.1843E-3
 CCX_2     .1247E-3   -.8192E-4
*******************************************************************************
```

We note that the squared canonical correlations available in Table 16.28 should be identical to the eigenvalues used in computing the trace and the maximal eigenvalue statistics normally used in the literature to test for cointegration. To check that this is in fact the case move to the Multivariate Menu and select option 1 (no intercepts or trends) from the Cointegrating VAR Menu. Set the lag order of the variables in the VAR to 1 and in the Commands and Data Transformations box type:

$$Y \quad X$$ RUN

Results, reported in Table 16.29, show that the eigenvalues for cointegration tests are identical to the squared canonical correlation reported in Table 16.28.

Table 16.29: Cointegration analysis between Y and X

```
             Cointegration with no intercepts or trends in the VAR
   Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix
******************************************************************************
 59 observations from 1953Q2 to 1967Q4. Order of VAR = 1, chosen r =1.
 List of variables included in the cointegrating vector:
 Y                X
 List of eigenvalues in descending order:
.74782    .0033638
******************************************************************************
 Null     Alternative     Statistic     95% Critical Value  90% Critical Value
 r = 0       r = 1         81.2799           11.0300               9.2800
 r<= 1       r = 2          .19880            4.1600               3.0400
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

             Cointegration with no intercepts or trends in the VAR
          Cointegration LR Test Based on Trace of the Stochastic Matrix
******************************************************************************
 59 observations from 1953Q2 to 1967Q4. Order of VAR = 1, chosen r =1.
 List of variables included in the cointegrating vector:
 Y                X
 List of eigenvalues in descending order:
.74782    .0033638
******************************************************************************
 Null     Alternative     Statistic     95% Critical Value  90% Critical Value
 r = 0       r>= 1         81.4787           12.3600              10.2500
 r<= 1       r = 2          .19880            4.1600               3.0400
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).

             Cointegration with no intercepts or trends in the VAR
Choice of the Number of Cointegrating Relations Using Model Selection Criteria
******************************************************************************
 59 observations from 1953Q2 to 1967Q4. Order of VAR = 1, chosen r =1.
 List of variables included in the cointegrating vector:
 Y                X
 List of eigenvalues in descending order:
.74782    .0033638
******************************************************************************
 Rank       Maximized LL        AIC              SBC              HQC
 r = 0       -818.7278       -818.7278        -818.7278        -818.7278
 r = 1       -778.0879       -781.0879        -784.2042        -782.3043
 r = 2       -777.9885       -781.9885        -786.1435        -783.6104
******************************************************************************
 AIC = Akaike Information Criterion      SBC = Schwarz Bayesian Criterion
 HQC = Hannan-Quinn Criterion
```

Now return to the Process window and in the Commands and Data Transformations box in the Process window type

$$\textbf{CCA} \quad DY \quad DX \quad \& \quad LY \quad LX \quad \& \quad INPT \quad TREND$$

RUN

The above instructions allow you to perform a canonical correlation analysis between DY, DX and LY, LX, after filtering out the effects of an intercept and a trend from these variables. Results from the output screen on canonical correlations, eigenvectors and the statistic for testing the independence of Y and X are reported in Table 16.30. As an exercise, check that the squared canonical correlations are identical to the eigenvalues used for computing cointegration tests in a $VAR(1)$ of Y and X. To this end, you should use option 5 (unrestricted intercepts and unrestricted trends) from the Cointegrating VAR Menu (in the System Estimation Menu).

Table 16.30: Canonical correlation analysis between DY, DX and LY, LX, after controlling for an intercept and a trend

```
                              Canonical Correlation Analysis
********************************************************************************
 Estimation period from 1953Q2 to 1967Q4, 59 observations.

 List of 2 Y-variables included in the canonical correlation analysis:
 DY              DX
 List of 2 X-variables included in the canonical correlation analysis:
 LY              LX

 The above variables have been filtered by the following 2 variables:
 INPT            TREND

 List of 2 non-zero squared canonical correlations in descending order:
 .70187   .096389

 Cumulative Squared Canonical Correlations:
 .70187    .79826

 Test statistic for testing the independence of Y and X variables distributed
 as chi-squred with (2-1)*(2-1)=1 degrees of freedom =   47.0972[.000]

 Percent Cumulative Variances:
 87.9250  100.0000

 The number of chosen canonical variates is 2.
********************************************************************************
 List of Y-eigenvectors associated with non-zero squared canonical correlations

              DY          DX
 CCY_1     .8118E-3   -.1938E-4
 CCY_2    -.6850E-4    .3705E-3

 List of X-eigenvectors associated with non-zero squared canonical correlations

              LY          LX
 CCX_1     .2011E-3   -.2067E-3
 CCX_2     .1688E-3    .1424E-4
********************************************************************************
```

16.9 Exercises in cointegration analysis

16.9.1 Exercise 16.1

Recompute the residual based ADF statistic for testing cointegration between real wages (WP) and labour productivity (YE) in Lesson 16.2 by running the reverse regression of YE on WP. Comment on your results.

16.9.2 Exercise 16.2

Use the time series observations in the file PHILLIPS.FIT to test the hypothesis that $WP = W - P$, and $YE = Y - E$ are cointegrated using the cointegrating VAR approach discussed in Lesson 16.3. Compare your results to the univariate approaches applied to the same problem in Lessons 16.1 and 16.2.

16.9.3 Exercise 16.3

In the first part of Lesson 16.3 we estimated a bivariate cointegrating VAR model in Y (Capital Expenditures), and X (Appropriations) for US Manufacturing. Check the robustness of the results by carrying out the exercise using logs of these variables instead of their levels. Which specification is preferable?

16.9.4 Exercise 16.4

Carry out tests of the Purchasing Power Parity hypothesis for the UK and the USA using the quarterly observation contained in the file G7EXCH.FIT. See Lesson 10.14 for details of the variables in this file and the batch file needed to construct the effective exchange rate and the foreign price indices.

16.9.5 Exercise 16.5

The file TERMUS.FIT contains monthly observations on eleven interest rates analysed by Hall, Anderson, and Granger (1992). Denote these interest rates by $Y1, Y2, ..., Y11$.

1. Analyse the cointegrating properties of the four interest rates $Y1, Y2, Y3$ and $Y4$.

2. Test the hypothesis that there are three cointegrating vectors among these four interest rates, and interpret the resultant cointegrating relations from the viewpoint of economic theory.

3. Test the hypothesis that $Y1 - Y2$, $Y2 - Y3$, and $Y1 - Y4$, are the long-run relations linking these interest rates together. Initially you need three restrictions on each of the three cointegrating vectors in order to just identify them. You should then impose the over-identifying restrictions (one by one and gradually) and test to determine whether these over-identifying restrictions are supported by the data.

Chapter 17

Lessons in VARX Modelling and Trend/Cycle Decomposition

In this chapter we show how the cointegrating $VARX$ option in the Multivariate Menu can be used to estimate structural vector error-correcting autoregressive models with exogenous variables ($VARX$). The lessons in this chapter also show how *Microfit* can be used to test for the existence of long-run relations, to compute multivariate forecasts and trend-cycle decompositions within the framework of cointegrating VAR and $VARX$ models.

The extension of the cointegrating VAR models to the case of exogenous variables is given in Pesaran, Shin, and Smith (2000). Review of the multivariate Beveridge-Nelson trend-cycle decomposition is in Garratt, Lee, Pesaran, and Shin (2006), Mills (2003), Robertson, Garratt, and Wright (2006), Johansen (1995), Evans and Reichlin (1994). See also Section 22.10 and 22.11.

17.1 Lesson 17.1: Testing the long-run validity of PPP and IRP hypotheses using UK data

In this lesson we re-examine the empirical evidence on the long-run validity of the Interest Rate Parity (IRP) and the Purchasing Parity Power (PPP) hypotheses studied in Johansen and Juselius (1992), Pesaran, Shin, and Smith (2000) and Pesaran and Shin (1996) (see also Lesson 16.3).

We use quarterly data on the UK economy over the period from 1966-1999, and consider a simple model containing three domestic variables, prices, interest rate and exchange rate, and two foreign variables, OECD prices and foreign interest rate. We carry a cointegration analysis on these five variables, treating the last two as $I(1)$ and weakly exogenous[1], and thus long-run forcing for the remaining variables. To this end, we will employ the cointegrating $VARX$ option provided in *Microfit* (see Section 22.10 for further details on $VARX$ models and reference to the literature).

[1] Weak exogeneity in this context means that the cointegrating vectors do not appear in the sub-system VECM for these exogenous variables (see Pesaran, Shin, and Smith (2000)).

Quarterly data on the UK economy over the period 1964q1 to 1999q4 (for a total of 144 observations) are contained in the special *Microfit* file UKCORE.FIT. Load this file into *Microfit* and in the Process window enter

BATCH UKCORE.BAT

Click the button, ensuring that the following variables are on the workspace

P	Logarithm of UK producer price index (seasonally adjusted)
PS	Logarithm of OECD producer price index (seasonally adjusted)
R	Domestic nominal interest rate. Computed as $0.25\ln(1+AR/100)$ where AR is the UK Treasury Bill annual rate
RS	Foreign nominal interest rate. Computed as $0.25\ln(1+ARSSRD/100)$ where $ARSSDR$ is an average of US, Japan, France and Germany Treasury Bill annual rate
E	Sterling effective exchange rate

For later use create a constant and a linear time trend by typing

$$INPT = 1; \quad T = \mathbf{CSUM}(1)$$

Choose option 3 from the System Estimation Menu (Multivariate Menu) and in the Commands and Data Transformations box enter

$$R \quad P \quad E; \quad RS \quad PS \quad \& \quad INPT \quad T$$

Alternatively, you can retrieve the file UKMOD.LST from the tutorial directory. The first three variables are the endogenous variables, followed by the exogenous and the deterministic variables, the latter two separated by the symbol '&'. Then specify the sample period (of 136 observations)

$$1966Q1 \quad 1999Q4$$

The selection of the order of the VAR can be based on selection criteria such as AIC and SBC available in the unrestricted VAR option 1 from the System Estimation Menu (Multivariate Menu) (see Lesson 16.3). As an exercise you should convince yourself that a $VAR(2)$ model can adequately capture the dynamic properties of the data. Hence, enter 2 as order of the lag of the endogenous and exogenous variables. Click on the 'Simulation of Critical Values' button; a selection window will appear. Check the 'Simulate critical values for cointegration tests' checkbox, and click to accept the default number of replications and significance levels. This allows you to obtain 95 and 90 per cent simulated critical values for cointegration rank test statistics (see Sections 22.8.1 and 22.8.2).

Click . Since we wish to avoid the possibility of quadratic trends in some of the variables, restrict the trend, and leave the intercept unrestricted. You are then asked to specify the marginal models for the exogenous variables PS and RS. In the box editor enter

$$1 \quad 1 \quad INPT;$$
$$1 \quad 1 \quad INPT$$

Numbers in the first column refer to the lag order of the first difference of endogenous variables, while those in the second column refer to the lag order of the first differences of exogenous variables. Hence, the above instructions imply the following marginal models for RS and PS:

$$\begin{aligned} dRS_t &= a_{01} + a_{r,1}dR_{t-1} + a_{p,1}dP_{t-1} + a_{e,1}dE_{t-1} + a_{rs,1}dRS_{t-1} + a_{ps,1}dPS_{t-1} + u_{rs,t} \\ dPS_t &= a_{02} + a_{r,2}dR_{t-1} + a_{p,2}dP_{t-1} + a_{e,2}dE_{t-1} + a_{rs,2}dRS_{t-1} + a_{ps,2}dPS_{t-1} + u_{ps,t} \end{aligned}$$

where $dRS_t = RS_t - RS_{t-1}$, and so on. Click RUN. *Microfit* starts the computations and yields the results reported in Table 17.1.

Table 17.1: Cointegration rank statistics for the UK model

```
  Cointegrating VAR Autoregression Model with Exogenous I(1) Variables
  Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix
******************************************************************************
 136 observations from 1966Q1 to 1999Q4. VARX(2,2), chosen r =1.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
 List of variables included in the cointegrating vector:
 R               P               E               RS              PS
 T
 List of I(1) exogenous variables included in the VAR:
 RS              PS
 List of unrestricted deterministic variables included in the VAR:
 INPT
 List of eigenvalues in descending order:
.22675    .086395    .049651
******************************************************************************
 Null     Alternative    Statistic     95% Critical Value  90% Critical Value
 r = 0       r = 1        34.9732          33.7259              30.8003
 r<= 1       r = 2        12.2885          27.5695              24.4638
 r<= 2       r = 3         6.9259          20.0140              17.0520
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).
 95% and 90% simulated critical values using 136 obs. and 1000 replications.

  Cointegrating VAR Autoregression Model with Exogenous I(1) Variables
         Cointegration LR Test Based on Trace of the Stochastic Matrix
******************************************************************************
 Null     Alternative    Statistic     95% Critical Value  90% Critical Value
 r = 0       r>= 1        54.1876          60.6062              56.6380
 r<= 1       r>= 2        19.2144          38.2391              34.9996
 r<= 2       r = 3         6.9259          20.0140              17.0520
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).
 95% and 90% simulated critical values using 136 obs. and 1000 replications.

  Cointegrating VAR Autoregression Model with Exogenous I(1) Variables
Choice of the Number of Cointegrating Relations Using Model Selection Criteria
******************************************************************************
 Rank       Maximized LL         AIC              SBC              HQC
 r = 0          1394.2          1370.2           1335.3           1356.0
 r = 1          1411.7          1379.7           1333.1           1360.8
 r = 2          1417.9          1379.9           1324.5           1357.4
 r = 3          1421.3          1379.3           1318.2           1354.5
******************************************************************************
 AIC = Akaike Information Criterion      SBC = Schwarz Bayesian Criterion
 HQC = Hannan-Quinn Criterion
```

The upper part of the table summarizes the model and gives the list of eigenvalues of the estimated stochastic matrix. The bottom of the table is divided in three panels and provides the output of three sets of test statistics for choosing the number of cointegrating relations together with their simulated critical values: the maximum eigenvalue statistic, the trace eigenvalue statistic, and the model selection criteria AIC, SBC and HQC (see Sections 22.8.1-22.8.3 for a description of these test statistics).

The maximum eigenvalue statistic rejects the hypothesis of no cointegration at the 5 per cent significance level, and indicates the existence of one cointegrating vector ($r = 1$)

between the five $I(1)$ variables under investigation; the trace statistic does not reject the hypothesis of no cointegration at the 5 per cent significance level.

However, economic theory predicts the following two long-run equilibrium relations:

$$R - RS \sim I(0) \tag{17.1}$$

$$P - E - PS \sim I(0) \tag{17.2}$$

Equation (17.1) is an Interest Rate Parity (IRP) relation, which captures the equilibrium outcome between domestic and foreign interest rates due to the effect of the arbitrage process between domestic and foreign bonds. Equation (17.2) is the Purchasing Parity Power (PPP), stating that, due to international trade in goods, domestic and foreign prices measured in a common currency equilibrate in the long-run.

In view of what economic theory suggests, we therefore continue the analysis assuming that there are two cointegrating relations among R, P, E, RS and PS. Under the assumption that $r = 2$, exact identification of the cointegrating relations requires the imposition of two independent restrictions on each of the two relations. We choose the following exactly identifying constraints:

$$\begin{array}{ll} \beta_{11} = 1, & \beta_{12} = 0 \\ \beta_{21} = 0, & \beta_{22} = 1 \end{array}$$

where

$$\begin{array}{rcl} \boldsymbol{\beta}_1 & = & (\beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}, \beta_{15}, \beta_{16})' \\ \boldsymbol{\beta}_2 & = & (\beta_{21}, \beta_{22}, \beta_{23,} \beta_{24}, \beta_{25}, \beta_{26})' \end{array}$$

are the two cointegrating vectors. associated with $\mathbf{z}_t = (R_t, P_t, E_t, RS_t, PS_t, t)'$. To estimate the $VARX$ model subject to the above restrictions, click to leave the output window, and in the Cointegrating VAR Post Estimation Menu choose option 2 to set the number of cointegrating vectors and enter

2

Then select option 6 to access to the Long-Run Structural Modelling Menu; choose option 4 and, when prompted, enter the just-identifying restrictions[2]

$$\begin{array}{ll} A1 = 1; & A2 = 0; \\ B1 = 0; & B2 = 1; \end{array} \tag{17.3}$$

The exactly identified ML estimates of the two cointegrating vectors and their asymptotic standard errors should now appear on the screen. These results are reproduced in Table 17.2.

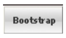
[2] Alternatively, you can retrieve the contents of the file UKEXID.EQU from the tutorial directory by clicking the button.

Table 17.2: ML estimates subject to exactly identifying restrictions for the UK model

```
            ML estimates subject to exactly identifying restriction(s)
        Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                           Converged after 1 iterations
     Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
136 observations from 1966Q1 to 1999Q4. VARX(2,2), chosen r =2.
Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
List of variables included in the cointegrating vector:
R                P                E                RS               PS
T
List of I(1) exogenous variables included in the VAR:
RS               PS
List of unrestricted deterministic variables included in the VAR:
INPT
Definition of the Marginal Model:
Equation 1 of the Marginal Model:
Dependent variable = dRS
List of regressors:
dRS(-1)          dPS(-1)          dR(-1)           dP(-1)           dE(-1)
INPT
Equation 2 of the Marginal Model:
Dependent variable = dPS

List of regressors:

dRS(-1)          dPS(-1)          dR(-1)           dP(-1)           dE(-1)
INPT
*******************************************************************************
List of imposed restriction(s) on cointegrating vectors:
                   Vector 1        Vector 2
R                   1.0000           0.00
                  (   *NONE*)     (    *NONE*)

P                     0.00         1.0000
                  (   *NONE*)     (    *NONE*)

E                  .026896        -1.5650
                  (  .023049)     (   1.0305)

RS                 .063019       -19.3425
                  (   .37180)     (  15.1371)

PS                -.049865        -.39857
                  (  .020434)     (   .87753)

T                 .6313E-3      -.0035132
                  ( .2193E-3)     ( .0091745)

*******************************************************************************
LL subject to exactly identifying restrictions=   1417.9
*******************************************************************************
```

Using the above exactly identified model, we can now test for a number of hypotheses. First, we test the co-trending hypothesis, namely whether the trend coefficients are zero in the two cointegrating relations, that is, if $\beta_{16} = \beta_{26} = 0$. Choose option 0 in the IR Analysis and Forecasting Menu and say that you wish to test (further) over-identifying restrictions. When presented with the box editor add the following restrictions

$$A6 = 0; \quad B6 = 0 \tag{17.4}$$

You are then presented with a window in which you can set the initial values for estimation; click the button to accept the default initial values, and then choose the new algorithm (B) in *Microfit 5.0* (option 2) for the estimation. The program starts the computations and, when finished, returns an output window with ML estimates subject to over-identifying restrictions. The LR statistic for testing the over-identifying restrictions is 9.69, equal to twice the difference between the maximized log-likelihood obtained from the estimation subject to just-identifying restrictions (see Table 17.2) and the maximized log-likelihood subject to the over-identifying restrictions. Given the small size of the sample, we suggest computation of the bootstrapped critical values for this statistic. Close the output window and press the 

button to accept the default number of replications and significance levels for the bootstrap. The program starts the computations and then presents you with the results reported in Table 17.3.

Table 17.3: Testing the co-trending hypothesis in the UK model

```
          ML estimates subject to over identifying restriction(s)
       Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                        Converged after 21 iterations
   Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
136 observations from 1966Q1 to 1999Q4. VARX(2,2), chosen r =2.
Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
List of variables included in the cointegrating vector:
R                P               E               RS              PS
T
List of I(1) exogenous variables included in the VAR:
RS               PS
List of unrestricted deterministic variables included in the VAR:
INPT
Definition of the Marginal Model:
Equation 1 of the Marginal Model:
Dependent variable = dRS
List of regressors:
dRS(-1)          dPS(-1)         dR(-1)          dP(-1)          dE(-1)
INPT
Equation 2 of the Marginal Model:
Dependent variable = dPS
List of regressors:
dRS(-1)          dPS(-1)         dR(-1)          dP(-1)          dE(-1)
INPT
*******************************************************************************
List of imposed restriction(s) on cointegrating vectors:
                  Vector 1       Vector 2
R                   1.0000         -.0000
                  (   *NONE*)    (   *NONE*)

P                    .0000         1.0000
                  (   *NONE*)    (   *NONE*)

E                  -.029875       -1.3492
                  (  .031079)    (   .49347)

RS                  -.78366       -14.7756
                  (   .36155)    (   5.8580)

PS                 .0078347        -.69036
                  (  .011469)    (   .18355)

T                    -.0000         -.0000
                  (   *NONE*)    (   *NONE*)

*******************************************************************************
LR Test of Restrictions             CHSQ(2)=    9.6866[.008]
95% Bootstrapped Critical Value =   12.1769
90% Bootstrapped Critical Value =    9.2611
Bootstrapped simulations based on 1000 SIMULATIONS.
DF=Total no of restrictions(6) - no of just-identifying restrictions(4)
LL subject to exactly identifying restrictions=    1417.9
LL subject to over-identifying restrictions=   1413.0
```

Since the LR statistic is below its bootstrapped critical value at the 5 per cent significance level (12.18), we do not reject the co-trending hypothesis on cointegrating relations (17.2)-(17.1).

We now turn to testing the IRP and PPP hypotheses (17.1)-(17.2), assuming that the co-trending hypothesis holds. Relation (17.1) implies that exchange rate (E), foreign interest rate (RS) and foreign prices (PS) enter in the long-run relation with coefficients $\beta_{13} = 0$, $\beta_{14} = -1$, and $\beta_{15} = 0$, respectively. Choose option 0 in the IR Analysis and Forecasting Menu and say that you wish to test (further) over-identifying restrictions. When presented with the box editor add the following three restrictions:

$$A3 = 0; \quad A4 = -1; \quad A5 = 0 \tag{17.5}$$

Repeat the above steps to compute bootstrapped critical values for the LR (with the default number of replications and significance levels); the program presents you with the results shown in Table 17.4.

Table 17.4: Testing the IRP hypothesis in the UK model

```
            ML estimates subject to over identifying restriction(s)
       Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                         Converged after 8 iterations
    Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
136 observations from 1966Q1 to 1999Q4. VARX(2,2), chosen r =2.
Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
List of variables included in the cointegrating vector:
R                P                E                RS               PS
T
List of I(1) exogenous variables included in the VAR:
RS               PS
List of unrestricted deterministic variables included in the VAR:
INPT
Definition of the Marginal Model:
Equation 1 of the Marginal Model:
Dependent variable = dRS
List of regressors:
dRS(-1)          dPS(-1)          dR(-1)           dP(-1)           dE(-1)
INPT
Equation 2 of the Marginal Model:
Dependent variable = dPS
List of regressors:
dRS(-1)          dPS(-1)          dR(-1)           dP(-1)           dE(-1)
INPT
*******************************************************************************
List of imposed restriction(s) on cointegrating vectors:
                  Vector 1       Vector 2
R                  1.0000         -.0000
                 (  *NONE*)     (  *NONE*)

P                  -.0000         1.0000
                 (  *NONE*)     (  *NONE*)

E                   .0000        -1.0058
                 (  *NONE*)     (   .27230)

RS                -1.0000       -16.6698
                 (  *NONE*)     (  4.2118)

PS                 -.0000        -.78246
                 (  *NONE*)     (   .10677)

T                   .0000         -.0000
                 (  *NONE*)     (  *NONE*)

*******************************************************************************
LR Test of Restrictions          CHSQ(5)=  14.0001[.016]
95% Bootstrapped Critical Value =  19.8297
90% Bootstrapped Critical Value =  16.1952
Bootstrapped simulations based on 1000 SIMULATIONS.
DF=Total no of restrictions(9) - no of just-identifying restrictions(4)
LL subject to exactly identifying restrictions=   1417.9
LL subject to over-identifying restrictions=   1410.9
```

Since the LR statistic (14.00) is below its bootstrapped critical value at the 5 per cent significance level (19.83), we do not reject the over-identifying restrictions, and conclude that there is evidence of a long-run relations between domestic and foreign interest rate.

We now test for the PPP hypothesis, again under the assumption that the cointegrating relations are co-trending. Relation (17.2) implies that exchange rate (E), foreign interest rate (RS) and foreign prices (PS) enter in the long-run relation with coefficients -1, 0, and -1, respectively. Choose, once again, option 0 in the IR Analysis and Forecasting Menu, and in the box editor replace restrictions (17.5) with the following:

$$B3 = -1; \quad B4 = 0; \quad B5 = -1 \tag{17.6}$$

Again, comparing the LR statistic with its boostrapped critical values (obtained with the default number of replications and significance levels), we are unable to reject the null hypothesis of PPP (jointly with the co-trending assumption) (see Table 17.5).

Table 17.5: Testing the PPP hypothesis in the UK model

```
          ML estimates subject to over identifying restriction(s)
        Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                        Converged after 58 iterations
    Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
136 observations from 1966Q1 to 1999Q4. VARX(2,2), chosen r =2.
Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
List of variables included in the cointegrating vector:
R                P                E                RS               PS
T
List of I(1) exogenous variables included in the VAR:
RS               PS
List of unrestricted deterministic variables included in the VAR:
INPT
Definition of the Marginal Model:
Equation 1 of the Marginal Model:
Dependent variable = dRS
List of regressors:
dRS(-1)          dPS(-1)          dR(-1)           dP(-1)           dE(-1)
INPT
Equation 2 of the Marginal Model:
Dependent variable = dPS
List of regressors:
dRS(-1)          dPS(-1)          dR(-1)           dP(-1)           dE(-1)
INPT
*******************************************************************************
List of imposed restriction(s) on cointegrating vectors:
                   Vector 1       Vector 2
R                   1.0000         -.0000
                  (   *NONE*)    (   *NONE*)

P                  -.0000          1.0000
                  (   *NONE*)    (   *NONE*)

E                 -.0087503       -1.0000
                  (  .018136)    (   *NONE*)

RS                  .19716          .0000
                  (  .55006)     (   *NONE*)

PS                 -.012162       -1.0000
                  ( .0077524)    (   *NONE*)

T                  -.0000           .0000
                  (   *NONE*)    (   *NONE*)

*******************************************************************************
LR Test of Restrictions             CHSQ(5)=  17.1030[.004]
95% Bootstrapped Critical Value =   24.0793
90% Bootstrapped Critical Value =   21.2427
Bootstrapped simulations based on 1000 SIMULATIONS.
DF=Total no of restrictions(9) - no of just-identifying restrictions(4)
LL subject to exactly identifying restrictions=   1417.9
LL subject to over-identifying restrictions=   1409.3
```

We finally report, in Table 17.6, the output when testing whether IRP and PPP long-run relations jointly hold, under the co-trending hypothesis. This is done by imposing both (17.5) and (17.6) and the co-trending restrictions (17.4). Results support the validity of these long-run relationships using UK data.

Table 17.6: Testing the IRP and PPP hypotheses in the UK model

```
              ML estimates subject to over identifying restriction(s)
          Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                            Converged after 2 iterations
       Cointegrating Vector Autoregressive Model with Exogenous I(1) Variables
*******************************************************************************
 136 observations from 1966Q1 to 1999Q4. VARX(2,2), chosen r =2.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
 List of variables included in the cointegrating vector:
 R               P               E               RS              PS
 T
 List of I(1) exogenous variables included in the VAR:
 RS              PS
 List of unrestricted deterministic variables included in the VAR:
 INPT
 Definition of the Marginal Model:
 Equation 1 of the Marginal Model:
 Dependent variable = dRS
 List of regressors:
 dRS(-1)         dPS(-1)         dR(-1)          dP(-1)          dE(-1)
 INPT
 Equation 2 of the Marginal Model:
 Dependent variable = dPS
 List of regressors:
 dRS(-1)         dPS(-1)         dR(-1)          dP(-1)          dE(-1)
 INPT
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                   Vector 1         Vector 2
 R                  1.0000            .0000
                 (   *NONE*)      (   *NONE*)

 P                  -.0000           1.0000
                 (   *NONE*)      (   *NONE*)

 E                  -.0000          -1.0000
                 (   *NONE*)      (   *NONE*)

 RS                -1.0000           -.0000
                 (   *NONE*)      (   *NONE*)

 PS                 -.0000          -1.0000
                 (   *NONE*)      (   *NONE*)

 T                   .0000           -.0000
                 (   *NONE*)      (   *NONE*)

*******************************************************************************
 LR Test of Restrictions          CHSQ(8)=  28.9008[.000]
 95% Bootstrapped Critical Value =   32.3792
 90% Bootstrapped Critical Value =   28.6967
 Bootstrapped simulations based on 1000 SIMULATIONS.
 DF=Total no of restrictions(12) - no of just-identifying restrictions(4)
 LL subject to exactly identifying restrictions=   1417.9
 LL subject to over-identifying restrictions=   1403.4
```

17.2 Lesson 17.2: A macroeconomic model for Indonesia

In this lesson we estimate a long-run structural macroeconometric model for Indonesia.[3] Indonesia can be considered a small open economy, where a subset of the long-run forcing variables, such as foreign income and prices, may be viewed as exogenous. Another characteristic of the Indonesian economy is that it has been subject to a major break due to the Asian financial crisis which occurred in 1997. To deal with these issues we adopt a vector error-correcting autoregressive model with exogenous variables ($VARX$), and introduce dummy variables to represent such breaks, trying to capture shifts both in the long-run and short-run relations.

The relevant data on the Indonesian economy can be found in the special *Microfit* file

[3] This Lesson and the following are based on analysis of the Indonesian economy by Affandi (2007) .

IDQ.FIT, containing the following variables:

Y	Logarithm of Indonesian real *per capita* GDP (1993=100)
P	Logarithm of Indonesian Consumer Price Index (1993=100)
R	Domestic nominal interest rate. Computed as $0.25\ln(1+R^m/100)$ where R^m is the 90-day money market rate per annum
$M1P$	Logarithm of Indonesian real *per capita* money stock (1993=100)
E	Logarithm of nominal Rupiah effective exchange rate (1993=100)
YS	Logarithm of foreign real *per capita* GDP (1993=100)
PS	Logarithm of foreign price index (1993=100)
RS	Foreign interest rate
PO	Logarithm of oil price (1993=100)
DP	Indonesian Consumer Price Index (CPI) inflation rate ($DP_t = P_t - P_{t-1}$)
$INPT$	Intercept term
$D97$	Dummy variable equal to 1 for the period from 1997q3 onwards and 0 otherwise
EP	Real exchange rate of the Indonesian Rupiah ($EP_t = E_t - P_t$)
T	Linear trend

This file contains quarterly data over the period 1979q1-2003q4, for a total of 100 observations: read it into *Microfit* and make sure that the above variables are in the workspace.

In the model, five variables will be treated as endogenous, namely domestic output and prices, interest rate, real exchange rate, and money stock, and four variables as exogenous, namely foreign output and prices, foreign interest rate and oil price. $D97$ is a dummy variable that we wish to include in the cointegrating vector, to investigate whether there exists a structural break in the long-run relations due to the 1997 Asian crisis.

Consider the following $VECM$ model:

$$\Delta \mathbf{z}_t = -\mathbf{\Pi}\mathbf{z}_{t-1} + \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta \mathbf{z}_{t-i} + \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{a}_2 D97_t + \mathbf{u}_t \tag{17.7}$$

where $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$, $\mathbf{y}_t$ is the vector of endogenous variables (containing Y_t, DP_t, R_t, EP_t, and $M1P_t$), $\mathbf{x}_t$ is the vector of weakly exogenous variables (containing YS_t, PS_t, RS_t, and PO_t), $\mathbf{a}_0$ is the intercept, t is the time trend, and $D97_t$ is the dummy variable. Furthermore, $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ is the matrix of long-run coefficients, $\mathbf{\Gamma}_i$ is the matrix of short-run coefficients, and $\mathbf{a}_1$ and $\mathbf{a}_2$ are trend coefficients. We assume the following restrictions on the trend coefficients:

$$\mathbf{a}_1 = \mathbf{\Pi}\mathbf{b}_1, \mathbf{a}_2 = \mathbf{\Pi}\mathbf{b}_2 \tag{17.8}$$

where $\mathbf{b}_1$ and $\mathbf{b}_2$ are arbitrary vectors of fixed constants. As shown in Pesaran, Shin, and Smith (2000), these restrictions ensure that the solution of the model in levels of $\mathbf{z}_t$ will not contain quadratic trends or cumulative effects of $D97_t$. If $\mathbf{\Pi}$ is full rank, then $\mathbf{b}_1$ and $\mathbf{b}_2$ are unrestricted, and can be computed as $\mathbf{b}_1 = \mathbf{\Pi}^{-1}\mathbf{a}_1$ and $\mathbf{b}_2 = \mathbf{\Pi}^{-1}\mathbf{a}_2$; if $\mathbf{\Pi}$ is rank deficient, $\mathbf{b}_1$ and $\mathbf{b}_2$ cannot be fully identified from $\mathbf{a}_1$ and $\mathbf{a}_2$.

Under (17.8), the $VECM$ in (17.7) becomes

$$\begin{aligned}
\Delta \mathbf{z}_t &= -\mathbf{\Pi}\left(\mathbf{z}_{t-1} - \mathbf{b}_1 t - \mathbf{b}_2 D97_t\right) + \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta \mathbf{z}_{t-i} + \mathbf{a}_0 + \mathbf{u}_t \\
&= -\mathbf{\Pi}\left[\mathbf{z}_{t-1} - \mathbf{b}_1 (t-1) - \mathbf{b}_2 \left(D97_{t-1} + DD97_t\right)\right] \\
&\quad + \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta \mathbf{z}_{t-i} + \mathbf{a}_0 + \mathbf{a}_1 + \mathbf{u}_t \\
&= -\mathbf{\Pi}_\diamond \mathbf{z}_{t-1}^\diamond + \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta \mathbf{z}_{t-i} + \mathbf{a}_0^\diamond + \mathbf{a}_2 DD97_t + \mathbf{u}_t \qquad (17.9)
\end{aligned}$$

where $\mathbf{\Pi}_\diamond = \boldsymbol{\alpha}\boldsymbol{\beta}_\diamond'$, $\boldsymbol{\beta}_\diamond' = \left(\boldsymbol{\beta}', -\boldsymbol{\beta}'\mathbf{b}_1, -\boldsymbol{\beta}'\mathbf{b}_2\right)$, $\mathbf{z}_{t-1}^\diamond = \left(\mathbf{z}_{t-1}', t-1, D97_{t-1}\right)'$, $\mathbf{a}_0^\diamond = \mathbf{a}_0 + \mathbf{a}_1$, and $DD97_t = D97_t - D97_{t-1}$. The new variable $DD97_t$ is a dummy variable introduced to capture changes in the short-run dynamics. Notice that the deterministic trend t and the dummy variable $D97_t$ are now part of the cointegrating relation $\boldsymbol{\beta}'\left(\mathbf{z}_{t-1} - \mathbf{b}_1 (t-1) - \mathbf{b}_2 D97_{t-1}\right) = \boldsymbol{\beta}_\diamond' \mathbf{z}_{t-1}^\diamond$. This implies that the cointegration relations could contain linear trends and may be subject to structural breaks.

Estimation of model (17.9) requires the construction of the dummy variable $DD97$. To create this variable, go to the Process window, and in the Commands and Data Transformations box type

$$DD97 = D97 - D97(-1)$$

A new variable entitled $DD97$ should appear in the list of your variables. Now select option 3 Cointegrating $VARX$ from the System Estimation Menu (Multivariate Menu), and in the Commands and Data Transformations window type the following specification:[4]

$$Y \quad DP \quad R \quad EP \quad M1P \quad ; \quad YS \quad PS \quad RS \quad PO \quad \& \quad INPT \quad T \quad D97 \quad DD97 \qquad (17.10)$$

For the estimation sample, enter

$$1981Q2 \quad 2003Q4$$

The selection of the order of the VAR can be based on selection criteria such as AIC and SBC available in the unrestricted VAR option 1 from the System Estimation Menu (Multivariate Menu) (see Lesson 16.3). On the basis of this preliminary analysis, we choose $p = 2$; hence, type 2 in the boxes entitled 'lag order of endogenous variables' and 'lag order of exogenous variables'.

Since in this application the estimation sample is relatively small (91 observations) and the dimension of the system is relatively large (9 variables), it is advisable to compute bootstrapped critical values for cointegration tests. Click on the button 'Simulation of Critical Values', check the box 'Simulate critical values for cointegration tests', and click OK to accept the default number of replications and significance level.

[4] Alternatively, you can retrieve the contents of the file IDQMOD.LST from the tutorial directory by clicking the Retrieve button.

Press the button. You will be presented with a window asking to select the variable(s) you wish to restrict. Since we wish to avoid the possibility of quadratic trends in some of the variables, restrict the time trend (T) and the shift dummy variable ($D97$), while leaving unrestricted the variable intercept ($INPT$) and the impulse dummy variable ($DD97$); click . A window appears in which you are asked to define the marginal models. Given that the $VECM$ has four exogenous variables (YS, PS, RS, PO), you need to specify four equations. In the box editor write[5]

$$\begin{array}{lll} 1 & 1 & INPT; \\ 1 & 1 & INPT; \\ 1 & 1 & INPT; \\ 1 & 1 & INPT \end{array}$$

Notice that the first line of instructions (namely 1 1 $INPT$) refers to the following marginal model:

$$\begin{aligned} dYS_t &= a_0 + a_y dY_{t-1} + a_{dp} dDP_{t-1} + a_r dR_{t-1} + a_{ep} dEP_{t-1} \\ &\quad + a_m dM1P_{t-1} + a_{ys} dYS_{t-1} + a_{ps} dPS_{t-1} + a_{rs} dRS_{t-1} \\ &\quad + a_{po} dPO_{t-1} + u_{ys,t} \end{aligned}$$

where $dYS_t = YS_t - YS_{t-1}$. Similar marginal models are specified for the other weakly exogenous variables. Following the order of the exogenous variables as appear in (17.10), the second line of instructions refers to the foreign price equation, and the third and fourth lines refer to the foreign interest rate and oil price equations, respectively. Click . *Microfit* starts the computation and, when finished, presents you with an output window that is reproduced in Table 17.7 .

[5] You can retrieve the contents of the file IDQMARG.LST from the tutorial directory by clicking the button.

Table 17.7: Cointegrating rank statistics for the Indonesian model

```
   Cointegrating Vector Autoregression Model with Exogenous I(1) Variables

   Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix
******************************************************************************
 91 observations from 1981Q2 to 2003Q4. VARX(2,2), chosen r =1.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
 List of variables included in the cointegrating vector:
 Y               DP              R               EP              M1P
 YS              PS              RS              PO              T
 D97
 List of I(1) exogenous variables included in the VAR:
 YS              PS              RS              PO
 List of unrestricted deterministic variables included in the VAR:
 INPT            DD97
 List of eigenvalues in descending order:
.61039     .55331     .41822     .25443     .22267
******************************************************************************
 Null     Alternative      Statistic      95% Critical Value  90% Critical Value
 r = 0       r = 1          85.7785            66.7810             62.7772
 r<= 1       r = 2          73.3357            56.8773             53.5272
 r<= 2       r = 3          49.2914            48.6080             45.0023
 r<= 3       r = 4          26.7182            39.7456             36.7546
 r<= 4       r = 5          22.9223            30.3324             26.4255
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).
 95% and 90% simulated critical values using 91 obs. and 1000 replications.

   Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
          Cointegration LR Test Based on Trace of the Stochastic Matrix
******************************************************************************
 Null     Alternative      Statistic      95% Critical Value  90% Critical Value
 r = 0       r>= 1         258.0461           178.7118            171.3283
 r<= 1       r>= 2         172.2676           135.2392            127.5695
 r<= 2       r>= 3          98.9319            92.8532             87.9143
 r<= 3       r>= 4          49.6405            60.0350             56.1207
 r<= 4       r = 5          22.9223            30.3324             26.4255
******************************************************************************
 Use the above table to determine r (the number of cointegrating vectors).
 95% and 90% simulated critical values using 91 obs. and 1000 replications.

   Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
:hoice of the Number of Cointegrating Relations Using Model Selection Criteria
******************************************************************************
 Rank       Maximized LL          AIC            SBC             HQC
 r = 0         1103.8            1028.8        934.6865        990.8572
 r = 1         1146.7            1056.7        943.7443          1011.1
 r = 2         1183.4            1080.4        951.0916          1028.2
 r = 3         1208.0            1094.0        950.9276          1036.3
 r = 4         1221.4            1098.4        943.9878          1036.1
 r = 5         1232.9            1102.9        939.6609          1037.0
******************************************************************************
 AIC = Akaike Information Criterion      SBC = Schwarz Bayesian Criterion
 HQC = Hannan-Quinn Criterion
```

Both the maximum eigenvalue statistic and the trace statistic indicate the existence of three cointegrating vectors ($r = 3$). On the basis of these results, and in view of what the economic theory suggests (see Lesson 17.3 and Affandi (2007)), in what follows we assume $r = 3$, that is, that there are three long-run relations.[6]

17.3 Lesson 17.3: Testing for over-identifying restrictions in the Indonesian model

In Lesson 17.2 we introduced a $VECM$ model with exogenous variables for the Indonesian economy; we carried out cointegration tests using the trace and the maximum eigenvalue statistics, and decided that there exist $r = 3$ cointegrating relations among the variables. We

[6] Also, some authors have pointed out that the trace statistic is in general more robust to model misspecification and deviations from normality than the maximum eigenvalue statistic (see, for example, Cheung and Lai (1993)).

now turn to testing for over-identifying restrictions on the cointegrating vectors, as predicted by economic theory. This analysis involves the maximum likelihood (ML) estimation of the model subject to the exactly and over-identifying restrictions (see Pesaran and Shin (2002) and Pesaran, Shin, and Smith (2000)). The tests of over-identifying restrictions will be in the form of chi-squared statistics with degrees of freedom equal to the number of the over-identifying restrictions (see Section 22.9, and, in particular, 22.9.3).

Affandi (2007) identifies the following three long-run equilibrium relationship:

$$EP_t + PS_t = b_{10} + \xi_{1,t+1}$$

$$DP_t - R_t = b_{20} + \xi_{2,t+1}$$

$$R_t - RS_t = b_{30} + \xi_{3,t+1}$$

where ξ_{it}, for $i = 1, 2, 3$, are stationary reduced form errors. The first equation describes the Purchasing Power Parity (PPP) relation, which predicts that domestic and foreign prices measured in a common currency equilibrate in the long-run. The second is the Fisher Inflation Parity (FIP) relation, stating that the nominal rate of interest should in the long-run equate the real rate of return plus the (expected) rate of inflation. Finally, the third equation is the Interest Rate Parity (IRP) relation, which captures the equilibrium outcome between domestic and foreign interest rates due to the effect of the arbitrage process between domestic and foreign bonds. The three relations can be compactly written as

$$\boldsymbol{\xi}_t = \boldsymbol{\beta}'\mathbf{z}_{t-1} - \mathbf{b}_0 - \mathbf{b}_1 (t-1) - \mathbf{b}_2 D97_{t-1} \tag{17.11}$$

where

$$\begin{aligned}
\mathbf{z}_t &= (Y_t, DP_t, R_t, EP_t, M1P_t, YS_t, PS_t, RS_t, PO_t)' \\
\mathbf{b}_0 &= (b_{10}, b_{20}, b_{30})', \mathbf{b}_1 = (0,0,0)', \mathbf{b}_2 = (0,0,0)', \boldsymbol{\xi}_t = (\xi_{1t}, \xi_{2t}, \xi_{3t})' \\
\boldsymbol{\beta} &= \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}
\end{aligned}$$

The matrix $\boldsymbol{\beta}$ is the over-identifying matrix, imposing all the restrictions as suggested by the long-run relations PPP, FIP and IRP.

In view of the restrictions implicit in the above three relations, we start our long-run analysis with the following just-identifying restrictions:

$$\begin{aligned}
\beta_{12} &= 0, \quad \beta_{14} = 1, \quad \beta_{15} = 0 \\
\beta_{21} &= 0, \quad \beta_{22} = 1, \quad \beta_{23} = -1 \\
\beta_{32} &= 0, \quad \beta_{33} = 1, \quad \beta_{38} = -1
\end{aligned}$$

Click 

to go to the Cointegrating VAR Post Estimation Menu, and select option 2, to specify the number of cointegrating vectors. Type

3 

Then select option 6 'Long-Run Structural Modelling, IR Analysis and Forecasting'. You will be presented with the Long-Run Structural Modelling Menu. Choose option 4, click and, when prompted, type the just-identifying restrictions[7]

$$
\begin{aligned}
A2 &= 0; \quad A4 = 1; \quad A5 = 0; \\
B1 &= 0; \quad B2 = 1; \quad B3 = -1; \\
C2 &= 0; \quad C3 = 1; \quad C8 = -1
\end{aligned}
$$

The exactly identified ML estimates of the three cointegrating vectors and their asymptotic standard errors should appear on the screen. These results are reproduced in Table 17.8 .

Table 17.8: ML estimates subject to exact identifying restrictions in the Indonesian model

```
              ML estimates subject to exactly identifying restriction(s)
         Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                         Converged after 1 iterations
      Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
 91 observations from 1981Q2 to 2003Q4. VARX(2,2), chosen r =3.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
 List of variables included in the cointegrating vector:
 Y                DP               R                EP               M1P
 YS               PS               RS               PO               T
 D97
 List of I(1) exogenous variables included in the VAR:
 YS               PS               RS               PO
 List of unrestricted deterministic variables included in the VAR:
 INPT             DD97
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                    Vector 1       Vector 2       Vector 3
 Y                    3.6481          .0000         .31257
                  (   4.6110)   (   *NONE*)   (    .54271)

 DP                     0.00         1.0000           0.00
                  (   *NONE*)   (   *NONE*)   (   *NONE*)

 R                  -41.3889        -1.0000         1.0000
                  (  41.5736)   (   *NONE*)   (   *NONE*)

 EP                   1.0000        .024489       -.043497
                  (   *NONE*)   (  .042042)   (  .064590)

 M1P                  -.0000       .0071140        -.11682
                  (   *NONE*)   (  .054510)   (   .16471)

 YS                 -11.7882        -.56785        .064650
                  (   9.0861)   (   .51348)   (   .54475)

 PS                   4.1765         .14681        -.22081
                  (   6.2966)   (   .28389)   (   .44285)

 RS                  90.1154         2.9075        -1.0000
                  (  77.2517)   (   1.2734)   (   *NONE*)

 PO                  -.86341       -.020479        .023821
                  (   1.0946)   (  .019172)   (  .025830)

 T                   .054605       .0048301      -.0019984
                  (  .058480)   ( .0032635)   ( .0047842)

 D97                 -.17344       -.050716        .091073
                  (   1.2145)   (  .039045)   (   .15447)

*******************************************************************************
 LL subject to exactly identifying restrictions=   1208.0
*******************************************************************************
```

The estimates in the first column refer to the PPP relation, while the second and the third column refer to FIP and IRP respectively.

[7] You can load the file IDQEXID.EQU with these restrictions by clicking the button.

Click the button and then select option 0. You are then asked whether you wish to test for over-identifying restrictions. Select YES and, when prompted, add the following restrictions:[8]

$$
\begin{aligned}
A7 &= 1; A1 = 0; A6 = 0; A3 = 0; A8 = 0; A9 = 0; A10 = 0; A11 = 0; \\
B4 &= 0; B5 = 0; B6 = 0; B7 = 0; B8 = 0; B9 = 0; B10 = 0; B11 = 0; \\
C1 &= 0; C4 = 0; C5 = 0; C6 = 0; C7 = 0; C9 = 0; C10 = 0; C11 = 0
\end{aligned}
$$

Notice that these restrictions, together with those given in the just-identifying case, specify the over-identifying matrix $\boldsymbol{\beta}$ in (17.11), and restrict the coefficients $\mathbf{b}_1$ and $\mathbf{b}_2$ for the trend and the lagged dummy variable $D97_{t-1}$ to be zero. In total, we have 33 restrictions, of which 24 are over-identifying. Click , and when prompted, press the button to accept the default initial values and choose option 2, the new algorithm (B) in *Microfit 5.0*.

The program returns an output window with ML estimates subject to over-identifying restrictions. Given the small size of the sample, we compute the bootstrapped critical values for the LR statistic. Click the button. You are asked whether you wish to bootstrap the critical values of the test for over-identifying restrictions. Press the button without changing the number of replications and the significance levels. *Microfit* starts the computations and then presents you with the results shown in Table 17.9.

[8] These restrictions are in the file IDQOVID.EQU that you can retrieve from the tutorial directory by clicking the . button.

Table 17.9: Testing for over-identifying restrictions in the Indonesian model

```
           ML estimates subject to over identifying restriction(s)
      Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                        Converged after 2 iterations
   Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
 91 observations from 1981Q2 to 2003Q4. VARX(2,2), chosen r =3.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
 List of variables included in the cointegrating vector:
 Y                DP               R                EP               M1P
 YS               PS               RS               PO               T
 D97
 List of I(1) exogenous variables included in the VAR:
 YS               PS               RS               PO
 List of unrestricted deterministic variables included in the VAR:
 INPT             DD97
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                   Vector 1        Vector 2        Vector 3
 Y                   -.0000          .0000          -.0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 DP                   .0000         1.0000          -.0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 R                   -.0000        -1.0000          1.0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 EP                  1.0000         -.0000           .0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 M1P                 -.0000          .0000           .0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 YS                   .0000          .0000           .0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 PS                  1.0000         -.0000           .0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 RS                  -.0000         -.0000         -1.0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 PO                  -.0000          .0000           .0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 T                    .0000          .0000          -.0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

 D97                 -.0000         -.0000          -.0000
                  (   *NONE*)     (   *NONE*)     (   *NONE*)

*******************************************************************************
 LR Test of Restrictions           CHSQ(24)=  74.9445[.000]
 95% Bootstrapped Critical Value =  75.8879
 90% Bootstrapped Critical Value =  70.7088
 Bootstrapped simulations based on 1000 SIMULATIONS.
 DF=Total no of restrictions(33) - no of just-identifying restrictions(9)
 LL subject to exactly identifying restrictions=   1208.0
 LL subject to over-identifying restrictions=   1170.6
```

Since the LR statistic (74.94) is below its bootstrapped critical value (75.90) at the 5 per cent significance level, we do not reject the 24 over-identifying restrictions. Notice that the linear trend does not enter in the cointegrating vectors, thus supporting the hypothesis that the long-run relations are cotrending. Furthermore, results on the tests for over-identifying restrictions indicate that the structural break in 1997, due to the Asian crisis, does not affect the long-run relations. This result, together with the absence of the time trend term T in the cointegration relations, suggests that the cointegrating relations on the Indonesian data are cobreaking and also cotrending.

Click 

and then CANCEL to return to the IR Analysis and Forecasting (Restricted

CVs) Menu, and choose option 8. You can now select in turn each variable to see the estimates of the corresponding individual error-correcting equation. For example, choose the variable inflation (DP). You will be presented with the results in Table 17.10.

Table 17.10: Error Correction Specifications for the Over-identified Model: 1981q2-2003q4

```
    ECM for variable DP estimated by OLS based on cointegrating VARX(2,2)
*******************************************************************************
 Dependent variable is dDP
 91 observations used for estimation from 1981Q2 to 2003Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 dY1                      -.30819             .11840            -2.6030[.011]
 dDP1                    .0053212            .088110            .060393[.952]
 dR1                       .67048             .23215             2.8882[.005]
 dEP1                      .10612            .017461             6.0774[.000]
 dM1P1                     .11035            .055369             1.9930[.050]
 dYS                       1.1346             .31086             3.6498[.000]
 dPS                       .78182             .47649             1.6408[.105]
 dRS                       .92961             1.3253             .70142[.485]
 dPO                     -.032388            .023366            -1.3861[.170]
 dYS1                     -1.3945             .28092            -4.9640[.000]
 dPS1                     -.35870             .42406            -.84588[.400]
 dRS1                     -2.0274             1.3958            -1.4524[.151]
 dPO1                    -.013221            .019687            -.67156[.504]
 ecm1(-1)                 .016510           .0075057             2.1997[.031]
 ecm2(-1)                 -.80015             .12664            -6.3185[.000]
 ecm3(-1)                 -.43801             .12953            -3.3814[.001]
 INPT                    -.085188            .039004            -2.1841[.032]
 DD97                   -.0035462            .016437            -.21574[.830]
*******************************************************************************
 List of additional temporary variables created:
 dDP = DP-DP(-1)
 dY1 = Y(-1)-Y(-2)
 dDP1 = DP(-1)-DP(-2)
 dR1 = R(-1)-R(-2)
 dEP1 = EP(-1)-EP(-2)
 dM1P1 = M1P(-1)-M1P(-2)
 dYS = YS-YS(-1)
 dPS = PS-PS(-1)
 dRS = RS-RS(-1)
 dPO = PO-PO(-1)
 dYS1 = YS(-1)-YS(-2)
 dPS1 = PS(-1)-PS(-2)
 dRS1 = RS(-1)-RS(-2)
 dPO1 = PO(-1)-PO(-2)
 ;ecm1 =    -.0000*Y +    .0000*DP    -.0000*R +   1.0000*EP    -.0000*M1P +
 .0000*YS +   1.0000*PS    -.0000*RS    -.0000*PO +    .0000*T    -.0000*D97;e
 cm2 =     .0000*Y +   1.0000*DP   -1.0000*R    -.0000*EP +    .0000*M1P +    .
 0000*YS    -.0000*PS    -.0000*RS +    .0000*PO +    .0000*T    -.0000*D97;ecm
 3 =    -.0000*Y    -.0000*DP +   1.0000*R +    .0000*EP +    .0000*M1P +    .0
 000*YS +    .0000*PS   -1.0000*RS +    .0000*PO    -.0000*T    -.0000*D97
*******************************************************************************
 R-Squared                         .80203   R-Bar-Squared                 .75593
 S.E. of Regression               .015763   F-Stat.    F(17,73)    17.3970[.000]
 Mean of Dependent Variable -.1220E-3   S.D. of Dependent Variable      .031906
 Residual Sum of Squares          .018137   Equation Log-likelihood     258.5655
 Akaike Info. Criterion          240.5655   Schwarz Bayesian Criterion  217.9678
 DW-statistic                      2.3110   System Log-likelihood         1170.6
*******************************************************************************

                              Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version         *           F Version        *
*******************************************************************************
*                     *                           *                            *
* A:Serial Correlation*CHSQ(4)  =    7.7663[.101]*F(4,69)      =    1.6096[.182]*
*                     *                           *                            *
* B:Functional Form   *CHSQ(1)  =    2.6774[.102]*F(1,72)      =    2.1826[.144]*
*                     *                           *                            *
* C:Normality         *CHSQ(2)  =    3.3914[.183]*       Not applicable        *
*                     *                           *                            *
* D:Heteroscedasticity*CHSQ(1)  =    6.0737[.014]*F(1,89)      =    6.3650[.013]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
*******************************************************************************
```

The lags of domestic interest rate ($dR1$), exchange rates ($dEP1$), and foreign output (dYS) are highly significant in explaining inflation rate. Note that all error-correcting coefficients are statistically significant.

As an exercise, obtain error-correction models for the remaining endogenous variables. A summary of results is reported in Table 17.11. These estimates show that the error-correction term for IRP and FIP are important in most equations, while the error-correcting coefficient of PPP is significant at 5 per cent significance level only in the inflation equation. Furthermore, the error-correction equations pass most of the diagnostic tests and fit the historical observations relatively well. In particular, the $\overline{R}^2$ of domestic output and interest rate equations, computed at 0.59 and 0.67 respectively, are quite high. The diagnostic statistics for tests of residual serial correlation and heteroscedasticity are within the 90 per cent critical values for the Y and $M1P$ equations. There is evidence of non-normal errors in the case of some of the error-correcting equations, and mis-specification for the EP equation.

Table 17.11: The reduced form error-correction specification for the Indonesian model

Equation	dY_t	dR_t	dEP_t	$dM1P_t$
$Intercept$	-0.022	-0.006	0.263	-0.043
$\widehat{\xi}_{1,t-1}$	0.007	0.003	-0.055	0.012
$\widehat{\xi}_{2,t-1}$	-0.096	0.117**	-2.506***	-1.396***
$\widehat{\xi}_{3,t-1}$	-0.500***	-0.190***	-0.344	-0.729**
dYS_t	0.076	0.024	1.422	-1.048
dPS_t	-0.345	0.200	1.439	-0.689
dRS_t	1.456	1.132**	6.763	3.125
dPO_t	0.026	-0.013	-0.167	0.006
dY_{t-1}	-0.110	-0.107**	-1.649**	-0.126
dDP_{t-1}	-0.306***	-0.027	2.148***	0.571***
dR_{t-1}	0.268	0.059	1.531	-0.094
dEP_{t-1}	-0.043***	0.029***	-0.052	-0.127***
$dM1P_{t-1}$	-0.018	0.014	0.082	0.014
dYS_{t-1}	0.740***	-0.381***	1.916	0.410
dPS_{t-1}	-0.048	0.115	-2.592	0.617
dRS_{t-1}	-0.508	-0.031	0.286	5.663*
dPO_{t-1}	-0.009	-0.008	0.081	-0.016
$DD97_t$	0.005	0.059***	0.191*	-0.092**
$\overline{R}^2$	0.5852	0.6728	0.1532	0.3853
$\widehat{\sigma}$	0.012	0.006	0.106	0.037
$\chi^2_{SC}[4]$	11.160**	5.020	14.93***	3.070
$\chi^2_{FF}[1]$	5.770**	4.980**	7.79**	1.490
$\chi^2_{N}[2]$	3.470	39.220***	11.33***	1.820
$\chi^2_{H}[1]$	0.080	0.220	25.660***	1.880

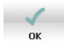
Notes: The asterisks '*', '**' and '***' indicate significance at the 10%, 5% and 1% levels, respectively. The diagnostic tests are chi-squared statistics for serial correlation (SC), functional form (FF), normality (N) and heteroscedasticity (H).

17.4 Lesson 17.4: Forecasting UK inflation

In this lesson we show how to use *Microfit* to compute dynamic forecasts in the context of a vector error-correcting autoregressive model with exogenous variables. We will consider the small model for the UK economy described in Lesson 17.1 and focus on forecasting the changes in domestic prices. To this end, we re-estimate the $VARX$ model over a shorter time period, and keep the remaining observations for forecasting purposes.

Microfit allows computing multivariate forecasts in two different ways, depending on how exogenous variables are treated (see Section 22.10.4). One possibility is to compute forecasts *conditional* on the values of the exogenous variables observed over the forecasting period. An alternative is to use the marginal models to forecast exogenous variables, and then employ these forecasts to compute *unconditional* forecasts for the variables. We will explore both possibilities in turn.

Follow the steps outlined in Lesson 17.1 to estimate a $VARX$ model with three endogenous variables (R, P and E), two exogenous variables (RS and PS), an intercept and a time trend, over the period $1977Q1$-$1997Q4$ (for a total of 84 observations), taking the IRP and PPP long-run relations as given. Hence, specify $r = 2$ and use option 4 from the Long-run Structural Modelling Menu to impose the exact-identifying restrictions represented in (17.1) and (17.2).[9]

Then go to the IR Analysis and Forecasting Menu, select option 6 to compute multivariate dynamic forecasts, and click . When prompted, enter the forecast period

$$1998Q1 \quad 1999Q4$$

and select the option 'Conditional Forecasts using the supplied values of exogenous variables', and click . You will be presented with the list of variables in the model. Choose P and then select 'change in P'. You enter the Multivariate Forecast Menu, where you can decide to see, plot or save the forecast. For example, choose option 1. The results in Table 17.12 should appear on the screen.

[9] For this you need to impose restrictions (17.3).

Table 17.12: Conditional forecasts of domestic price changes in the UK model

```
                 Multivariate dynamic forecasts for the change in P
      Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
 Conditional forecasts computed using the future observations of the weakly
 exogenous variables.
*******************************************************************************
 84 observations from 1977Q1 to 1997Q4. VARX(2,2), chosen r =2.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
 List of variables included in the cointegrating vector:
 R               P               E                RS               PS
 T
 List of I(1) exogenous variables included in the VAR:
 RS              PS
 List of unrestricted deterministic variables included in the VAR:
 INPT

 Definition of the Marginal Model:

 Equation 1 of the Marginal Model:

 Dependent variable = dRS

 List of regressors:

 dRS(-1)         dPS(-1)         dR(-1)           dP(-1)           dE(-1)
 INPT
 Equation 2 of the Marginal Model:
 Dependent variable = dPS
 List of regressors:
 dRS(-1)         dPS(-1)         dR(-1)           dP(-1)           dE(-1)
 INPT
*******************************************************************************
 Observation              Actual              Prediction              Error
  1998Q1                -.0023802             -.4259E-3             -.0019543
  1998Q2                 .0014434              .9550E-4              .0013479
  1998Q3                 .0022262              .0017672              .4590E-3
  1998Q4                -.0012894              .0013555             -.0026449
  1999Q1                -.4590E-3             -.0015161              .0010571
  1999Q2                 .0090551              .0027419              .0063132
  1999Q3                 .0069724              .0082221             -.0012497
  1999Q4                 .0053846              .0087039             -.0033193
*******************************************************************************
             Summary Statistics for Residuals and Forecast Errors
*******************************************************************************
                      Estimation Period             Forecast Period
                      1977Q1 to 1997Q4             1998Q1 to 1999Q4
*******************************************************************************
 Mean                           .0000                  .1132E-5
 Mean Absolute                .0038800                 .0022932
 Mean Sum Squares             .2490E-4                 .8299E-5
 Root Mean Sum Squares        .0049897                 .0028809
*******************************************************************************
```

The forecasts are very close to the actual values, especially in the first four quarters. The plot of actual and forecast values is presented in Figure 17.1.

Repeat the earlier process with the same variable P, but this time select the option 'Unconditional Forecasts based on the forecasts of the exogenous variables using the Marginal Model'. The forecast and errors are displayed in Table 17.13.

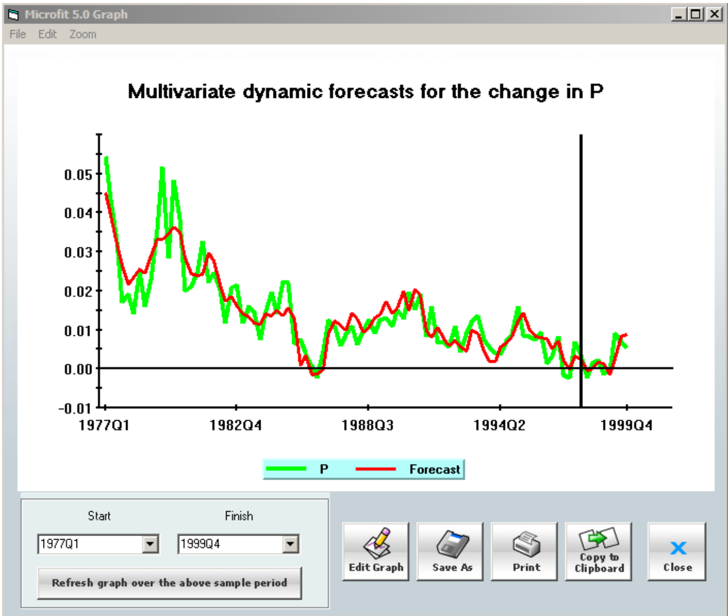


Figure 17.1: Plot of in-sample fitted values and out-of-sample dynamic conditional forecasts of GDP growth in Indonesia

Table 17.13: Unconditional forecasts of domestic price changes in the UK model

```
              Multivariate dynamic forecasts for the change in P
   Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
 Unconditional forecasts using the marginal model specified for the weakly
 exogenous variables.
*******************************************************************************
 84 observations from 1977Q1 to 1997Q4. VARX(2,2), chosen r =2.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 2.
 List of variables included in the cointegrating vector:
 R                P                E                RS               PS
 T
 List of I(1) exogenous variables included in the VAR:
 RS               PS
 List of unrestricted deterministic variables included in the VAR:
 INPT

 Definition of the Marginal Model:
 Equation 1 of the Marginal Model:
 Dependent variable = dRS
 List of regressors:
 dRS(-1)          dPS(-1)          dR(-1)           dP(-1)           dE(-1)
 INPT
 Equation 2 of the Marginal Model:
 Dependent variable = dPS
 List of regressors:
 dRS(-1)          dPS(-1)          dR(-1)           dP(-1)           dE(-1)
 INPT
*******************************************************************************
 Observation              Actual              Prediction              Error
  1998Q1                -.0023802              .0036589             -.0060391
  1998Q2                 .0014434              .0050490             -.0036056
  1998Q3                 .0022262              .0058507             -.0036245
  1998Q4                -.0012894              .0064149             -.0077043
  1999Q1                -.4590E-3              .0068988             -.0073578
  1999Q2                 .0090551              .0073354              .0017197
  1999Q3                 .0069724              .0077301             -.7577E-3
  1999Q4                 .0053846              .0080846             -.0026999
*******************************************************************************
            Summary Statistics for Residuals and Forecast Errors
*******************************************************************************
                        Estimation Period             Forecast Period
                        1977Q1 to 1997Q4             1998Q1 to 1999Q4
*******************************************************************************
 Mean                          .0000                      -.0037587
 Mean Absolute               .0038800                      .0041886
 Mean Sum Squares            .2490E-4                      .2337E-4
 Root Mean Sum Squares       .0049897                      .0048338
*******************************************************************************
```

Notice that forecasts errors are larger than those in Table 17.12. The root mean sum of squares of forecast errors computed over the forecast period, equal to 0.005, is larger than that displayed in Table 17.12, computed using conditional forecasts (0.003). The plot of actual and forecasted values is presented in Figure 17.2. Note that in-sample fitted values are identical using either conditional or unconditional forecasts.

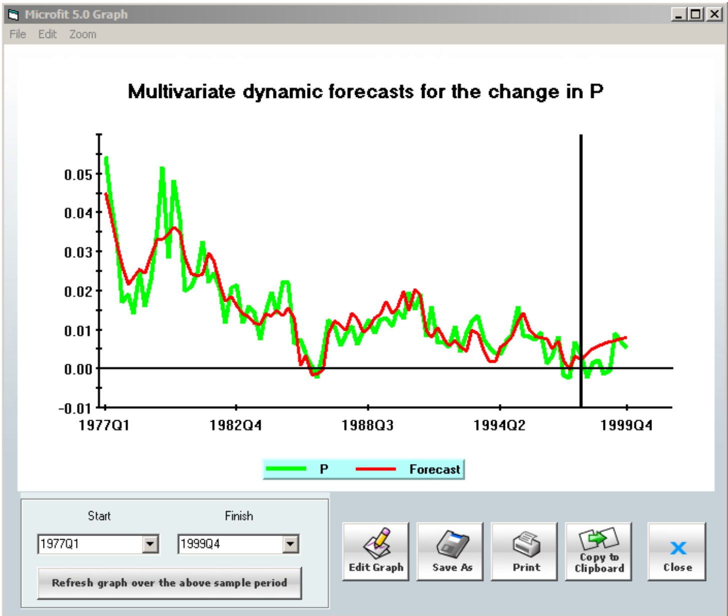


Figure 17.2: Plot of in-sample fitted values and out-of-sample dynamic unconditional forecasts of GDP growth in Indonesia

17.5 Lesson 17.5: Permanent and transitory components of output and consumption in a small model of the US economy

In this lesson we consider a simple model in the logarithms of (private) real output (Y), real consumption (C), and real investment (I) for the US economy, using data from the file KPSW.FIT in the tutorial directory. In Lesson 16.6 we concluded that these variables could be $I(1)$, and that a $VAR(2)$ model can adequately capture the dynamic properties of data. In this lesson we start from these results and employ the multivariate Beveridge-Nelson representation to compute permanent and transitory components of output, consumption and investments in the US.

The multivariate BN procedure allows you to extract from non-stationary series a permanent component and a transitory component (or cycle), where the permanent component is further sub-divided into a trend or deterministic part and a stochastic component (see Section 22.11). More specifically, let $\mathbf{z}_t = (C_t, I_t, Y_t)'$ be the vector of variables and consider the decomposition of $\mathbf{z}_t$ as

$$\mathbf{z}_t = \mathbf{z}_t^P + \mathbf{z}_t^C$$

where $\mathbf{z}_t^C$ is the cycle and $\mathbf{z}_t^P$ is the permanent component satisfying

$$\mathbf{z}_t^P = \mathbf{z}_{dt}^P + \mathbf{z}_{st}^P$$

with $\mathbf{z}_{dt}^P$ and $\mathbf{z}_{st}^P$ being the trend and stochastic parts. The multivariate BN decomposition in *Microfit* differs from the classic BN decomposition since it allows taking into account restrictions on intercepts and/or trends, as well as incorporating the long-run relations between variables. For a description of the multivariate trend-cycle decomposition, see Section 22.11.

Load into *Microfit* the file KPSW.FIT (containing quarterly observations over the period 1947(1)-1988(4) for the US economy), and in the Process window create a constant ($INPT$) and a linear time trend (T). Choose option 3 (Cointegrating $VARX$) from the System Estimation Menu (Multivariate Menu) and in the Commands and Data Transformations box type[10]

$$C \quad I \quad Y \quad \& \quad INPT \quad T$$

Then specify the sample period

$$1954Q1 \quad 1988Q4$$

Enter 2 as the order of the lag of the endogenous variable[11], and click 

. Since we wish to avoid the possibility of quadratic trends in some of the variables, restrict the trend, while leaving the intercept unrestricted.

[10] Since in this application there are no exogenous variables, identical results can be obtained choosing option 2 from the System Estimation Menu (Multivariate Menu) and selecting the unrestricted intercepts, restricted trends option 4.

[11] Use the AIC or SBC selection criteria available in option 1 in the System Estimation Menu (Multivariate Menu) to check that the order 2 is adequate.

The balanced growth hypothesis implies that consumption and investment should be cointegrated with output, each of them with unit cointegrating vector (see King, Plosser, Stock, and Watson (1991) and also the discussion in Lesson 16.6). In particular, we expect the following two long-run relations among these variables (recall that all the three variables are in logs):

$$C - Y \sim I(0) \tag{17.12}$$

$$I - Y \sim I(0) \tag{17.13}$$

Following the same line of reasoning as in Lessons 17.1 and 17.2, we now test for the over-identifying restrictions implied by (17.12) and (17.13). In the Cointegrating Var Post Estimation Menu, specify the number of cointegrating vectors $r = 2$; go to option 6 and then 4, and type in the box editor the following exact-identifying restrictions:

$$A1 = 1; \qquad A2 = 0;$$
$$B1 = 0; \qquad B2 = 1;$$

Microfit returns an output window with the ML estimates together with their asymptotic standard errors, which is reproduced in Table 17.14.

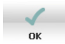
Table 17.14: ML estimates under exactly identifying restrictions for the US model

```
           ML estimates subject to exactly identifying restriction(s)
       Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                          Converged after 1 iterations
    Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
 140 observations from 1954Q1 to 1988Q4. VARX(2,0), chosen r =2.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 0.
 List of variables included in the cointegrating vector:
 C               I               Y               T
 List of unrestricted deterministic variables included in the VAR:
 INPT
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                  Vector 1       Vector 2
 C                   1.0000           0.00
                  (    *NONE*)    (    *NONE*)

 I                     0.00         1.0000
                  (    *NONE*)    (    *NONE*)

 Y                  -1.0756        -1.5473
                  (    .16332)    (    .34556)

 T                  .1030E-4       .0019219
                  ( .6380E-3)    ( .0013448)

*******************************************************************************
 LL subject to exactly identifying restrictions=    1353.4
*******************************************************************************
```

Notice that for both cointegrating vectors the coefficients of the variable Y are very close to -1, which is the value we would expect under the two long-run relations (17.12) and (17.13). Furthermore, the coefficients on the time trends are nearly zero in both relations, suggesting that the co-trending restrictions are likely to hold in the case of both long-run relations. Click 

to leave the output screen, and select option 0 to test for over/identifying restrictions. On the basis of the above results, we now impose restrictions on the coefficients for Y and on the linear trend T. In the box editor add the following over-identifying restrictions:

$$
\begin{aligned}
A3 &= -1; \quad A4 = 0; \\
B3 &= -1; \quad B4 = 0;
\end{aligned}
$$

Click 

to accept the default initial values, and choose the back substitution Algorithm B. When the output appears on the screen, close it and click

to compute the bootstrapped critical values, using the default number of replications and significance levels. *Microfit* starts the computations and, when finished, returns the output reproduced in Table 17.15.

Table 17.15: Maximum likelihood estimates subject to over-identifying restrictions

```
              ML estimates subject to over identifying restriction(s)
         Estimates of Restricted Cointegrating Relations (SE'sin Brackets)
                          Converged after 2 iterations
      Cointegrating Vector Autoregression Model with Exogenous I(1) Variables
*******************************************************************************
 140 observations from 1954Q1 to 1988Q4. VARX(2,0), chosen r =2.
 Lag order of endogenous variables = 2. Lag order of exogenous variables = 0.
 List of variables included in the cointegrating vector:
 C                I                Y                T
 List of unrestricted deterministic variables included in the VAR:
 INPT
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
                  Vector 1       Vector 2
 C                   1.0000          .0000
                  (    *NONE*)    (    *NONE*)

 I                    .0000         1.0000
                  (    *NONE*)    (    *NONE*)

 Y                  -1.0000        -1.0000
                  (    *NONE*)    (    *NONE*)

 T                   -.0000          .0000
                  (    *NONE*)    (    *NONE*)

*******************************************************************************
 LR Test of Restrictions          CHSQ(4)=    9.3073[.054]
 95% Bootstrapped Critical Value =  15.9068
 90% Bootstrapped Critical Value =  13.5887
 Bootstrapped simulations based on 1000 SIMULATIONS.
 DF=Total no of restrictions(8) - no of just-identifying restrictions(4)
 LL subject to exactly identifying restrictions=   1353.4
 LL subject to over-identifying restrictions=   1348.7
*******************************************************************************
```

Note that the LR test (9.31) is below the 95 per cent bootstrapped critical value (15.91). Hence, we do not reject the over-identifying restrictions listed above, and in the following analysis of the trend-cycle properties of the model we adopt the cointegrating vectors displayed in Table 17.15.

Click Close and then Cancel to go to the IR Analysis and Forecasting Menu. Option 5 in this menu enables you to compute the multivariate Beveridge-Nelson (BN) decomposition. You will be presented with a list of initial values and growth rates for the long-run variables. Leave them unchanged, and click the OK button. You are then asked to choose a variable for which you wish to inspect the trend-cycle decomposition. Choose, for example, the variable output (Y), and click OK, then close the output window to go to the Trend/Cycle Decomposition Result Menu. Hence, select option 3 to save the decomposition for all variables in a CSV file. When the Save as dialogue appears, choose the drive and the directory in which you want to save the data, and type in a filename in the usual way. Click OK, and then specify the sample period

$$1947Q1 \quad 1988Q4$$ OK

Then click as many times as required to move back to the Commands and Data Transformations box, go to the File Menu and open the CSV file you have just saved. Click the Data button to make sure that the variables have been correctly loaded in your workspace. Also, notice that the values for the variables in the workspace are missing until 1953q4, since we estimated the model over the period 1954q1-1988q4.

Go to the Process window, clear the Commands and Data Transformations box, and type

SAMPLE $1954Q1 \quad 1988Q4$; **PLOT** $Y \quad P\_Y$ RUN

The output is reproduced in Figure 17.3.

Notice that the permanent component of output rises quite rapidly in the 1960s, and then more slowly in the 1970s, consistent with the much debated productivity slowdown that occurred over the course of this decade (see Attfield and Temple (2003)). In the 1980s, the growth of the permanent component is more rapid. Our analysis indicates that the rate of output growth observed in the 1960s was higher than the rate of growth of the permanent component, reflecting favourable transitory shocks. To investigate the effect of transitory shocks, we can add the cycle series of output (C_Y) to the graph. Close the graph to return to the Commands and Data Transformation box, and type

SAMPLE $1954Q1 \quad 1988Q4$; **PLOT** $Y \quad P\_Y \quad \& \quad C\_Y$ RUN

Microfit produces a new line graph of Y, P_Y, and C_Y in which is added a secondary y-axis reporting values for variable C_Y (see Section for more information on the **PLOT** command). The graph is shown in Figure 17.4.

Notice that the C_Y series is subject to significant downward, as well as upward, shifts at various points in the sample. Its behaviour reflects positive transitory shocks during the 1960s and the 1980s.

We now consider the relationship between output and consumption. Close the graph, and in the Commands and Data Transformations type

SAMPLE $1954Q1 \quad 1988Q4$; **PLOT** $C\_Y \quad C\_C$ RUN

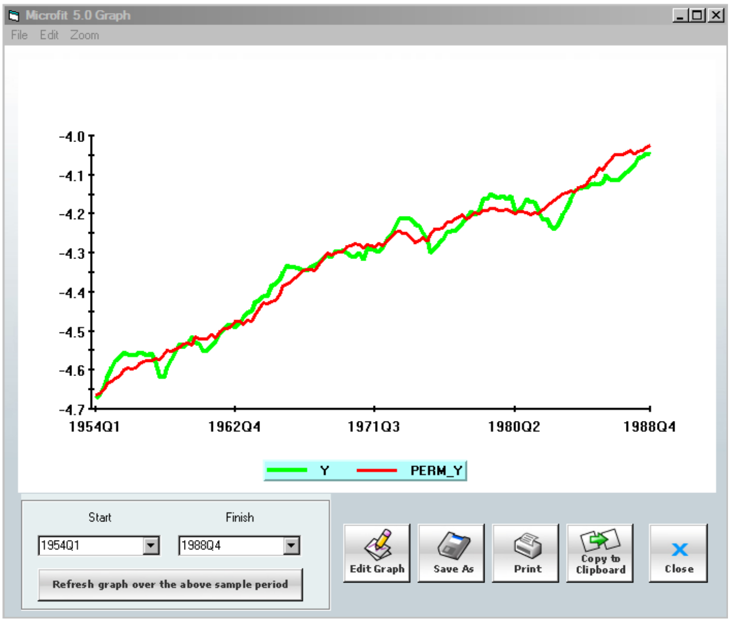


Figure 17.3: Plot of the US output and the permanent component

The graph is shown in Figure 17.5. We note that given the multivariate approach to detrending, the two series incorporate the effect of the interactions between output, consumption and investments.

As expected, the degree of co-movement between the transitory components of output and consumption is very high, yielding a correlation coefficient of 0.87 (you can obtain the correlation coefficient by using the command **COR**; see Section 4.4.8 for the use of this command). However, the variations of transitory consumption are smaller in absolute value than those in transitory output, which is in line with the excess smoothness puzzle discussed in the consumption literature.

The plot of transitory output and transitory investments is provided in Figure 17.6. Again, there is a high degree of synchronization between the two cyclical components, and not surprisingly the transitory components of investment show a much more pronounced cyclical fluctuations as compared to the fluctuations in the transitory component of real output.

17.6 Lesson 17.6: The trend-cycle decomposition of interest rates

In this lesson we consider a bivariate VAR model in two interest rates and derive the *BN* permanent/transitory decompositions of these variables analytically and then show that they do in fact coincide with the *BN* decompositions obtained using *Microfit*. In the empirical

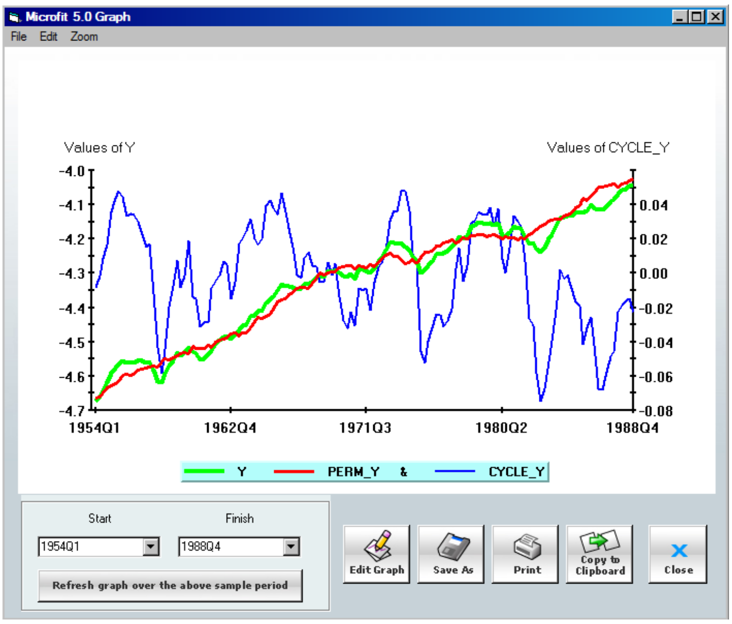


Figure 17.4: Plot of the US output, the permanent component and the cycle

part we use data on UK domestic and foreign interest rates (for further details on this exercise, see Dees, di Mauro, Pesaran, and Smith (2007), and for more information on the BN decomposition see Section 22.11).

Let r_t and r_t^* be the domestic (UK) and foreign interest rates respectively, and consider the following simple error-correction models:

$$\begin{aligned}\Delta r_t &= a(r_{t-1} - r_{t-1}^*) + \varepsilon_t \\ \Delta r_t^* &= b(r_{t-1} - r_{t-1}^*) + \varepsilon_t^*\end{aligned}$$

The above two equations can be written more compactly as

$$\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} + \mathbf{u}_t \tag{17.14}$$

where $\mathbf{z}_t = (r_t, r_t^*)'$,$\mathbf{u}_t = (\varepsilon_t, \varepsilon_t^*)'$ and

$$\mathbf{A} = \begin{pmatrix} 1+a & -a \\ b & 1-b \end{pmatrix}$$

Solving the difference equation (17.14) by recursive substitution we have

$$\mathbf{z}_{t+h} = \mathbf{A}^h\mathbf{z}_t + \mathbf{A}^{h-1}\mathbf{u}_{t+1} + \mathbf{A}^{h-2}\mathbf{u}_{t+2} + ... + \mathbf{u}_{t+h}$$

and hence

$$E_t\left(\mathbf{z}_{t+h}\right) = \mathbf{A}^h\mathbf{z}_t$$

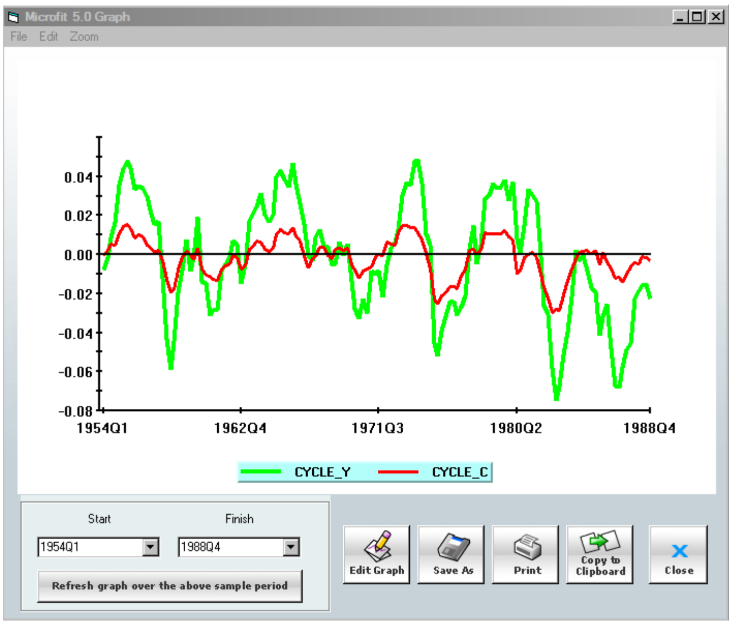


Figure 17.5: Plot of the transitory components of output and consumption in the US

where $E_t(\cdot)$ denotes the expectation operator conditional on the information at time t. Since in this example there are no deterministic variables such as intercept or trend, the permanent component of $\mathbf{z}_t$ is given by

$$\mathbf{z}_t^P = \mathbf{z}_{st}^P = \lim_{h\to\infty} E_t\left(\mathbf{z}_{t+h}\right) = \left(\lim_{h\to\infty} \mathbf{A}^h\right)\mathbf{z}_t = \mathbf{A}^\infty \mathbf{z}_t \tag{17.15}$$

If we instead use the common component moving average representation of $\mathbf{z}_t$, we have

$$\mathbf{z}_t = \widetilde{\mathbf{z}}_0 + \mathbf{C}(1)\mathbf{s}_{ut} + \mathbf{C}^*(L)\mathbf{u}_t$$

where $\widetilde{\mathbf{z}}_0 = \mathbf{z}_0 - \mathbf{C}^*(L)\mathbf{u}_0$, $\mathbf{s}_{ut} = \sum_{i=1}^t \mathbf{u}_i$, and

$$\mathbf{C}(1) = \sum_{i=0}^{\infty} \mathbf{C}_i$$

$$\mathbf{C}^*(L) = \sum_{i=0}^{\infty} \mathbf{C}_i^* L^i$$

with

$$\begin{aligned} \mathbf{C}_0 &= \mathbf{I}_2,\ \mathbf{C}_i = -(\mathbf{I}_2 - \mathbf{A})\mathbf{A}^{i-1} \quad \text{for } i = 1, 2, \ldots \\ \mathbf{C}_i^* &= \mathbf{C}_{i-1}^* + \mathbf{C}_i \end{aligned}$$

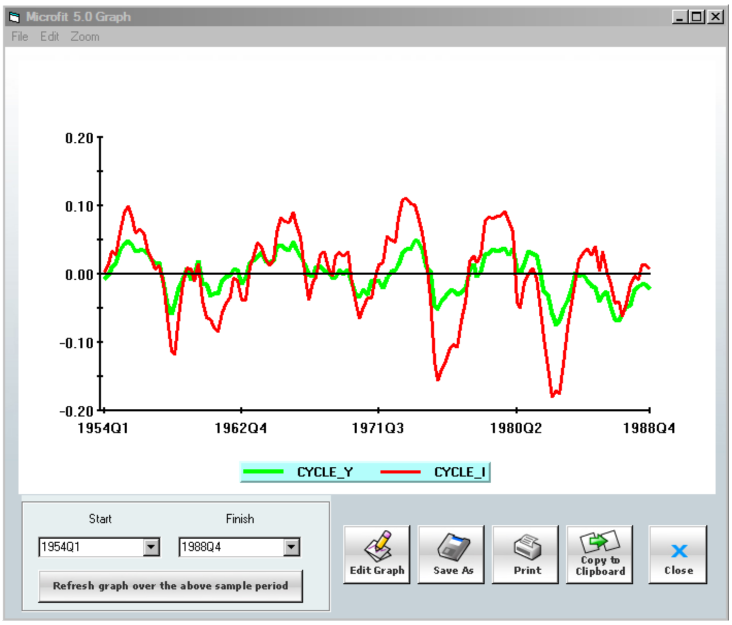


Figure 17.6: Plot of transitory output and investments in the US

Also, recall that $\mathbf{C}^*(L)\mathbf{u}_t$ is the stationary component of $\mathbf{z}_t$. Hence,

$$\begin{aligned} E_t(\mathbf{z}_{t+h}) &= \mathbf{z}_0 + E_t[\mathbf{C}(1)\mathbf{s}_{u,t+h}] + E_t[\mathbf{C}^*(L)\mathbf{u}_{t+h}] \\ &= \widetilde{\mathbf{z}}_0 + \mathbf{C}(1)\mathbf{s}_{ut} + E_t[\mathbf{C}^*(L)\mathbf{u}_{t+h}] \end{aligned}$$

and since $\mathbf{C}^*(L)\mathbf{u}_{t+h}$ is stationary, then

$$\lim_{h\to\infty} E_t(\mathbf{z}_{t+h}) = \mathbf{z}_0 + \mathbf{C}(1)\mathbf{s}_{ut}$$

but noting that

$$\mathbf{C}(1) = \mathbf{I}_2 - (\mathbf{I}_2 - \mathbf{A})(\mathbf{I}_2 + \mathbf{A} + \mathbf{A}^2 + ...) = \lim_{h\to\infty} \mathbf{A}^h = \mathbf{A}^\infty$$

Hence,

$$\lim_{h\to\infty} E_t(\mathbf{z}_{t+h}) = \widetilde{\mathbf{z}}_0 + \mathbf{A}^\infty \mathbf{s}_{ut} = \mathbf{z}_0 + \mathbf{A}^\infty(\mathbf{u}_1 + \mathbf{u}_2 + ... + \mathbf{u}_t) \tag{17.16}$$

This result looks very different from that in (17.15) obtained using the direct method. However, notice that

$$\mathbf{z}_t = \mathbf{A}^t\mathbf{z}_0 + \sum_{j=0}^{t-1} \mathbf{A}^{j-1}\mathbf{u}_{t-j}$$

Premultiplying both sides by $\mathbf{A}^h$ and letting $h \to \infty$ we have

$$\left(\lim_{h\to\infty} \mathbf{A}^h\right)\mathbf{z}_t = \left(\lim_{h\to\infty} \mathbf{A}^{t+h}\right)\mathbf{z}_0 + \sum_{j=0}^{t-1}\left(\lim_{h\to\infty} \mathbf{A}^{h+j-1}\right)\mathbf{u}_{t-j}$$

and since $\lim_{h\to\infty}\mathbf{A}^{t+h}=\mathbf{A}^{\infty}$, for any t, then

$$\begin{aligned}\mathbf{A}^{\infty}\mathbf{z}_t &= \mathbf{A}^{\infty}\mathbf{z}_0+\mathbf{A}^{\infty}\left(\sum_{j=0}^{t-1}\mathbf{u}_{t-j}\right)\\ &= \widetilde{\mathbf{z}}_0+\mathbf{A}^{\infty}(\mathbf{u}_1+\mathbf{u}_2+...+\mathbf{u}_t)\end{aligned}$$

implying that (17.15) and (17.16) are equivalent.

In this example $\mathbf{A}^{\infty}$ can be obtained explicitly. It is easily seen that the eigenvalues of $\mathbf{A}$ are $\lambda_1=1$ and $\lambda_2=1+a-b$. Hence, the Jordan form of $\mathbf{A}$ is given by

$$\mathbf{A}^h=\mathbf{Q}\begin{pmatrix}1 & 0\\ 0 & a-b+1\end{pmatrix}^h\mathbf{Q}^{-1}$$

where

$$\mathbf{Q}=\begin{pmatrix}1 & 1\\ 1 & \frac{1}{a}b\end{pmatrix},\ \mathbf{Q}^{-1}=\begin{pmatrix}\frac{b}{-a+b} & -\frac{a}{-a+b}\\ -\frac{a}{-a+b} & \frac{a}{-a+b}\end{pmatrix}$$

Assuming that $0<b-a<2$, then $|\lambda_2|<1$, and we have

$$\begin{aligned}\lim_{h\to\infty}\mathbf{A}^h &= \mathbf{Q}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\mathbf{Q}^{-1}\\ &= \frac{1}{b-a}\begin{pmatrix}b & -a\\ b & -a\end{pmatrix}\end{aligned}$$

Therefore, the stochastic component of $\mathbf{z}_t$, $\mathbf{z}_t^P$, is given by

$$\mathbf{z}_t^P=\frac{1}{b-a}\begin{pmatrix}br_t-ar_t^*\\ br_t-ar_t^*\end{pmatrix}\tag{17.17}$$

Clearly, r_t and r_t^* have the same stochastic components. Furthermore, the cycles for r_t and r_t^* are given by

$$\tilde{r}_t = r_t-\frac{br_t-ar_t^*}{b-a}=\frac{-a(r_t-r_t^*)}{b-a}\tag{17.18}$$

$$\tilde{r}_t^* = r_t^*-\frac{br_t-ar_t^*}{b-a}=\frac{b(r_t-r_t^*)}{b-a}\tag{17.19}$$

where, using UK data over the period 1979-2003, $a=-0.13647$ and $b=0.098014$ (Dees, di Mauro, Pesaran, and Smith (2007)). Note that $b-a=0.2345$ which falls in the range $(0,2)$, as required.

To check that the above results in (17.17) and (17.18)-(17.19) coincide with their counter parts computed using *Microfit*, load the file UKM07.FIT from the tutorial directory, and make sure that the variables R and RS are in the workspace. Then select option 1 (no intercept or trends) from the Cointegrating VAR Menu, and in the Commands and Data Transformation box type

R RS

Choose 1 as lag of the VAR, and select the estimation period

$$1979Q2 \quad 2003Q4$$

Close the output window, and use option 2 to set the number of cointegrating relations $r = 1$. Then choose option 6 and then 4, and type the following restriction:

$$A1 = 1$$

Close the output window, choose option 1 to test for over-identifying restrictions, and in the box editor add the restriction

$$A2 = -1$$

Click twice to see the results, and check that the LR test does not reject the over-identifying restrictions. Hence, make the necessary steps to return to the IR Analysis and Forecasting Menu, and select option 5 to compute the BN trend/cycle decomposition of the variable R. Close the output window, and in the Trend/Cycle Decomposition Menu choose to save the decomposition for all the variables in a CSV file (option 3). When prompted, type in a filename and click OK to save the decomposition over the whole time period. Then, click as many times as required to move back to the Commands and Data Transformation box. Go to the File Menu and choose to load the CSV file that you have just saved. Notice that in the workspace there are the variables R, RS and their BN decompositions (Click on the button to view the variable names and their descriptions).

Go to the Process window and in the Commands and Data Transformations box type

$$\begin{aligned}
AHAT &= -0.13647;\ BHAT = 0.098014; \\
PERM &= (BHAT * R - AHAT * RS)/(BHAT - AHAT); \\
CYCLER &= R - (BHAT * R - AHAT * RS)/(BHAT - AHAT); \\
CYCLERS &= RS - (BHAT * R - AHAT * RS)/(BHAT - AHAT)
\end{aligned}$$

Alternatively you can retrieve the file UKM07.EQU from the tutorial directory. The above commands construct the variables provided by the equations (17.17) and (17.18)-(17.19). Now clear the Commands and Data Transformations box, and inspect the correlation between the variables just constructed and the decomposition provided by *Microfit*. For example, type

COR $PERM \quad P\_R \quad P\_RS$

Notice that all correlations are equal to 1. Then try with

COR $CYCLER \quad C\_R \quad C\_RS$

Again, all correlations should be equal to 1. Therefore, the directly computed series $PERM$, $CYCLER$, and $CYCLERS$ are perfectly correlated with the ones (respectively, P_R C_R and C_RS) computed using *Microfit*.

17.7 Lesson 17.7: The US equity market and the UK economy

This lesson investigates the time profile of the effects on the UK economy of shocks to US real equity prices, by using the generalized impulse response methodology in the context of a $VARX$ model (see Section 22.10.4). In this lesson the US economy is considered exogenous to the UK economy, and plays the role of the rest of the world. See Dees, di Mauro, Pesaran, and Smith (2007) and Pesaran, Schuermann, and Weiner (2004) for further details on cointegrating VARs for multiple countries.

We set up a $VARX$ model for the UK economy with five endogenous variables, namely domestic output, interest rate, exchange rate, inflation and equity price index, and three exogenous variables, given by real oil price (in the UK), and inflation rate, interest rate and equity prices in the US. These variables for the period 1979q2-2003q4 are contained in the CSV file entitled UKUSEK.CSV in the tutorial directory:

Y	UK real output
DP	UK inflation rate
$EPEPS$	Sterling effective exchange rate
Q	UK real equity price index
R	UK rate of interest (0.25*log(1+R/100))
$RPOIL$	UK real oil price
US_DP	US inflation
US_Q	US real equity price index
US_R	US interest rate (0.25*log(1+US_R/100))
$INPT$	Intercept
T	Time trend

Select option 3 from the Multivariate Menu, and switch on the 'Simulation of critical values' button. Then choose the period 1979Q2-2003Q4 as an estimation sample, select 1 as lags for the both endogenous and exogenous variables,[12] and in the Commands and Data Transformation type

$Y \quad R \quad EPEPS \quad DP \quad Q; \quad US\_DP \quad US\_R \quad US\_Q \quad RPOIL \quad \& \quad INPT \quad T$ RUN

Alternatively, you can retrieve the LST file UKUSMOD.LST from the tutorial directory. Check the checkbox corresponding to the variable T to restrict the time trend, and as marginal models for the four exogenous variables specify

$$1 \quad 1 \quad INPT;$$
$$1 \quad 1 \quad INPT;$$
$$1 \quad 1 \quad INPT;$$
$$1 \quad 1 \quad INPT$$ RUN

[12] This choice was based on the SBC criterion obtained from option 1 (unrestricted VAR) in the Multivariate Menu.

Notice that both trace and maximum eigenvalue statistics in the output window show the presence of two cointegrating relations. Hence, close the output and use option 2 to specify the number of cointegrating relations $r = 2$. Then go to the Long-Run Structural Modelling Menu and choose option 4 to test for restrictions. In the box editor type the following restrictions:

$$\begin{aligned} A2 &= 1; A4 = -1; \\ B2 &= 1; B7 = -1 \end{aligned}$$

The above restrictions represent the Interest Rate Parity (IRP) and the Fisher Inflation Parity (FIP) (see Lesson 17.1 and 17.3 for a description of these relations). Close the output window, and in the IR Analysis and Forecasting Menu choose option 1 to perform an impulse response analysis of shocks to equations. Select the generalized impulse response analysis (option 2) and choose to shock the equation for the US equity price index (US_Q). Then define the horizon for impulse response as an interval of 25 time periods. You can inspect the results that follow, or use option 2 in the following Impulse Response Results Menu to obtain a graphic display of these impulse responses. By using option 0 you can also compute the bootstrapped confidence intervals. Choose, for example, to inspect results for real output (Y) and then interest rate (R), and their 2.5 and 97.5 percentiles. The results are displayed in Figure 17.7 and 17.8, respectively. Notice that the shock is accompanied by a statistically significant increase in the UK real output. Conversely, the increase in the interest rate is not statistically significant, as is evident in Figure 17.8.

Use option 3 from the IR Analysis Menu to inspect the effect of a shock to the equation for US_Q on the cointegrating vectors. Using the same steps outlined above, choose to see the effect on $CV1$, which describes the IRP relation. Results are displayed in Figure 17.9. Notice that the response of the first cointegrating relation to a shock on US real equity prices is very small, becomes statistically significant after four quarters.

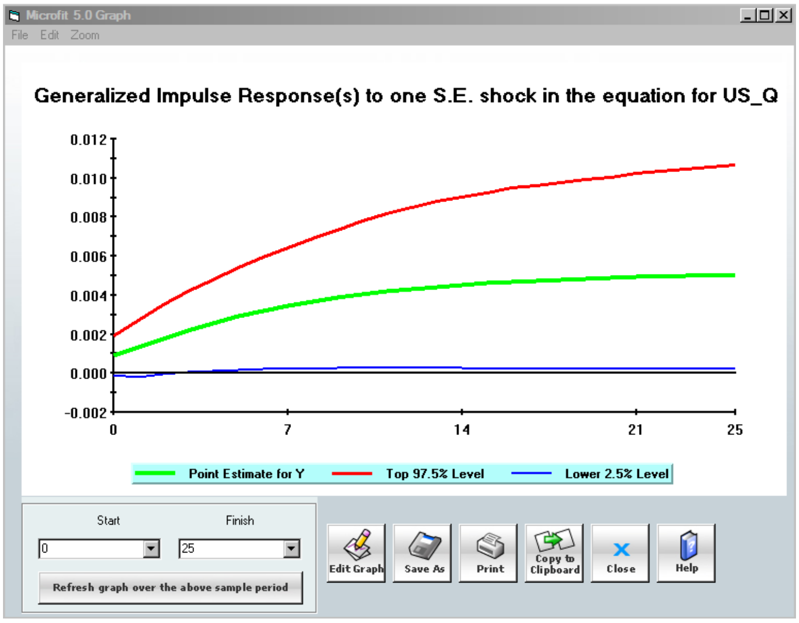


Figure 17.7: Generalized impulse response of real output to a shock in the equation for US_Q

17.8 Exercises in VARX modelling

Exercise 17.1

Load into *Microfit* the file UKCORE.FIT (see Lesson 17.1 for a description of the variables) and consider the following long-run relations:

$$\begin{aligned} PPP &: \quad P - E - PS \sim I(0) \\ OG &: \quad Y - YS \sim I(0) \end{aligned}$$

where Y and YS are domestic and foreign outputs. The first relation describes a purchasing power parity (PPP) relation, which predicts that in the long-run domestic and foreign prices measured in a common currency equilibrate. The second relation is an output gap (OG) relation, which states that production at home and abroad equilibrate in the long-run. Use option 3 from the Multivariate Menu to estimate a model with Y, YS, P, E and PS, where YS and PS are assumed to be weakly exogenous, over the period 1966q1-1999q4. Test the PPP and OG. Initially you need two restrictions on each of the two cointegrating vectors in order to just identify them. You should then impose the over-identifying restrictions (one by one and gradually) and test to see if these over-identifying restrictions are supported by the data.

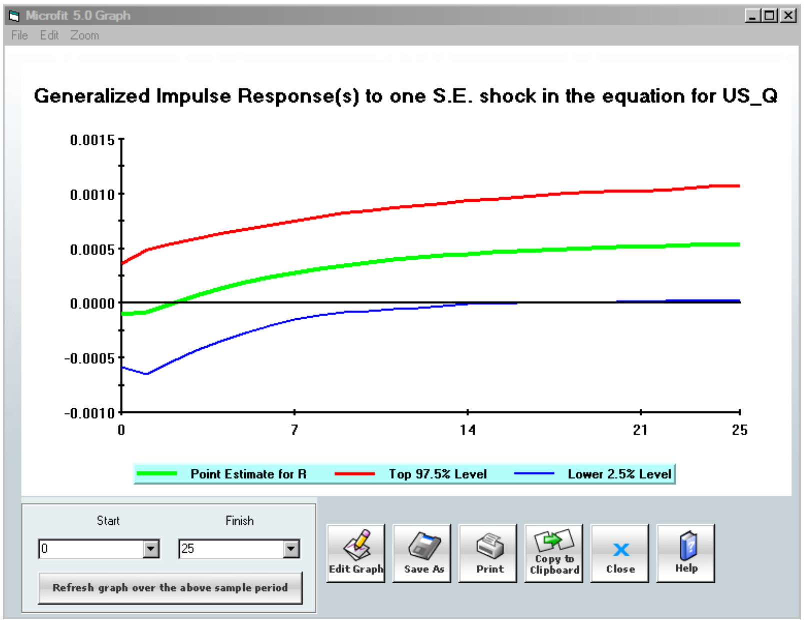


Figure 17.8: Generalized impulse response of inflation to a shock in the equation for US_Q

Exercise 17.2

Repeat the steps of Exercise 17.2 and then use the impulse response analysis to study the effect of a system-wide shock on the two cointegrating relations.

Exercise 17.3

Repeat Lesson 17.4, and compute conditional and unconditional forecasts for changes in foreign prices.

17.8.1 Exercise 17.4

Consider the effects of a unit shock to US interest rates on real output and inflation in the UK in the context of a VARX model, using the data analysed in Lesson 17.6 (17.7).

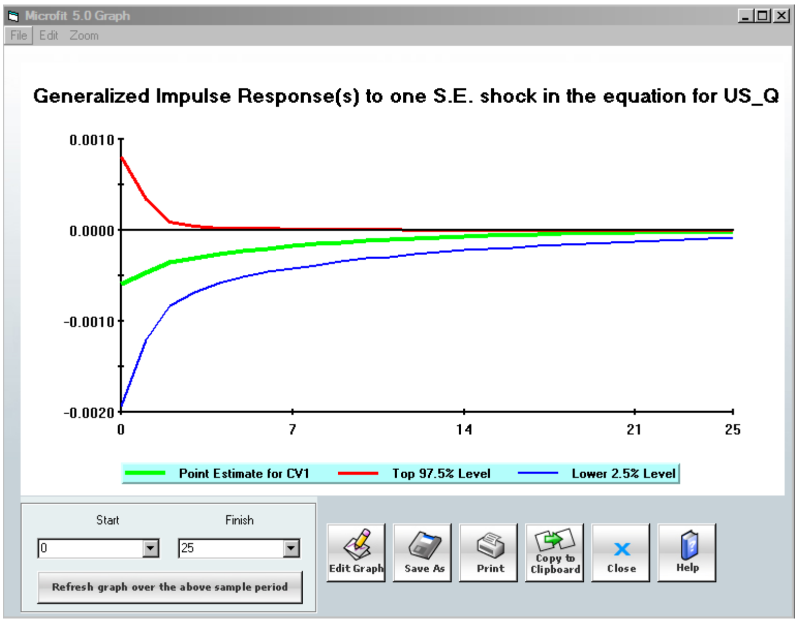


Figure 17.9: Generalized impulse response of $CV1$ to a shock in the equation for US_Q

Chapter 18

Lessons in SURE Estimation

The lessons in this chapter are concerned with the options in *Microfit* for the maximum likelihood (ML) estimation of Seemingly Unrelated Regression Equations ($SURE$) models. They can be used for estimation of VAR (or vector error correction) models subject to restrictions on their short-run parameters, or to estimate systems of equations. The restricted $SURE$ option can also be used to estimate systems of equations subject to general linear restrictions, including linear cross-equation restrictions. This option is particularly useful for estimation of systems of budget shares and factor demand share equations subject to homogeneity and/or symmetry restrictions. Another important use of the restricted $SURE$ option is in the area of pooling of time-series across a relatively small number of groups (countries, firms, and so on). For an account of the $SURE$ options and how they can be used see Sections 7.7.1 and 7.7.2. The econometrics that underlie the analysis of $SURE$ models and the numerical algorithms used to compute them can be found in Section 22.1.

18.1 Lesson 18.1: A restricted bivariate VAR model of patents and output growth in the US

In this lesson we show how to use the $SURE$ option to estimate restricted VAR models. For this purpose we consider the following bivariate model in the rates of change of US GDP output, denoted by $DLYUSA$ and patents ($DLQUSA$) granted to US firms by the US Patent Office:

$$\begin{aligned} DLYUSA_t &= a_1 + \sum_{j=1}^{2} b_{1j} DLYUSA_{t-j} + \sum_{j=9}^{10} c_{1j} DLQUSA_{t-j} + d_1 D74_t + u_{1t} \\ DLQUSA_t &= a_2 + \sum_{j=1}^{2} b_{2j} DLYUSA_{t-j} + \sum_{j=1}^{2} c_{2j} DLQUSA_{t-j} + u_{2t} \end{aligned}$$

where $D74$ is an oil shock dummy variable which takes the value of 1 in the four quarters of 1974 and zero elsewhere. This model assumes that changes in patents (intended to proxy technological innovations) only start to affect output growth after at least two years. It also assumes that only output growth was immediately affected by the oil crisis. This model is a restricted version of a $VAR(10)$ model in $DLYUSA_t$ and $DLQUSA_t$, and needs to be estimated using the $SURE$ option.

The relevant data is available in the special *Microfit* file G7GDP.FIT, and is described in detail in Chapter 15 in Lesson 15.1. Read this file into *Microfit,* and choose the unrestricted $SURE$ option 4, in the System Estimation Menu (Multivariate Menu: see Section 7.3). Specify the restricted VAR model set out above by typing

$$\begin{array}{lllll} dlyusa & const & dlyusa\{1-2\} & dlqusa\{9-10\} & d74; \\ dlqusa & const & dlyusa\{1-2\} & dlqusa\{1-2\} & \end{array}$$

Choose the whole sample for estimation, and click  to carry out the computations. You will be presented with the Post System Estimation Menu. To see the estimates of the output growth equation choose option 1 and then select the variable $DLYUSA$. You should obtain the results in Table 18.1.

Table 18.1: Relationship between US output growth and patents

```
                    Seemingly Unrelated Regressions Estimation
                      The estimation method converged after 2 iterations
*******************************************************************************
 Dependent variable is DLYUSA
 113 observations used for estimation from 1965Q4 to 1993Q4
*******************************************************************************
 Regressor               Coefficient        Standard Error          T-Ratio[Prob]
 CONST                     .0045452            .0011482              3.9584[.000]
 DLYUSA(-1)                  .22006             .093830              2.3453[.021]
 DLYUSA(-2)                  .12892             .093340              1.3812[.170]
 DLQUSA(-9)                 .011251            .0061970              1.8155[.072]
 DLQUSA(-10)                .012662            .0062412              2.0288[.045]
 D74                     -.0094370            .0046040             -2.0497[.043]
*******************************************************************************
 R-Squared                       .17183   R-Bar-Squared                    .13313
 S.E. of Regression            .0088589   F-Stat.    F(5,107)       4.4400[.001]
 Mean of Dependent Variable    .0065667   S.D. of Dependent Variable    .0095149
 Residual Sum of Squares       .0083974   Equation Log-likelihood       376.8179
 DW-statistic                    2.0116   System Log-likelihood         450.6412
 System AIC                    439.6412   System SBC                    424.6406
*******************************************************************************
```

The lagged patent variables are marginally significant. A test of the joint hypothesis that $c_{1,9} = 0$ and $c_{1,10} = 0$ can be carried out using option 3 in the Post System Estimation Menu. Click  and then  twice to reach this menu, choose option 3, and then type

$$A4 = 0; \quad A5 = 0$$ 

The test results should appear on the screen. The Wald statistic for testing the joint hypothesis that lagged patent growths have no impact on output growth is rejected at 5 per cent level. Other hypothesis of interest, such as Granger non-causality of output growth, can also be tested using option 4 in the Post System Estimation Menu.

18.2 Lesson 18.2: Estimation of Grunfeld-Griliches investment equations

In an important study of investment demand, Grunfeld (1960) and Grunfeld and Griliches (1960) estimated investment equations for ten firms in the US economy over the period 1935-1954. In this Lesson we estimate investment equations for five of these firms by the $SURE$ method. This smaller data set is also analysed in Greene (2002). The *Microfit* file GGSURE.FIT contains annual observations over the period 1935-1954 (inclusive) on the variables

I_{it}	Gross investment
F_{it}	Market value of the firm at the end of the previous year
C_{it}	Value of the stock of plant and equipment at the end of the previous year

The five firms indexed by i are General Motors (GM), Chrysler (CH), General Electric (GE), Westinghouse (WE) and US Steel (USS). In the file, these variables are denoted by adding the prefixes GM, CH, WE and USS to the variable names. For example, GMI refers to General Motors' gross investment, and WEF to the market value of Westinghouse.

The $SURE$ model to be estimated is given by

$$I_{it} = \beta_{i1} + \beta_{i2}F_{it} + \beta_{i3}C_{it} + u_{it} \tag{18.1}$$

for $i = GM, CH, GE, WE$, and USS, and $t = 1935, 1936, ..., 1954$. Read the file GGSURE.FIT and choose option 4 (unrestricted $SURE$ method) in the System Estimation Menu (see Section 7.3). You will now be asked to specify all the equations in the $SURE$ model, separating them by semicolons. In the box editor type

$$\begin{array}{llll}
GMI & const & gmf & gmc; \\
CHI & const & chf & chc; \\
GEI & const & gef & gec; \\
WEI & const & wef & wec; \\
USSI & const & ussf & ussc
\end{array}$$

Notice that upper- and lower-case letters have the same effects in *Microfit*, and here we have used upper case letters for the left-hand-side variables simply for expositional convenience. Also it is often much simpler to specify the different equations in the model one after another without starting on a new line. For an example, see Figure 18.1 (the equations are saved in the file GGSURE.LST, which can be retrieved using the 

button).

Estimate the model over the whole sample period. The program carries out the computations and then presents you with the Post System Estimation Menu. You can use the various options in this menu to see the $SURE$ estimates, test restrictions on the coefficients, and compute forecasts. For example, if you wish to see the results for Chrysler, choose Option 1 and, when prompted, select the variable CHI. The results in Table 18.2 should now appear on the screen.

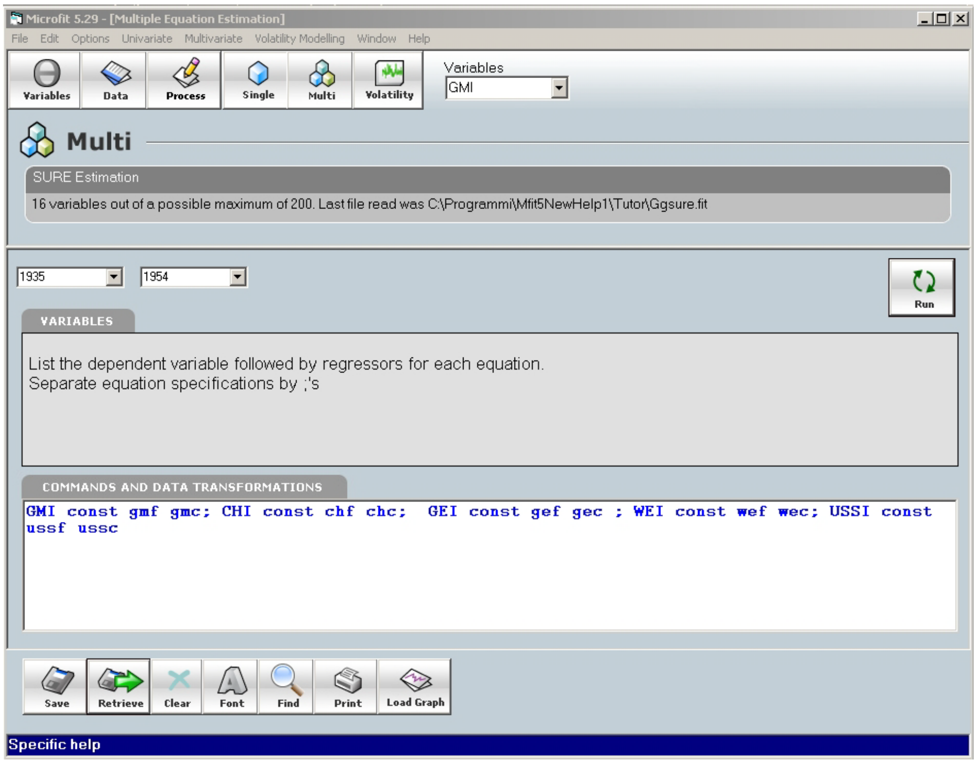


Figure 18.1: *SURE* estimation

Table 18.2: *SURE* estimates of the investiment equation for the Chrysler company

```
                    Seemingly Unrelated Regressions Estimation
                     The estimation method converged after 18 iterations
*******************************************************************************
 Dependent variable is CHI
 20 observations used for estimation from 1935 to 1954
*******************************************************************************
 Regressor              Coefficient        Standard Error         T-Ratio[Prob]
 CONST                       2.3783               12.6160           .18852[.853]
 CHF                        .067451               .018550           3.6362[.002]
 CHC                         .30507               .028274          10.7898[.000]
*******************************************************************************
 R-Squared                     .91057   R-Bar-Squared                    .90004
 S.E. of Regression           13.5081   F-Stat.    F(2,17)      86.5413[.000]
 Mean of Dependent Variable   86.1235   S.D. of Dependent Variable      42.7256
 Residual Sum of Squares       3102.0   Equation Log-likelihood        -78.8193
 DW-statistic                  1.8851   System Log-likelihood         -459.0922
 System AIC                 -474.0922   System SBC                    -481.5602
*******************************************************************************
```

Except for the intercept term, the results in this table are comparable with the *SURE* estimates for the same equations reported in Table 14.3 in Greene (2002). Similarly, the estimates for the other investment equations in the model can be inspected.

Since the $SURE$ estimation is appropriate under a non-diagonal error covariance matrix, it may now be of interest to test this hypothesis. For this purpose we need to estimate all the five individual equations separately by the OLS method, and then employ the log-likelihood ratio procedure discussed in Section 22.2.2 (see also Lesson 15.4). The maximized log-likelihood values for the five equations estimated separately are for General Motors (−117.1418), Chrysler (−78.4766), General Electric (−93.3137), Westinghouse (−73.2271) and US Steel (−119.3128), respectively, yielding the restricted log-likelihood value of

$$-481.472\,(= -117.1418 - 78.4766 - 93.3137 - 73.2271 - 119.3128)$$

The maximized log-likelihood value for the unrestricted system (namely, when the error covariance matrix is *not* restricted) is given at the bottom right-hand corner of Table 18.2, under 'System Log-likelihood' (= −459.0922). Therefore, the log-likelihood ratio statistic for testing the diagonality of the error covariance matrix is given by $LR = 2(-459.0922 + 481.472) = 44.76$, which is asymptotically distributed as a chi-squared variate with $5\,(5-1)\,/2 = 10$ degrees of freedom. The 95 per cent critical value of the chi-squared distribution with 10 degrees of freedom is 19.31. Hence, we reject the hypothesis that the error covariance matrix of the five investment equations is diagonal, which provides support for the application of the $SURE$ technique to this problem. To see the magnitude of the off-diagonal elements of the estimated error covariance matrix you need to choose option 2 in the Post System Estimation Menu, by which you should obtain the results in Table 18.3. As you can see, the covariance estimates on the off-diagonal elements are quite large relative to the respective diagonal elements.

Table 18.3: Estimated system covariance matrix of errors for Grunfeld-Griliches investment equations

	GMI	CHI	GEI	WEI	USSI
GMI	8600.8	-389.2322	644.4398	139.1155	-3394.4
CHI	-389.2322	182.4680	13.6558	22.1336	544.8885
GEI	644.4398	13.6558	873.1736	259.9662	1663.1
WEI	139.1155	22.1336	259.9662	121.7357	868.3544
USSI	-3394.4	544.8885	1663.1	868.3544	11401.0

18.3 Lesson 18.3: Testing cross-equation restrictions after SURE estimation

In the previous lesson we estimated investment equations for five US firms, and found that the $SURE$ procedure was an appropriate estimation method to apply. Suppose you are now interested in testing the hypothesis that the coefficients of F_{it}, the market value of the firms, are the same across all the five companies. In terms of the coefficients of the equations in (18.1), the relevant null hypothesis is

$$H_0 : \beta_{i2} = \beta_2, \quad \text{for} \quad i = GM,\ CH,\ GE,\ WE,\ USS$$

These four restrictions clearly involve coefficients from all the five equations. To implement this test, read the file GGSURE.FIT and then choose the Unrestricted *SURE* option (option 4) in the System Estimation Menu (Multivariate Menu). Retrieve the file GGSURE.LST into the box editor on the screen by clicking Retrieve. Estimate the equations over the whole sample period and choose option 3 in the Post System Estimation Menu. Type the following four restrictions in the box editor that appears on the screen:

$$A2 = B2; \quad B2 = C2; \quad C2 = D2; \quad D2 = E2$$ RUN

You should now obtain the test results in Table 18.4. The LR statistic for testing these restrictions is 20.46 which is well above the 95 per cent critical value of the chi-squared distribution with 4 degrees of freedom, and we therefore strongly reject the slope homogeneity hypothesis.

Table 18.4: Testing the slope homogeneity hypothesis

```
                Wald test of restriction(s) imposed on parameters
*******************************************************************************
 The underlying estimated model is:
 GMI const gmf gmc; CHI const chf chc;  GEI const gef gec ; WEI const wef wec;
 USSI const ussf ussc
 20 observations used for estimation from 1935 to 1954
*******************************************************************************

 List of restriction(s) for the Wald test:
 A2=B2;  B2=C2;  C2=D2;  D2=E2
*******************************************************************************
 Wald Statistic                  CHSQ(4)=  20.4580[.000]
*******************************************************************************
```

18.4 Lesson 18.4: Estimation of a static almost ideal demand system

In this lesson we show how to use the restricted *SURE* option in *Microfit* to estimate a system of demand equations subject to homogeneity and symmetry restrictions. For this purpose we consider the static version of the almost ideal demand model of Deaton and Muellbauer (1980), which postulates that the expenditure share of the ith commodity group, w_{it}, $i = 1, 2, ..., m$ is determined by

$$w_{it} = \alpha_i + \sum_{j=1}^{m} \gamma_{ij} \log P_{jt} + \delta_i \log (Y_t/P_t) + u_{it} \tag{18.2}$$

where P_{jt} is the price deflator of the commodity group j, Y_t is the *per capita* expenditure on all the goods, and P_t is the general price index, approximated using the Stone formula:

$$\log P_t = \sum_{j=1}^{m} w_{j0} \log P_{jt} \tag{18.3}$$

The weights w_{j0} refer to the budget shares in the base year, which is taken to be 1990. Consumer theory imposes the following restrictions on the parameters of the share equation:

$$\textit{Homogeneity restrictions}: \quad \sum_{j=1}^{m} \gamma_{ij} = 0, \qquad i = 1, 2, ..., m \tag{18.4}$$

$$\textit{Symmetry restrictions}: \quad \gamma_{ij} = \gamma_{ji}, \qquad \text{for all} \qquad i, j \tag{18.5}$$

We also have $\sum_{i=1}^{m} w_{it} = 1$, and hence only $m-1$ of the shares can be independently explained/estimated.

The share equations in (18.2) are unlikely to hold in each and every period, and are best incorporated within a dynamic framework which freely allows for *short-run* departures from the 'equilibrium' budget share equations given by (18.2). Such an exercise is carried out in Pesaran and Shin (2002). Here we consider the above static formulation both to demonstrate how the restricted *SURE* option in *Microfit* can be used, and also to highlight the importance of appropriately allowing for dynamics in estimation of demand equations.

We estimate a three-commodity system on the UK quarterly seasonally adjusted data over the period 1956(1)-1993(2). The three commodity groups are

(1) Food, drink and tobacco
(2) Services (including rents and rates)
(3) Energy and other nondurable

The relevant data are in the special *Microfit* file CONG3.FIT, containing the following variables:[1]

$INPT$	Intercept term
$LP1$	Log of implicit price deflator for the Food, Drink and Tobacco group
$LP2$	Log of implicit price deflator for the Services group
$LP3$	Log of implicit price deflator for the Energy group
LRY	Log of *per capita* real total expenditure
$W1$	Share of expenditure on Food, Drink and Tobacco group
$W2$	Share of expenditure on the Services group
$W3$	Share of expenditure on the Energy group

Load this file, and move to the System Estimation Menu. Initially choose option 4 to estimate the unrestricted version of the equations in (18.2), and type

$$\begin{array}{llllll} W1 & INPT & LP1 & LP2 & LP3 & LRY; \\ W2 & INPT & LP1 & LP2 & LP3 & LRY \end{array}$$

[1]For the source of this data set and other details see Pesaran and Shin (2002).

Notice that we have decided to work with the budget shares $W1$ and $W2$, but we could have equally considered any one of the pairs ($W2$, $W3$) or ($W1$, $W3$). The results are invariant to this choice.

For the estimation period type

$$1956Q1 \quad 1993Q2$$

Click . You can now inspect the estimation results by choosing option 1 in the Post System Estimation Menu. The results seem quite satisfactory, except for the low value of the Durbin-Watson statistics obtained for both of the share equations, thus highlighting the importance of the missing dynamics.

To test the homogeneity restrictions given in (18.4), return to the Post System Estimation Menu, choose option 3, and type

$$A2 + A3 + A4 = 0;$$
$$B2 + B3 + B4 = 0$$

The Wald statistic, 4.0755, for testing these restrictions should now appear on the screen, and suggests that the homogeneity restrictions cannot be rejected.

We now re-estimate the share equations subject to the homogeneity restrictions. For this purpose, return to the System Estimation Menu, and choose the restricted $SURE$ option 5. Click to accept the (unrestricted) specification of the share equations, still retained in the box editor from previously. You will now be asked to specify your parameter restrictions. Type

$$A2 + A3 + A4 = 0;$$
$$B2 + B3 + B4 = 0$$

The program now carries out the necessary computations and presents you with the restricted $SURE$ estimates. For example, to see the estimates for the $W1$ equation (the share of Food, Drink and Tobacco) choose option 1 in the Post System Estimation Menu, and select $W1$. The results in Table 18.5 should appear on the screen.

Table 18.5: *SURE* estimates of the budget share equation for food subject to the homogeneity restrictions

```
                    Restricted Seemingly Unrelated Regressions
                     The estimation method converged after 0 iterations
*******************************************************************************
 Dependent variable is W1
 150 observations used for estimation from 1956Q1 to 1993Q2
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                       .29913            .0013826          216.3456[.000]
 LP1                        .24282             .010707           22.6792[.000]
 LP2                       -.13966             .012037          -11.6027[.000]
 LP3                       -.10315             .012448           -8.2866[.000]
 LRY                       -.23062            .0097290          -23.7049[.000]
*******************************************************************************
 R-Squared                          .99391   R-Bar-Squared                   .99374
 S.E. of Regression               .0059939   F-Stat.    F(4,145)       5915.0[.000]
 Mean of Dependent Variable         .41741   S.D. of Dependent Variable     .075761
 Residual Sum of Squares          .0052093   Equation Log-likelihood       557.2548
 DW-statistic                       .41831   System Log-likelihood           1140.5
 System AIC                         1132.5   System SBC                      1120.5
*******************************************************************************
```

Notice that the sum of the estimated coefficients on the price variables do in fact add up to zero, as required by the homogeneity hypothesis.

We now consider testing the symmetry restriction. In the present case where $m = 2$, there is only one restriction implied by the symmetry hypothesis, namely $\gamma_{12} = \gamma_{21}$. Press the Esc key twice to return to the Post System Estimation Menu, and choose the Wald test option 3. Click Clear to clear the box editor, and type

$$A3 = B2$$ RUN

to obtain the test results reproduced in Table 18.6.

Table 18.6: Wald test of the symmetry restriction

```
                Wald test of restriction(s) imposed on parameters
*******************************************************************************
 The underlying estimated restricted model is:
 W1  INPT  LP1  LP2  LP3  LRY;W2  INPT  LP1  LP2  LP3  LRY
 List of restrictions imposed on the model:
 150 observations used for estimation from 1956Q1 to 1993Q2
*******************************************************************************

 List of restriction(s) for the Wald test:
 A3=B2
*******************************************************************************
 Wald Statistic                 CHSQ(1)=  21.9498[.000]
*******************************************************************************
```

The Wald statistic for testing the symmetry restriction is equal to 21.95, which is well in excess of 3.84, the 95 per cent critical value of the chi-squared distribution with one degree of freedom. Therefore, in the present static formulation the symmetry hypothesis is decisively rejected. But in view of the very low DW statistics of the estimated equations and the problem of dynamic mis-specifications that surround these share equations, this may not be a valid conclusion. In fact, using the long-run structural modelling approach where the share-equations are embodied within an unrestricted VAR model, Pesaran and Shin (2002) cannot reject the homogeneity or the symmetry restrictions using the same data set.

18.5 Lesson 18.5: Estimation of a New Keynesian three equation model

In this exercise we illustrate the estimation of a simple three-equation New Keynesian (NK) macroeconomic model for the US economy by restricted three-stage least squares (see Section 22.2). The model explains the dynamics of three variables: the rate of inflation, DP_t; the output gap, Y_C_t; and the nominal short-term rate of interest, R_t. The output gap is measured as the deviation of GDP (in logs) from its long-horizon forecast, calculated using the Cointegrating VARX option from the Multivariate Menu. Specifically, the NK model is defined by

$$
\begin{aligned}
DP_t &= a_1 + a_2 Y\_C_t + a_3 DP_{t-1} + a_4 DPS_t + \varepsilon_{1t} && (18.6)\\
Y\_C_t &= b_1 + b_2 DP_t + b_3 R_{t-1} + b_4 Y\_C_{t-1} + \varepsilon_{2t} && (18.7)\\
R_t &= c_1 + c_2 DP_t + c_3 Y\_C_t + c_4 R_{t-1} + \varepsilon_{3t} && (18.8)
\end{aligned}
$$

where DPS_t is the foreign inflation. The above equations describe a Phillips Curve, an IS or aggregate demand curve, and a Taylor Rule for monetary policy, respectively. The system is augmented by foreign inflation to act as a cost shock in the Phillips Curve. Notice that in a perfect foresight economy the ex-ante real interest rate in period $t-1$ is $(r_{t-1} - \pi_t)$ so that we would expect $b_2 = -b_3$. The relevant data are in the special *Microfit* file 3EQMODUS.FIT. Load this file, and move to the System Estimation Menu (Multivariate Menu). Choose option 9 to estimate by restricted three stage least squares system (18.6)-(18.8), and type

$$
\begin{aligned}
&DP \quad INPT \quad Y\_C \quad DP(-1) \quad DPS;\\
&Y\_C \quad INPT \quad DP \quad R(-1) \quad Y\_C(-1);\\
&R \quad INPT \quad DP \quad Y\_C \quad R(-1);
\end{aligned}
$$

Alternatively, you can retrieve the file 3EQM.EQU from the tutorial directory. With this option you do not need to specify the instruments, as *Microfit* automatically identifies them from the system, and in this case are $INPT$, DPS_t, DP_{t-1}, Y_C_{t-1}, and R_{t-1}.[2] Notice that the system is over-identified. A window appears in which you are requested to impose restrictions on the parameters. In the box editor type

$$B2 + B3 = 0$$

[2] *Microfit* sets as exogenous all variables that do not appear on the left hand side of equation system (18.6)-(18.8).

Now select option 1 from the Post System Estimation Menu, and choose to display results from estimation for each equation in turn. The output for the three equations is reproduced in Table 18.7.

Table 18.7: Three-stage least squares estimation of the New Keynesian model

```
                    Restricted Three Stage Least Squares Estimation
                         The estimation method converged after 9 iterations
*******************************************************************************
 Dependent variable is DP
 96 observations used for estimation from 1980Q1 to 2003Q4
*******************************************************************************
 Regressor               Coefficient        Standard Error         T-Ratio[Prob]
 INPT                      .0038861            .0017837            2.1787[.032]
 Y_C                        .065764             .031743            2.0717[.041]
 DP(-1)                      .59532             .072680            8.1911[.000]
 DPS                         .11781             .033618            3.5043[.001]
*******************************************************************************
 R-Squared                      .63646   R-Bar-Squared                   .62461
 S.E. of Regression           .0041881   F-Stat.    F(3,92)      53.6893[.000]
 Mean of Dependent Variable   .0092465   S.D. of Dependent Variable    .0068356
 Residual Sum of Squares      .0016137   Equation Log-likelihood       391.4729
 DW-statistic                   2.1070   System Log-likelihood           1200.6
 System AIC                     1189.6   System SBC                      1175.5
*******************************************************************************
*******************************************************************************
 Dependent variable is Y_C
 96 observations used for estimation from 1980Q1 to 2003Q4
*******************************************************************************
 Regressor               Coefficient        Standard Error         T-Ratio[Prob]
 INPT                     -.0075227            .0021569           -3.4877[.001]
 DP                          .52934              .16347            3.2382[.002]
 R(-1)                      -.52934              .16347           -3.2382[.002]
 Y_C(-1)                     .78019             .044362           17.5868[.000]
*******************************************************************************
 R-Squared                      .83670   R-Bar-Squared                   .83319
 S.E. of Regression           .0074022   F-Stat.    F(3,92)     158.8394[.000]
 Mean of Dependent Variable  -.050461    S.D. of Dependent Variable     .018124
 Residual Sum of Squares      .0050958   Equation Log-likelihood       336.2792
 DW-statistic                   1.7599   System Log-likelihood           1200.6
 System AIC                     1189.6   System SBC                      1175.5
*******************************************************************************
*******************************************************************************
 Dependent variable is R
 96 observations used for estimation from 1980Q1 to 2003Q4
*******************************************************************************
 Regressor               Coefficient        Standard Error         T-Ratio[Prob]
 INPT                     -.0017055            .0012166           -1.4018[.164]
 DP                          .21908              .12044            1.8189[.072]
 Y_C                       -.040554             .023025           -1.7613[.082]
 R(-1)                       .83988             .074834           11.2232[.000]
*******************************************************************************
 R-Squared                      .92600   R-Bar-Squared                   .92358
 S.E. of Regression           .0023396   F-Stat.    F(3,92)     383.7202[.000]
 Mean of Dependent Variable    .016384   S.D. of Dependent Variable    .0084635
 Residual Sum of Squares      .5036E-3   Equation Log-likelihood       447.3698
 DW-statistic                   1.7817   System Log-likelihood           1200.6
 System AIC                     1189.6   System SBC                      1175.5
*******************************************************************************
```

Notice that in the equations for DP, and for R and all estimated parameters, are statistically significant and have the expected signs. In the Y_C equation, the parameter estimate attached to $Y\_C(-1)$ is also significant with the correct sign.

18.6 Lesson 18.6: 2SLS and 3SLS estimation of an exactly identified system

In this application we estimate an exactly identified system of equations by two-stage least squares and three-stage least squares. Exact identification arises when, in the equation to be estimated, the number of excluded exogenous variables is exactly equal to the number of included endogenous variables minus 1. (see Section 22.2 for more details). Consider the following system of equations:

$$\begin{aligned}
DP_t &= a_1 + a_2 Y_t + DPS_t + \varepsilon_{1t} \\
Y_t &= b_1 + b_2 DP_t + b_3 R_{t-1} + \varepsilon_{2t} \\
R_t &= c_1 + c_2 DP_t + c_3 Y_t + \varepsilon_{3t}
\end{aligned}$$

where DP_t and DPS_t denote domestic and foreign inflation respectively, Y_t is the real output, and R_t is the nominal interest rate. The exogenous/predetermined variables are DPS_t, R_{t-1} and the unit vector (the intercept). The endogenous variables are identified by *Microfit* as the left-hand-side variables, namely DP_t, Y_t and R_t. Notice that all the equations in this system are exactly identified. For example, the DP_t equation has two endogenous variables with one excluded predetermined variable, R_{t-1}, and so is exactly identified. The same applies to the other two equations. The R_t equation has three endogenous variables and two excluded exogenous/pre-determined variables, and hence it is also exactly identified since the number of included endogenous variables minus 1 is equal to the number of excluded exogenous/pre-determined variables.

The relevant data are in the special *Microfit* file UKM07.FIT (see Lesson 18.2). Load this file into *Microfit* and in the Process window create an intercept ($INPT$). Select option 6 from the Multivariate Menu, and in the Commands and Data Transformation box type

$$\begin{array}{llll}
DP & INPT & Y & DPS; \\
Y & INPT & DP & R(-1); \\
R & INPT & DP & Y;
\end{array}$$

Table 18.8 reports results on the estimation of each individual equation. Make the necessary steps to return to the Process and Data Transformation box, select option 8 (3SLS), and in the Commands and Data Transformation box type the same system specification as above. Inspect each equation in turn and verify that the output is identical to that reported in Table 18.8. Clearly, in the exact identified case there is no gain in efficiency by using the 3SLS method, and the two-stage least squares method can be used.

Table 18.8: Two-stage least squares estimation

```
                         Two Stage Least Squares Estimation
*******************************************************************************
 Dependent variable is DP
 98 observations used for estimation from 1979Q3 to 2003Q4
*******************************************************************************
 Regressor               Coefficient       Standard Error         T-Ratio[Prob]
 INPT                        .014240            .088994            .16001[.873]
 Y                         -.0036804            .019209           -.19160[.848]
 DPS                          1.0344             .38694            2.6733[.009]
*******************************************************************************
 R-Squared                       .44090   R-Bar-Squared                  .42913
 S.E. of Regression             .0093709  F-Stat.    F(2,95)     37.4582[.000]
 Mean of Dependent Variable      .012282  S.D. of Dependent Variable     .012403
 Residual Sum of Squares        .0083423  Equation Log-likelihood       320.1419
 DW-statistic                     2.0217  System Log-likelihood         650.2572
 System AIC                     641.2572  System SBC                    629.6249
*******************************************************************************
*******************************************************************************
 Dependent variable is Y
 98 observations used for estimation from 1979Q3 to 2003Q4
*******************************************************************************
 Regressor               Coefficient       Standard Error         T-Ratio[Prob]
 INPT                         4.6867            .043654          107.3618[.000]
 DP                         -11.3843             2.7216           -4.1830[.000]
 R(-1)                       -8.2205             2.7351           -3.0055[.003]
*******************************************************************************
 R-Squared                       .28741   R-Bar-Squared                  .27240
 S.E. of Regression              .15198   F-Stat.    F(2,95)     19.1578[.000]
 Mean of Dependent Variable      4.3717   S.D. of Dependent Variable      .17817
 Residual Sum of Squares         2.1943   Equation Log-likelihood        47.0993
 DW-statistic                    1.0415   System Log-likelihood         650.2572
 System AIC                     641.2572  System SBC                    629.6249
*******************************************************************************
*******************************************************************************
 Dependent variable is R
 98 observations used for estimation from 1979Q3 to 2003Q4
*******************************************************************************
 Regressor               Coefficient       Standard Error         T-Ratio[Prob]
 INPT                         .45303             .13740            3.2971[.001]
 DP                          -.88365             .54635           -1.6174[.109]
 Y                          -.096289            .029964           -3.2135[.002]
*******************************************************************************
 R-Squared                      -1.6352   R-Bar-Squared                 -1.6907
 S.E. of Regression              .013687  F-Stat.    F(2,95)            *NONE*
 Mean of Dependent Variable      .021229  S.D. of Dependent Variable    .0083440
 Residual Sum of Squares         .017797  Equation Log-likelihood       283.0161
 DW-statistic                     .75009  System Log-likelihood         650.2572
 System AIC                     641.2572  System SBC                    629.6249
*******************************************************************************
```

18.7 Exercises in SURE Estimation

18.7.1 Exercise 18.1

In Lesson 18.1 we assumed that growth of patents only influences output growth with a lag of two years. Estimate similar bivariate models for Japan and Germany, and comment on your results. The relevant data can be found in the special *Microfit* file G7GDP.FIT.

18.7.2 Exercise 18.2

Test the hypothesis that the error covariance matrix of the restricted bivariate model estimated in Lesson 18.1 is diagonal.

18.7.3 Exercise 18.3

Re-estimate the Grunfeld-Griliches investment equations using a log-linear specification. Compute the values of Akaike and Schwarz criteria to discriminate between the linear and the log-linear specifications.

18.7.4 Exercise 18.4

Estimate a simple dynamic version of the Grunfeld-Griliches investment equations by including a first-order lag of the investments variable among the regressors.

18.7.5 Exercise 18.5

Repeat the estimation and testing exercises in Lesson 18.4 with ($W1$, $W3$) and ($W2$, $W3$) as the budget shares to be explained. Check that you do in fact obtain identical estimation results and make the same inferences.

18.7.6 Exercise 18.6

Use the data in CONG3.FIT to estimate the following dynamic version of the share equations given by (18.2) over the period 1956(1)-1993(2):

$$w_{1t} = \alpha_1 + \lambda_{11} w_{1,t-1} + \lambda_{12} w_{2,t-1} + \sum_{j=1}^{3} \gamma_{1j} \log P_{jt} + \delta_1 \log (Y_t/P_t) + u_{1t}$$

$$w_{2t} = \alpha_2 + \lambda_{21} w_{1,t-1} + \lambda_{22} w_{2,t-1} + \sum_{j=1}^{3} \gamma_{2j} \log P_{jt} + \delta_2 \log (Y_t/P_t) + u_{2t}$$

Derive the parameter restrictions for testing the *long-run* homogeneity and the symmetry restrictions. Test the validity of these restrictions and compare your results with those obtained in Lesson 18.4.

Chapter 19

Lessons in Univariate GARCH Modelling

Lessons in this section show how to estimate linear regression models under a variety of specifications for conditional error variances. The relevant estimation options, the underlying econometric methods and the computational algorithms are described in Sections 8.6 and 23.1. For a comprehensive review of the application of the $GARCH$ modelling to financial data see Pagan (1996).

19.1 Lesson 19.1: Testing for ARCH effects in monthly \$/£ exchange rates

In this lesson we consider the following $AR(1)$ model for the rate of change of the \$/£ exchange rate:

$$\Delta \log(USD_t) = \beta_0 + \beta_1 \Delta \log(USD_{t-1}) + u_t \tag{19.1}$$

and test for $ARCH$ effects in the conditional variance of u_t, namely $h_t^2 = Var(u_t \,|\Omega_{t-1})$. The $ARCH(q)$ specification for h_t^2 is given by

$$h_t^2 = \rho_0 + \rho_1 u_{t-1}^2 + \rho_2 u_{t-2}^2 + \cdots + \rho_q u_{t-q}^2 \tag{19.2}$$

The null hypothesis of 'no $ARCH$ effect', is

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_q = 0$$

to be tested against the alternative hypothesis that

$$H_0 : \rho_1 \neq 0, \quad \rho_2 \neq 0, \ \ldots, \quad \rho_q \neq 0$$

The test involves running a regression of squared OLS residuals from the regression (19.1) on lagged squared residuals (see Section 23.1.7). In what follows we apply the test to monthly observations on the US Dollar/Sterling rate available in the special *Microfit* file

EXMONTH.FIT. This file contains monthly observations on the following exchange rates over the period 1973(1)-1995(6)

CAN	Canadian Dollar to one British sterling
DM	German Deutschmark to one British sterling
EP	Spanish peseta to one British sterling
FF	French franc to one British sterling
ITL	Italian lira to one British sterling
SF	Swiss franc to one British sterling
USD	US dollar to one British sterling
YEN	Japanese yen to one British sterling

Read this file into *Microfit*, and in the Commands and Data Transformations window of the Process window type

$$DLUSD = \mathbf{LOG}(USD/USD(-1)); \quad ONE = 1$$ 

to create the rate of change of \$/£ in the variable $DLUSD$. ONE is a vector of ones. To specify the regression equation (19.1) choose option 1 from the Single Equation Estimation Menu (Univariate Menu), selecting the OLS option. Type

$$DLUSD \quad ONE \quad DLUSD(-1)$$

For the estimation period enter

$$1973M1 \quad 1994M12$$

thus keeping the observations for the first six months of 1995 for volatility predictions (see Lessons 19.4). Click 

, and when the results appear click 

to move to the Post Estimation Menu. Choose option 2 to move to the Hypothesis Testing Menu (see Section 6.23) and choose option 2 to carry out the $ARCH$ test. You will be asked to specify the order (from 1 to 12) of the test. Type

$$1$$ 

The test results should now appear on the screen. The LM version of the test yields a statistic of 8.65, which is well above the 95 per cent critical value of χ_1^2, and hence rejects the hypothesis that there are no $ARCH$ effects in (19.1): see Table 19.1. The same conclusion is reached if one considers the F-version of the test. Tests of higher-order $ARCH$ effects also yield similar results, although as the order of the test increases the power of the test tends to decline. For example, the value of the $ARCH(12)$ test statistic is 11.89, which is well below the 95 per cent critical value of χ_{12}^2, and does not reject the hypothesis that $\rho_1 = \rho_2 = \cdots = \rho_{12} = 0$. In practice, it is prudent to carry out the test for different orders and then make an overall judgement.

Table 19.1: Testing for the $ARCH$ effect in the monthly dollar/sterling rate

```
Autoregressive Conditional Heteroscedasticity Test of Residuals (OLS Case)
*******************************************************************
Dependent variable is DLUSD
List of the variables in the regression:
ONE          DLUSD(-1)
262 observations used for estimation from 1973M3 to 1994M12
*******************************************************************
Lagrange Multiplier Statistic      CHSQ(   1)    = 8.6475[.003]
F Statistic                        F(   1,   259) = 8.8403[.003]
*******************************************************************
```

As with all diagnostic tests, rejection of the null hypothesis that there are no $ARCH$ effects does not *necessarily* imply that conditional variance of $\Delta\log(USD)$ is variable. This can happen particularly if the disturbances u_t in (19.1) are serially correlated. But in the present application, using option 1 in the Hypothesis Testing Menu, the hypothesis that u_ts are serially uncorrelated cannot be rejected, and therefore there may well be important $ARCH$ effects in the data. We estimate various $ARCH$ type models of exchange rate changes in Section 19.2.

19.2 Lesson 19.2: Estimating GARCH models for monthly \$/£ exchange rate

In the previous lesson we found some evidence of an $ARCH$ effect in the rate of change of the \$/£ exchange rate, $\Delta\log(USD_t)$; see equation (19.1). Here we show how to use *Microfit* to estimate $GARCH$ models for h_t^2, the conditional variance of $\Delta\log(USD_t)$. We assume that the mean equation for the exchange rate changes follows the $AR(1)$ process,

$$x_t = \beta_0 + \beta_1 x_{t-1} + u_t$$

where $x_t = \Delta\log(USD_t)$, and u_t is a white noise process,

$$E(u_t) = 0;\ E(u_t u_{t-j}) = \left\{ \begin{array}{cc} \sigma^2 & \text{for } j = 0 \\ 0 & \text{for } j \neq 0 \end{array} \right\}$$

and $|\beta_1| < 1$. Therefore, x_t is a covariance-stationary process. In this case the optimal linear forecast of x_t is given by

$$E(x_t|x_{t-1}, x_{t-2}, ...) = \beta_0 + \beta_1 x_{t-1}$$

where $E(x_t|x_{t-1}, x_{t-2}, ...)$ denotes the linear projection of x_t on constant and the past observations $(x_{t-1}, x_{t-2}, ...)$. While the conditional mean of x_t changes over time, if the process is covariance-stationary, the *unconditional* mean of x_t is constant:

$$E(x_t) = \frac{\beta_0}{1-\beta_1}$$

We may be interested in forecasting not only the level of x_t, but also its variance. Although the unconditional variance of u_t is constant, σ^2, its conditional variance could be time-varying. A simple example of such a process is the autoregressive conditional heteroscedastic process of order 1, denoted $u_t \sim ARCH(1)$ or $GARCH(0,1)$:

$$V(x_t \,|\Omega_{t-1}) = V(u_t \,|\Omega_{t-1}) = h_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$$

Furthermore, if $\alpha_1 < 1$, then the unconditional variance of u_t is given by

$$\sigma^2 = \frac{\alpha_0}{1-\alpha_1}$$

which is constant. In general, if h_t^2 evolves according to

$$h_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + ... + \alpha_q u_{t-q}^2 + \phi_1 h_{t-1}^2 + \cdots + \phi_p h_{t-p}^2$$

we have the generalized autoregressive conditional heteroscedastic model, denoted by $u_t \sim GARCH(p,q)$. For further details and other models of volatility see Sections 8.6 and 23.1.

In this lesson we consider estimating the following $GARCH(1,1)$ models:

$$V(u_t \,|\Omega_{t-1}) = h_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \phi_1 h_{t-1}^2 \tag{19.3}$$

For a well-defined $GARCH(1,1)$ process we must have $\alpha_0 > 0$, $|\phi_1| < 1$, and $1 - \alpha_1 - \phi_1 > 0$. These restrictions ensure that the unconditional variance of u_t given by $V(u_t) = \alpha_0/(1 - \alpha_1 - \phi_1)$ is positive. In a number of econometric applications the coefficient α_1 in (19.3) is also assumed to be positive. *Microfit* does not impose this restriction in estimation of the model, but if a large negative estimate of α_1 is encountered, or if $\alpha_0/(1 - \alpha_1 - \phi_1)$ becomes negative in the course of the iterations, the program produces an error message.

To estimate the $GARCH(1,1)$ model (19.3) on the \$/£ exchange rate variable, $\Delta\log(USD_t)$, first follow the steps in Lesson 19.1, and specify the regression equation (19.1) to be estimated over the sample period 1973(1)-1994(12). Move to the Volatility Modelling Menu and choose option 1. In the $GARCH$ Estimation Menu select option 1. You are then asked to choose between a normal and a t-distribution for the conditional distribution of the errors, u_t. Choose the normal distribution option by choosing option 1 and clicking 

. In the box editor, type the orders of the $GARCH(p,q)$ model to be estimated as follows:

1; 1 

You will now be presented with another box editor to list any additional variables that you may wish to include in the model for the conditional variance, (19.3): see Section 8.6 for more details. In the present application there are no additional regressors in the $GARCH$ model, so click 

to move to the next screen, where you will be asked to provide initial estimates for the parameters of the $GARCH$ model, namely α_0, α_1 and ϕ_1, respectively. *Microfit* automatically suggests using the OLS estimate of the unconditional variance, $\hat{\sigma}^2$, as the initial estimate for α_0. You must, however, provide initial estimates for α_1 and ϕ_1. For α_1 (the 'MA lag 1' coefficient) type 0.1, and for ϕ_1 (the 'AR lag 1' coefficient) type 0.4

and click to accept these initial estimates. *Microfit* starts the computations, which converge after 36 iterations. The results reproduced in Table 19.2 should now appear on the screen.

Table 19.2: Modelling conditional heteroscedasticity of the dollar/sterling rate

```
                    GARCH(1,1) assuming a normal distribution

                                converged after 36 iterations
*******************************************************************************
 Dependent variable is DLUSD
 262 observations used for estimation from 1973M3 to 1994M12
*******************************************************************************
 Regressor                Coefficient       Standard Error         T-Ratio[Prob]
 ONE                       -.0010864          .0019866            -.54686[.585]
 DLUSD(-1)                   .073352           .068311             1.0738[.284]
*******************************************************************************
 R-Squared                    .0098739   R-Bar-Squared                  -.0016392
 S.E. of Regression            .033712   F-Stat.     F(3,258)      .85763[.464]
 Mean of Dependent Variable -.0016091   S.D. of Dependent Variable      .033684
 Residual Sum of Squares        .29321   Equation Log-likelihood       520.8315
 Akaike Info. Criterion       516.8315   Schwarz Bayesian Criterion    509.6948
 DW-statistic                   1.9360
*******************************************************************************

            Parameters of the Conditional Heteroscedastic Model
        Explaining H-SQ, the Conditional Variance of the Error Term
*******************************************************************************
                          Coefficient       Asymptotic Standard Error
        Constant           .6839E-3               .2590E-3
        E-SQ(-1)            .19664                 .11384
         H-SQ(-1)            .19941                 .26253
*******************************************************************************
 H-SQ stands for the conditional variance of the error term.
 E-SQ stands for the square of the error term.
```

The ML estimates $\hat{\alpha}_1 = 0.19664(0.1138)$ and $\hat{\phi}_1 = 0.19941(0.2625)$ have the correct signs, but neither are statistically significant at the 95 per cent level. The bracketed figures are asymptotic standard errors.

Microfit also allows you to test linear or non-linear restrictions on the coefficients of the $GARCH$ model. Suppose you are interested in testing the joint hypothesis that

$$H_0 : \alpha_1 = \phi_1 = 0$$

against

$$H_1 : \alpha_1 \neq 0, \ \phi_1 \neq 0,$$

Choose option 2 in the Post Regression Menu (after the $GARCH$ estimation), and then option 7 in the Hypothesis Testing Menu. Click and type

$$B2 = 0; \quad B3 = 0$$

The results in Table 19.3 should appear on the screen. The Wald statistic for testing the joint hypothesis that $\alpha_1 = \phi_1 = 0$ is equal to 4.94, which is significant at the 91.5 per cent level, but not at the conventional 95 per cent level.

Table 19.3: Testing joint restrictions on the parameters of the $GARCH$ model for the dollar/sterling rate

```
              Wald test of restriction(s) imposed on parameters
******************************************************************************
 Based on GARCH regression of DLUSD on:
 ONE             DLUSD(-1)
 262 observations used for estimation from 1973M3 to 1994M12
******************************************************************************
 Coefficients A1 to A2 are assigned to the above regressors respectively.
 Coeffs. B1 to B3 are assigned to ARCH parameters respectively.
 List of restriction(s) for the Wald test:
 B2=0;  B3=0
******************************************************************************
 Wald Statistic                 CHSQ(2)=   4.9391[.085]
******************************************************************************
```

Suppose now that you wish to estimate the $GARCH(1,1)$ model (19.3) using the Student-t density function for the conditional distribution of the errors. Choose Option 0 to return to the $GARCH$ Estimation Menu. Estimate the $GARCH(1,1)$ model (19.3) again, and when asked to specify the underlying distribution, choose option 2. Specify the order of the $GARCH$ as

1; 1 RUN

and click RUN again when the second editor appears. You will now be asked to give initial estimates for the α_1 and ϕ_1. For the degrees of freedom of the t-distribution the default initial value of 10 is suggested by *Microfit*. Type 0.1 and 0.4 for the 'MA lag 1', and 'AR lag 1', coefficients and click OK to accept. *Microfit* starts the computations, and converges after 31 iterations. The estimates in Table 19.4 should appear on the screen.

Table 19.4: Modelling conditional heteroscedasticity of the dollar/sterling rate

```
                         GARCH(1,1) assuming a t distribution

                               converged after 31 iterations
*******************************************************************************
 Dependent variable is DLUSD
 262 observations used for estimation from 1973M3 to 1994M12
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 ONE                     -.0011084          .0019593            -.56572[.572]
 DLUSD(-1)                  .10873           .066423             1.6369[.103]
*******************************************************************************
 R-Squared                     .010930   R-Bar-Squared                  -.5706E-3
 S.E. of Regression            .033694   F-Stat.    F(3,258)      .95038[.417]
 Mean of Dependent Variable -.0016091   S.D. of Dependent Variable     .033684
 Residual Sum of Squares        .29290   Equation Log-likelihood      523.8620
 Akaike Info. Criterion       518.8620   Schwarz Bayesian Criterion   509.9412
 DW-statistic                   2.0098
*******************************************************************************

            Parameters of the Conditional Heteroscedastic Model
        Explaining H-SQ, the Conditional Variance of the Error Term
*******************************************************************************
                           Coefficient       Asymptotic Standard Error
        Constant            .3440E-3              .8610E-4
        E-SQ(-1)            .13414                .088329
         H-SQ(-1)            .55995                 .11063
        D.F. of t-Dist.      9.8958                 5.0507
*******************************************************************************
 H-SQ stands for the conditional variance of the error term.
 E-SQ stands for the square of the error term.
```

The estimate of 9.896 obtained for ν, the parameter of the t-distribution, suggests only a mild departure from the normality. However, the ML estimate of ϕ_1 given by 0.5600(0.1107) now suggests a significant $GARCH$ effect. The Wald statistic for the test of the joint hypothesis $\alpha_1 = \phi_1 = 0$ is now 49.61, which is highly significant.

A comparison of the Akaike and Schwarz criteria across the above two specification of the conditional error distribution tends to favour the t-distribution, although as to be expected, the SBC heavily penalizes the t-distribution for the additional degrees of freedom parameter, ν, which is estimated: see Tables 19.2 and 19.4.

19.3 Lesson 19.3: Estimating EGARCH models for monthly $/£ exchange rate

The exponential $GARCH(1,1)$ (or $EGARCH(1,1)$) model is defined by

$$\begin{aligned}\log h_t^2 &= \alpha_0 + \alpha_1\left(\frac{u_{t-1}}{h_{t-1}}\right) + \alpha_1^*\left(\left|\frac{u_{t-1}}{h_{t-1}}\right| - \mu\right) \\ &\quad +\phi_1 \log h_{t-1}^2\end{aligned}$$

where $\mu = E\left(|u_{t-1}/h_{t-1}|\right)$. Unlike the $GARCH(1,1)$ model estimated in Lesson 19.2, the above specification has a well-defined conditional variance $h_t^2 = V(u_t\,|\Omega_{t-1})$, for all parameter values, $\alpha_0, \alpha_1, \alpha_1^*$. But for the process to be stable it is still required that $|\phi_1| < 1$. For further details and references to the literature see Section 23.1.

The exchange rate model to be estimated is the same as in Lessons 19.1 and 19.2. Read the *Microfit* file EXMONTH.FIT and specify the regression equation (19.1) to be estimated over the period 1973(1)-1994(12), then choose option 1 in the Volatility Modelling Menu to move to the $GARCH$ Estimation Menu (see Section 8.6). Then choose option 5, click , and enter

1; 1

for the initial estimate of α_1 ('MA lag 1'), α_1^* ('$ABS\ MA$ 1'), and ϕ_1 ('AR lag 1') type 0.1, 0.2 and 0.4, respectively. Click . The program starts the computations, and after 39 iterations you will be presented with the results in Table 19.5. According to these estimates only α_1^* is significantly different from zero at the 95 per cent level.

Table 19.5: Modelling conditional heteroscedasticity of the dollar/sterling rate

```
 Exponential GARCH(1,1) assuming a normal distribution

                                  converged after 39 iterations
*******************************************************************************
 Dependent variable is DLUSD
 262 observations used for estimation from 1973M3 to 1994M12
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 ONE                     -.0014268          .0018712              -.76251[.446]
 DLUSD(-1)                .080856           .066757               1.2112[.227]
*******************************************************************************
 R-Squared                     .010449   R-Bar-Squared                -.0049528
 S.E. of Regression            .033768   F-Stat.    F(4,257)      .67842[.607]
 Mean of Dependent Variable -.0016091   S.D. of Dependent Variable    .033684
 Residual Sum of Squares        .29304   Equation Log-likelihood      522.6913
 Akaike Info. Criterion       517.6913   Schwarz Bayesian Criterion   508.7705
 DW-statistic                   1.9520
*******************************************************************************

             Parameters of the Conditional Heteroscedastic Model
 Explaining the Logarithm of H-SQ, the Conditional Variance of the Error Term
*******************************************************************************
                           Coefficient        Asymptotic Standard Error
        Constant             -5.8206                  3.0714
        (E/H)(-1)            -.13235                 .094098
        ABS(E/H)(-1)-MEU      .33411                  .15824
         LOG(H-SQ(-1))        .14818                  .44828
*******************************************************************************
 H stands for the conditional standard error of the error term.
 E stands for the error term.
 MEU stands for the expectation of the absolute value of the standardized
 disturbance term. MEU = SQRT(2/3.14159) =   .79788
```

Consider now the estimation of the $EGARCH(1,1)$ model with a t-distribution for the conditional distribution of the errors. Move to the $GARCH$ Estimation Menu and choose option 5, then choose option 2. Click twice. You should see the screen editor for the specification of the initial estimates. Choose the ML estimates of α_1, α_1^* and ϕ_1 in Table 19.5 as initial values (namely -0.13, 0.33 and 0.15) and click . You should see the results in Table 19.6 on the screen.

Table 19.6: Modelling conditional heteroscedasticity of the dollar/sterling rate

```
                    Exponential GARCH(1,1) assuming a t distribution

                                      converged after 34 iterations
*******************************************************************************
 Dependent variable is DLUSD
 262 observations used for estimation from 1973M3 to 1994M12
*******************************************************************************
 Regressor               Coefficient       Standard Error         T-Ratio[Prob]
 ONE                      -.0012211           .0018680           -.65370[.514]
 DLUSD(-1)                 .11649             .065918             1.7672[.078]
*******************************************************************************
 R-Squared                     .010881   R-Bar-Squared                -.0045137
 S.E. of Regression            .033760   F-Stat.    F(4,257)     .70680[.588]
 Mean of Dependent Variable -.0016091   S.D. of Dependent Variable     .033684
 Residual Sum of Squares        .29292   Equation Log-likelihood      525.2848
 Akaike Info. Criterion       519.2848   Schwarz Bayesian Criterion   508.5797
 DW-statistic                   2.0260
*******************************************************************************

            Parameters of the Conditional Heteroscedastic Model
 Explaining the Logarithm of H-SQ, the Conditional Variance of the Error Term
*******************************************************************************
                            Coefficient       Asymptotic Standard Error
        Constant              -2.6069                  2.6109
        (E/H)(-1)            -.098542                  .10721
        ABS(E/H)(-1)-MEU       .29783                  .13410
         LOG(H-SQ(-1))         .61908                  .38102
        D.F. of t-Dist.       10.8345                  5.8475
*******************************************************************************
 H stands for the conditional standard error of the error term.
 E stands for the error term.
 MEU stands for the expectation of the absolute value of the standardized
 disturbance term. The maximum likelihood estimate of MEU =   .77566
```

A comparison of the results in Tables 19.5 and 19.6 shows that according to the AIC, there is a clear evidence in favour of the t-distribution, but the evidence is much less clear cut, if the SBC is used.

19.4 Lesson 19.4: Forecasting volatility

In this lesson we use the $GARCH(1,1)$ model estimated for the monthly observations on \$/£ exchange rate in Lesson 19.2 to compute the conditional standard errors, h_t, over the estimation period 1973(1)-1994(12), and then obtain (multi)-step ahead forecasts of h_t over the period 1995(1)-1995(6). In Lesson 19.2 we found that the t-distribution performs slightly better than the normal. So, in what follows we base our estimation on the exchange rate model

$$\begin{aligned} \Delta \log USD_t &= \beta_0 + \beta_1 \Delta \log(USD_{t-1}) + u_t, \\ V(u_t \,|\Omega_{t-1}) &= h_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \phi_1 h_{t-1}^2 \end{aligned}$$

assuming that conditional on Ω_{t-1}, the errors have the standardized Student t-distribution.

First follow the steps in Lesson 19.2 and choose option 1 in the $GARCH$ Estimation Menu to estimate the above model. Once you have successfully estimated the model, choose option 8 in the Post Regression Menu after the $GARCH$ result screen, to compute dynamic forecasts of $\Delta \log USD_t$ over the six months from 1995(1) to 1995(6). The program presents you with the forecasts of $\Delta \log USD_t$, $t = 95(1), ..., 95(6)$. To obtain the forecasts of h_t return to the Post Regression Menu, and choose option 3 to move to the Display/Save Residuals and Fitted Values Menu (see Section 6.21). Option 9 in this menu allows you to save the values of $\hat{h}_t$ over the estimation and the forecast period. You can also see the plot of $\hat{h}_t$ over the sample period 1973(1)-1994(12) if you choose option 8 in this menu. See Figure 19.1.

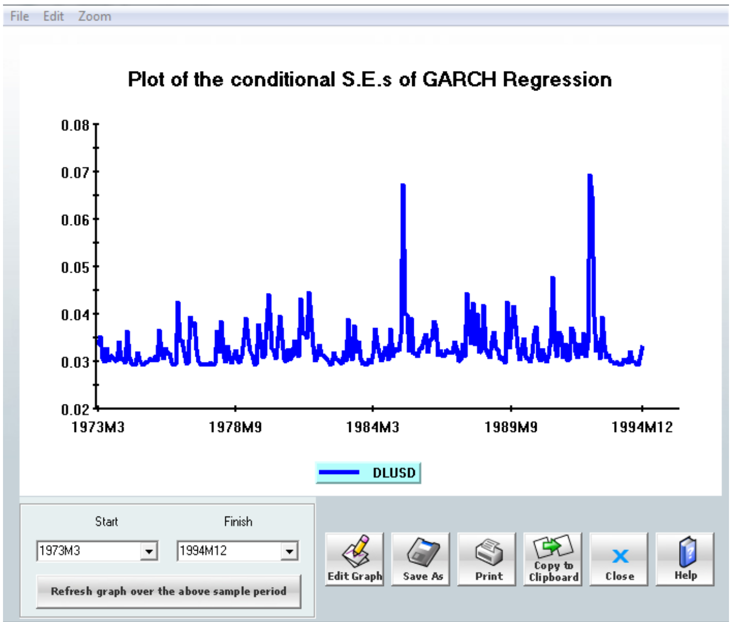
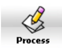


Figure 19.1: Estimated values of $\hat{h}_t$ for the dollar/sterling exchange rate

To list or plot the forecasts of h_t over the period 1995(1)-1995(6), choose option 9 in the Display/Save Residuals and Fitted Values Menu and type

$HHAT$ Estimates of the Conditional Standard Errors

click  to retun to the Post Regression Menu, choose option 0, and in the GARCH Estimation Menu choose option 0 again. Return to the Single Equation Estimation window, and click  to move to the Commands and Data Transformations box. Clear it, and type

SAMPLE $1995M1$ $1995M6$; **LIST** $HHAT$ 

The volatility forecasts in Table 19.7 should now appear on the screen.

Table 19.7: Volatility forecasts

Month	Forecasts of $\hat{h}_t$
1995M1	.035793
1995M2	.034515
1995M3	.033995
1995M4	.033787
1995M5	.033704
1995M6	.033672

Note that these are multiple-step ahead forecasts of h_t computed using the formulae in Section 23.1.9.

19.5 Lesson 19.5: Modelling volatility in daily exchange rates

The degree of volatility tends to increase with the frequency with which observations are sampled. This can be seen clearly as one moves from monthly to daily observations on exchange rates. As an example, consider the daily £/\$ exchange rates in the file EXDAILY.FIT covering the period from 2-Jan-1985 to 28-July-1993. In total this file contains 2,168 daily observations. Load this file into *Microfit,* and in the Commands and Data Transformations box create the rate of change of the £/\$ exchange rate in the variable $DEUS$:

$$DEUS = \mathbf{LOG}(EUS/EUS(-1)); \quad INPT = 1$$

To fit a $GARCH(1,2)$ model to the errors of this regression equation, choose option 1 in the Volatility Modelling Menu and specify the following AR(2) specification for the conditional mean of $DEUS$:

$$DEUS \quad INPT \quad DEUS(-1) \quad DEUS(-2)$$

Specify the start and end dates as

$$02\text{-}Jan\text{-}90 \quad 30\text{-}Jun\text{-}93$$

keeping the remaining 20 observations for forecast analysis. Click . Choose option 1 in the $GARCH$ Estimation Menu, click  twice and enter

1; 1 2 

For the initial values type 0.1 (for 'MA lag 1'), 0.2 (for 'MA lag 2') and 0.6 (for 'AR lag 1'). Notice that the sum of these initial estimates cannot exceed unity. Click  to start the computations. The process converges after 26 iterations, and you will be presented with the results in Table 19.8.

Table 19.8: Modelling conditional heteroscedasticity of the daily dollar/sterling rate

```
                      GARCH(1,2) assuming a normal distribution

                                   converged after 26 iterations
*******************************************************************************
 Dependent variable is DEUS
 884 observations used for estimation from 02-Jan-90 to 30-Jun-93
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INPT                    .3093E-3          .2449E-3               1.2631[.207]
 DEUS(-1)                 .12092            .037494               3.2250[.001]
 DEUS(-2)                -.047880           .036073              -1.3273[.185]
*******************************************************************************
 R-Squared                    .021038   R-Bar-Squared                    .015463
 S.E. of Regression          .0076251   F-Stat.    F(5,878)      3.7737[.002]
 Mean of Dependent Variable -.8143E-4   S.D. of Dependent Variable    .0076847
 Residual Sum of Squares      .051048   Equation Log-likelihood         3081.0
 Akaike Info. Criterion        3075.0   Schwarz Bayesian Criterion      3060.7
 DW-statistic                  1.9178
*******************************************************************************

             Parameters of the Conditional Heteroscedastic Model
         Explaining H-SQ, the Conditional Variance of the Error Term
*******************************************************************************
                           Coefficient       Asymptotic Standard Error
        Constant            .1642E-5               .5545E-5
        E-SQ(-1)            .084435                .041144
        E-SQ(-2)           -.031427                .041737
         H-SQ(-1)            .91959                 .020592
*******************************************************************************
 H-SQ stands for the conditional variance of the error term.
 E-SQ stands for the square of the error term.
```

The second part of this table clearly shows the importance of the AR component in the $GARCH(1,2)$ specification. However, the second-order coefficient of the MA part of the process is not statistically significant. Therefore, once again a $GARCH(1,1)$ model seems to fit the observations reasonably well. Re-estimating the exchange rate equation assuming a $GARCH(1,1)$ model yields the results in Table 19.9. (To obtain these results we started the iterations with 0.1 and 0.8 for MA and AR coefficients).

Table 19.9: Modelling conditional heteroscedasticity of daily dollar/sterling rate

```
                         GARCH(1,1) assuming a normal distribution

                                converged after 25 iterations
*******************************************************************************
 Dependent variable is DEUS
 884 observations used for estimation from 02-Jan-90 to 30-Jun-93
*******************************************************************************
 Regressor              Coefficient        Standard Error         T-Ratio[Prob]
 INPT                    .2933E-3            .2454E-3              1.1950[.232]
 DEUS(-1)                 .12536              .036193              3.4636[.001]
 DEUS(-2)                -.046199             .036109             -1.2794[.201]
*******************************************************************************
 R-Squared                   .021585   R-Bar-Squared                    .017132
 S.E. of Regression         .0076186   F-Stat.    F(4,879)      4.8479[.001]
 Mean of Dependent Variable -.8143E-4  S.D. of Dependent Variable     .0076847
 Residual Sum of Squares     .051020   Equation Log-likelihood          3084.8
 Akaike Info. Criterion       3079.8   Schwarz Bayesian Criterion       3067.9
 DW-statistic                 1.9272
*******************************************************************************

            Parameters of the Conditional Heteroscedastic Model
        Explaining H-SQ, the Conditional Variance of the Error Term
*******************************************************************************
                         Coefficient        Asymptotic Standard Error
        Constant          .1760E-5               .5023E-5
        E-SQ(-1)          .056053                .020252
         H-SQ(-1)          .91452                 .018822
*******************************************************************************
 H-SQ stands for the conditional variance of the error term.
 E-SQ stands for the square of the error term.
```

One hypothesis of interest involving the coefficients of the $GARCH(1,1)$ model is whether or not the sum of the coefficients of this model is unity. When the coefficients add up to unity, the model is known as the Integrated $GARCH$ or $IGARCH$ model, and implies that the shocks to the conditional variance are persistent. From the results in Table 19.9 it is clear that the sum of the estimates $\hat{\alpha}_1 = 0.0560$ and $\hat{\phi}_1 = 0.9146$ is very close to unity. To test the hypothesis that $\alpha_1 + \phi_1 = 1$, choose option 7 in the Hypothesis Testing Menu and type in the box editor

$$B2 + B3 = 1$$

(Recall that in *Microfit* the coefficients of the $GARCH$ model are denoted by $B1$, $B2$ ...). The Wald statistic, distributed asymptotically as a χ^2 with one degree of freedom, is equal 42.78, and strongly rejects the hypothesis that the $GARCH$ model is integrated.

19.6 Lesson 19.6: Estimation of GARCH-in-mean models of US excess returns

The regressions in Lesson 11.11 show that a statistically significant fraction of the variance of excess returns can be predicted by *ex ante* data variables, such as lagged dividend yields and lagged interest rates. This evidence (also replicated using other portfolios in other stock markets) rejects the joint hypothesis of market efficiency and risk neutrality. However, in situations where market participants are risk averse, standard efficient market models do not rule out the possibility that excess returns on stocks can be predicted. One important class of asset pricing models predicts a positive relationship between conditional expectations of excess returns and their conditional variances (see, for example, Merton (1980)). If $ERSP_t$ is the excess return on the SP500 defined in Lesson 11.11, then a generalized version of the Merton (1980) mean-variance model can be written as

$$ERSP_t = \boldsymbol{\beta}' \mathbf{x}_{t-1} + \gamma V\left(ERSP_t \left| \Omega_{t-1}\right.\right) + u_t \tag{19.4}$$

where

$$V\left(ERSP_t \left| \Omega_{t-1}\right.\right) = V\left(u_t \left| \Omega_{t-1}\right.\right) = h_t^2$$

Ω_{t-1} is the publicly available information at time $t-1$, and $\mathbf{x}_{t-1}$ is a vector of *ex ante* dated variables. In what follows we assume that $\mathbf{x}_{t-1}$ includes the variables YSP_{t-1}, $PI12_{t-2}$, $DI11_{t-1}$ and $DIP12_{t-2}$, defined in Lesson 11.11, and that h_t^2 has the $GARCH(1,1)$ specification:

$$h_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \phi_1 h_{t-1}^2 \tag{19.5}$$

This model can be readily estimated using the $GARCH$-in-mean option in *Microfit*. See Section 8.6. Load the special *Microfit* file PTMONTH.FIT and then run the batch file PTMONTH.BAT on it in the Commands and Data Transformations box. Choose option 1 in the Volatility Modelling Menu, retrieve the file PTMONTH.LST into the box editor, and click . Choose option 2 (the $GARCH$-in-mean), followed by the normal distribution option, and enter

1; 1 

Choose the values of 0.5, 0.1 and 0.2 for the initial estimates of γ ('in mean'), and α_1 ('MA lag 1') and ϕ_1 ('AR lag 1'), respectively, and click  to start the computations. The iterative procedure converges after 34 iterations, and the results in Table 19.10 should appear on the screen.

Table 19.10: Excess return regression for SP500 portfolio with $GARCH$-in mean effect

```
                 GARCH(1,1) in mean assuming a normal distribution

                              converged after 34 iterations
*******************************************************************************
 Dependent variable is ERSP
 468 observations used for estimation from 1954M1 to 1992M12
*******************************************************************************
 Regressor               Coefficient       Standard Error         T-Ratio[Prob]
 INPT                      -.027861            .014131            -1.9717[.049]
 YSP(-1)                    12.7055             3.2047             3.9647[.000]
 PI12(-2)                  -.29479             .068527            -4.3018[.000]
 DI11(-1)                -.0067114            .0027614            -2.4304[.015]
 DIP12(-2)                 -.14717             .038211            -3.8514[.000]
 H-Squared                  5.8897              6.7439              .87334[.383]
*******************************************************************************
 R-Squared                      .091374   R-Bar-Squared                 .077547
 S.E. of Regression             .040750   F-Stat.    F(7,460)      6.6084[.000]
 Mean of Dependent Variable   .0059055   S.D. of Dependent Variable     .042428
 Residual Sum of Squares         .76386   Equation Log-likelihood       841.5218
 Akaike Info. Criterion        833.5218   Schwarz Bayesian Criterion    816.9279
 DW-statistic                    2.0030
*******************************************************************************

            Parameters of the Conditional Heteroscedastic Model
        Explaining H-SQ, the Conditional Variance of the Error Term
*******************************************************************************
                          Coefficient       Asymptotic Standard Error
        Constant           .1383E-3              .1152E-4
        E-SQ(-1)           .067967               .043356
         H-SQ(-1)           .84898                .032462
*******************************************************************************
 H-SQ stands for the conditional variance of the error term.
 E-SQ stands for the square of the error term.
```

According to these results the ML estimates of γ is 5.89 (6.74) and has the correct sign, but is not statistically significant. Therefore, there does not seem to be any evidence of $GARCH$-in-mean effect in this model. Notice, however, that the other variables in the excess return regression continue to be highly significant. The evidence on the volatility of the conditional variance of u_t (i.e. h_t^2) is rather mixed. The ML estimate of ϕ_1 at 0.84898 (0.0325) is highly significant, but the ML estimate of α_1 is near zero and is not statistically significant.

Now consider estimating the above $GARCH$-in-mean model (19.4) and (19.5), assuming the conditional distribution of the errors to be t-distributed. Return to the $GARCH$ Estimation Menu, choose option 2, and then select the t-distribution option. Click  to accept the content of the first box editor, and then  again. For the initial estimates of γ, α_1, and ϕ_1, choose the values 0.1 ('in mean'), 0.2 ('MA lag 1') and 0.4 ('AR lag 1'), respectively. Click  to accept these initial estimates. The program starts the computations, and yields the results in Table 19.11 after 25 iterations.

Table 19.11: Excess return for SP500 portfolio with *GARCH*-in-mean effect

GARCH(1,1) in mean assuming a t distribution

converged after 25 iterations

Dependent variable is ERSP
468 observations used for estimation from 1954M1 to 1992M12

Regressor	Coefficient	Standard Error	T-Ratio[Prob]
INPT	-.028471	.016612	-1.7139[.087]
YSP(-1)	11.7720	3.1699	3.7137[.000]
PI12(-2)	-.30635	.067573	-4.5337[.000]
DI11(-1)	-.0045283	.0026368	-1.7174[.087]
DIP12(-2)	-.14492	.037070	-3.9094[.000]
H-Squared	8.7211	8.9383	.97570[.330]

R-Squared	.087992	R-Bar-Squared	.074114
S.E. of Regression	.040826	F-Stat. F(7,460)	6.3402[.000]
Mean of Dependent Variable	.0059055	S.D. of Dependent Variable	.042428
Residual Sum of Squares	.76671	Equation Log-likelihood	848.2294
Akaike Info. Criterion	839.2294	Schwarz Bayesian Criterion	820.5613
DW-statistic	2.0006		

Parameters of the Conditional Heteroscedastic Model
Explaining H-SQ, the Conditional Variance of the Error Term

	Coefficient	Asymptotic Standard Error
Constant	.1917E-3	.2273E-4
E-SQ(-1)	.061568	.043320
H-SQ(-1)	.82076	.033397
D.F. of t-Dist.	8.6483	3.2139

H-SQ stands for the conditional variance of the error term.
E-SQ stands for the square of the error term.

Comparing these results with those in Tables 19.10 clearly suggests that the t-distribution fits the data much better than the normal. Even the Schwarz criterion unambiguously selects the model with conditionally t-distributed errors. Nevertheless $\hat{\gamma} = 8.72$ (8.94), and hence the hypothesis that $\gamma = 0$ cannot be rejected. The same conclusion also applies to α_1 (we have $\hat{\alpha}_1 = 0.0616(0.043)$).

The extent to which the t-distribution has been successful in dealing with the non-normal errors can be assessed graphically. Click 

to move to the Post Regression Menu, then choose option 3, followed by option 5 in the Display/Save Residuals and Fitted Values Menu. The histogram of the scaled residuals (defined by $\hat{u}_t/\hat{h}_t$) should now appear on the screen (See Figure 19.2).

Except for a possible 'outlier' to the left of the graph, the t-distribution seems to provide a reasonable fit for the distribution of the scaled residuals. A closer inspection of the results shows, perhaps not surprisingly, that the apparent outlier refers to the 1987 October crash.

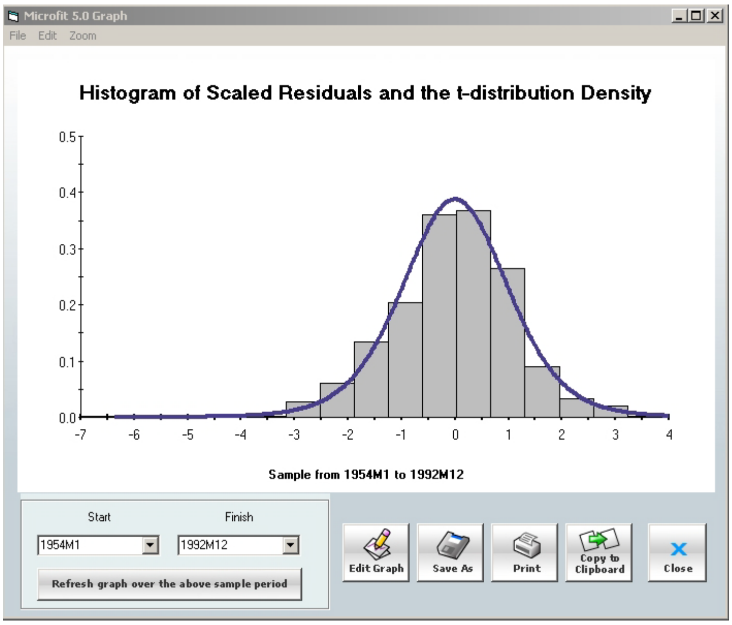


Figure 19.2: Histogram of the scaled residuals ($\hat{u}_t/\hat{h}_t$) for the SP500 excess return regression

19.7 Exercises in GARCH modelling

19.7.1 Exercise 19.1

Use monthly observations in EXMONTH.FIT to estimate a first-order autoregressive model in the rate of change of the Deutschemark/Sterling exchange rate. Is there any evidence of $ARCH$ effects in this regression?

19.7.2 Exercise 19.2

Use the data in PTMONTH.FIT to estimate the $GARCH$-in-mean regression (19.4), assuming that $V(u_t\,|\Omega_{t-1})$ has the exponential specification. Compare your result with those obtained in Lesson 19.6.

Chapter 20

Lessons in Multivariate GARCH Modelling

The lessons in this chapter are concerned with the estimation of multivariate generalized autoregressive conditional heteroscedastic ($MGARCH$) models. They show how *Microfit* can be used to estimate dynamic conditional correlation (DCC) models, how to compute the VaR of a portfolio, and how to calculate forecasts of conditional volatilities and correlations. The relevant estimation options, the underlying econometric methods, and the computational algorithms are described in Section 23.2. For a review of the application of dynamic conditional correlation models to financial data see Engle (2002), Pesaran and Pesaran (2007), Pesaran, Schleicherc, and Zaffaroni (2009), Andersen, Bollerslev, Diebold, and Ebens (2001), and Andersen, Bollerslev, Diebold, and Labys (2001).

All the lessons in this chapter use the futures data analyzed in Pesaran and Pesaran (2007).

20.1 Lesson 20.1: Estimating DCC models for a portfolio of currency futures

In this lesson we demonstrate how to use the multivariate $GARCH$ option to estimate dynamic conditional correlation (DCC) models for a portfolio composed of returns on six currency futures: Australian dollar (AD), British pound (BP), Canadian dollar (CD), Swiss franc (CH), Euro (EU), and Japanese yen (JY). We will initially assume that returns are normally distributed, and then consider the case of t-distributed returns.

Daily data on currency futures are available in the file FUTURESDATA.FIT, and cover the period from 31-Dec-93 to 01-Jan-07.[1] In the present application we will use data from 1995 to 2005 (for a total of $2,610$ observations) for estimation, and keep the observations from 2005 to 2007 for evaluation of the model and for forecasting purposes (see Lessons 20.4

[1] In the tutorial directory the CSV and *Excel* versions of the file FUTURESDATA.FIT are also available. Notice that to load the *Excel* version of this file you need to set in *Microfit* the 'European/US date' option to the European date format. See Section 3.2.10 for further details on how to set the date format in *Microfit*.

and 20.5).

Read this file into *Microfit*, choose option 2 from the Volatility Modelling Menu, and in the Commands and Data Transformations box enter the list of variables to be included in the $MGARCH$ model:[2]

$$AD \quad BP \quad CD \quad CH \quad EU \quad JY$$

Then specify the sample period:

$$02\text{-}Jan\text{-}95 \quad 31\text{-}Dec\text{-}04$$

Click 

. You will be presented with the Multivariate $GARCH$ window, shown in Figure 20.1.

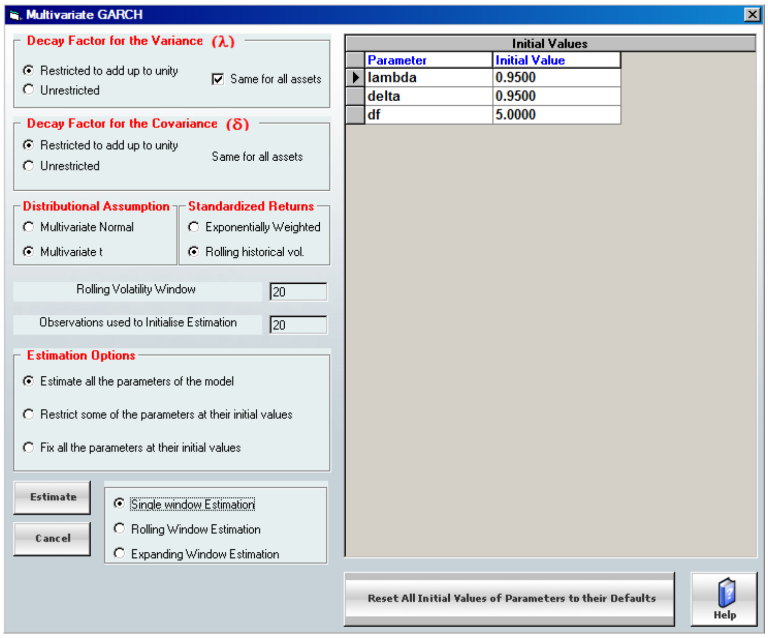


Figure 20.1: The multivariate $GARCH$ window

In the 'Decay factor for the variance' panel, select the 'Unrestricted' option and ensure that the option 'Same for all assets' is unchecked. This enables you to compute two parameters (λ_{1i} and λ_{2i}) for each asset $i = 1, .., 6$. In the 'Decay factor for the covariance' panel, select the 'Unrestricted' option, which allows for mean-reverting conditional correlations. Choose the multivariate normal distribution, and select the 'Rolling historical volatility' option with rolling window equal to 20.[3] Do not change the the default number of observations

[2] Alternatively you can retrieve the content of the CFUTURES.LST from the tutorial directory by using the 

button.

[3] In the present application, a choice of p well above 20 does not allow the jumps in returns to become adequately reflected in the estimated realized volatility (see Pesaran and Pesaran (2007)).

(20) used to initialize the estimation. For further information on the initialization sample, see Section 23.2.1.

The right panel of the multivariate *GARCH* window displays the default initial values of parameters, which are set to 0.95 for the parameters λ_{1i}, for $i = 1, ..., 6$, and δ_1 and 0.05 for λ_{2i}, for $i = 1, ..., 6$, and δ_2.

Press the 

button to accept these default values and estimate *all* parameters of the model. *Microfit* starts the computations and, after 51 iterations, presents you with results reported in Table 20.1.

Table 20.1: ML estimates of the Gaussian DCC model on futures daily returns

```
     Multivariate GARCH with underlying multivariate Normal distribution
                       Converged after 51 iterations
******************************************************************************
 Based on  2610 observations from 02-Jan-95 to 31-Dec-04.
 The variables (asset returns) in the multivariate GARCH model are:
 AD  BP  CD  CH  EU  JY
 Volatility decay factors unrestricted, different for each variable.
 Correlation decay factors unrestricted, same for all variables.
******************************************************************************
 Parameter                  Estimate       Standard Error         T-Ratio[Prob]
 lambda1_AD                  .95917          .0065101           147.3345[.000]
 lambda1_BP                  .93141           .020478            45.4825[.000]
 lambda1_CD                  .95005          .0066782           142.2621[.000]
 lambda1_CH                  .95761          .0064245           149.0559[.000]
 lambda1_EU                  .96532          .0053634           179.9831[.000]
 lambda1_JY                  .95352          .0072213           132.0424[.000]
 lambda2_AD                 .027573          .0038418             7.1772[.000]
 lambda2_BP                 .034280          .0076195             4.4990[.000]
 lambda2_CD                 .042579          .0052590             8.0965[.000]
 lambda2_CH                 .029089          .0036401             7.9911[.000]
 lambda2_EU                 .025664          .0033657             7.6252[.000]
 lambda2_JY                 .039369          .0052836             7.4512[.000]
 delta1                      .96964          .0023264           416.7932[.000]
 delta2                     .019732          .0012327            16.0068[.000]
******************************************************************************
 Maximized Log-Likelihood =    -9602.1
******************************************************************************

                  Estimated Unconditional Volatility Matrix
    Unconditional Volatilities (Standard Errors) on the Diagonal Elements
           Unconditional Correlations on the Off-Diagonal Elements

******************************************************************************
                  AD        BP        CD        CH        EU        JY
 AD            .68066    .27696    .36365    .26443    .30291    .23641

 BP            .27696    .50936    .13144    .59205    .60306    .28641

 CD            .36365    .13144    .38607    .11979    .14352    .10871

 CH            .26443    .59205    .11979    .69833    .92238    .41880

 EU            .30291    .60306    .14352    .92238    .63419    .37862

 JY            .23641    .28641    .10871    .41880    .37862    .75003

******************************************************************************
 For the time-varying conditional volatilities and correlations see the Post
 Estimation Menu.
```

The upper panel of the table presents the maximum likelihood estimates of λ_{i1}, λ_{i2} for the six currencies futures returns, and δ_1, δ_2. Observe that the asset-specific estimates of the volatility decay parameters are all highly significant, with the estimates of λ_{i1}, $i = 1, 2, ..., 6$ very close to unity.

The lower panel of the table reports the estimated unconditional volatilities and correlations of the vector of assets. Notice the high correlation between Euro and Swiss franc (0.922), and between Euro and British pound (0.603) futures returns.

Now click 

to leave the estimation results, and choose option 0 to re-specify the *DCC* model. In the initial window select the multivariate Student t-distribution option. Notice that a new parameter, the degrees of freedom (df) of the t-distribution, appears in the list of parameters to be initialized on the right panel of the multivariate *GARCH* window. Its default value is set to 5. Click the 

button to accept these initial values and start the computations. The results reported in Table 20.2 appear after 23 iterations.

Table 20.2: ML estimates of the t-DCC model on futures daily returns

Multivariate GARCH with underlying multivariate t-distribution
Converged after 23 iterations

Based on 2610 observations from 02-Jan-95 to 31-Dec-04.
The variables (asset returns) in the multivariate GARCH model are:
AD BP CD CH EU JY
Volatility decay factors unrestricted, different for each variable.
Correlation decay factors unrestricted, same for all variables.

Parameter	Estimate	Standard Error	T-Ratio[Prob]
lambda1_AD	.96558	.0064669	149.3113[.000]
lambda1_BP	.96360	.0086567	111.3120[.000]
lambda1_CD	.96829	.0056984	169.9231[.000]
lambda1_CH	.95854	.0067738	141.5068[.000]
lambda1_EU	.96529	.0055849	172.8398[.000]
lambda1_JY	.95147	.0091225	104.3002[.000]
lambda2_AD	.024122	.0037590	6.4171[.000]
lambda2_BP	.025943	.0049584	5.2321[.000]
lambda2_CD	.029057	.0047125	6.1660[.000]
lambda2_CH	.029370	.0041348	7.1032[.000]
lambda2_EU	.026119	.0036595	7.1375[.000]
lambda2_JY	.039971	.0066522	6.0087[.000]
delta1	.97102	.0023837	407.3686[.000]
delta2	.019970	.0013523	14.7672[.000]
df	5.9053	.22626	26.0993[.000]

Maximized Log-Likelihood = -8848.4

df is the degrees of freedom of the multivariate t distribution

Estimated Unconditional Volatility Matrix
Unconditional Volatilities (Standard Errors) on the Diagonal Elements
Unconditional Correlations on the Off-Diagonal Elements

	AD	BP	CD	CH	EU	JY
AD	.68066	.27696	.36365	.26443	.30291	.23641
BP	.27696	.50936	.13144	.59205	.60306	.28641
CD	.36365	.13144	.38607	.11979	.14352	.10871
CH	.26443	.59205	.11979	.69833	.92238	.41880
EU	.30291	.60306	.14352	.92238	.63419	.37862
JY	.23641	.28641	.10871	.41880	.37862	.75003

For the time-varying conditional volatilities and correlations see the Post Estimation Menu.

The maximized log-likelihood value (−8848.4) is significantly larger than that obtained under the normality assumption (−9602.1). The estimated degrees of freedom for the t-distribution is 5.9, well below 30, and any other value one would expect for a multivariate normal distribution. This suggests that the t-distribution is more appropriate in capturing the fat-tailed nature of the distribution of asset returns.

20.2 Lesson 20.2: Plotting the estimated conditional volatilities and correlations

In this lesson we wish to inspect the temporal dynamics of conditional volatilities and correlations of currency futures returns. Read the special *Microfit* file FUTURESDATA.FIT, and follow the steps outlined in Lesson 20.1 to estimate an unrestricted *t-DCC* model with asset-specific volatility parameters, common conditional correlation parameters over the period from 02-*Jan*-95 to 31-*Dec*-04. Click Close to leave the output window with the estimation results, and choose option 2 to list/plot/save estimated conditional volatilities and correlations. You will be presented with the Conditional Volatilities, Correlations and Eigenvalues Menu, select option 2, and click OK. A list of conditional volatilities and correlations, and the eigenvalues of the covariance and correlation matrices appears on the screen.[4] For example, select the conditional volatility of the six currencies, namely $Vol(AD), Vol(BP), Vol(CD), Vol(CH), Vol(EU)$ and $Vol(JY)$, and click OK. To reduce the impact on the graph of initialization of estimates, we suggest excluding from the plot the values relative to the initial period from 02-*Jan*-95 to 01-*Jan*-96. Hence, once the graph is displayed, use the Start and the Finish drop-down lists to select sample period from 02-*Jan*-96 to 31-*Dec*-04, and click the REFRESH button. You will be presented with Figure 20.2. It can be seen from the graph that over time, conditional volatilities of currency futures returns tend to move more closely together. The convergence of volatilities could reflect the advent of Euro and a closer financial integration, particularly in the Euro area, as well as a decline in the importance of idiosyncratic factors.

Now return to the list of conditional volatilities and correlations and select the correlations of Euro with all other currencies, namely $corr(EU, AD)$, $corr(EU, BP)$, $corr(EU, CD)$, $corr(EU, CH)$, and $corr(JY, EU)$. Use the Start and the Finish drop-down lists to select sample period from 02-*Jan*-96 to 31-*Dec*-04. You will be presented with Figure 20.3. It can be clearly seen from the graph that conditional correlations of returns on Euro futures with the other currencies have been rising over time. Also, notice the high correlation between returns on Euro and Swiss franc futures.

[4] Cov_eigen_max, cov_eigen_2, cov_eigen_3, ..., cov_eig_min and corr_eigen_max, corr_eigen_2, corr_eigen_3, and so on, are the eigenvalues of the covariance and correlation matrix respectively, in decreasing order.

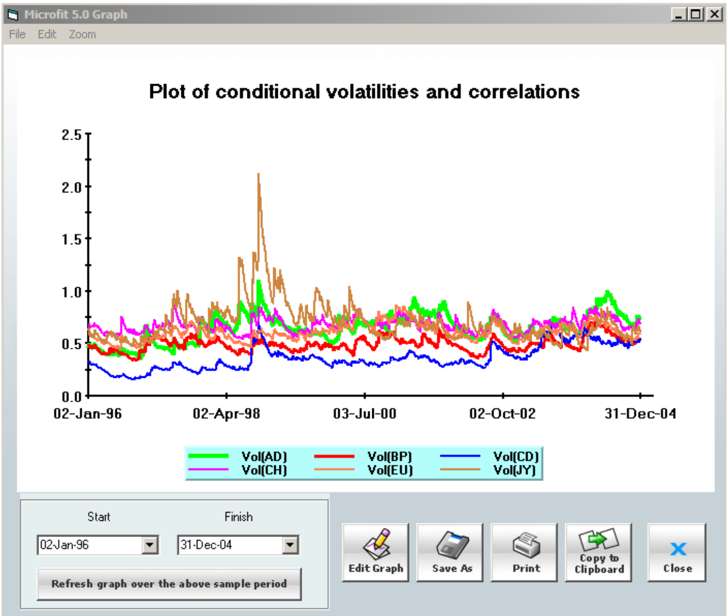


Figure 20.2: Conditional volatilities of currency futures returns over the period 02-Jan-96 to 31-Dec-04

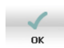
20.3 Lesson 20.3: Testing for linear restrictions

Consider the *t-DCC* model of currency futures returns estimated in Lesson 20.1. In this lesson we shall focus on the problem of testing the hypothesis that one of the assets returns has non-mean reverting volatility. That is, let λ_{i1} and λ_{i2} be the parameters for the conditional volatility equation of the ith asset we wish to test:

$$H_0 : \lambda_{i1} + \lambda_{i2} = 1$$

Under H_0 the process is non-mean reverting, and the unconditional variance for this asset does not exists (see Section 23.2 for further details on mean-reverting processes).

Load into *Microfit* the file FUTURESDATA.FIT, and follow the steps outlined in Lesson 20.1 to estimate a *t-DCC* model with asset-specific volatility parameters, and common conditional correlation parameters δ_1 and δ_2 (see Table 20.2). Click 

to leave the output window, and in the Multivariate $GARCH$ Post Estimation Menu select option 4 to estimate/test functions of the parameters of the model. Click 

. You will be asked to specify the restrictions that you wish to test. Suppose that initially we are interested in testing whether conditional volatilities of Euro returns are mean-reverting. Then in the box editor type

$$ZEROS = 1 - LAMBDA1\_EU - LAMBDA2\_EU$$

You should now see the test results on the screen, which we have reproduced in 20.3. As $1 -$

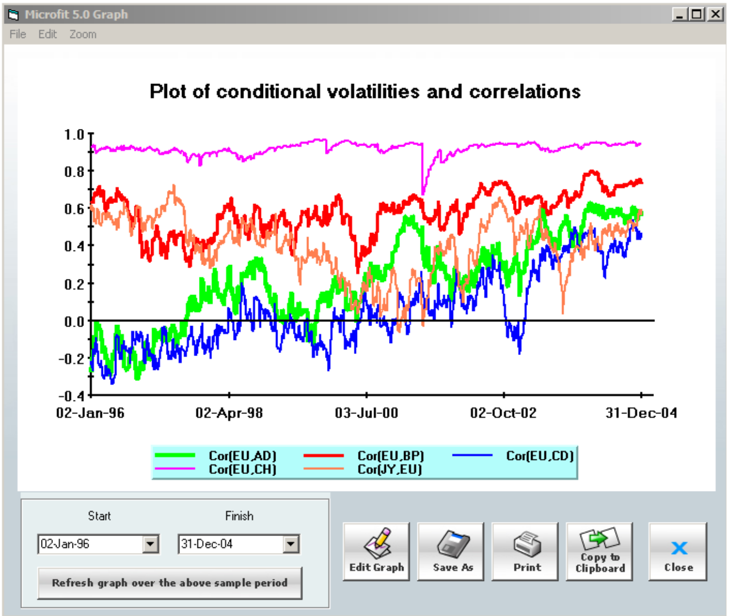


Figure 20.3: Conditional correlations of Euro futures returns with other currencies over the period 02-Jan-96 to 31-Dec-04

$\lambda_{i1} - \lambda_{i2} = 0.0086$, with a standard error of 0.0026, results suggest very slow but statistically significant mean-reverting volatility for Euro futures returns.

Table 20.3: Testing for mean revertion of volatility of Euro futures returns

```
                   Analysis of Function(s) of Parameter(s)
******************************************************************************
 The variables (asset returns) in the multivariate GARCH model are:
 AD    BP    CD    CH    EU    JY
 Volatility decay factors unrestricted, different for each variable.
 Correlation decay factors unrestricted, same for all variables.
 2610 observations used for estimation from 02-Jan-95 to 31-Dec-04
******************************************************************************

 List of specified functional relationship(s):
 ZEROS=1-LAMBDA1_EU-LAMBDA2_EU
******************************************************************************
 Function                  Estimate       Standard Error         T-Ratio[Prob]
 ZEROS                    .0085882          .0026249              3.2718[.001]
******************************************************************************

        Estimated Variance Matrix of the Function(s) of the Parameters

******************************************************************************
              ZEROS
 ZEROS       .6890E-5

******************************************************************************
```

As an exercise, test for mean reversion of volatility using other assets. A summary of results for various assets is reported in Table 20.4.

Table 20.4: Testing for mean revertion of volatility of currency futures returns

Asset	$1-\hat{\lambda}_1-\hat{\lambda}_2$	Std.errors	t-ratio
Australian dollar	0.01029	0.00407	2.5261
British pound	0.01046	0.00479	2.1812
Canadian dollar	0.00266	0.00199	1.3375
Swiss franc	0.01209	0.00355	3.4092
Yen	0.00855	0.00314	2.7240

20.4 Lesson 20.4: Testing the validity of the t-DCC model

In this lesson we test the validity of the *t-DCC* model estimated in Lessons 20.1, using the VaR and distribution free diagnostics available in *Microfit* (see Section 23.2.3 for further details on these test statistics).

Load the file FUTURESDATA.FIT, and follow the steps outlined in Lesson 20.1 to obtain the estimates of *t-DCC* model presented in Table 20.2. Click 

to leave the estimation results, return to the Multivariate *GARCH* Post Estimation Menu, and choose option 5. You will be required to specify the forecast/evaluation period that is used to evaluate the

$MGARCH$. Enter

$$03\text{-}Jan\text{-}05 \quad 01\text{-}Jan\text{-}07$$

You are then asked to list the variables that contain the portfolio weights to be attached to each of the assets returns. In the box editor type

$$C \quad C \quad C \quad C \quad C \quad C$$

where C is a variable with all values being equal to unity. This implies that each asset has equal weight in the portfolio. The next screen is a menu entitled Testing the Validity of Multivariate $GARCH$ Menu; choose option 1. You will be presented with a list of available diagnostic tests based on the probability integral transforms. If you select option 1, *Microfit* produces an output with the Lagrange Multiplier (LM) test statistic of serial correlation on the probability integral transform variable (see Section 23.2.3 for further information on these diagnostics). The results are shown in Table 20.5. Under the null hypothesis of correct specification of the t-DCC model, the probability transform estimates are serially uncorrelated and uniformly distributed over the range $(0, 1)$. The LM test, equal to 17.46 with a p-value of 0.133, is not statistically significant, thus supporting the validity of the t-DCC model.

Table 20.5: Testing serial correlation on the probability transform estimates

Test of Serial Correlation of Residuals (OLS case)

Dependent variable is U-Hat
List of variables in OLS regression:
Intercept
521 observations used for estimation from 03-Jan-05 to 01-Jan-07

Regressor	Coefficient	Standard Error	T-Ratio[Prob]
OLS RES(-1)	-.023600	.044353	-.53209[.595]
OLS RES(-2)	.0063146	.044100	.14319[.886]
OLS RES(-3)	.039020	.044078	.88523[.376]
OLS RES(-4)	-.061237	.044075	-1.3894[.165]
OLS RES(-5)	.021169	.043959	.48156[.630]
OLS RES(-6)	-.027219	.043969	-.61905[.536]
OLS RES(-7)	.0060669	.044016	.13783[.890]
OLS RES(-8)	.099379	.044018	2.2577[.024]
OLS RES(-9)	.049738	.044157	1.1264[.261]
OLS RES(-10)	-.047538	.044286	-1.0734[.284]
OLS RES(-11)	-.11116	.044335	-2.5073[.012]
OLS RES(-12)	-.028538	.044677	-.63876[.523]

Lagrange Multiplier Statistic	CHSQ(12)=	17.4553[.133]
F Statistic	F(12,508)=	1.4675[.132]

U-Hat denotes the probability integral transform.
Under the null hypothesis, U-Hat should not display any serial correlation.

If you select option 2, Figure 20.4 should appear on the screen. The graph compares the empirical cumulative distribution function of the probability integral transform variable with that of a uniform. The Kolmogorov-Smirnov test statistic (0.045) is smaller than its

5 per cent critical value (0.596), and hence it does not reject the null hypothesis that the probability integral transforms are uniformly distributed.[5]

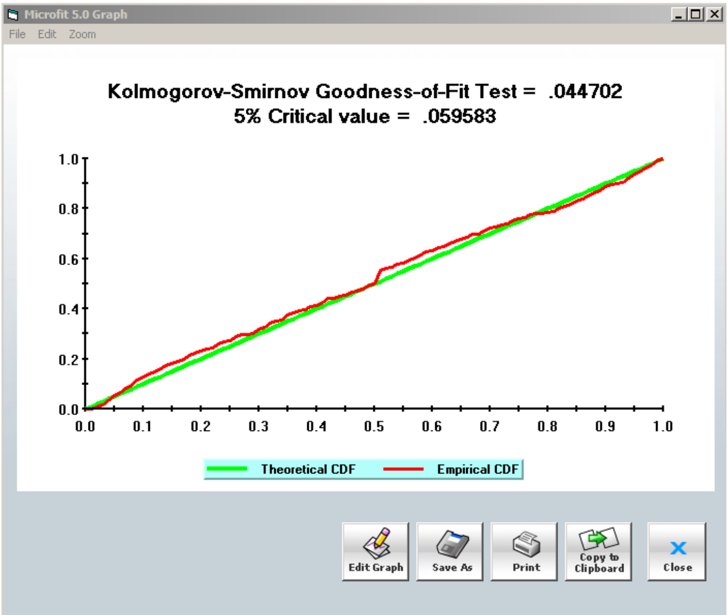
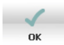


Figure 20.4: Kolmogorov-Smirnov test of normality

Close the graph and select option 3 to view the histograms of the probability integral transform variable. You will be presented with Figure 20.5.

Choose option 0 to return to the Testing the Validity of Multivariate *GARCH* Menu, and select option 2 to test for *VaR* violations. Enter 0.01 as tolerance probability of the *VaR*, and press the 

button. You can now list, plot or save the estimated Value at Risk under the selected tolerance probability, or compute the diagnostic statistics $\hat{\pi}_N$ and z_π for the *t*-*DCC* model, based on the *VaR* (see Section 23.2.3). Notice that the *VaR* is computed in percentage points. The plot of the Value at Risk of the portfolio for the forecasting period from 03-*Jan*-05 to 01-*Jan*-07 is obtained choosing option 2 and is shown in Figure 20.6. Notice the decreasing pattern indicating a diminishing portfolio risk over time.

By choosing option 5 you can obtain the mean hit rate $\hat{\pi}_N$ (equal to 0.998) level, and z_π (1.854) under the tolerance probability 0.01 (see Table 20.6). Notice that $\hat{\pi}_N$ is very close to its expected value (0.990), and that the test statistic z_π is not significant, both supporting the validity of the *t*-*DCC* model. Finally note that the probability integral transform, the indicator variable, and the *VaR* can be saved to the workspace for later use.

[5] For details of the Kolmogorov-Smirnov test and its critical values see, for example, Neave and Worthington (1992), pp.89-93, and Massey (1951).

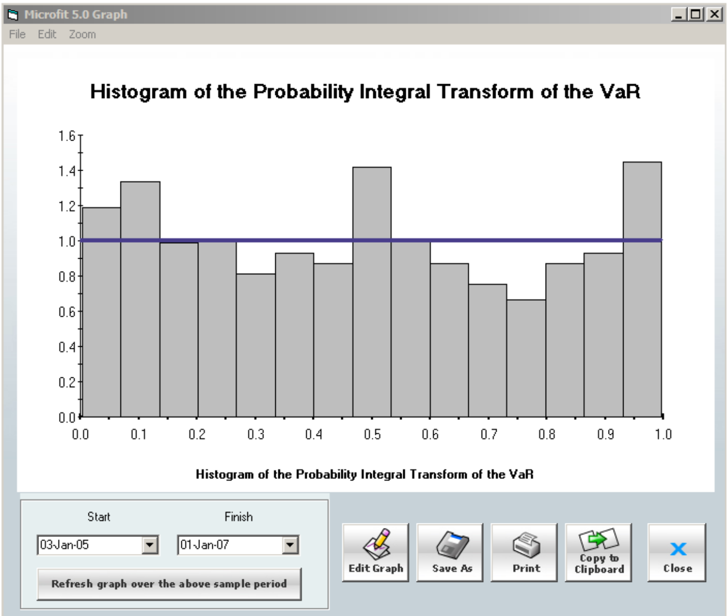


Figure 20.5: Histogram of the probability integral transform of the VaR

Table 20.6: VaR violations and the associated test statistics

```
      Mean VaR Violations and the Associated Diagnostic Test Statistics
*******************************************************************************
 Mean Hit Rate (pihat statistic) =    .99808 with expected value of    .99000
 Standard Normal Test Statistic=   1.8537[.064]
*******************************************************************************
```

20.5 Lesson 20.5: Forecasting conditional correlations

In this lesson we use the *t-DCC* model estimated above to compute forecasts of conditional correlations of Euro futures returns with other currency futures returns. Read the special *Microfit* file FUTURESDATA.FIT, and follow the steps in Lesson 20.1 to obtain the estimates of an unrestricted *t-DCC* model with asset-specific volatility parameters, common conditional correlation parameters, over the period from 02-*Jan*-95 to 31-*Dec*-04. Select option 7 from the Multivariate *GARCH* Post Estimation Menu. When prompted, choose as the forecast period the interval from 03-*Jan*-05 to 31-*Jan*-05 and click OK. Now select option 2 to plot the forecasts of conditional volatilities and correlations and click OK. Select all correlations of Euro with other currency futures returns, and use the Start and the Finish drop-down lists to select sample period from 02-*Jan*-03 to 31-*Jan*-05. The plot is reproduced in Figure 20.7.

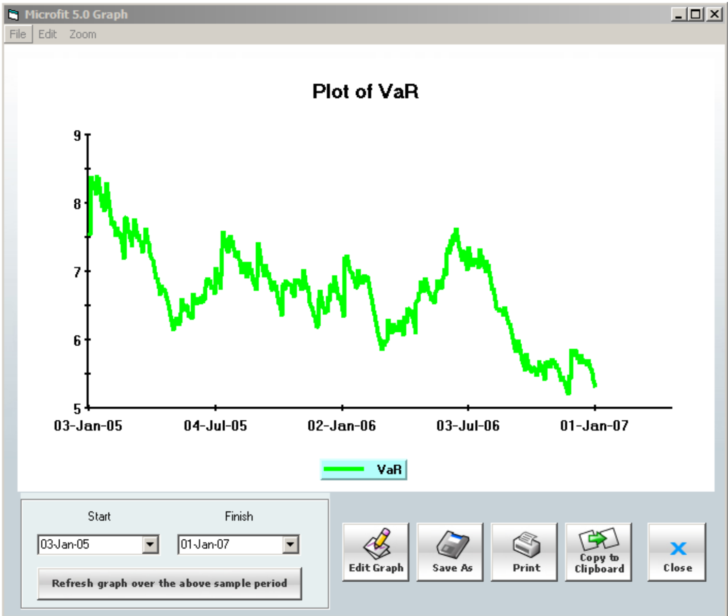


Figure 20.6: Plot of the Value at Risk of the portfolio for $\alpha = 0.01$, over the period from 03-Jan-05 to 01-Jan-07

20.6 Lesson 20.6: MGARCH applied to a set of OLS residuals

In this lesson we use the FUTURESDATA.FIT file to estimate a t-DCC model on residuals obtained from a regression of returns on currency futures on their past values. To this end, we use option 3 '$MGARCH$ applied to the OLS residuals of a set of regressions' from the Volatility Modelling Menu. This option performs a two-step estimation method, where in the first step residuals are obtained by running separate OLS regressions for each variable, and in the second step the DCC model is applied to these residuals.

Read the special *Microfit* file FUTURESDATA.FIT, and in the Commands and Data Transformation box write[6]

$$\begin{array}{lll} AD & AD(-1) & C; \\ BP & BP(-1) & C; \\ CD & CD(-1) & C; \\ CH & CH(-1) & C; \\ EU & EU(-1) & C; \\ JY & JY(-1) & C \end{array}$$

where C is a variable with all values being equal to unity. The above instructions specify a first-order autoregressive process for each of the six currency futures returns. Set the sample

[6] Alternatively you can retrieve the content of the OLSFUTURES.EQU from the tutorial directory using the 

button.

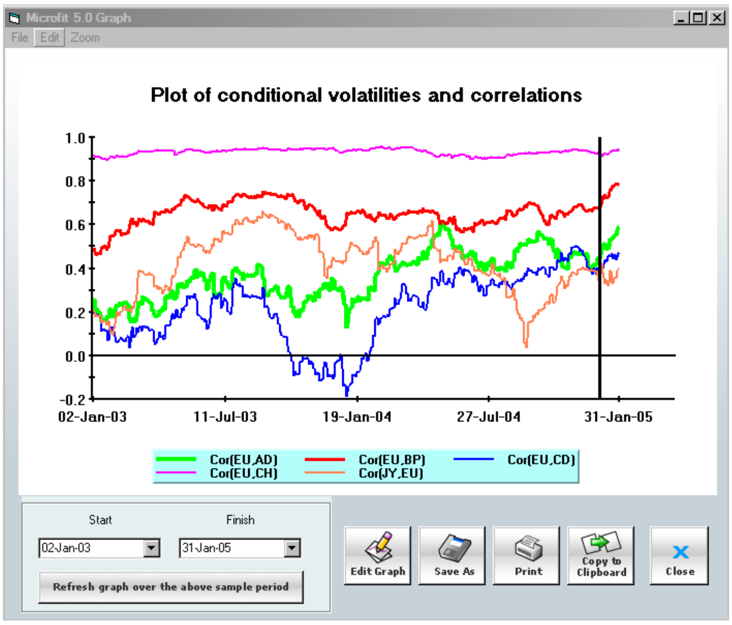


Figure 20.7: Forecast of conditional correlations of Euro with other currency futures returns (02-Jan-03 to 31-Jan-05)

period as

$$02\text{-}Jan\text{-}95 \quad 31\text{-}Dec\text{-}04$$

and click . In the Multivariate $GARCH$ window select the 'Unrestricted' option in the panels 'Decay factor for the variance' and 'Decay factor for the covariance'. Make sure that the option 'Same for all assets' is unchecked, choose the multivariate t-distribution, and select the 'Rolling historical volatility' option with rolling window equal to 20. Click to accept the default number of observations (20) used to initialize the estimation, and the default initial values for the parameters. The results are reported in Table 20.7.

Table 20.7: ML estimates of t-DCC model on OLS residuals

```
        Multivariate GARCH with underlying multivariate t-distribution
                       Converged after 23 iterations
*******************************************************************************
 Based on  2610 observations from 02-Jan-95 to 31-Dec-04.
 The underlying multivariate GARCH model is:
 AD AD(-1) C;    BP BP(-1) C;CD CD(-1) C;CH CH(-1) C;EU EU(-1) C;JY J
 Y(-1) C
 Volatility decay factors unrestricted, different for each variable.
 Correlation decay factors unrestricted, same for all variables.
*******************************************************************************
 Parameter                 Estimate       Standard Error        T-Ratio[Prob]
 lambda1_AD                 .96586          .0063843           151.2865[.000]
 lambda1_BP                 .96298          .0088887           108.3369[.000]
 lambda1_CD                 .96867          .0056633           171.0441[.000]
 lambda1_CH                 .95929          .0066766           143.6802[.000]
 lambda1_EU                 .96534          .0055704           173.2986[.000]
 lambda1_JY                 .95081          .0092300           103.0133[.000]
 lambda2_AD                .023984          .0037262             6.4364[.000]
 lambda2_BP                .026111          .0050162             5.2054[.000]
 lambda2_CD                .028653          .0046781             6.1249[.000]
 lambda2_CH                .028962          .0040915             7.0785[.000]
 lambda2_EU                .026124          .0036538             7.1498[.000]
 lambda2_JY                .040500          .0067245             6.0228[.000]
 delta1                     .97110          .0023914           406.0736[.000]
 delta2                    .019912          .0013541            14.7049[.000]
 df                         5.8783            .22390            26.2540[.000]
*******************************************************************************
 Maximized Log-Likelihood =     -8838.3
*******************************************************************************
df is the degrees of freedom of the multivariate t distribution

                  Estimated Unconditional Volatility Matrix
    Unconditional Volatilities (Standard Errors) on the Diagonal Elements
          Unconditional Correlations on the Off-Diagonal Elements

*******************************************************************************
                   AD        BP        CD        CH        EU        JY
 AD             .68054    .27724    .36486    .26361    .30311    .23695

 BP             .27724    .50887    .13172    .59119    .60307    .28617

 CD             .36486    .13172    .38567    .11827    .14284    .10872

 CH             .26361    .59119    .11827    .69750    .92233    .41934

 EU             .30311    .60307    .14284    .92233    .63368    .37887

 JY             .23695    .28617    .10872    .41934    .37887    .74985

*******************************************************************************
 For the time-varying conditional volatilities and correlations see the Post
 Estimation Menu.
```

The asset-specific estimates of the volatility and correlation decay parameters are highly significant and very close to the estimates reported in Table 20.2. You can now inspect regression results for each equation by closing the output window and choosing option 8. Select, for example, the equation for the Euro futures returns (EU). You will be presented with Table 20.8. The estimated parameter for the temporal lag of the variable EU is close to zero (-0.042), though significant at the 5 per cent significance level.

Table 20.8: Estimation of the *EU* equation

```
                        Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is EU
 2630 observations used for estimation from 05-Dec-94 to 31-Dec-04
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 EU(-1)                   -.041792            .019494            -2.1438[.032]
 C                        .0020226            .012351             .16375[.870]
*******************************************************************************
 R-Squared                    .0017458   R-Bar-Squared                 .0013659
 S.E. of Regression             .63342   F-Stat.    F(1,2628)     4.5960[.032]
 Mean of Dependent Variable   .0019350   S.D. of Dependent Variable      .63385
 Residual Sum of Squares        1054.4   Equation Log-likelihood        -2529.9
 Akaike Info. Criterion        -2531.9   Schwarz Bayesian Criterion     -2537.8
 DW-statistic                   1.9986
*******************************************************************************

                               Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *           F Version        *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(1)  =    .22012[.639]*F(1,2627)    =    .21989[.639]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(1)  =    1.1465[.284]*F(1,2627)    =    1.1457[.285]*
*                     *                          *                            *
* C:Normality         *CHSQ(2)  = 302.7795[.000]*       Not applicable       *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(1)  =   13.5764[.000]*F(1,2628)    =   13.6365[.000]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

20.7 Exercises in Multivariate GARCH Estimation

20.7.1 Exercise 20.1

Repeat Lesson 20.1 using exponentially weighted returns. Check to see if the conclusions obtained are robust to the way returns are standardized.

20.7.2 Exercise 20.2

Use the special *Microfit* file FUTURESDATA.FIT to test whether conditional correlations of Euro with other currency futures returns are mean reverting, both in a Gaussian and *t-DCC* model (see Lesson 20.3).

20.7.3 Exercise 20.3

Repeat the lessons in this chapter using the daily returns on bonds (BU, BE, BG, BJ) and equities (SP, FTSE, DAX, CAC, NK) futures included in FUTURESDATA.FIT. What are the main similarities and differences in the *MGARCH* models for the three types of assets?

Part V

Econometric Methods

Chapter 21

Econometrics of Single Equation Models

This chapter provides the technical details of the econometric methods and the algorithms that underlie the computation of the various estimators and test statistics in the case of single equation models. It complements Chapter 6, which describes the estimation options in *Microfit* for single-equation models. Text-book treatments of some of the topics covered here can be found in Harvey (1981), Amemiya (1985), Judge, Griffiths, Hill, Lütkepohl, and Lee (1985), Godfrey (1988), Maddala (1988), Chatfield (2003), Greene (2002), Davidson and MacKinnon (1993), and Hamilton (1994).

21.1 Summary statistics and autocorrelation coefficients

The **COR** command applied to the observation x_t, $t = 1, 2, .., n$, computes the following statistics:

$$\text{Sample mean} = \bar{x} = \sum_{t=1}^{n} x_t/n,$$

$$\text{Standard deviation} = S_x = \sqrt{\left(\sum_{t=1}^{n}(x_t - \bar{x})^2 \,/ (n-1)\right)}$$

$$\text{Coefficient of variation} = \text{Absolute value of } (S_x/\bar{x})$$

$$\text{Skewness } = \sqrt{b_1} = m_3/m_2^{3/2}$$

$$\text{Kurtosis} = b_2 = m_4/m_2^2$$

where

$$m_k = \sum_{i=1}^{n}(x_t - \bar{x})^k/n, \qquad k = 2, 3, 4$$

The program displays $\sqrt{b_1}$ and b_2-3. These estimates can be used to construct different tests of departures from normality. A non-normal distribution which is asymmetrical has a value of

$\sqrt{\beta_1}$ (the population value of $\sqrt{b_1}$), which is non-zero. $\sqrt{\beta_1} > 0$ indicates the skewness to the right, and $\sqrt{\beta_1} < 0$ indicates skewness to the left. The measure of Kurtosis (or curvature), usually denoted by β_2, is equal to 3 for the normal distribution. For unimodal non-normal distributions with thicker tails than normal, we have $\beta_2 - 3 > 0$, and for distributions with thinner tails than normal, we have $\beta_2 - 3 < 0$. (See, for example, D'Agostino et al. (1990) for further details.)

The Jarque-Bera test of the normality of regression residuals in Section 21.6.2 can also be computed using the **COR** command, and is given by

$$\chi^2_N(2) = n\left\{\tfrac{1}{6}b_1 + \tfrac{1}{24}(b_2 - 3)^2\right\}$$

where $\sqrt{b_1}$ and b_2 are computed by applying the **COR** command to the OLS residuals. (The above expression assumes that an intercept term is included in the regression.)

$$\text{The } l\text{th order autocorrelation coefficient} = R_l = C_l/C_0, \qquad l = 1, 2, ..., [n/3]$$

$$C_l = n^{-1}\sum_{t=l+1}^{n}(x_t - \bar{x})(x_{t-l} - \bar{x})$$

The program also computes an approximate estimate of the standard error of R_l, using Bartlett (1946) approximation reported in Kendall, Stuart, and Ord (1983), Chapter 48, p. 549:

$$\text{Standard Error of } R_l = \sqrt{\left[\frac{1}{n}\left(1 + 2\sum_{j=1}^{l-1} R_j^2\right)\right]}, \qquad l = 1, 2, ..., [n/3]$$

The Box and Pierce (1970) Q-statistic (of order p):

$$Q = n\sum_{j=1}^{p} R_j^2 \overset{a}{\sim} \chi^2_p$$

The Ljung and Box (1978) statistic (of order p):[1]

$$Q^* = n(n+2)\sum_{j=1}^{p} R_j^2/(n-j) \overset{a}{\sim} \chi^2_p$$

21.1.1 Box-Pierce and Ljung-Box tests

Under the assumption that x_t are serially uncorrelated, the Box-Pierce and the Ljung-Box statistics are both distributed asymptotically as χ^2 variates with p degrees of freedom. The two tests are asymptotically equivalent, although the Ljung-Box statistic is likely to perform better in small samples. See, for example, Harvey (1981) p. 211, and Chapters 48 and 50 in Kendall, Stuart, and Ord (1983), for more details.

[1]The symbol $\overset{a}{\sim}$ denotes 'approximately distributed as', and χ^2_p denotes the chi-squared variate with p degrees of freedom.

When the **COR** command is followed by more than one variable, say $x_{1t}, x_{2t}, ..., x_{kt}$, $t = 1, 2, ..., n$, the program computes the correlation matrix of these variables over the specified sample period using the formula

$$\begin{aligned}\hat{\rho}_{ij} &= \text{estimated (or sample) correlation coefficient of } x_i \text{ and } x_j \\ &= \textstyle\sum_{t=1}^{n}(x_{it}-\bar{x}_i)(x_{jt}-\bar{x}_j)/(n-1)S_iS_j, \qquad i,j=1,2,...,k\end{aligned}$$

where $\bar{x}_i$ and S_i are the mean and the standard deviation of x_{it}, $t = 1, 2, ..., n$, respectively.

21.2 Non-parametric estimation of the density function

The non-parametric functions available in *Microfit 5.0* provide estimates of the density function $f(\cdot)$ of a set of data, $x_1, x_2, ..., x_n$, using various kernel functions and window widths. All functions generate the estimate $\hat{f}$ at the point x using the formula

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{N}\frac{1}{h_n}K\left(\frac{x-x_i}{h_n}\right)$$

where $K(\cdot)$ is the kernel function, and h_n is the window width, also called the smoothing parameter or bandwidth. The following kernel functions are available in *Microfit*:

1. *Gaussian*
$$K(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

2. *Epanechnikov*
$$K(x) = \begin{cases} \frac{3}{4}\left(1-\frac{1}{5}x^2\right)/\sqrt{5}, & \text{if } |x| < \sqrt{5} \\ 0, & \text{otherwise}\end{cases}$$

As for the choice of h_n, *Microfit* allows the following options:

1. *Silverman rule of thumb:*
$$h_{srot} = 0.9An^{-\frac{1}{5}} \tag{21.1}$$
where $A = \min(\sigma, R/1.34)$, σ is the standard deviation of the variable x, R is the interquartile range, and n is the number of observations. See Silverman (1986), p.47.

2. *Least squares cross-validation*: the window width is the value, h_{lscv}, that minimizes the following criterion
$$ISE(h_n) = \frac{1}{n^2h_n}\sum_{i\neq j}^{n}K_2\left(\frac{x_i-x_j}{h_n}\right) - \frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(x_i), \tag{21.2}$$
where $K_2(.)$ is the convolution of the kernel with itself, defined by
$$K_2(x) = \int_{-\infty}^{+\infty}K(t)K(x-t)\,dt.$$

and $\hat{f}_{-i}(x_i)$ is the density estimator obtained after omitting the i^{th} observation. We have

$$\frac{1}{n}\sum_{i=1}^{n}\hat{f}_{-i}(x_i) = \frac{1}{n(n-1)h_n}\sum_{i\neq j}^{n}K\left(\frac{x_i - x_j}{h_n}\right).$$

If K is the Gaussian kernel, then K_2 is $N(0,2)$, or

$$K_2(x) = (4\pi)^{-1/2}e^{-x^2/4},$$

for the *Epanechnikov kernel* we have

$$K_2(x) = \begin{cases} \frac{3\sqrt{5}}{100}\left(4 - x^2\right), & \text{if } |x| < \sqrt{5} \\ 0, & \text{otherwise.} \end{cases}$$

For the Gaussian kernel the expression for $ISE(h_n)$ simplifies to (see Bowman and Azzalini (1997), p.37)

$$\begin{aligned} ISE(h_n) &= \frac{1}{(n-1)}\phi\left(0, \sqrt{2}h_n\right) + \frac{n-2}{n(n-1)^2}\sum_{i\neq j}^{n}\phi\left(x_i - x_j, \sqrt{2}h_n\right) \\ &\quad - \frac{2}{n(n-1)}\sum_{i\neq j}^{n}\phi(x_i - x_j, h_n), \end{aligned}$$

where $\phi(x,\sigma)$ denotes the normal density function with mean 0 and standard deviation σ:

$$\phi(x,\sigma) = (2\pi\sigma^2)^{-1/2}\exp\left(\frac{-x^2}{2\sigma^2}\right).$$

In cases where local minima is encountered we select the bandwith that corresponds to the local minimum with the largest value for h_n. See Bowman and Azzalini (1997), p.33-34.

See also Sheather (2004), Pagan and Ullah (1999), Silverman (1986) and Jones, Marron, and Sheather (1996) for further details.

21.3 Estimation of spectral density

The **SPECTRUM** command computes the estimates of the standardized spectrum of x_t multiplied by π, for the n observations $x_1, x_2, ..., x_n$, using the formula

$$\hat{f}(\omega_j) = \left(\lambda_0 + 2\sum_{k=1}^{m}\lambda_k R_k\cos(k\omega_j)\right)$$

where $\omega_j = j\pi/m,\ j = 0, 1, ..., m$, m is the 'window size', R_k is the autocorrelation coefficient of order k defined by

$$R_k = \left(\sum_{t=k+1}^{n}(x_t - \bar{x})(x_{t-k} - \bar{x})\right)\Bigg/\left(\sum_{t=1}^{n}(x_t - \bar{x})^2\right)$$

and $\{\lambda_k\}$ is a set of weights called the 'lag window'. The program computes estimates of $\hat{f}(\omega_j)$, $j = 0, 1, ..., m$ for the following lag windows:

$$\begin{array}{rcll} \textit{Bartlett window} & \lambda_k & = & 1 - k/m, \qquad 0 \leq k \leq m \\ \textit{Tukey window} & \lambda_k & = & \frac{1}{2}\left\{1 + \cos(\pi k/m)\right\}, \qquad 0 \leq k \leq m \\ \textit{Parzen window} & \lambda_k & = & \left\{ \begin{array}{ll} 1 - 6(k/m)^2 + 6(k/m)^3, & 0 \leq k \leq \frac{m}{2} \\ 2(1 - k/m)^3, & \frac{m}{2} \leq k \leq m \end{array} \right\} \end{array}$$

The default value for m is set equal to $2\sqrt{n}$.

The standard errors reported for the estimates of the standardized spectrum are calculated according to the following formulae, which are valid asymptotically:

$$\begin{aligned} \widehat{S.E.}(\hat{f}(\omega_j)) &= \sqrt{\tfrac{2}{v}}\hat{f}(\omega_j), \qquad \text{for } j = 1, 2, ..., m-1 \\ &= \sqrt{\tfrac{4}{v}}\hat{f}(\omega_j), \qquad \text{for } j = 0, m \end{aligned}$$

where $v = 2n/\sum_{k=-m}^{m}(\lambda_k^2)$. For the three different windows, v is given by

$$\begin{array}{rcll} \textit{Bartlett window} & v & = & 3n/m \\ \textit{Tukey window} & v & = & 8n/3m \\ \textit{Parzen window} & v & = & 3.71n/m \end{array}$$

The spectrum for the residuals is estimated using the Parzen window and does not display the standard error bands for the estimates.

For an introductory text on the estimation of the spectrum see Chatfield (2003), Chapter 7. For more advanced treatments of the subject see Priestley (1981) Chapter 6 and Brockwell and Davis (1991), Chapter 10.

21.4 Hodrick-Prescott (HP) filter

The HP filter is a curve fitting procedure proposed by Hodrick and Prescott (1997), to estimate the trend path, $\{y_t^*, \ t = 1, 2, ..., n\}$ of a series $\{y_t, \ t = 1, 2, ..., n\}$ subject to the constraint that the sum of the squared second differences of the trend series is not too large. More specifically, $\{y_t^*, \ t = 1, 2, ..., n\}$ is computed from $\{y_t, \ t = 1, 2, ..., n\}$ by solving the following optimization problem:

$$\min_{y_1^*, y_2^*, \ldots y_n^*} \left\{ \sum_{t=1}^{n}(y_t - y_t^*)^2 + \lambda \sum_{t=2}^{n-1}(\Delta^2 y_{t+1}^*)^2 \right\}$$

The 'smoothing' parameter λ is usually chosen by trial and error, and for quarterly observations is set to 1,600. For a discussion of the statistical properties of the HP filter, see, for example, Cogley (1995), Harvey and Jaeger (1993) and Söderlind (1994). In particular, Harvey and Jaeger show that the use of the HP filter can generate spurious cyclical patterns.

21.5 Pesaran-Timmermann non-parametric test of predictive performance

Let $x_t = \hat{E}(y_t \,|\Omega_{t-1})$ be the predictor of y_t found with respect to the information set Ω_{t-1}, and suppose that the observations (y_1, x_1), $(y_2, x_2), ..., (y_n, x_n)$ are available on these variables. The test proposed by Pesaran and Timmermann (1992), and Pesaran and Timmermann (1994) is based on the proportion of times that the direction of change in y_t is correctly predicted by x_t. The test statistic is computed as

$$S_n = \frac{\hat{P} - \hat{P}_*}{\left\{\hat{V}(\hat{P}) - \hat{V}(\hat{P}_*)\right\}^{\frac{1}{2}}} \overset{a}{\sim} N(0,1)$$

where

$$\begin{aligned}
\hat{P} &= n^{-1}\sum_{t=1}^{n} \text{Sign}(y_t x_t), \ \ \hat{P}_* = \hat{P}_y \hat{P}_x + (1-\hat{P}_y)(1-\hat{P}_x) \\
\hat{V}(\hat{P}) &= n^{-1}\hat{P}_*(1-\hat{P}_*) \\
\hat{V}(\hat{P}_*) &= n^{-1}(2\hat{P}_y - 1)^2 \hat{P}_x (1-\hat{P}_x) + n^{-1}(2\hat{P}_x - 1)^2 \hat{P}_y (1-\hat{P}_y) \\
&\quad +4n^{-2}\hat{P}_y \hat{P}_x (1-\hat{P}_y)(1-\hat{P}_x) \\
\hat{P}_y &= n^{-1}\sum_{t=1}^{n} \text{Sign}(y_t), \ \ \hat{P}_x = n^{-1}\sum_{t=1}^{n} \text{Sign}(x_t)
\end{aligned}$$

and $\text{Sign}(A)$ is the sign (or the indicator) function that takes the value of unity if $A > 0$ and zero otherwise.

Under the null hypothesis that y_t and x_t are distributed independently (namely x_t has no power in prediction y_t), S_n is asymptotically distributed as a standard normal. This test does not require quantitative information on y and uses only information on the signs of y_t and x_t. The test statistic is undefined when $\hat{P}_y$ or $\hat{P}_x$ take the extreme values of zero or unity.

21.6 Ordinary least squares estimates

The estimates computed using the *OLS* option are based on the following linear regression model:

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + u_t, \ t = 1, 2, ..., n \tag{21.3}$$

where y_t is the dependent variable, $\boldsymbol{\beta}$ is a $k \times 1$ vector of unknown coefficients, $\mathbf{x}_t$ is the $k \times 1$ vector of regressors, and u_t is a disturbance term, assumed here to satisfy the classical normal assumptions (A1 to A5), set out in Section 6.1. Writing the n relations in (21.3) in vector and matrix notations we have

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{21.4}$$

where $\mathbf{y}$ is the $n \times 1$ vector of observations on the dependent variable, $\mathbf{X}$ is the $n \times k$ matrix of observations on the regressors (usually including an intercept term), and $\mathbf{u}$ is the $n \times 1$ vector of disturbances (errors). We shall also assume that $\mathbf{X}$ has a full column rank, and therefore $\mathbf{X}'\mathbf{X}$ has an inverse.

21.6.1 Regression results

The program computes the OLS estimates using the following formulae:

$$\hat{\boldsymbol{\beta}}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{21.5}$$

$$\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_{OLS}) = \hat{\sigma}^2_{OLS}(\mathbf{X}'\mathbf{X})^{-1} \tag{21.6}$$

where $\hat{\sigma}_{OLS}$ is the Standard Error (SE) of the regression given by

$$\hat{\sigma}^2_{OLS} = (y - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS})/(n-k) \tag{21.7}$$

$$\text{Fitted values} = \hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS} \tag{21.8}$$

$$\text{Residuals} = \mathbf{e}_{OLS} = \mathbf{y} - \hat{\mathbf{y}} \tag{21.9}$$

A typical element of $\mathbf{e}_{OLS}$ is given by

$$e_t = y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}_{OLS}, \qquad t = 1, 2, ..., n \tag{21.10}$$

where $\mathbf{x}_t$ is the $k \times 1$ vector of the regressors observed at time t

$$RSS = \text{residual sum of squares} = \sum_{t=1}^{n} e_t^2 = \mathbf{e}'_{OLS}\mathbf{e}_{OLS}$$

$$\text{mean of the dependent variable} = \bar{y} = \sum_{t=1}^{n} y_t/n.$$

$$TSS = \text{total sum of squares} = S_{yy} = \sum_{t=1}^{n}(y_t - \bar{y})^2$$

$$\text{standard deviation } (SD) \text{ of the dependent variable} = \hat{\sigma}_y$$

$$= \sqrt{S_{yy}/(n-1)} \tag{21.11}$$

$$R\text{-squared} = R^2 = 1 - (RSS/TSS) \tag{21.12}$$

$$R\text{-bar-squared} = \bar{R}^2 = 1 - (\hat{\sigma}_{OLS}/\hat{\sigma}_y)^2 \tag{21.13}$$

Notice also that we have the following relationship between R^2 and $\bar{R}^2$:

$$1 - \bar{R}^2 = \left(\frac{n-1}{n-k}\right)(1 - R^2)$$

$$\text{Maximized Value of the Log-Likelihoood Function} \tag{21.14}$$
$$= LL_{OLS} = \frac{-n}{2}\left\{1 + \log(2\pi\tilde{\sigma}^2)\right\}$$

where $\widetilde{\sigma}^2 = \mathbf{e}'\mathbf{e}/n$, is the Maximum Likelihood (ML) estimator of σ^2. Notice that unlike $\hat{\sigma}^2_{OLS}$, which is an unbiased estimator of σ^2 under the classical assumptions, $\widetilde{\sigma}^2$ is biased. But for large enough sample sizes, the two estimators ($\hat{\sigma}^2_{OLS}$ and $\tilde{\sigma}^2$) are equivalent.

In addition to the above statistics, the regression results also include the following test statistics and model selection criteria:

$$F\text{-statistic} = \left(\frac{R^2}{1-R^2}\right)\left(\frac{n-k}{k-1}\right) \sim F(k-1, n-k) \tag{21.15}$$

which is appropriate only when the regression equation includes an intercept (or a constant) term. Under classical assumptions the F-statistic can be used to test the statistical significance of the included regressors other than the intercept term. The F-statistic is distributed as F with $k-1$ and $n-k$ degrees of freedom under the null hypothesis that all the regression coefficients, except for the intercept terms, are zero.

$$DW \text{ statistic} = DW = \sum_{t=2}^{n}(e_t - e_{t-1})^2 \Big/ \sum_{t=1}^{n} e_t^2 \tag{21.16}$$

This is the familiar Durbin and Watson (1950), Durbin and Watson (1951) statistic for testing residual serial correlation. It is valid only when lagged values of the dependent variable are *not* included amongst the regressors.

In the case where the regressors include a single one-period lag of the dependent variable (y_{t-1}), and this is specified explicitly at the estimation stage, we have

$$y_t = \lambda y_{t-1} + \boldsymbol{\beta}'\mathbf{x}_t + u_t \tag{21.17}$$

the program also reports the Durbin (1970) h-statistic

$$h\text{-statistic} = h = \left(1 - \frac{DW}{2}\right)\sqrt{\left\{n/[1 - n\hat{V}(\hat{\lambda})]\right\}} \tag{21.18}$$

where $\hat{V}(\hat{\lambda})$ is the estimated variance of the OLS estimator of the coefficient of the lagged dependent variable, λ.[2] In situations where $n\hat{V}(\hat{\lambda}) \geq 1$, the h-statistic is not defined, and it will be set equal to $*NONE*$ by the program. Under the null hypothesis of non-autocorrelated disturbances, the h-statistic is asymptotically distributed as a standardized normal variate. For a two-sided test, a value of h, exceeding 1.96 (the 5 per cent critical value of the standard normal distribution), in absolute value indicates rejection of the null hypothesis that the disturbances u_t in (21.17) are serially uncorrelated.

[2] Notice also that $1 - (DW/2) \approx r_1$, where r_1 is the first order autocorrelation of the OLS residuals. For more detail see, for example, Harvey (1981), pp. 274-275.

21.6.2 Diagnostic test statistics (the OLS case)

In the case of the *OLS* option the diagnostic statistics are computed according to the following formulae.

Godfrey's test of residual serial correlation

The Lagrange multiplier (LM) version of the test statistic is computed using the formula (see Godfrey (1978b), Godfrey (1978c))

$$\chi^2_{SC}(p) = n\left(\frac{\mathbf{e}'_{OLS}\mathbf{W}(\mathbf{W}'\mathbf{M}_x\mathbf{W})^{-1}\mathbf{W}'\mathbf{e}_{OLS}}{\mathbf{e}'_{OLS}\mathbf{e}_{OLS}}\right) \overset{a}{\sim} \chi^2_p \tag{21.19}$$

where

$$\mathbf{M}_x = \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

$$\mathbf{e}_{OLS} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS} = (e_1, e_2, ..., e_n)'$$

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & ... & 0 \\ e_1 & 0 & ... & 0 \\ e_2 & e_1 & ... & 0 \\ \vdots & e_2 & & \vdots \\ \vdots & \vdots & & e_{n-p-1} \\ e_{n-1} & e_{n-2} & ... & e_{n-p} \end{bmatrix} \tag{21.20}$$

and p is the order of the error process.

The F-version of (21.19) is given by[3]

$$F_{SC}(p) = \left(\frac{n-k-p}{p}\right)\left(\frac{\chi^2_{SC}(p)}{n-\chi^2_{SC}(p)}\right) \overset{a}{\sim} F_{p,n-k-p} \tag{21.21}$$

The expression for $\chi^2_{SC}(p)$ is already given by (21.19). The above statistic can also be computed as the F-statistic for the (joint) test of zero restrictions on the coefficients of $\mathbf{W}$ in the auxiliary regression

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{W}\boldsymbol{\delta} + \mathbf{v}$$

Harvey (1981) p. 173, refers to the F-version of the LM statistic (21.19) as the 'modified LM' statistic. The two versions of the test of residual serial correlation, namely $\chi^2_{SC}(p)$ and $F_{SC}(p)$, are asymptotically equivalent.

[3]For a derivation of the relationship between the LM-version, and the F-version of the test statistics see, for example, Pesaran (1981) pp. 78-80.

Ramsey's RESET test of functional form

The form of Ramsey's $RESET$ test statistic is the same as those given by (21.19)and (21.21), for the LM- and the F-versions, respectively (Anscombe (1961), Ramsey (1969), and Ramsey (1970)). In the case of the $RESET$ test, the columns of $\mathbf{W}$ are set equal to the powers of fitted values, $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS}$. In the Diagnostic Tests Table, the statistics reported for the $RESET$ test are computed for the simple case where the elements of $\mathbf{W}$ are specified to be equal to the square of fitted values. That is

$$\mathbf{W} = (\hat{y}_1^2, \hat{y}_2^2, ..., \hat{y}_n^2)'$$

Higher-order $RESET$ tests can be computing using the variable addition test in the Hypothesis Testing Menu, using $\hat{y}_t^2, \hat{y}_t^3, ..., \hat{y}_t^p$ as the additional variables.

Jarque-Bera's test of the normality of regression residuals

The LM version of the statistic for the normality test is given by

$$\begin{aligned} \chi_N^2(2) &= n\left\{\mu_3^2/(6\mu_2^3) + (1/24)(\mu_4/\mu_2^2 - 3)^2\right\} \\ &\quad + n\left\{3\mu_1^2/(2\mu_2) - \mu_3\mu_1/\mu_2^2\right\} \overset{a}{\sim} \chi^2(2) \end{aligned} \tag{21.22}$$

where

$$\mu_j = \sum_{t=1}^{n} e_t^j/n, \qquad j = 1, 2, ...$$

Notice that in situations where an intercept term is included in the regression, $\mu_1 = 0$ (see Jarque and Bera (1980) and Bera and Jarque (1981)).

Test of homoscedasticity

The statistics reported for this test (the equality of error-variances) are based on the auxiliary regression

$$e_t^2 = constant + \alpha\hat{y}_t^2 \tag{21.23}$$

and gives the LM- and the F-statistics for the test of $\alpha = 0$, against $\alpha \neq 0$.

Predictive failure test

Consider the following linear regression models specified for each of two sample periods:

$$\mathbf{y}_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{u}_1; \qquad \mathbf{u}_1 \sim N(0, \sigma_1^2\mathbf{I}_{n1}) \tag{21.24}$$

$$\mathbf{y}_2 = \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{u}_2; \qquad \mathbf{u}_2 \sim N(0, \sigma_2^2\mathbf{I}_{n2}) \tag{21.25}$$

where $\mathbf{y}_i, \mathbf{X}_i, i = 1, 2$, are $n_i \times 1$ and $n_i \times k$ observation matrices on the dependent variable and the regressors for the two sample periods, and $\mathbf{I}_{n_1}$and $\mathbf{I}_{n_2}$ are identity matrices of order

n_1 and n_2, respectively. Combining (21.24) and (21.25) by stacking the observations on the two sample periods now yields

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{X}_2 & \mathbf{I}_{n2} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\delta} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

The above system of equations may also be written more compactly as

$$\mathbf{y}_0 = \mathbf{X}_0\boldsymbol{\beta}_1 + \mathbf{S}_2\boldsymbol{\delta} + \mathbf{u}_0 \tag{21.26}$$

where $\mathbf{y}_0 = (\mathbf{y}_1', \mathbf{y}_2')', \mathbf{X}_0 = (\mathbf{X}_1', \mathbf{X}_2')'$, and $\mathbf{S}_2$ represents the $(n_1 + n_2) \times n_2$ matrix of n_2 dummy variables, one dummy variable for each observation in the second period. The predictive failure test can now be carried out by testing the hypothesis of $\boldsymbol{\delta} = 0$ against $\boldsymbol{\delta} \neq 0$ in (21.26). This yields the following F-statistic:

$$F_{PF} = \frac{(\mathbf{e}_0'\mathbf{e}_0 - \mathbf{e}_1'\mathbf{e}_1)/n_2}{\mathbf{e}_1'\mathbf{e}_1/(n_1 - k)} \sim F(n_2, n_1 - k) \tag{21.27}$$

where

$\mathbf{e}_0$ is the OLS residual vector of the regression of $\mathbf{y}_0$ on $\mathbf{X}_0$ (based on the first and the second sample periods together).

$\mathbf{e}_1$ is the OLS residual vector of the regression of $\mathbf{y}_1$ on $\mathbf{X}_1$ (based on the first sample period).

Under the classical normal assumptions, the predictive failure test statistic, F_{PF}, has an exact F-distribution with n_2 and $n_1 - k$ degrees of freedom.

The LM version of the above statistic is computed as

$$\chi^2_{PF} = n_2 F_{PF} \overset{a}{\sim} \chi^2(n_2). \tag{21.28}$$

which is distributed as a chi-squared with n_2 degrees of freedom for large n_1 (see Chow (1960), Salkever (1976), Dufour (1980) and Pesaran, Smith, and Yeo (1985) Section III).

A test of the stability of the regression coefficients: the Chow test

This is the first test proposed by Chow (1960) and is aimed at testing the hypothesis that in (21.24) and (21.25) $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$, conditional on equality of variances ($\sigma_1^2 = \sigma_2^2$). The Chow test is also known as the analysis of covariance test (see Scheffe (1959)). The F-version of the Chow test statistic is defined by

$$F_{SS} = \frac{(\mathbf{e}_0'\mathbf{e}_0 - \mathbf{e}_1'\mathbf{e}_1 - \mathbf{e}_2'\mathbf{e}_2)/k}{(\mathbf{e}_1'\mathbf{e}_1 + \mathbf{e}_2'\mathbf{e}_2)/(n_1 + n_2 - 2k)} \sim F(k, n_1 + n_2 - 2k) \tag{21.29}$$

where

$\mathbf{e}_0$ is the OLS residual vector for the first two sample periods together

$\mathbf{e}_1$ is the OLS residual vector for the first sample period

$\mathbf{e}_2$ is the OLS residual vector for the second sample period

The LM version of this test statistic is computed as

$$\chi^2_{SS} = kF_{SS} \overset{a}{\sim} \chi^2(k) \tag{21.30}$$

For more details see, for example, Pesaran, Smith, and Yeo (1985), p. 285.

21.7 Statistical model selection criteria

Model selection in econometric analysis involves both statistical and non-statistical considerations. It depends on the objective(s) of the analysis, the nature and the extent of economic theory used, and the statistical adequacy of the model under consideration compared with other econometric models. (For a discussion of the general principles involved in model selection see Pesaran and Smith (1985)). The various choice criteria reported by *Microfit* are only concerned with the issue of 'statistical fit' and provide different approaches to trading-off 'fit' and 'parsimony' of a given econometric model.

The program automatically computes Theil's $\bar{R}^2$ criterion for choosing between linear (and non-linear) regression models estimated by least squares, and the GR^2 criterion proposed by Pesaran and Smith (1994) for choosing between linear and non-linear single equation regression models estimated by the instrumental variables method. See Sections 21.6.1 and 21.10.4 for details.

In addition *Microfit* computes the values of the criterion function proposed by Akaike (1973), Akaike (1974), Schwarz (1978), and Hannan and Quinn (1979), both for single and multi-equation models. All these three model selection criteria measure the 'fit' of a given model by its maximised value of the log-likelihood function, and then use different penalty functions to take account of the fact that different number of unknown parameters may have been estimated for different models under consideration.

21.7.1 Akaike information criterion (AIC)

Let $\ell_n(\widetilde{\boldsymbol{\theta}})$ be the maximized value of the log-likelihood function of an econometric model, where $\widetilde{\boldsymbol{\theta}}$ is the maximum likelihood estimator of $\widetilde{\boldsymbol{\theta}}$, based on a sample of size n. The Akaike information criterion (AIC) for this model is defined as

$$AIC_\ell = \ell_n(\widetilde{\boldsymbol{\theta}}) - p \tag{21.31}$$

where

$$p \equiv \text{ Dimension } (\boldsymbol{\theta}) \equiv \text{The number of freely estimated parameters}$$

In the case of single-equation linear (or non-linear) regression models the AIC_ℓ can also be written equivalently as

$$AIC_\sigma = \log(\tilde{\sigma}^2) + \frac{2p}{n} \tag{21.32}$$

where $\tilde{\sigma}^2$ is the ML estimator of the variance of regression disturbances, u_t, given by $\tilde{\sigma}^2 = \mathbf{e}'\mathbf{e}/n$ in the case of linear regression models (see Section 21.6.1). The two versions of the AIC in (21.31) and (21.32) yield identical results. When using (21.31), the model with the highest value of AIC_ℓ is chosen. But when using the criterion based on the estimated standard errors (21.32), the model with the lowest value for AIC_σ is chosen.[4]

21.7.2 Schwarz Bayesian criterion (SBC)

The SBC provides a large sample approximation to the posterior odds ratio of models under consideration. It is defined by

$$SBC_\ell = \ell_n(\widetilde{\boldsymbol{\theta}}) - \tfrac{1}{2}p\log n \tag{21.33}$$

In applying the SBC across models, the model with the highest SBC value is chosen. For regression models an alternative version of (21.33), based on the estimated standard error of the regression, $\tilde{\sigma}$, is given by

$$SBC_\sigma = \log(\tilde{\sigma}^2) + \left(\frac{\log n}{n}\right)p$$

According to this criterion, a model is chosen if it has the lowest SBC_σ value.

21.7.3 Hannan and Quinn criterion (HQC)

This criterion has been primarily proposed for selection of the order of autoregressive-moving average or vector autoregressive models, and is defined by

$$HQC_\ell = \ell_n(\widetilde{\boldsymbol{\theta}}) - (\log\log n)p$$

or equivalently (in the case of regression models)

$$HQC_\sigma = \log\tilde{\sigma} + \left(\frac{2\log\log n}{n}\right)p$$

21.7.4 Consistency properties of the different model selection criteria

Among the above three model selection criteria, the SBC selects the most parsimonious model (a model with the least number of freely estimated parameters) if $n \geq 8$, and the AIC selects the least parsimonious model. The HQC lies somewhere between the other

[4] For linear regression models the equivalence of (21.31) and (21.32) follows by substituting for $\ell_n(\hat{\theta})$ given by (21.14) in (21.31):

$$AIC_\ell = -\frac{n}{2}(1+\log 2\pi) - \frac{n}{2}\log\tilde{\sigma}^2 - p,$$

hence using (21.32)

$$AIC_\ell = -\frac{n}{2}(1+\log 2\pi) - \frac{n}{2}AIC_\sigma.$$

Therefore, in the case of regression models estimated on the same sample period, the same preference ordering across models will result, irrespective of whether AIC_ℓ or AIC_σ criteria are used.

two criteria. Useful discussion of these and other model selection criteria can be found in Amemiya (1980), Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) Chapter 21, and Lütkepohl (2005), Section 4.3. The last reference is particularly useful for selecting the order of the vector autoregressive models, and contains some discussion of the consistency property of the above three model-selection criteria. Under certain regularity conditions it can be shown that SBC and HQC are consistent, in the sense that for large enough samples they lead to the correct model choice, assuming of course that the 'true' model does in fact belong to the set of models over which one is searching. The same is not true of the AIC or Theil's $\bar{R}^2$ criteria. This does not, however, mean that the SBC (or HQC) is necessarily preferable to the AIC or the $\bar{R}^2$ criterion, bearing in mind that it is rarely the case that one is sure that the 'true' model is one of the models under consideration.

21.8 Non-nested tests for linear regression models

Consider the following two linear regression models:

$$M_1 \quad : \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta}_1 + \mathbf{u}_1, \qquad \mathbf{u}_1 \sim N(0, \sigma^2\mathbf{I}_n) \tag{21.34}$$

$$M_2 \quad : \quad \mathbf{y} = \mathbf{Z}\boldsymbol{\beta}_2 + \mathbf{u}_2, \qquad \mathbf{u}_2 \sim N(0, \omega^2\mathbf{I}_n) \tag{21.35}$$

where $\mathbf{y}$ is the $n \times 1$ vector of observations on the dependent variable, $\mathbf{X}$ and $\mathbf{Z}$ are $n \times k_1$ and $n \times k_2$ observation matrices for the regressors of models M_1 and M_2, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are the $k_1 \times 1$ and $k_2 \times 1$ unknown regression coefficient vectors, and $\mathbf{u}_1$ and $\mathbf{u}_2$ are the $n \times 1$ disturbance vectors.

Broadly speaking, models M_1 and M_2 are said to be non-nested if the regressors of M_1 (respectively M_2) cannot be expressed as an exact linear combination of the regressors of M_2 (respectively M_1). For a formal definition of the concepts of nested and non-nested models, see Pesaran (1987a). A review of the literature of non-nested hypothesis testing can be found in McAleer and Pesaran (1986).

The program computes the following statistics for the test of M_1 against M_2 and vice versa.

The N-test

This is the Cox (1961) and Cox (1962) test originally derived in Pesaran (1974) pp. 157-8. The Cox statistic for the test of M_1 against M_2 is computed as

$$N_1 = \left\{\frac{n}{2}\log\left(\hat{\omega}^2/\hat{\omega}_*^2\right)\right\} \Big/ \hat{V}_1 \tag{21.36}$$

where

$$\begin{aligned}
\hat{\omega}^2 &= \mathbf{e}_2'\mathbf{e}_2/n \\
\hat{\omega}_*^2 &= (\mathbf{e}_1'\mathbf{e}_1 + \hat{\boldsymbol{\beta}}_1'\mathbf{X}'\mathbf{M}_2\mathbf{X}\hat{\boldsymbol{\beta}}_1)/n \\
\hat{V}_1^2 &= (\hat{\sigma}^2/\hat{\omega}_*^4)\hat{\boldsymbol{\beta}}_1'\mathbf{X}'\mathbf{M}_2\mathbf{M}_1\mathbf{M}_2\mathbf{X}\hat{\boldsymbol{\beta}}_1 \\
\hat{\sigma}^2 &= \mathbf{e}_1'\mathbf{e}_1/n, \quad \hat{\boldsymbol{\beta}}_1 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y \\
\mathbf{M}_1 &= \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'; \qquad \mathbf{M}_2 = \mathbf{I}_n - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'
\end{aligned}$$

Similarly, the Cox statistic N_2 is also computed for the test of M_2 against M_1.

The NT-test

This is the *adjusted* Cox test derived in Godfrey and Pesaran (1983), p. 138, which is referred to as the $\tilde{N}$-test (or the NT-test). The NT-statistic for the test of M_1 against M_2 is given by the following (see also equations (20) and (21) in Godfrey and Pesaran (1983))

$$\tilde{N}_1 = \tilde{T}_1 \Big/ \sqrt{\left\{\tilde{V}_1(\tilde{T}_1)\right\}} \tag{21.37}$$

where

$$\begin{aligned}
\tilde{T}_1 &= \tfrac{1}{2}(n-k_2)\log(\tilde{\omega}^2/\tilde{\omega}_*^2) \\
\tilde{\omega}^2 &= \mathbf{e}_2'\mathbf{e}_2/(n-k_2), \quad \tilde{\sigma}^2 = \mathbf{e}_1'\mathbf{e}_1/(n-k_1) \\
\tilde{\omega}_*^2 &= \left\{\tilde{\sigma}^2 Tr(\mathbf{M}_1\mathbf{M}_2) + \hat{\boldsymbol{\beta}}_1'\mathbf{X}'\mathbf{M}_2\mathbf{X}\hat{\boldsymbol{\beta}}_1\right\}/(n-k_2) \\
\tilde{V}_1(\tilde{T}_1) &= (\tilde{\sigma}^2/\tilde{\omega}_*^4)\left\{\hat{\boldsymbol{\beta}}_1'\mathbf{X}'\mathbf{M}_2\mathbf{M}_1\mathbf{M}_2\mathbf{X}\hat{\boldsymbol{\beta}}_1 + \tfrac{1}{2}\tilde{\sigma}^2 Tr(\mathbf{B}^2)\right\} \\
Tr(\mathbf{B}^2) &= k_2 - Tr(\mathbf{A}_1\mathbf{A}_2)^2 - \frac{\{k_2 - Tr(\mathbf{A}_1\mathbf{A}_2)\}^2}{n-k_1} \\
\mathbf{A}_1 &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}', \qquad \mathbf{A}_2 = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'
\end{aligned} \tag{21.38}$$

Similarly, the $\tilde{N}$-test statistic, $\tilde{N}_2$, is also computed for the test of M_2 against M_1.

The W-test

This is the Wald-type test of M_1 against M_2 proposed in Godfrey and Pesaran (1983), and is based on the statistic

$$W_1 = \frac{(n-k_2)(\tilde{\omega}^2 - \tilde{\omega}_*^2)}{\left\{2\tilde{\sigma}^4 Tr(\mathbf{B}^2) + 4\tilde{\sigma}^2\hat{\boldsymbol{\beta}}_1'\mathbf{X}'\mathbf{M}_2\mathbf{M}_1\mathbf{M}_2\mathbf{X}\hat{\boldsymbol{\beta}}_1\right\}^{1/2}} \tag{21.39}$$

All the notations are as above. The program also computes a similar statistic, W_2, for the test of M_2 against M_1.

The J-test

This test is due to Davidson and MacKinnon (1981), and for the test of M_1 against M_2 is based on the t-ratio of λ in the 'artificial' OLS regression

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_1 + \lambda(\mathbf{Z}\hat{\boldsymbol{\beta}}_2) + \mathbf{u}$$

The relevant statistic for the J-test of M_2 against M_1 is the t-ratio of μ in the OLS regression

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\beta}_2 + \mu(\mathbf{X}\hat{\boldsymbol{\beta}}_1) + \mathbf{v}$$

where $\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$, and $\hat{\boldsymbol{\beta}}_2 = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$. The J-test is asymptotically equivalent to the above non-nested tests, but as demonstrated by extensive Monte Carlo experiments in Godfrey and Pesaran (1983), for small samples the $\tilde{N}$-test, and the W-test, defined above, are preferable to it.

The JA-test

This test is due to Fisher and McAleer (1981), and for the test of M_1 against M_2 is based on the t-ratio of λ in the OLS regression

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_1 + \lambda(\mathbf{A}_2\mathbf{X}\hat{\boldsymbol{\beta}}_1) + \mathbf{u}$$

The relevant statistic for the JA-test of M_2 against M_1 is the t-ratio of μ in the OLS regression

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\beta}_2 + \mu(\mathbf{A}_1\mathbf{Z}\hat{\boldsymbol{\beta}}_2) + \mathbf{v}$$

The matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ are already defined by (21.38).

The encompassing test

This test has been proposed in the literature by Deaton (1982), Dastoor (1983), Gourierous, Holly, and Monfort (1982), and Mizon and Richard (1986). In the case of testing M_1 against M_2, the encompassing test is the same as the classical F-test, and is computed as the F-statistic for testing $\boldsymbol{\delta} = 0$ in the combined OLS regression

$$\mathbf{y} = \mathbf{X}\mathbf{a}_0 + \mathbf{Z}^*\boldsymbol{\delta} + \mathbf{u}$$

where $\mathbf{Z}^*$ denotes the variables in M_2 that cannot be expressed as exact linear combinations of the regressors of M_1. Similarly, the program computes the F-statistic for the test of M_2 against M_1. The encompassing test is asymptotically equivalent to the above non-nested tests under the null hypothesis, but in general it is less powerful for a large class of alternative non-nested models (see Pesaran (1982a)).

A Monte Carlo study of the relative performance of the above non-nested tests in small samples can be found in Godfrey and Pesaran (1983).

Choice criteria

Let the maximized log-likelihood functions of models M_1 and M_2 be LL_1 and LL_2 respectively.

The Akaike information criterion (AIC) for the choice between models M_1 and M_2 is computed as (Akaike (1973), and Akaike (1974)):

$$AIC(M_1 : M_2) = LL_1 - LL_2 - (k_1 - k_2)$$

Model M_1 is preferred to M_2 if $AIC(M_1 : M_2) > 0$, otherwise M_2 is preferable to M_1.

The Schwarz Bayesian criterion (SBC) for the choice between models M_1 and M_2 is computed as (Schwarz (1978)):

$$SBC(M_1 : M_2) = LL_1 - LL_2 - \tfrac{1}{2}(k_1 - k_2)\log(n)$$

Model M_1 is preferred to M_2 if $SBC(M_1 : M_2) > 0$, otherwise M_2 is preferable to M_1 (see also Section 21.7).

21.9 Non-nested tests for models with different transformations of the dependent variable

The program computes four non-nested test statistics and two choice criteria for pairwise testing and choice between non-nested models, where their right-hand-side variables are *different* known functions of a given underlying dependent variable. More specifically, *Microfit* enables you to consider the following non-nested models:

$$M_f \;:\; \mathbf{f}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}_1 + \mathbf{u}_1, \qquad \mathbf{u}_1 \sim N(0, \sigma^2\mathbf{I}_n) \tag{21.40}$$

$$M_g \;:\; \mathbf{g}(\mathbf{y}) = \mathbf{Z}\boldsymbol{\beta}_2 + \mathbf{u}_2, \qquad \mathbf{u}_2 \sim N(0, \omega^2\mathbf{I}_n) \tag{21.41}$$

where $\mathbf{f}(\mathbf{y})$ and $\mathbf{g}(\mathbf{y})$ are known transformations of the $n \times 1$ vector of observations on the underlying dependent variable of interest, $\mathbf{y}$. You can either specify your particular choice for the functions $\mathbf{f}(\mathbf{y})$ and $\mathbf{g}(\mathbf{y})$, or you can select one of the following specifications:

1.	Linear form	$\mathbf{f}(\mathbf{y}) = \mathbf{y}$
2.	Logarithmic form	$\mathbf{f}(\mathbf{y}) = \log(\mathbf{y})$
3.	Ratio form	$\mathbf{f}(\mathbf{y}) = \mathbf{y}/\mathbf{z}$
4.	Difference form	$\mathbf{f}(\mathbf{y}) = \mathbf{y} - \mathbf{y}(-1)$
5.	Log-difference form	$\mathbf{f}(\mathbf{y}) = \log \mathbf{y} - \log \mathbf{y}(-1)$

where $\mathbf{z}$ is a variable in the workspace. Notice that $\log(\mathbf{y})$ refers to a vector of observations with elements equal to $\log(y_t)$, $t = 1, 2, ..., n$. Also $\mathbf{y} - \mathbf{y}(-1)$ refers to a vector with a typical element equal to $y_t - y_{t-1}$, $t = 1, 2, ..., n$.

21.9.1 The $\mathbf{P_E}$ Test Statistic

This statistic is proposed by MacKinnon, White, and Davidson (1983) and in the case of testing M_f against M_g is given by the t-ratio of α_f in the auxiliary regression

$$\mathbf{f}(\mathbf{y}) = \mathbf{X}\mathbf{b} + \alpha_f[\mathbf{Z}\hat{\boldsymbol{\beta}}_2 - \mathbf{g}\{\mathbf{f}^{-1}(\mathbf{X}\hat{\boldsymbol{\beta}}_1)\}] + \text{Error} \tag{21.42}$$

Similarly, the P_E statistic for testing M_g against M_f is given by the t-ratio of α_g in the auxiliary regression

$$\mathbf{g}(\mathbf{y}) = \mathbf{Z}\mathbf{d} + \alpha_g[\mathbf{X}\hat{\boldsymbol{\beta}}_1 - \mathbf{f}\{\mathbf{g}^{-1}(\mathbf{Z}\hat{\boldsymbol{\beta}}_2)\}] + \text{Error} \tag{21.43}$$

Functions $f^{-1}(\cdot)$ and $g^{-1}(\cdot)$ represent the inverse functions for $f(\cdot)$ and $g(\cdot)$, respectively, such that $f(f^{-1}(y)) = y$, and $g(g^{-1}(y)) = y$. For example, in the case where M_f is linear (i.e., $f(y) = y$) and M_g is log-linear (i.e. $g(y) = \log y$), we have

$$\begin{aligned} f^{-1}(y_t) &= y_t \\ g^{-1}(y_t) &= \exp(y_t) \end{aligned}$$

In the case where M_f is in first-differences ($f(y_t) = y_t - y_{t-1}$) and M_g is in log-differences ($g(y_t) = \log(y_t/y_{t-1})$) we have

$$\begin{aligned} f^{-1}(y_t) &= f(y_t) + y_{t-1} \\ g^{-1}(y_t) &= y_{t-1} \exp\{g(y_t)\} \end{aligned}$$

$\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ are the OLS estimators of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ under M_f and M_g, respectively.

21.9.2 The Bera-McAleer test statistic

The statistic proposed by Bera and McAleer (1989) is for testing linear versus log-linear models, but can be readily extended to general known one-to-one transformations of the dependent variable of interest, namely y_t. To compute the Bera-McAleer (BM) statistic for testing M_f against M_g, the program first computes the residuals $\hat{\boldsymbol{\eta}}_g$ from the regression of $\mathbf{g}[\mathbf{f}^{-1}(\mathbf{X}\hat{\boldsymbol{\beta}}_1)]$ on $\mathbf{Z}$. It then computes the BM statistic for testing M_f against M_g as the t-ratio of θ_f in the auxiliary regression

$$\mathbf{f}(\mathbf{y}) = \mathbf{Xb} + \theta_f \hat{\boldsymbol{\eta}}_g + \text{Error} \tag{21.44}$$

The BM statistic for the test of M_g against M_f is given by the t-ratio of θ_g in the auxiliary regression

$$\mathbf{g}(\mathbf{y}) = \mathbf{Zd} + \theta_g \hat{\boldsymbol{\eta}}_f + \text{Error} \tag{21.45}$$

where $\hat{\boldsymbol{\eta}}_f$ is the residual vector of the regression of $\mathbf{f}\{\mathbf{g}^{-1}(\mathbf{Z}\hat{\boldsymbol{\beta}}_2)\}$ on $\mathbf{X}$.

21.9.3 The double-length regression test statistic

The double-length (DL) regression statistic was proposed by Davidson and MacKinnon (1984), and for the test of M_f against M_g is given by

$$DL_f = 2n - SSR_f \tag{21.46}$$

where SSR_f denotes the sums of squares of residuals from the DL regression

$$\begin{bmatrix} \mathbf{e}_1/\hat{\sigma} \\ \boldsymbol{\tau} \end{bmatrix} = \begin{bmatrix} -\mathbf{X} \\ \mathbf{0} \end{bmatrix} \mathbf{b} + \begin{bmatrix} \mathbf{e}_1/\hat{\sigma} \\ -\boldsymbol{\tau} \end{bmatrix} \mathbf{c} + \begin{bmatrix} -\mathbf{e}_2 \\ \hat{\sigma}\hat{\mathbf{v}} \end{bmatrix} \mathbf{d} + \text{Error} \tag{21.47}$$

where

$$\begin{aligned} &\mathbf{e}_1 = \mathbf{f}(\mathbf{y}) - \mathbf{X}\hat{\boldsymbol{\beta}}_1, && \hat{\sigma}^2 = \mathbf{e}_1'\mathbf{e}_1/(n-k_1) \\ &\mathbf{e}_2 = \mathbf{g}(\mathbf{y}) - \mathbf{Z}\hat{\boldsymbol{\beta}}_2, && \hat{\omega}^2 = \mathbf{e}_2'\mathbf{e}_2/(n-k_2) \\ &\hat{\mathbf{v}} = (\hat{v}_1, \hat{v}_2, ..., \hat{v}_n)', && \hat{v}_t = g'(y_t)/f'(y_t) \end{aligned}$$

$\boldsymbol{\tau} = (1, 1, ..., 1)'$ is an $n \times 1$ vector of ones, and $g'(y_t)$ and $f'(y_t)$ denote the derivatives of $g(y_t)$ and $f(y_t)$ with respect to y_t.

To compute the SSR_f statistic we first note that

$$SSR_f = \tilde{\mathbf{y}}'\tilde{\mathbf{y}} - \tilde{\mathbf{y}}'\tilde{\mathbf{X}}(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{y}}$$

where

$$\tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{e}_1/\hat{\sigma} \\ \boldsymbol{\tau} \end{bmatrix}, \qquad \tilde{\mathbf{X}} = \begin{bmatrix} -\mathbf{X} & \mathbf{e}_1/\hat{\sigma} & -\mathbf{e}_2 \\ \mathbf{0} & -\boldsymbol{\tau} & \hat{\sigma}\hat{\mathbf{v}} \end{bmatrix}$$

but $\tilde{\mathbf{y}}'\tilde{\mathbf{y}} = \mathbf{e}_1'\mathbf{e}_2/\hat{\sigma}^2 + n = 2n - k_1$

$$\tilde{\mathbf{y}}'\tilde{\mathbf{X}} = \left[0,\ -k_1, \hat{\sigma}\boldsymbol{\tau}'\hat{\mathbf{v}} - \frac{\mathbf{e}_1\mathbf{e}_2}{\hat{\sigma}}\right]$$

and

$$\tilde{\mathbf{X}}'\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{0} & \mathbf{X}'\mathbf{e}_2 \\ \mathbf{0} & 2n - k_1 & \dfrac{-\mathbf{e}_1'\mathbf{e}_2}{\hat{\sigma}} - \hat{\sigma}(\boldsymbol{\tau}'\hat{\mathbf{v}}) \\ \mathbf{e}_2'\mathbf{X} & \dfrac{-\mathbf{e}_1'\mathbf{e}_2}{\hat{\sigma}} - \hat{\sigma}(\boldsymbol{\tau}'\hat{\mathbf{v}}) & \mathbf{e}_2'\mathbf{e}_2 + \hat{\sigma}^2\hat{\mathbf{v}}'\hat{\mathbf{v}} \end{bmatrix}$$

Using these results, and after some algebra, we obtain

$$DL_f = \frac{1}{D}\left(k_1^2 R_1 + (2n - k_1)R_3^2 - 2k_1 R_2 R_3\right) \tag{21.48}$$

where

$$\begin{aligned} R_1 &= (\mathbf{e}_2'\mathbf{M}_1\mathbf{e}_2)/\hat{\sigma}^2 + \hat{\mathbf{v}}'\hat{\mathbf{v}} \\ R_2 &= (\boldsymbol{\tau}'\hat{\mathbf{v}}) + (\mathbf{e}_1'\mathbf{e}_2)/\hat{\sigma}^2 \\ R_3 &= (\boldsymbol{\tau}'\hat{\mathbf{v}}) - (\mathbf{e}_1'\mathbf{e}_2)/\hat{\sigma}^2 \\ D &= (2n - k_1)R_1 - R_2^2 \end{aligned}$$

A similar statistic is also computed for the test of M_g against M_f.

21.9.4 The Cox non-nested statistics computed by simulation

The simulated Cox test statistics were introduced in Pesaran and Pesaran (1993) and subsequently applied to tests of linear versus log-linear models, and first-difference versus log-difference stationary models in Pesaran and Pesaran (1995) (or PP(95)).

Three versions of the simulated Cox statistic is considered by PP(95). The three test statistics have the same numerator and differ by the choice of the estimator of the variance used to standardize the Cox statistic. In *Microfit 5.0* we have only programmed the SC_c statistic which seems to have much better small sample properties than the other two test statistics (namely SC_a and SC_b) considered by PP(95).[5] (In the program this is referred to

[5] The Monte Carlo results reported in $PP(95)$ also clearly show that the SC_c and the DL tests are more powerful than the PE or BM tests discussed in Section 21.9.1 and 21.9.2.

as the SC_c test statistic.) The numerator of the SC_c statistic for testing M_f against M_g is computed as

$$\begin{aligned} n^{1/2}T_f(R) &= -\tfrac{1}{2}n^{1/2}\log(\hat{\sigma}^2/\hat{\omega}^2) + n^{-1/2}\sum_{t=1}^{n}\log\left(\left|f^{'}(y_t)/g'(y_t)\right|\right) \\ &\quad +\tfrac{1}{2}n^{-1/2}(k_1-k_2) - n^{1/2}C_R(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\gamma}}_*(R)) \end{aligned} \tag{21.49}$$

where $\hat{\boldsymbol{\theta}} = \left(\widehat{\boldsymbol{\beta}}_1', \hat{\sigma}^2\right)'$, R is the number of replications, and $\hat{\boldsymbol{\gamma}}_*(R)$ is the simulated pseudo-ML estimator of $\boldsymbol{\gamma} = \left(\boldsymbol{\beta}_2', \omega^2\right)'$ under M_f

$$\hat{\boldsymbol{\gamma}}_*(R) = R^{-1}\sum_{j=1}^{R}\hat{\boldsymbol{\gamma}}_j \tag{21.50}$$

where $\hat{\boldsymbol{\gamma}}_j$ is the ML estimator of $\boldsymbol{\gamma}$ computed using the artificially simulated independent observations $\mathbf{Y}_j = (\mathbf{Y}_{j1}, \mathbf{Y}_{j2}, ..., \mathbf{Y}_{jn})$ obtained under M_f with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. $C_R(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\gamma}}_*(R))$ is the simulated estimator of the 'closeness' measure of M_f with respect of M_g (see Pesaran (1987a)):

$$C_R(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\gamma}}_*(R)) = R^{-1}\sum_{j=1}^{R}[L_f(\mathbf{Y}_j,\hat{\boldsymbol{\theta}}) - L_g\{\mathbf{Y}_j,\hat{\boldsymbol{\gamma}}_*(R)\}] \tag{21.51}$$

where $L_f(\mathbf{Y},\boldsymbol{\theta})$ and $L_g(\mathbf{Y},\boldsymbol{\gamma})$ are the average log-likelihood functions under M_f and M_g, respectively:

$$\begin{aligned} L_f(\mathbf{Y},\boldsymbol{\theta}) &= -\tfrac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left[\sum_{t=1}^{n}\{f(y_t)-\boldsymbol{\beta}_1'\mathbf{x}_t\}^2/n\right] \\ &\quad +n^{-1}\sum_{t=1}^{n}\log\left|f'(y_t)\right| \end{aligned} \tag{21.52}$$

$$\begin{aligned} L_g(\mathbf{Y},\boldsymbol{\gamma}) &= -\tfrac{1}{2}\log(2\pi\omega^2) - \frac{1}{2\omega^2}\left[\sum_{t=1}^{n}\{g(y_t)-\boldsymbol{\beta}_2'\mathbf{z}_t\}^2/n\right] \\ &\quad +n^{-1}\sum_{t=1}^{n}\log\left|g'(y_t)\right| \end{aligned} \tag{21.53}$$

The denominator of the SC_c statistic is computed as

$$V_{*d}^2(R) = (n-1)^{-1}\sum_{t=1}^{n}(d_{*t}-\bar{d}_*)^2 \tag{21.54}$$

where $\bar{d}_* = n^{-1}\sum_{t=1}^{n} d_{*t}$, and

$$\begin{aligned} d_{*t} &= -\tfrac{1}{2}\log\left(\hat{\sigma}^2/\hat{\omega}_*^2(R)\right) - \frac{1}{2\hat{\sigma}^2}e_{t1}^2 \\ &\quad +\frac{1}{2\hat{\omega}_*^2(R)}\left[g(y_t)-\mathbf{z}_t'\hat{\beta}_{*2}(R)\right]^2 + \log\left(\left|f'(y_t)/g'(y_t)\right|\right) \end{aligned}$$

and

$$e_{t1} = f(y_t) - \mathbf{x}_t' \widehat{\boldsymbol{\beta}}_1$$

Recall also that $\widehat{\boldsymbol{\beta}}_{*2}(R)$ and $\hat{\omega}_*^2(R)$ are given by (21.50), where $\hat{\boldsymbol{\gamma}}_*(R) = \left(\widehat{\boldsymbol{\beta}}_{*2}'(R), \hat{\omega}_*^2(R)\right)'$.

The standardized Cox statistic reported by *Microfit 5.0* for the test of M_f against M_g is given by

$$SC_c(R) = n^{\frac{1}{2}} T_f(R) / V_{*d}(R)$$

where $n^{\frac{1}{2}} T_f(R)$ is defined by (21.49) and $V_{*d}(R)$ by (21.54). A similar statistic is also computed for the test of M_g against M_f.

21.9.5 Sargan and Vuong's likelihood criteria

The Sargan (1964) likelihood criterion simply compares the maximized values of the log-likelihood functions under M_f and M_g:[6]

$$LL_{fg} = n\{L_f(\mathbf{Y}, \hat{\boldsymbol{\theta}}) - L_g(\mathbf{Y}, \hat{\boldsymbol{\gamma}})\}$$

or using (21.52) and (21.53)

$$LL_{fg} = -\frac{n}{2} \log(\hat{\sigma}^2/\hat{\omega}^2) + \sum_{t=1}^{n} \log \left| f'(y_t)/g'(y_t) \right| + \tfrac{1}{2}(k_1 - k_2) \tag{21.55}$$

Known model selection criteria such as AIC and SBC could also be applied to the models M_f and M_g (see Section 21.7). For example, in the case of the AIC we have

$$AIC(M_f : M_g) = LL_{fg} - (k_1 - k_2)$$

Vuong's criterion is motivated in the context of testing the hypothesis that M_f and M_g are equivalent, using the Kullback and Leibler (1951) information criterion as a measure of goodness of fit. The Vuong (1989) test criterion for the comparison of M_f and M_g is computed as

$$V_{fg} = \frac{\sum_{t=1}^{n} d_t}{\left(\sum_{t=1}^{n} (d_t - \bar{d})^2 \right)^{1/2}} \tag{21.56}$$

where $\bar{d} = n^{-1} \sum_{t=1}^{n} d_t$, and

$$\begin{aligned} d_t &= -\tfrac{1}{2} \log(\hat{\sigma}^2/\hat{\omega}^2) - \tfrac{1}{2} \left(\frac{e_{t1}^2}{\hat{\sigma}^2} - \frac{e_{t2}^2}{\hat{\omega}^2} \right) + \log \left(\left| f'(y_t)/g'(y_t) \right| \right) \\ e_{t1} &= f(y_t) - \widehat{\boldsymbol{\beta}}_1' \mathbf{x}_t, \qquad e_{t2} = g(y_t) - \widehat{\boldsymbol{\beta}}_2' \mathbf{z}_t \end{aligned}$$

Under the null hypothesis that 'M_f and M_g are equivalent', V_{fg} is approximately distributed as a standard normal variate.

[6] Notice that throughout, $\hat{\sigma}^2 = \mathbf{e}_1'\mathbf{e}_1/(n - k_1)$ and $\hat{\omega}^2 = \mathbf{e}_2'\mathbf{e}_2/(n - k_2)$ are used as estimators of σ^2 and ω^2, respectively.

21.10 The generalized instrumental variable method

Consider the linear regression model:

$$\underset{n\times 1}{\mathbf{y}} = \underset{n\times k}{\mathbf{X}}\ \underset{k\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\mathbf{u}}$$

and suppose that there exists an $n \times s$ matrix $\mathbf{Z}$ containing observations on s instrumental variables ($s \geq k$). Then the Generalized Instrumental Variable Estimator ($GIVE$)[7] of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}}_{IV} = (\mathbf{X}'\mathbf{P}_z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_z\mathbf{y} \tag{21.57}$$

where $\mathbf{P}_z$ is the $n \times n$ projection matrix

$$\mathbf{P}_z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}' \tag{21.58}$$

The estimator of the variance matrix of $\hat{\boldsymbol{\beta}}_{IV}$ is given by

$$\hat{V}(\hat{\boldsymbol{\beta}}_{IV}) = \hat{\sigma}^2_{IV}(\mathbf{X}'\mathbf{P}_z\mathbf{X})^{-1} \tag{21.59}$$

where $\hat{\sigma}^2_{IV}$ is the IV estimator of σ^2 (the variance of u_t):

$$\hat{\sigma}^2_{IV} = (n-k)^{-1}\hat{\mathbf{e}}'_{IV}\hat{\mathbf{e}}_{IV} \tag{21.60}$$

where $\hat{\mathbf{e}}_{IV}$ is the IV residuals given by

$$\hat{\mathbf{e}}_{IV} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{IV} \tag{21.61}$$

The estimator $\hat{\boldsymbol{\beta}}_{IV}$ can also be derived by minimizing the weighted quadratic form

$$Q(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{P}_z(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{21.62}$$

with respect to $\boldsymbol{\beta}$. The program reports the minimized value of $Q(\boldsymbol{\beta})$ (namely $Q(\hat{\boldsymbol{\beta}}_{IV})$) under the heading 'value of IV minimand'. Notice that $Q(\hat{\boldsymbol{\beta}}_{IV})$ will be identically equal to zero when the number of instruments is exactly equal to the number of the regressors (when $s = k$).

21.10.1 Two-stage least squares

The IV estimator, $\hat{\boldsymbol{\beta}}_{IV}$, can also be computed using a two-step procedure, known as the two-stage least squares ($2SLS$), where in the first step the fitted values of the OLS regression of $\mathbf{X}$ on $\mathbf{Z}$, $\hat{\mathbf{X}} = \mathbf{P}_z\mathbf{X}$ are computed. Then $\hat{\boldsymbol{\beta}}_{IV}$ is obtained by the OLS regression of $\mathbf{y}$ on $\hat{\mathbf{X}}$. Notice, however, that such a two-step procedure does not, in general, produce a correct estimator of σ^2, and hence of $\hat{V}(\hat{\boldsymbol{\beta}}_{IV})$. This is because the IV residuals, $\hat{\mathbf{e}}_{IV}$, defined by

[7] The idea of the generalized IV estimator is due to Sargan (1958).

(21.61), used in the estimation of $\hat{\sigma}^2_{IV}$ is not the same as the residuals obtained at the second stage of the 2SLS method. To see this denote the $2SLS$ residuals by $\mathbf{e}_{2SLS}$ and note that

$$\begin{aligned} \mathbf{e}_{2SLS} &= \mathbf{y} - \hat{\mathbf{X}}\hat{\boldsymbol{\beta}}_{IV} \\ &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{IV}) + (\mathbf{X} - \hat{\mathbf{X}})\hat{\boldsymbol{\beta}}_{IV} \\ &= \mathbf{e}_{IV} + (\mathbf{X} - \hat{\mathbf{X}})\hat{\beta}_{IV} \end{aligned} \tag{21.63}$$

Where $\mathbf{X} - \hat{\mathbf{X}}$ are the residual matrix $(n \times k)$ of the regressions of $\mathbf{X}$ on $\mathbf{Z}$. Only in the case where $\mathbf{Z}$ is an exact predictor of $\mathbf{X}$, the two sets of residuals will be the same.

21.10.2 Generalized $\mathbf{R}^2$ for IV regressions

The use of R^2 and $\bar{R}^2$ as measures of goodness of fit in the case of IV regressions is not valid. As is well known, there is no guarantee that R^2 of a regression model estimated by the IV method is positive, and this result does not depend on whether or not an intercept term is included in the regression. (See, for example, Maddala (1988) p. 309). An appropriate measure of fit for IV regressions is the Generalized R^2, or GR^2, measure proposed by Pesaran and Smith (1994). In the case of IV regressions, *Microfit* reports this measure along with the other summary statistics. The GR^2 is computed as

$$GR^2 = 1 - \left(\mathbf{e}'_{2SLS}\mathbf{e}_{2SLS}\right) \Bigg/ \left(\sum_{t=1}^{n}(y_t - \bar{y})^2\right) \tag{21.64}$$

where $\mathbf{e}_{2SLS}$, given by (21.63), is the vector of residuals from the second stage in the $2SLS$ procedure. Notice also that

$$\mathbf{e}_{2SLS} = \mathbf{e}_{IV} + (\mathbf{X} - \hat{\mathbf{X}})\hat{\beta}_{IV} \tag{21.65}$$

A degrees-of-freedom adjusted Generalized R^2 measure is given by

$$\overline{GR}^2 = 1 - \left(\frac{n-1}{n-k}\right)\left(1 - GR^2\right) \tag{21.66}$$

Pesaran and Smith (1994) show that under reasonable assumptions the use of GR^2 is a valid discriminator for models estimated by the IV method, asymptotically.

21.10.3 Sargan's general mis-specification test

This test is proposed in Sargan (1964) pp. 28-9, as a general test of misspecification in the case of IV estimation, and is based on the statistic $(s > k)$

$$\chi^2_{SM} = Q(\hat{\boldsymbol{\beta}}_{IV})/\hat{\sigma}^2_{IV} \overset{a}{\sim} \chi^2(s-k) \tag{21.67}$$

where $Q(\hat{\beta}_{IV})$ is the value of IV minimand given by

$$\begin{aligned} Q\left(\hat{\beta}_{IV}\right) &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{IV})'\mathbf{P}_z(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{IV}) \\ &= \mathbf{y}'[\mathbf{P}_z - \mathbf{P}_z\mathbf{X}(\mathbf{X}'\mathbf{P}_z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_z]\mathbf{y} \\ &= \hat{\mathbf{y}}'[\mathbf{I} - \hat{\mathbf{X}}(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1}\mathbf{X}']\hat{\mathbf{y}} \end{aligned} \tag{21.68}$$

Under the null hypothesis that the regression equation (21.4) is correctly specified, and that the s ($s > k$) instrumental variables $\mathbf{Z}$ are valid instruments, Sargan's mis-specification statistic, χ^2_{SM}, is asymptotically distributed as a chi-squared variate with $s - k$ degrees of freedom. The J-statistic proposed by Hansen (1982) is a generalization of Sargan's mis-specification statistic.

21.10.4 Sargan's test of residual serial correlation for IV regressions

The statistic underlying this test is given in Appendix B of Breusch and Godfrey (1981), and can be written as

$$\chi^2_{SC}(p) = n\mathbf{e}'_{IV}\mathbf{W}(\mathbf{W}'\mathbf{H}\mathbf{W})^{-1}\mathbf{W}'\mathbf{e}_{IV}/\mathbf{e}'_{IV}\mathbf{e}_{IV} \overset{a}{\sim} \chi^2(p) \tag{21.69}$$

where $\mathbf{e}_{IV}$ is the vector of IV residuals defined by (21.61):

$$\mathbf{e}_{IV} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{IV} = (e_{1,IV}, e_{2,IV}, ..., e_{n,IV})'$$

$\mathbf{W}$ is the $n \times p$ matrix consisting of the p lagged valued of $\mathbf{e}_{IV}$, namely

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ e_{1,IV} & 0 & \dots & 0 \\ e_{2,IV} & e_{1,IV} & \dots & 0 \\ . & . & & \vdots \\ \vdots & \vdots & & e_{n-p-1,IV} \\ e_{n-1,IV} & e_{n-2,IV} & \dots & e_{n-p,IV} \end{bmatrix} \tag{21.70}$$

and

$$\mathbf{H} = \mathbf{I}_n - \mathbf{X}(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1}\hat{\mathbf{X}}' - \hat{\mathbf{X}}(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1}\mathbf{X}' + \mathbf{X}(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1}\mathbf{X}'$$

in which $\hat{\mathbf{X}} = \mathbf{P}_z\mathbf{X}$.[8] Notice that when $\mathbf{Z}$ includes $\mathbf{X}$, then $\hat{\mathbf{X}} = \mathbf{X}$, and (21.69) reduces to (21.19). Under the null hypothesis that the disturbances in (21.4) are serially uncorrelated, $\chi^2_{SC}(p)$ in (21.69) is asymptotically distributed as a chi-squared variate with p degrees of freedom.

21.11 Exact ML/AR estimators

This estimation method (option 3 in the Linear Regression Estimation Menu, 6.5), provides exact Maximum Likelihood (ML) estimates of the parameters of (21.3) under the assumption that the disturbances u_t follow stationary

$$AR(1) : u_t = \rho u_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma^2_\epsilon), \quad t = 1, 2, ..., n \tag{21.71}$$

or

$$AR(2) : u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma^2_\epsilon), \quad t = 1, 2, ..., n \tag{21.72}$$

[8] See Breusch and Godfrey (1981) p. 101, for further details. The statistic in (21.69) is derived from the results in Sargan (1976).

processes with 'stochastic initial values'. This estimation procedure assumes that the underlying AR error processes are started a long-time prior to the first observation date ($t = 1$) and are stationary. This implies that the initial values (u_1 for the $AR(1)$ process, and u_1 and u_2 for the $AR(2)$ process) are normally distributed with zero means and a constant variance given by

$$AR(1)\ Case: \qquad V(u_1) = \frac{\sigma_\epsilon^2}{1-\rho^2}$$

$$AR(2)\ Case: \qquad \begin{cases} V(u_1) = V(u_2) = \dfrac{\sigma_\epsilon^2(1-\rho_2)}{(1+\rho_2)^3 - \rho_1^2(1+\rho_2)} \\ \\ Cov(u_1,\ u_2) = \dfrac{\sigma_\epsilon^2 \rho_1}{(1+\rho_2)^3 - \rho_1^2(1+\rho_2)} \end{cases}$$

The exact ML estimation procedure then allows for the effect of initial values on the parameter estimates by adding the logarithm of the density function of the initial values to the log-density function of the remaining observations obtained conditional on the initial values. For example, in the case of the $AR(1)$ model the log-density function of $(u_2,\ u_3, ..., u_n)$ conditional on the initial value, u_1, is given by:

$$\log\left\{f(u_2, u_3, ..., u_n\,|u_1)\right\} = -\frac{(n-1)}{2}\log(2\pi\sigma_\epsilon^2) - \frac{1}{2\sigma_\epsilon^2}\left(\sum_{t=2}^{n}\sigma_t^2\right) \tag{21.73}$$

and

$$\log\left\{f(u_1)\right\} = -\frac{1}{2}\log(2\pi\sigma_\epsilon^2) + \frac{1}{2}\log(1-\rho^2) - \frac{(1-\rho^2)}{2\sigma_\epsilon^2}u_1^2$$

Combining the above log-densities yields the full (unconditional) log-density function of $(u_1, u_2, .., u_n)$

$$\begin{aligned} \log\left\{f(u_1, u_2, ..., u_n)\right\} &= -\frac{n}{2}\log(2\pi\sigma_\epsilon^2) + \frac{1}{2}\log(1-\rho^2) \\ &\quad -\frac{1}{2\sigma_\epsilon^2}\left(\sum_{t=2}^{n}(u_t - \rho u_{t-1})^2 + (1-\rho^2)u_1^2\right) \end{aligned} \tag{21.74}$$

Asymptotically, the effect of the distribution of the initial values on the ML estimators is negligible, but it could be important in small samples where $\mathbf{x}_t$s are trended and ρ is suspected to be near but *not* equal to unity. See Pesaran (1972), and Pesaran and Slater (1980) Chapters 2 and 3, for further details. Also see Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) Section 8.2, Davidson and MacKinnon (1993) Section 10.6, and the papers by Hildreth and Dent (1974), and Beach and MacKinnon (1978). Strictly speaking, the ML estimation will be exact if lagged values of y_t are not included amongst the regressors. For a discussion of the exact ML estimation of models with lagged dependent variable and serially correlated errors, see Pesaran (1981).

21.11.1 The AR(1) case

For this case, the ML estimators are computed by maximizing the log-likelihood function[9]

$$\begin{aligned} LL_{AR1}(\boldsymbol{\theta}) &= -\frac{n}{2}\log(2\pi\sigma_\epsilon^2) + \tfrac{1}{2}\log(1-\rho^2) \\ &\quad -\frac{1}{2\sigma_\epsilon^2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{R}(\rho)(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) \end{aligned} \tag{21.75}$$

with respect to the unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_\epsilon^2, \rho)'$, where $\mathbf{R}(\rho)$ is the $n \times n$ matrix

$$R(\rho) = \begin{bmatrix} 1 & -\rho & 0 & 0 & & \ldots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & & \ldots & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \ldots\ldots & & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & \ldots\ldots & & 0 & -\rho & 1 \end{bmatrix} \tag{21.76}$$

and $|\rho| < 1$.

The computations are carried out by the 'inverse interpolation' method, which is *certain to converge*. See Pesaran and Slater (1980) pp. 36-38, for further details.

The concentrated log-likelihood function in this case is given by

$$\begin{aligned} LL_{AR1}(\rho) &= -\frac{n}{2}[1+\log(2\pi)] + \tfrac{1}{2}\log(1-\rho^2) \\ &\quad -\frac{n}{2}\log\{\tilde{\mathbf{u}}'R(\rho)\tilde{\mathbf{u}}/n\}, \qquad |\rho| < 1 \end{aligned} \tag{21.77}$$

where $\tilde{\mathbf{u}}$ is the $n \times 1$ vector of ML residuals:

$$\tilde{\mathbf{u}} = \mathbf{y} - \mathbf{X}[\mathbf{X}'\mathbf{R}(\rho)\mathbf{X}]^{-1}\mathbf{X}'\mathbf{R}(\rho)\mathbf{y}$$

21.11.2 The AR(2) case

For this case, the ML estimators are obtained by maximizing the log-likelihood function

$$\begin{aligned} LL_{AR2}(\boldsymbol{\theta}) &= -\frac{n}{2}\log(2\pi\sigma_\epsilon^2) + \log(1+\rho_2) \\ &\quad +\tfrac{1}{2}\log\left[(1-\rho_2)^2 - \rho_1^2\right] \\ &\quad -\frac{1}{2\sigma_\epsilon^2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{R}(\boldsymbol{\rho})(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) \end{aligned} \tag{21.78}$$

[9] This result follows readily from (21.74), and can be obtained by substituting $u_t = y_t - \boldsymbol{\beta}'\mathbf{x}_t$ in (21.74).

with respect to $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_\epsilon^2, \boldsymbol{\rho})'$, where $\boldsymbol{\rho} = (\rho_1, \rho_2)'$

$$R(\boldsymbol{\rho}) = \begin{bmatrix} 1 & -\rho_1 & -\rho_2 & 0 & 0 & \ldots & 0 & 0 \\ -\rho_1 & 1+\rho_1^2 & -\rho_1+\rho_1\rho_2 & -\rho_2 & 0 & \ldots & 0 & 0 \\ -\rho_2 & -\rho_1+\rho_1\rho_2 & 1+\rho_1^2+\rho_2^2 & -\rho_1+\rho_1\rho_2 & -\rho_2 & \ldots & 0 & 0 \\ 0 & -\rho_2 & -\rho_1+\rho_1\rho_2 & 1+\rho_1^2+\rho_2^2 & \ldots & & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & & \ldots & 1+\rho_1^2 & -\rho_1 \\ 0 & 0 & 0 & 0 & & \ldots & -\rho_1 & 1 \end{bmatrix} \tag{21.79}$$

The estimation procedure imposes the restrictions

$$\left.\begin{array}{l} 1+\rho_2 > 0, \\ 1-\rho_2+\rho_1 > 0 \\ 1-\rho_2-\rho_1 > 0 \end{array}\right\} \tag{21.80}$$

needed if the $AR(2)$ process, (21.72), is to be stationary.

21.11.3 Covariance matrix of the exact ML estimators for the AR(1) and AR(2) options

The estimates of the covariance matrix of the exact ML estimators defined in the above sub-sections are computed on the assumption that the regressors $\mathbf{x}_t$ do not include lagged values of the dependent variable.[10]

For the $AR(1)$ case we have

$$\tilde{V}(\tilde{\boldsymbol{\beta}}) = \hat{\sigma}_\epsilon^2[\mathbf{X}'\mathbf{R}(\tilde{\rho})\mathbf{X}]^{-1} \tag{21.81}$$

$$\tilde{V}(\tilde{\rho}) = n^{-1}(1-\tilde{\rho}^2) \tag{21.82}$$

where $\mathbf{R}(\tilde{\rho})$ is already defined by (21.76), and $\hat{\sigma}_\epsilon^2$ is given below by (21.91).

For the $AR(2)$ case we have

$$\tilde{V}(\tilde{\boldsymbol{\beta}}) = \hat{\sigma}_\epsilon^2[\mathbf{X}'\mathbf{R}(\tilde{\rho}_1, \tilde{\rho}_2)\mathbf{X}]^{-1} \tag{21.83}$$

$$\tilde{V}(\tilde{\rho}_1) = \tilde{V}(\tilde{\rho}_2) = n^{-1}(1-\tilde{\rho}_2^2) \tag{21.84}$$

$$\widetilde{Cov}(\tilde{\rho}_1, \tilde{\rho}_2) = -n^{-1}\tilde{\rho}_1(1+\tilde{\rho}_2) \tag{21.85}$$

where $\mathbf{R}(\tilde{\rho}_1, \tilde{\rho}_2)$ is defined by (21.79). Here the ML estimators are designated by $\sim$.

[10] When the regression contains lagged values of the dependent variable, the Cochrane-Orcutt or the Gauss-Newton options 4 or 5 in the Linear Regression Estimation Menu should be used (see Sections 21.12 and 21.13).

21.11.4 Adjusted residuals, $\mathbf{R}^2$, $\bar{\mathbf{R}}^2$, and other statistics

In the case of the exact ML estimators, the 'adjusted' residuals are computed as follows (see Pesaran and Slater (1980), pp. 49, 136):

$$\tilde{\epsilon}_1 = \tilde{u}_1 \sqrt{\left\{[(1-\tilde{\rho}_2)^2 - \tilde{\rho}_1^2](1+\tilde{\rho}_2)/(1-\tilde{\rho}_2)\right\}} \tag{21.86}$$

$$\tilde{\epsilon}_2 = \tilde{u}_2 \sqrt{(1-\tilde{\rho}_2^2)} - \tilde{u}_1\ \tilde{\rho}_1 \sqrt{[(1+\tilde{\rho}_2)/(1-\tilde{\rho}_2)]} \tag{21.87}$$

$$\tilde{\epsilon}_t = \tilde{u}_t - \tilde{\rho}_1 \tilde{u}_{t-1} - \tilde{\rho}_2 \tilde{u}_{t-2}, \qquad t = 3, 4, ..., n. \tag{21.88}$$

where

$$\tilde{u}_t = y_t - \mathbf{x}_t' \tilde{\boldsymbol{\beta}}, \qquad t = 1, 2, ..., n$$

are the 'unadjusted' residuals, and

$$\tilde{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{R}(\tilde{\boldsymbol{\rho}})\mathbf{X}\right)^{-1} \mathbf{X}'\mathbf{R}(\tilde{\boldsymbol{\rho}})\mathbf{y} \tag{21.89}$$

Recall that $\tilde{\boldsymbol{\rho}} = (\tilde{\rho}_1,\ \tilde{\rho}_2)'$. The program also takes account of the specification of the AR-error process in computations of the fitted values. Denoting these adjusted (or conditional) fitted values by $\tilde{y}_t$, we have

$$\tilde{y}_t = \tilde{E}(y_t \,|y_{t-1}, y_{t-2}, ...;\ \mathbf{x}_t, \mathbf{x}_{t-1}, ...) = y_t - \tilde{\epsilon}_t, \quad t = 1, 2, ..., n \tag{21.90}$$

The standard error of the regression is computed using the formula

$$\hat{\sigma}_\epsilon^2 = \tilde{\mathbf{u}}'\mathbf{R}(\tilde{\boldsymbol{\rho}})\tilde{\mathbf{u}}/(n-k-m) \tag{21.91}$$

where $m = 1$, for the $AR(1)$ case, and $m = 2$ for the $AR(2)$ case. Given the way the adjusted residuals $\tilde{\epsilon}_t$ are defined above, we also have

$$\hat{\sigma}_\epsilon^2 = \tilde{\mathbf{u}}'\mathbf{R}(\tilde{\boldsymbol{\rho}})\tilde{\mathbf{u}}/(n-k-m) = \sum_{t=1}^{n} \tilde{\epsilon}_t^2/(n-k-m) \tag{21.92}$$

Notice that this estimator of σ_ϵ^2 differs from the ML estimator given by

$$\tilde{\sigma}_\epsilon^2 = \sum_{t=1}^{n} \tilde{\epsilon}_t^2/n$$

and the estimator adopted in Pesaran and Slater (1980). The difference lies in the way the sum of squares of residuals, $\sum_{t=1}^{n} \tilde{\epsilon}_t^2$, is corrected for the loss in degrees of freedom arising from the estimation of the regression coefficients, $\boldsymbol{\beta}$, and the parameters of the error process, $\boldsymbol{\rho} = (\rho_1, \rho_2)'$.

The $R^2, \bar{R}^2$, and the F-statistic are computed from the adjusted residuals:

$$\begin{aligned} R^2 &= 1 - \left(\sum_{t=1}^{n} \tilde{\epsilon}_t^2 \Bigg/ \sum_{t=1}^{n} (y_t - \bar{y})^2 \right) \\ \bar{R}^2 &= 1 - (\hat{\sigma}_\epsilon^2 / \hat{\sigma}_y^2) \end{aligned} \tag{21.93}$$

where $\hat{\sigma}_y$ is the standard deviation of the dependent variable, defined as before by $\hat{\sigma}_y^2 = \sum_{t=1}^{n} (y_t - \bar{y})^2/(n-1)$.

The Durbin-Watson statistic is also computed using the adjusted residuals, $\tilde{\epsilon}_t$:

$$\widetilde{DW} = \frac{\sum\limits_{t=2}^{n} (\tilde{\epsilon}_t - \tilde{\epsilon}_{t-1})^2}{\sum\limits_{t=1}^{n} \tilde{\epsilon}_t^2}$$

The F-statistics reported following the regression results are computed according to the formula

$$F\text{-statistic} = \left(\frac{R^2}{1 - R^2} \right) \left(\frac{n-k-m}{k+m-1} \right) \overset{a}{\sim} F(k+m-1,\ n-k-m) \tag{21.94}$$

with

$$m = 1, \qquad \text{under } AR(1) \text{ error specification}$$

and

$$m = 2, \qquad \text{under } AR(2) \text{ error specification}$$

Notice that R^2 in (21.94) is given by (21.93). The above F-statistic can be used to test the joint hypothesis that except for the intercept term, all the other regression coefficients *and* the parameters of the AR-error process are zero. Under this hypothesis the F-statistic is distributed *approximately* as F with $k + m - 1$ and $n - k - m$ degrees of freedom. The chi-squared version of this test can be based on $nR^2/(1 - R^2)$, which under the null hypothesis of zero slope and AR coefficients is asymptotically distributed as a chi-squared variate with $k + m - 1$ degrees of freedom.

21.11.5 Log-likelihood ratio statistics for tests of residual serial correlation

The log-likelihood ratio statistic for the test of $AR(1)$ against the non-autocorrelated error specification is given by

$$\chi^2_{AR1,OLS} = 2(LL_{AR1} - LL_{OLS}) \overset{a}{\sim} \chi^2_1$$

The log-likelihood ratio statistic for the test of the $AR(2)$-error specification against the $AR(1)$-error specification is given by

$$\chi^2_{AR1,OLS} = 2(LL_{AR2} - LL_{AR1}) \overset{a}{\sim} \chi^2_1$$

Both of the above statistics are asymptotically distributed, under the null hypothesis, as a chi-squared variate with one degree of freedom.

The log-likelihood values, LL_{OLS}, LL_{AR1} and LL_{AR2}, represent the maximized values of the log-likelihood functions defined by (21.14), (21.75) and (21.78), respectively.

21.12 The Cochrane-Orcutt iterative method

This estimation method employs the Cochrane and Orcutt (1949) iterative procedure to compute ML estimators of (21.4) under the assumption that the disturbances, u_t, follow the $AR(m)$ process

$$u_t = \sum_{i=1}^{m} \rho_i u_{t-i} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2), \quad t = 1, 2, ..., n \tag{21.95}$$

with 'fixed initial' values. The 'fixed initial value' assumption is the same as treating the values, $y_1, y_2, ..., y_m$ as given or non-stochastic. This procedure in effect ignores the possible contribution of the distribution of the initial values to the overall log-likelihood function of the model. Once again the primary justification of treating initial values as fixed is asymptotic, and is plausible only when (21.95) is stationary and n is reasonably large (see Pesaran and Slater (1980) Section 3.2, and Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) Section 8.2.1c) for further discussion).

The log-likelihood function for this case is defined by

$$LL_{CO}(\boldsymbol{\theta}) = -\frac{(n-m)}{2} \log(2\pi\sigma_\epsilon^2) - \frac{1}{2\sigma_\epsilon^2} \sum_{t=m+1}^{n} \epsilon_t^2 + c \tag{21.96}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_\epsilon^2, \boldsymbol{\rho}')'$ with $\boldsymbol{\rho} = (\rho_1, \rho_2, ..., \rho_m)'$. Notice that the constant term c in (21.96) is undefined, and is usually set equal to zero. The Cochrane-Orcutt (CO) method maximizes $LL_{CO}(\boldsymbol{\theta})$, or equivalently minimizes $\Sigma_{t=m+1}^{n} \epsilon_t^2$ with respect to $\boldsymbol{\theta}$ by the iterative method of 'successive substitution'. Each iteration involves two steps: in the first step, LL_{CO} is maximized with respect to $\boldsymbol{\beta}$, taking $\boldsymbol{\rho}$ as given. In the second step $\boldsymbol{\beta}$ is taken as given, and the log-likelihood function is maximized with respect to $\boldsymbol{\rho}$. In each of these steps the optimization problem is solved by running OLS regressions. To start the iterations, ρ is initially set equal to zero. The iterations are terminated if

$$\sum_{i=1}^{m} \left| \tilde{\rho}_{i,(j)} - \tilde{\rho}_{i,(j-1)} \right| < m/1000 \tag{21.97}$$

where $\boldsymbol{\rho}_{(j)} = (\tilde{\rho}_{1,(j)}, \tilde{\rho}_{2,(j)}, ..., \tilde{\rho}_{m(j)})'$, and $\boldsymbol{\rho}_{(j-1)}$ denotes estimators of $\boldsymbol{\rho}$ in the jth and $(j-1)$th iterations, respectively. The estimator of σ_ϵ^2 is computed as

$$\hat{\sigma}_\epsilon^2 = \sum_{t=m+1}^{n} \tilde{\epsilon}_t^2 / (n - 2m - k) \tag{21.98}$$

where $\tilde{\epsilon}_t$, the adjusted residuals, are given by

$$\tilde{\epsilon}_t = \tilde{u}_t - \sum_{i=1}^{m} \tilde{\rho}_i \tilde{u}_{t-i}, \qquad t = m+1, m+2, ..., n \tag{21.99}$$

where

$$\tilde{u}_t = y_t - \sum_{i=1}^{k} \tilde{\beta}_i x_{it}, \qquad t = 1, 2, ..., n \tag{21.100}$$

As before, the symbol $\sim$ on top of an unknown parameter stands for ML estimators (now under fixed initial values). The estimator of σ_ϵ^2 in (21.98) differs from the ML estimator, given by $\tilde{\sigma}_\epsilon^2 = \sum_{t=m+1}^{n} \tilde{\epsilon}_t^2/(n-m)$. The estimator $\hat{\sigma}_\epsilon^2$ allows for the loss of the degrees of freedom associated with the estimation of the unknown coefficients, $\boldsymbol{\beta}$, and the parameters of the AR process, $\boldsymbol{\rho}$. Notice also that the estimator of σ_ϵ^2 is based on $n-m$ adjusted residuals, since the initial values $y_1, y_2, ..., y_m$ are treated as fixed.

The adjusted fitted values, $\tilde{y}_t$, in the case of this option are computed as

$$\tilde{y}_t = \hat{E}(y_t \,| y_{t-1}, y_{t-2}, ...; \mathbf{x}_t, \mathbf{x}_{t-1}, ...) = y_t - \tilde{\epsilon}_t \tag{21.101}$$

for $t = m+1, m+2, ..., n$. Notice that the initial values $\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_m$, are not defined.

In the case where $m = 1$, the program also provides a plot of the concentrated log-likelihood function in terms of ρ_1, defined by

$$LL_{CO}(\tilde{\rho}_1) = -\frac{(n-1)}{2}[1 + \log(2\pi\tilde{\sigma}_\epsilon^2)] \tag{21.102}$$

where

$$\tilde{\sigma}_\epsilon^2 = \sum_{t=2}^{n} \tilde{\epsilon}_t^2/(n-1)$$

and $\tilde{\epsilon}_t = \tilde{u}_t - \tilde{\rho}_1 \tilde{u}_{t-1}$.

21.12.1 Covariance matrix of the CO estimators

The estimator of the asymptotic variance matrix of $\tilde{\boldsymbol{\phi}} = (\tilde{\boldsymbol{\beta}}', \tilde{\boldsymbol{\rho}}')'$ is computed as

$$\hat{V}(\tilde{\boldsymbol{\phi}}) = \hat{\sigma}_\epsilon^2 \begin{bmatrix} \tilde{\mathbf{X}}_*'\tilde{\mathbf{X}}_* & \tilde{\mathbf{X}}_*'\mathbf{S} \\ \mathbf{S}'\tilde{\mathbf{X}}_* & \mathbf{S}'\mathbf{S} \end{bmatrix}^{-1} \tag{21.103}$$

where $\tilde{\mathbf{X}}_*$ is the $(n-m) \times k$ matrix of transformed regressors[11]

$$\tilde{\mathbf{X}}_* = \sum_{i=1}^{m} \tilde{\rho}_i \mathbf{X}_{-i} \tag{21.104}$$

[11] A typical element of $\tilde{X}_*$ is given by

$$\tilde{x}_{jt}^* = x_{jt} - \sum_{i=1}^{m} \tilde{\rho}_i x_{j,t-i} \qquad t = m+1, m+2, ..., n, \qquad j = 1, 2, ..., k.$$

and $\mathbf{S}$ is an $(n-m) \times m$ matrix containing the m lagged values of the CO residuals, $\tilde{u}_t$, namely

$$\mathbf{S} = \begin{bmatrix} \tilde{u}_m & \tilde{u}_{m-1} & \dots & \tilde{u}_1 \\ \tilde{u}_{m+1} & \tilde{u}_m & \dots & \tilde{u}_2 \\ \vdots & \vdots & & \vdots \\ \tilde{u}_{n-1} & \tilde{u}_{n-2} & \dots & \tilde{u}_{n-m} \end{bmatrix} \tag{21.105}$$

The unadjusted residuals, $\tilde{u}_t$, are already defined by (21.100). The above estimator of the variance matrix of $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\rho}}$ is asymptotically valid even if the regression model (21.4) contains lagged dependent variables.

21.13 ML/AR estimators by the Gauss-Newton method

This method provides an alternative numerical procedure for the maximization of the log-likelihood function (21.96). In cases where this log-likelihood function has a unique maximum, the Gauss-Newton and the CO iterative methods should converge to nearly identical results. But in general this need not be the case. However, the Gauss-Newton method is likely to perform better than the CO method when the regression equation contains lagged dependent variables.

The computations for the Gauss-Newton procedure are based on the following iterative relations:

$$\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\rho}} \end{pmatrix}_j = \begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\rho}} \end{pmatrix}_{j-1} + \begin{bmatrix} \tilde{\mathbf{X}}_*'\tilde{\mathbf{X}}_* & \tilde{\mathbf{X}}_*'\mathbf{S} \\ \mathbf{S}'\tilde{\mathbf{X}}_* & \mathbf{S}'\mathbf{S} \end{bmatrix}_{j-1}^{-1} \begin{bmatrix} \tilde{\mathbf{X}}_*'\tilde{\boldsymbol{\epsilon}} \\ \mathbf{S}'\tilde{\boldsymbol{\epsilon}} \end{bmatrix}_{j-1} \tag{21.106}$$

where the subscripts j and $j-1$ refer to the jth and the $(j-1)$th iterations; and $\tilde{\boldsymbol{\epsilon}} = (\tilde{\epsilon}_{m+1}, \tilde{\epsilon}_{m+2}, ..., \tilde{\epsilon}_n)'$, $\tilde{\mathbf{X}}_*$, and $\mathbf{S}$ have the same expressions as those already defined by (21.99), (21.104), and (21.105) respectively. The program starts the iterations with

$$\begin{aligned} \tilde{\boldsymbol{\beta}}_{(0)} &= \tilde{\boldsymbol{\beta}}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ \tilde{\boldsymbol{\rho}}_{(0)} &= 0 \end{aligned}$$

and ends them if either the number of iterations exceeds 20, or if the condition (21.97) is satisfied.

On exit from the iterations the program computes a number of statistics including estimates of σ_ϵ^2, the variance matrices of $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\rho}}$, $R^2, \bar{R}^2$, and so on, using the results already set out in Sections 21.12.1 and 21.11.4.

21.13.1 AR(m) error process with zero restrictions

The program applies the Gauss-Newton iterative method to compute estimates of the regression equation when the AR(m) error process (21.95) is subject to zero restrictions (see option 5 in the Linear Regression Estimation Menu, and Section 6.10). Notice that in the restricted case the estimator of the standard error of the regression is given by

$$\hat{\sigma}_\epsilon^2 = \sum_{t=m+1}^{n} \tilde{\epsilon}_t^2/(n-m-r-k), \qquad n > m+r+k \tag{21.107}$$

where r represents the number of non-zero parameters of the AR(m) process.

Similarly, the appropriate formula for the F-statistic (21.15) is now given by

$$F = \left(\frac{R^2}{1-R^2}\right)\left(\frac{n-m-k-r}{k+r-1}\right) \overset{a}{\sim} F(k+r-1, n-m-k-r) \tag{21.108}$$

The chi-squared version of this statistic can, as before, be computed by $nR^2/(1-R^2)$, which is asymptotically distributed (under the null hypothesis) as a chi-squared variate with $k+r-1$ degrees of freedom.

21.14 The IV/AR estimation method

This procedure provides estimates of the following linear regression model with $AR(m)$ errors by the instrumental variable method:

$$\begin{aligned} y_t &= \mathbf{x}_t'\boldsymbol{\beta} + u_t, \qquad t = 1, 2, ..., n \\ u_t &= \sum_{i=1}^{m} \rho_i u_{t-i} + \epsilon_t, \qquad t = m+1, m+2, ..., n \end{aligned}$$

The method assumes that there exists an $n \times s$ matrix Z containing observations on the s instrumental variables $\mathbf{z}_t (s \geq k+m)$. Then the IV/AR estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\rho}$ are computed by minimizing the criterion function (see Scheffe (1959)):

$$Q(\boldsymbol{\beta}, \boldsymbol{\rho}) = \boldsymbol{\epsilon}'\mathbf{P}_z\boldsymbol{\epsilon} \tag{21.109}$$

with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\rho}$, where $\mathbf{P}_z$ is the projection matrix defined by (21.58), and $\boldsymbol{\epsilon} = (\epsilon_{m+1}, \epsilon_{m+1}, ..., \epsilon_n)'$, is the $(n-m) \times 1$ vector of 'adjusted' residuals.

Application of the Gauss-Newton method to solve the above minimization problem yields the following iterative relations that are a generalization of (21.106):

$$\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\rho}} \end{pmatrix}_j = \begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\rho}} \end{pmatrix}_{j-1} + \begin{bmatrix} \tilde{\mathbf{X}}_*'\mathbf{P}_z\tilde{\mathbf{X}}_* & \tilde{\mathbf{X}}_*'\mathbf{P}_z\mathbf{S} \\ \mathbf{S}'\mathbf{P}_z\tilde{\mathbf{X}}_* & \mathbf{S}'\mathbf{P}_z\mathbf{S} \end{bmatrix}_{j-1}^{-1} \begin{bmatrix} \tilde{\mathbf{X}}_*'\mathbf{P}_z\tilde{\boldsymbol{\epsilon}} \\ \mathbf{S}'\mathbf{P}_z\tilde{\boldsymbol{\epsilon}} \end{bmatrix}_{j-1} \tag{21.110}$$

The notations are as before (see Section 21.12).

The program starts the iterations with

$$\tilde{\boldsymbol{\beta}}_{(0)} = \tilde{\boldsymbol{\beta}}_{IV} = (\mathbf{X}'\mathbf{P}_z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_z\mathbf{y}$$

and

$$\tilde{\boldsymbol{\rho}}_{(0)} = 0$$

See Section 21.13 for details of the iterative process.

In addition to the usual summary statistics, the program also reports the minimized value of $Q(\boldsymbol{\beta}, \boldsymbol{\rho})$, namely $Q(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\rho}})$, where here $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\rho}}$ represent the IV/AR estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\rho}$, respectively. The program reports $Q(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\rho}})$ as the 'value of IV minimand'.

21.14.1 Sargan's general mis-specification test in the case of the IV/AR option

This is the generalization of the Sargan (1964) test described in Section 21.10.3, and is based on the statistic

$$\chi^2_{SM} = Q(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\rho}})/\hat{\sigma}^2_\epsilon \overset{a}{\sim} \chi^2(s-k-r) \tag{21.111}$$

where, as before,

$$\hat{\sigma}^2_\epsilon = \sum_{t=m+1}^{n} \tilde{\epsilon}^2_t/(n-m-r-k)$$

and $Q(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\rho}})$ is the minimized value of the IV minimand (21.109).[12] Under the null hypothesis that the regression equation (21.4) and the AR-error process (21.95) are correctly specified, and that the $s(s > k+r)$ instrumental variables $\mathbf{z}_t$ are valid instruments, Sargan's mis-specification statistic χ^2_{SM} in (21.111) is asymptotically distributed as a chi-squared variate with $s-k-r$ degrees of freedom.

21.14.2 $\mathbf{R^2}$,$\mathbf{\bar{R}^2}$,$\mathbf{GR^2}$,$\mathbf{\overline{GR}^2}$, and other statistics: AR options

The summary statistics reported in the case of the AR and IV/AR options are computed using the adjusted residuals $\tilde{\epsilon}_t$, defined by (21.99). The R^2 of the regression is computed as

$$R^2 = 1 - \left\{ \sum_{t=m+1}^{n} \tilde{\epsilon}^2_t \Bigg/ \sum_{t=m+1}^{n} (y_t - \bar{y}_m)^2 \right\} \tag{21.112}$$

where

$$\bar{y}_m = \sum_{t=m+1}^{n} y_t/(n-m) \tag{21.113}$$

Notice that in the case of the CO and IV/AR options the initial values $y_1, y_2, ..., y_m$ are assumed fixed.

The $\bar{R}^2$ is computed as

$$\overline{R}^2 = 1 - \left(\frac{n-m-1}{n-m-r-k}\right)(1-R^2) \tag{21.114}$$

The F-statistic in the case of the AR options is computed as

$$F\text{-statistic} = \left(\frac{R^2}{1-R^2}\right)\left(\frac{n-m-r-k}{k+r-1}\right) \overset{a}{\sim} F(k+r-1, n-m-k-r) \tag{21.115}$$

On the null hypothesis that all the regression coefficients other than the intercept term are zero, and that $\boldsymbol{\rho} = 0$, the above F-statistic is approximately distributed as F with $k+r-1$, and $n-m-k-r$ degrees of freedom. Notice that in the case of the CO option, $r = m$. Also asymptotically,

$$nR^2/(1-R^2) \overset{a}{\sim} \chi^2_{k+r-1}$$

[12] Recall that r is the number of non-zero coefficients of (21.95).

The DW statistic is computed for this option using the adjusted residuals, $\tilde{\epsilon}_t$; namely

$$DW = \frac{\sum_{t=m+2}^{n} (\tilde{\epsilon}_t - \tilde{\epsilon}_{t-1})^2}{\sum_{t=m+1}^{n} \tilde{\epsilon}_t^2}$$

The computation of the GR^2 and $\overline{GR}^2$ statistics are based on the one-step ahead (in-sample) prediction errors defined by

$$\tilde{\xi}_t = \tilde{\epsilon}_t + \tilde{\boldsymbol{\beta}}'(\mathbf{x}_t - \hat{\mathbf{x}}_t), \; t = m+1, m+2, ..., n \tag{21.116}$$

where $\hat{\mathbf{x}}_t$, $t = 1, 2, ..., n$ are the fitted values from the regression of $\mathbf{x}_t$ on the instrumental variables $\mathbf{z}_t$. Relation (21.116) is a generalization of (21.65) to the case of serially correlated errors. More specifically we have

$$\begin{aligned} GR^2 &= 1 - \left(\sum_{t=m+1}^{n} \tilde{\xi}_t^2 \right) \Big/ \left(\sum_{t=m+1}^{n} (y_t - \bar{y}_m)^2 \right) \\ \overline{GR}^2 &= 1 - \left(\frac{n-m-1}{n-m-r-k} \right) \left(1 - GR^2\right) \end{aligned} \tag{21.117}$$

where $\bar{y}_m$ is defined by (21.113). Clearly, these measures reduce to the corresponding least squares measures in (21.112) and (21.114) when $\mathbf{z}_t$ is a sub-set of $\mathbf{x}_t$. See also Section 21.10.2.

21.15 Exact ML/MA estimators

The moving-average (MA) estimation option in the Linear Regression Estimation Menu (see Section 6.5) computes estimates of the parameters of the regression model:

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t, \qquad t = 1, 2, ..., n \tag{21.118}$$

where

$$u_t = \sum_{i=0}^{q} \gamma_i \epsilon_{t-i}, \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2), \qquad \gamma_0 \equiv 1 \tag{21.119}$$

by maximizing the log-likelihood function

$$LL_{MA}(\boldsymbol{\theta}) = -\frac{n}{2}(2\pi\sigma_\epsilon^2) - \tfrac{1}{2}\log|\boldsymbol{\Omega}| - \frac{1}{2\sigma_\epsilon^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Omega}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{21.120}$$

where $\mathbf{u} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$, and $E(\mathbf{u}\mathbf{u}') = \sigma_\epsilon^2\boldsymbol{\Omega}$. This yields *exact ML* estimates of the unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}', \gamma_1, \gamma_2, ..., \gamma_q, \sigma_\epsilon^2)'$, when the regressors $\mathbf{x}_t$ do not include lagged values of y_t.

The numerical method used to carry out the above maximization problem is similar to the Kalman filter procedure and is described in Pesaran (1988a).[13]

The method involves a Cholesky decomposition of the variance-covariance matrix $\boldsymbol{\Omega}$. For the $MA(1)$ error specification we have:

$$\boldsymbol{\Omega}=\begin{bmatrix} 1+\gamma_1^2 & \gamma_1 & 0 & \dots & 0 & 0 \\ \gamma_1 & 1+\gamma_1^2 & \gamma_1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ & & & & 1+\gamma_1^2 & \gamma_1 \\ 0 & 0 & 0 & \dots & \gamma_1 & 1+\gamma_1^2 \end{bmatrix}=\mathbf{HWH}' \tag{21.121}$$

where $\mathbf{W}$ is a diagonal matrix with elements w_t, $t=1,2,...,n$, and $\mathbf{H}$ is the upper triangular matrix

$$H=\begin{bmatrix} 1 & h_1 & 0 & \dots & & 0 \\ & 1 & h_2 & \dots & & 0 \\ & & \cdot & & & \vdots \\ & & & \cdot & & \\ & \mathbf{0} & & & \cdot & 0 \\ & & & & 1 & h_{n-1} \\ & & & & & 1 \end{bmatrix} \tag{21.122}$$

The elements w_t and h_t satisfy the following *forward* recursions:[14]

$$\begin{aligned} h_t &= \gamma_1/w_{t+1}, & \text{for} \quad t=n-1,n-2,...,1 \\ w_t &= 1+\gamma_1^2-w_{t+1}h_t^2, & \text{for} \quad t=n-1,n-2,...,1 \end{aligned}$$

starting with the terminal value of $w_n=1+\gamma_1^2$.

Using (21.121) in (21.120) now yields (notice that $\boldsymbol{\Omega}^{-1}=\mathbf{H}'^{-1}\mathbf{W}^{-1}\mathbf{H}^{-1}$ and $|\boldsymbol{\Omega}|=|\mathbf{W}|=w_1w_2,...,w_n$)

$$LL_{MA}(\boldsymbol{\theta})=-\tfrac{n}{2}\log(2\pi\sigma_\epsilon^2)-\tfrac{1}{2}\sum_{t=1}^{n}\log w_t-\frac{1}{2\sigma_\epsilon^2}(\mathbf{y}^*-\mathbf{X}^*\boldsymbol{\beta})'(\mathbf{y}^*-\mathbf{X}^*\boldsymbol{\beta}) \tag{21.123}$$

where

$$\begin{aligned} y_t^* &= w_t^{-1/2}y_t^f \\ \mathbf{x}_t^* &= w_t^{-1/2}\mathbf{x}_t^f \end{aligned}$$

[13] For a description of the Kalman filter algorithm and its use in the estimation of MA processes see, for example, Harvey (1989).

[14] In Pesaran (1988a), $\mathbf{H}$ is chosen to be a lower triangular matrix, and the resultant recursions are consequently backward in w_t and h_t. In the case of the MA option both the backward and the forward recursion methods yield identical results, but forward recursion is the appropriate method to use for estimation of certain classes of rational expectations models. See Section 21.16.

and y_t^f and $\mathbf{x}_t^f$ represent the forward filtered values of y_t and $\mathbf{x}_t$ defined by

$$\begin{aligned} y_t^f &= y_t - h_t y_{t+1}^f, \qquad \text{for} \quad t = n-1, n-2, ..., 1 \\ \mathbf{x}_t^f &= \mathbf{x}_t - h_t \mathbf{x}_{t+1}^f, \qquad \text{for} \quad t = n-1, n-2, ..., 1 \end{aligned}$$

with the terminal values $y_n^f = y_n$, $\mathbf{x}_n^f = \mathbf{x}_n$. The ML estimators of $\boldsymbol{\beta}$ and σ_ϵ^2 are given by

$$\begin{aligned} \tilde{\boldsymbol{\beta}} &= (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\mathbf{X}^{*\prime}\mathbf{y}^* \\ \tilde{\sigma}_\epsilon^2 &= (\mathbf{y}^* - \mathbf{X}^*\tilde{\boldsymbol{\beta}})'(\mathbf{y}^* - \mathbf{X}^*\tilde{\boldsymbol{\beta}})/n \end{aligned}$$

The estimation of γ_1 is carried out iteratively using the modified Powell conjugate direction algorithm that does not require derivatives (see below for further details and references).

The above procedure can be readily extended to higher-order MA processes. For a general $MA(q)$ process the generalization of (21.122) is given by

$$\mathbf{H} = \begin{bmatrix} 1 & h_{11} & h_{21} & \dots & h_{q1} & \dots & & 0 \\ & 1 & h_{12} & h_{22} & \dots & h_{q2} & & 0 \\ & & \cdot & \cdot & \cdot & & \ddots & \vdots \\ & & & \cdot & \cdot & \cdot & & h_{q,n-q} \\ & & & & \cdot & \cdot & & \vdots \\ & & \mathbf{0} & & & \cdot & \cdot & h_{2,n-2} \\ & & & & & & 1 & h_{1,n-1} \\ & & & & & & & 1 \end{bmatrix}$$

$$\begin{aligned} w_t &= \delta_0 - \sum_{i=1}^{q} h_{it}^2 w_{t+i}, && t = n-1, n-2, ..., 1 \\ h_{jt} &= w_{t+j}^{-1}\left(\delta_j - \sum_{i=j+1}^{q} h_{it} h_{i-j,t+j} w_{t+i}\right), && \begin{array}{l} t = n-j, n-j-1, ..., 1; \\ j = q-1, q-2, ..., 1 \end{array} \\ h_{qt} &= w_{t+q}^{-1}\delta_q, \\ \delta_s &= \begin{cases} \sum\limits_{i=1}^{q} \gamma_i \gamma_{i-s} & 0 \le s \le q, \\ 0 & s > q. \end{cases} \end{aligned}$$

The forward filters on y_t and $\mathbf{x}_t$ are given by

$$\begin{aligned} y_t^f &= y_t - \sum_{i=1}^{q} h_{it} y_{t+i}^f, \qquad \text{for} \quad t = n-1, n-2, ..., 1 \\ \mathbf{x}_t^f &= \mathbf{x}_t - \sum_{i=1}^{q} h_{it} \mathbf{x}_{t+i}^f, \qquad \text{for} \quad t = n-1, n-2, ..., 1 \end{aligned}$$

and as before, $y_t^* = w_t^{-1/2} y_t^f$, $\mathbf{x}_t^* = w_t^{-1/2}\mathbf{x}_t^f$, The terminal values for the above recursions are given by

$$\begin{aligned} w_n &= \delta_0 = 1 + \gamma_1^2 + \gamma_2^2 + \ldots + \gamma_q^2 \\ h_{jn} &= h_{j,n-1} = \cdots = h_{j,n-j+1} = 0 \\ y_n^f &= y_n \\ \mathbf{x}_n^f &= \mathbf{x}_n \end{aligned}$$

For a given value of $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, ..., \gamma_q)'$, the estimator of $\boldsymbol{\beta}$ is computed by the OLS regression of y_t^* on $\mathbf{x}_t^*$. The estimation of $\boldsymbol{\gamma}$ is carried out iteratively by the modified Powell's method of conjugate directions that *does not* require derivatives of the log-likelihood function. See Powell (1964), Brent (1973) Chapter 7, and Press, Flannery, Teukolsky, and Vetterling (1989) Section 10.5. The application of the Gauss-Newton method to the present problem requires derivatives of the log-likelihood function which are analytically intractable, and can be very time-consuming if they are to be computed numerically.

21.15.1 Covariance matrix of the unknown parameters in the MA option

In the case of the MA option the covariance matrix of $\tilde{\boldsymbol{\psi}} = (\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\gamma}})$ is computed as

$$\left[\frac{-\partial^2 LL_{MA}(\boldsymbol{\theta})}{\partial\boldsymbol{\psi}\partial\boldsymbol{\psi}'}\right]^{-1}_{\boldsymbol{\psi}=\tilde{\boldsymbol{\psi}}}$$

where the Hessian matrix, $-\partial^2 LL_{MA}(\boldsymbol{\theta})/\partial\boldsymbol{\psi}\partial\boldsymbol{\psi}'$ is computed by taking second numerical derivatives of the log-likelihood function defined by (21.123) at the ML estimators $\tilde{\boldsymbol{\psi}} = (\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\gamma}})$. Notice that this estimator of the variance matrix of $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\gamma}}$ is asymptotically valid even if the regression model (21.118) contains lagged values of the dependent variable.[15]

21.16 The IV/MA estimators

Consider the regression model (21.118) with the moving average errors (21.119), and suppose that there exists an $n \times s$ matrix, $\mathbf{Z}$, containing observations on the s instrumental variables, $\mathbf{z}_t$. Then the IV/MA estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are computed by minimizing the criterion function

$$S(\boldsymbol{\beta}, \boldsymbol{\gamma}) = (n/2)\mathbf{u}^{*'}\mathbf{P}_z\mathbf{u}^* + \tfrac{1}{2}\log|\boldsymbol{\Omega}| \tag{21.124}$$

where $\mathbf{u}^*$ are the forward filtered values of $\mathbf{u} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$, $E(\mathbf{uu}') = \sigma_\epsilon^2\boldsymbol{\Omega}$, and $\mathbf{P}_z$ is the projection matrix defined by (21.58). The forward filter procedure applied to the u_ts is the

[15] The program reports standard errors and probability values for the MA parameters only when none of the roots of

$$\sum_{i=0}^{q} \gamma_i \mathbf{z}^i = 0, \qquad \gamma_0 \equiv 1,$$

fall on the unit circle. The Lahmer-Schur algorithm is used to check whether any of the roots of the above polynomial equation falls on the unit circle. See, for example, Acton (1970) Chapter 7.

same as that described in Section 21.15. The term $\mathbf{u}^{*'}\mathbf{P}_z\mathbf{u}^*$ is the same as the IV criterion used in the Hayashi and Sims (1983) estimation procedure. Notice that only y_t and x_{it}s are forward filtered, and not the instruments. The second term in (21.124) is asymptotically negligible when the MA process in (21.119) is invertible, but as argued in Pesaran (1990), its inclusion in the criterion function helps ensure that in small samples $S(\boldsymbol{\beta},\boldsymbol{\gamma})$ achieves a minimum when the roots of the MA process are close to the unit circle.

For given values of γ_i, $i = 1, 2, ..., q$, the regression coefficients $\boldsymbol{\beta}$ are estimated by the IV regression of the forward filtered variables $\mathbf{y}^*$ on $\mathbf{X}^*$ (see Section 21.15 for the relevant expressions for the forward filtering procedure). The estimation of $\boldsymbol{\gamma}$ is carried out iteratively by the modified Powell's conjugate direction algorithm that does not require derivatives of the IV minimand defined in (21.124) (see Section 21.15 and Pesaran (1990) for further details.) The computation of the variance-covariance matrix of the parameter estimates, and the various summary statistics reported by the program for this option, are also carried out along the lines set out in Sections 21.15.1 and 21.16.1.

21.16.1 $R^2, \bar{R}^2, GR^2, \overline{GR}^2$, and other statistics: MA options

The summary statistics reported in the case of the MA and the MA/IV options are computed using the adjusted residuals $\tilde{\epsilon}_t$ defined by

$$\tilde{\epsilon}_t = -\sum_{i=1}^{q} \gamma_i \tilde{\epsilon}_{t-i} + \tilde{u}_t, \qquad t = 1, 2, ..., n$$

where the initial values $\tilde{\epsilon}_0, \tilde{\epsilon}_{-1}, ..., \tilde{\epsilon}_{-q+1}$ are set equal to zero, and

$$\tilde{u}_t = y_t - \sum_{i=1}^{k} \tilde{\beta}_i \mathbf{x}_{it}, \qquad t = 1, 2, ..., n$$

Assuming that the $MA(q)$ process (21.119) is estimated with $q - r$ zero restrictions, σ_ϵ^2, the variance of ϵ_t is estimated by $(\gamma_q \neq 0)$

$$\hat{\sigma}_\epsilon^2 = \sum_{t=1}^{n} \tilde{\epsilon}_t^2/(n - r - k), \qquad n > r + k$$

The other statistics included in the result table are computed as

$$\begin{aligned} R^2 &= 1 - \left(\sum_{t=1}^{n} \tilde{\epsilon}_t^2\right) \Bigg/ \left(\sum_{t=1}^{n} (y_t - \bar{y})^2\right) \\ \bar{R}^2 &= 1 - \left(\frac{n-1}{n-r-k}\right)\left(1 - R^2\right), \end{aligned} \tag{21.125}$$

$$F = \left(\frac{R^2}{1-R^2}\right)\left(\frac{n-k-r}{k+r-1}\right) \overset{a}{\sim} F(k+r-1, n-k-r) \tag{21.126}$$

and

$$DW = \sum_{t=2}^{n}(\tilde{\epsilon}_t - \tilde{\epsilon}_{t-1})^2 \Bigg/ \sum_{t=1}^{n}\tilde{\epsilon}_t^2 \tag{21.127}$$

Under the null hypothesis that all the regression coefficients other than the intercept term are zero and $\boldsymbol{\gamma} = 0$, the F-statistic (21.126) is approximately distributed as F with $k+r-1$ and $n-k-r$ degrees of freedom.

As in the case of the IV/AR option, the computation of the GR^2 and $\overline{GR}^2$ statistics are based on the one-step ahead (in-sample) prediction errors defined by

$$\tilde{\xi}_t = \tilde{\epsilon}_t + \tilde{\boldsymbol{\beta}}'(\mathbf{x}_t - \hat{\mathbf{x}}_t),\ t = 1, 2, ..., n \tag{21.128}$$

where $\hat{\mathbf{x}}_t$, $t = 1, 2, ..., n$ are the fitted values from the regression of $\mathbf{x}_t$ on the instrumental variables $\mathbf{z}_t$. More specifically we have

$$\begin{aligned} GR^2 &= 1 - \left(\sum_{t=1}^{n}\tilde{\xi}_t^2\right) \Bigg/ \left(\sum_{t=1}^{n}(y_t - \bar{y})^2\right) \\ \overline{GR}^2 &= 1 - \left(\frac{n-1}{n-r-k}\right)\left(1 - GR^2\right) \end{aligned} \tag{21.129}$$

21.17 Recursive regressions

The econometric model underlying the recursive regressions is given by

$$\underset{1\times 1}{\mathbf{y}_t} = \underset{1\times k}{\mathbf{x}_t'}\ \underset{k\times 1}{\boldsymbol{\beta}_t} + \underset{1\times 1}{\mathbf{u}_t}, \qquad t = 1, 2, ..., n \tag{21.130}$$

where the coefficients $\boldsymbol{\beta}_t$, and the variances of the disturbance terms, σ_t^2, are now allowed to vary with t, typically a time subscript. (Notice that for this option it is necessary that the observations are ordered.)

21.17.1 The CUSUM test

The Cumulative Sum ($CUSUM$) test is described in Brown, Durbin, and Evans (1975), and is based on the $CUSUM$ of recursive residuals defined by

$$W_r = \frac{1}{\hat{\sigma}_{OLS}} \sum_{j=k+1}^{r} v_j, \qquad r = k+1, k+2, ..., n \tag{21.131}$$

where v_t is the recursive residual based on the first j observations given below by (21.136), and $\hat{\sigma}_{OLS}$ is already defined by (21.7).

The test employs a graphic technique and involves plotting W_r and a pair of straight lines for values of $r = k+1, k+2, ..., n$. The straight lines are drawn assuming 5 per cent significance level.

The equations of the lines are given by

$$W = \pm\left\{0.948\sqrt{(n-k)} + 1.896(r-k)\sqrt{(n-k)}\right\} \tag{21.132}$$

for $r = k+1, k+2, ..., n$.

For further details see Brown, Durbin, and Evans (1975) Section 2.3, and Harvey (1981) pp. 151-154.

21.17.2 The CUSUM of squares test

This test is described fully in Brown, Durbin, and Evans (1975), and employs the squared recursive residuals v_j^2. It is based on the quantities

$$WW_r = \sum_{j=k+1}^{r} v_j^2 \Big/ \sum_{j=k+1}^{n} v_j^2, \qquad r = k+1, k+2, ..., n \tag{21.133}$$

and involves plotting WW_r and a pair of lines whose equations are given by

$$WW = \pm c_0 + (r-k)/(n-k), \qquad r = k+1, k+2, ..., n \tag{21.134}$$

where c_0 is determined by the significance level chosen for the test. The program uses the values of c_0 appropriate for a five per cent significance level. These are based on the critical values in Harvey (1981) Table C, pp. 364-5.

21.17.3 Recursive coefficients: the OLS option

Let

$$\begin{aligned} \mathbf{X}_r &= (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_r)' \\ \mathbf{y}_r &= (y_1, y_2, ..., y_r)' \end{aligned}$$

then the recursive coefficients are defined by

$$\hat{\boldsymbol{\beta}}_r = (\mathbf{X}_r'\mathbf{X}_r)^{-1}\mathbf{X}_r'\mathbf{y}_r, \qquad r = k+1, k+2, ..., n \tag{21.135}$$

The program computes $\hat{\boldsymbol{\beta}}_r$ recursively, using the results (3) and (4) in Brown, Durbin, and Evans (1975) p. 152.

21.17.4 Standardized recursive residuals: the OLS option

The standardized recursive residuals based on the first r observations are defined by[16]

$$v_r = (y_r - \mathbf{x}_r'\widehat{\boldsymbol{\beta}}_{r-1})/d_r, \qquad r = k+1, k+2, ..., n \tag{21.136}$$

[16] Notice that under

$$H_0: \left\{ \begin{array}{l} \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_n = \boldsymbol{\beta} \\ \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2 \end{array} \right.$$

the recursive residuals v_r, $r = k+1, ..., n$ are independent, $N(0, \sigma^2)$. See Lemma 1 in Brown, Durbin, and Evans (1975).

where $\hat{\boldsymbol{\beta}}_r$ are defined by (21.135), and

$$d_r = \sqrt{\left\{1 + \mathbf{x}_r'(\mathbf{X}_{r-1}'\mathbf{X}_{r-1})^{-1}\mathbf{x}_r\right\}} \tag{21.137}$$

21.17.5 Recursive standard errors: the OLS option

Denoting the estimator of σ_t^2 based on the first r observations by $\hat{\sigma}_r^2$, we have

$$\widehat{\sigma}_r^2 = S_r/(r-k), \qquad r = k+1, k+2, ..., n \tag{21.138}$$

where

$$S_r = S_{r-1} + v_r^2, \qquad r = k+1, k+2, ..., n,$$

with $S_k = 0$. Equivalently $S_r = (\mathbf{y}_r - \mathbf{X}_r\hat{\boldsymbol{\beta}}_r)'(\mathbf{y}_r - \mathbf{X}_r\hat{\boldsymbol{\beta}}_r)$.

21.17.6 Recursive estimation: the IV option

Recursive coefficients In the IV case the recursive coefficients are computed using the relations

$$\hat{\boldsymbol{\beta}}_{r,IV} = (\mathbf{X}_r'\mathbf{P}_r\mathbf{X}_r)^{-1}\mathbf{X}_r'\mathbf{P}_r\mathbf{y}_r \tag{21.139}$$

$$\hat{V}(\hat{\boldsymbol{\beta}}_{r,IV}) = \hat{\sigma}_{r,IV}^2(\mathbf{X}_r'\mathbf{P}_r\mathbf{X}_r)^{-1} \tag{21.140}$$

for $r = k+1, k+2, ..., n$, where

$$\begin{aligned} \mathbf{P}_r &= \mathbf{Z}_r(\mathbf{Z}_r'\mathbf{Z}_r)^{-1}\mathbf{Z}_r' \\ \mathbf{Z}_r &= (\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_r)' \end{aligned}$$

$\hat{\sigma}_{r,IV}^2$ is defined by (21.143) below and $\mathbf{z}_t$, $t = 1, 2, ..., n$ are the $s \times 1$ vectors of observations on the s instrumental variables

Standardized recursive residuals

$$v_{r,IV} = (y_r - \mathbf{x}_r'\hat{\boldsymbol{\beta}}_{r-1,IV})/d_{r,IV}, \qquad r = k+1, k+2, ..., n \tag{21.141}$$

where $\hat{\boldsymbol{\beta}}_{r,IV}$ is defined by (21.139) and

$$d_{r,IV} = \left\{1 + \mathbf{x}_r'(\mathbf{X}_{r-1}'\mathbf{P}_{r-1}\mathbf{X}_{r-1})^{-1}\mathbf{x}_r\right\}^{1/2} \tag{21.142}$$

Recursive standard errors

$$\hat{\sigma}_{r,IV}^2 = (\mathbf{y}_r - \mathbf{X}_r\hat{\boldsymbol{\beta}}_{r,IV})'(\mathbf{y}_r - \mathbf{X}_r\hat{\boldsymbol{\beta}}_{r,IV})/(r-k) \tag{21.143}$$

for $r = k+1, k+2, ..., n$.

21.17.7 Adaptive coefficients in expectations formation models under incomplete learning

Starting with the parameter varying model (21.130), as shown in Pesaran (1987b) Section 9.3.2, the augmented adaptive learning model under incomplete information can be written as

$$_{t}y_{t+1}^{*} -_{t-1} y_{t}^{*} = \mu_t(y_t -_{t-1} y_t^{*}) + (\mathbf{\Delta x}_t)'\hat{\boldsymbol{\beta}}_{t-1} + \text{error}$$

where $_{t}y_{t+1}^{*}$ denotes the expectations of y_{t+1} formed at time t. The adaptive coefficients in this model are computed as

$$\mu_t = \mathbf{x}_t' \left(\sum_{j=1}^{t} \mathbf{x}_t\mathbf{x}_j' \right)^{-1} \mathbf{x}_t = \mathbf{x}_t'(\mathbf{X}_t'\mathbf{X}_t)^{-1}\mathbf{x}_t, \qquad t = k+1, k+2, ..., n \tag{21.144}$$

21.17.8 Recursive predictions

Two sets of recursive predictions are computed by the program:

1. conditional on the actual values of $\mathbf{x}_t$
2. conditional on the k variables $\mathbf{w}_t$ (typically predictors of $\mathbf{x}_t$)

In case 1, recursive predictions of y_t are computed as

$$\hat{y}_{Rt} = \mathbf{x}_t'\widehat{\boldsymbol{\beta}}_{t-1}, \qquad t = k+1, k+2, ..., n \tag{21.145}$$

In case 2 predictions of y_t are computed as

$$\hat{y}_{Rt} = \mathbf{w}_t'\widehat{\boldsymbol{\beta}}_{t-1}, \qquad t = k+1, k+2, ..., n \tag{21.146}$$

where $\hat{\boldsymbol{\beta}}_t$ is defined by (21.135) in the case of the *OLS* option, and by (21.139) in the case of the IV option.

The standard errors of the recursive forecasts are computed using the results

$$\textit{OLS} \text{ option} \qquad \hat{V}\left(\hat{y}_{Rt}\right) = \hat{\sigma}_{t-1}^2\mathbf{w}_t'(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}\mathbf{w}_t \tag{21.147}$$

where $\hat{\sigma}_{t-1}^2$ is defined by (21.138), and

$$\text{IV option} \qquad \hat{V}\left(\hat{y}_{Rt}\right) = \hat{\sigma}_{t-1}^2\mathbf{w}_t'(\mathbf{X}_{t-1}'\mathbf{P}_{t-1}\mathbf{X}_{t-1})^{-1}\mathbf{w}_t \tag{21.148}$$

where $\hat{\sigma}_{t-1}^2$ is defined by (21.143).

In the case where the forecasts are based on the actual values of the regressors, in the above formula $\mathbf{w}_t$ is replaced by $\mathbf{x}_t$. The matrices $\mathbf{X}_t$ and $\mathbf{P}_t$ are already defined in Sections 21.17.3 and 21.17.6.

21.18 Phillips-Hansen fully modified OLS estimators

This estimator was proposed by Phillips and Hansen (1990), and is appropriate for estimation and inference when there exists a *single* cointegrating relation between a set of $I(1)$ variables. Consider the following linear regression model:

$$y_t = \beta_0 + \boldsymbol{\beta}_1' \mathbf{x}_t + u_t,\ t = 1, 2, ..., n \tag{21.149}$$

where the $k \times 1$ vector of $I(1)$ regressors are not themselves cointegrated. Therefore, $\mathbf{x}_t$ has a first-difference stationary process given by

$$\Delta \mathbf{x}_t = \boldsymbol{\mu} + \mathbf{v}_t,\ \ t = 2, 3, ..., n \tag{21.150}$$

in which $\boldsymbol{\mu}$ is a $k \times 1$ vector of drift parameters and $\mathbf{v}_t$ is a $k \times 1$ vector of $I(0)$, or stationary variables. It is also assumed that $\boldsymbol{\xi}_t = (u_t, \mathbf{v}_t')'$ is strictly stationary with zero mean and a finite positive-definite covariance matrix, Σ.

The computation of the FM-OLS estimator of $\boldsymbol{\beta}$ is carried out in two stages. In the first stage y_t is corrected for the long-run interdependence of u_t and $\mathbf{v}_t$. For this purpose let $\hat{u}_t$ be the OLS residual vector in (21.149), and write

$$\hat{\boldsymbol{\xi}}_t = \begin{pmatrix} \hat{u}_t \\ \hat{\mathbf{v}}_t \end{pmatrix}, t = 2, 3, ..., n \tag{21.151}$$

where $\hat{\mathbf{v}}_t = \Delta \mathbf{x}_t - \hat{\boldsymbol{\mu}}$, for $t = 2, 3, ..., n$, and $\hat{\boldsymbol{\mu}} = (n-1)^{-1} \sum_{t=2}^{n} \Delta \mathbf{x}_t$.

A consistent estimator of the long-run variance of $\boldsymbol{\xi}_t$ is given by

$$\widehat{\boldsymbol{\Omega}} = \widehat{\boldsymbol{\Sigma}} + \widehat{\boldsymbol{\Lambda}} + \widehat{\boldsymbol{\Lambda}}' = \begin{bmatrix} \underset{1 \times 1}{\widehat{\Omega}_{11}} & \underset{1 \times k}{\widehat{\boldsymbol{\Omega}}_{12}} \\ \underset{k \times 1}{\widehat{\boldsymbol{\Omega}}_{21}} & \underset{k \times k}{\widehat{\boldsymbol{\Omega}}_{22}} \end{bmatrix} \tag{21.152}$$

where

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n-1} \sum_{t=2}^{n} \hat{\boldsymbol{\xi}}_t \hat{\boldsymbol{\xi}}_t' \tag{21.153}$$

and

$$\widehat{\boldsymbol{\Lambda}} = \sum_{s=1}^{m} \omega(s, m) \widehat{\boldsymbol{\Gamma}}_s \tag{21.154}$$

$$\widehat{\boldsymbol{\Gamma}}_s = n^{-1} \sum_{t=1}^{n-s} \hat{\boldsymbol{\xi}}_t \hat{\boldsymbol{\xi}}_{t+s}' \tag{21.155}$$

and $\omega(s, m)$ is the lag window with horizon (or truncation) m.

Now let

$$\widehat{\boldsymbol{\Delta}} = \widehat{\boldsymbol{\Sigma}} + \widehat{\boldsymbol{\Lambda}} = \begin{bmatrix} \widehat{\boldsymbol{\Delta}}_{11} & \widehat{\boldsymbol{\Delta}}_{12} \\ \widehat{\boldsymbol{\Delta}}_{21} & \widehat{\boldsymbol{\Delta}}_{22} \end{bmatrix} \tag{21.156}$$

$$\widehat{\mathbf{Z}} = \widehat{\mathbf{\Delta}}_{21} - \widehat{\mathbf{\Delta}}_{22}\widehat{\mathbf{\Omega}}_{22}^{-1}\widehat{\mathbf{\Omega}}_{21}, \tag{21.157}$$

$$\hat{y}_t^* = y_t - \widehat{\mathbf{\Omega}}_{12}\widehat{\mathbf{\Omega}}_{22}^{-1}\hat{\mathbf{v}}_t, \tag{21.158}$$

$$\underset{(k+1)\times k}{\mathbf{D}} = \begin{bmatrix} \underset{1\times k}{\mathbf{0}} \\ \\ \underset{k\times k}{\mathbf{I}_k} \end{bmatrix} \tag{21.159}$$

In the second stage the $FM-OLS$ estimator of β is given by

$$\widehat{\boldsymbol{\beta}}_* = (\mathbf{W}'\mathbf{W})^{-1}(\mathbf{W}'\hat{\mathbf{y}}^* - n\mathbf{D}\widehat{\mathbf{Z}}) \tag{21.160}$$

where $\hat{\mathbf{y}}^* = (\hat{y}_1^*, \hat{y}_2^*, ..., \hat{y}_n^*)'$, $\mathbf{W} = (\boldsymbol{\tau}_n, \mathbf{X})$, and $\boldsymbol{\tau}_n = (1, 1, ..., 1)'$.

21.18.1 Choice of lag windows ω(s,m)

The computation of the *FM-OLS* estimators can be carried out in *Microfit* for the following four choices of lag windows:

Uniform window

$$\omega(s,m) = 1, \qquad 0 \le s \le m$$

Bartlett window

$$\omega(s,m) = 1 - s/m, \qquad 0 \le s \le m$$

Tukey window

$$\omega(s,m) = \tfrac{1}{2}\left\{1 + \cos(\pi s/m)\right\}, \qquad 0 \le s \le \tfrac{m}{2}$$

Parzen window

$$\omega(s,m) = \begin{cases} 1 - 6(s/m)^2 + 6(s/m)^3, & 0 \le s \le \frac{m}{2}, \\ 2(1 - s/m)^3, & \frac{m}{2} < s \le m. \end{cases}$$

Notice, however, that the use of the uniform window may lead to an estimate of $\widehat{\mathbf{\Omega}}$ which is not a positive-definite matrix. The other three lag windows generally result in estimates of $\mathbf{\Omega}$ that are positive definite.

21.18.2 Estimation of the variance matrix of the FM-OLS estimator

A consistent estimator of the variance matrix of $\widehat{\boldsymbol{\beta}}_*$defined in (21.160) is given by

$$\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}}_*) = \hat{\omega}_{11.2}(\mathbf{W}'\mathbf{W})^{-1} \tag{21.161}$$

where

$$\hat{\omega}_{11.2} = \widehat{\Omega}_{11} - \widehat{\mathbf{\Omega}}_{12}\widehat{\mathbf{\Omega}}_{22}^{-1}\widehat{\mathbf{\Omega}}_{21} \tag{21.162}$$

The t-ratios of the FM-OLS estimators reported by *Microfit* are computed as the ratio of $\widehat{\beta}_{*i}$ to the square-root of the ith diagonal element of the matrix defined by (21.161).

The program also computes FM-OLS fitted values, $\hat{\mathbf{y}}_{FM-OLS} = \mathbf{W}\widehat{\boldsymbol{\beta}}_*$, FM-OLS residuals, $e_{FM\text{-}OLS} = \mathbf{y} - \mathbf{W}\widehat{\boldsymbol{\beta}}_*$, autocorrelation coefficients of the FM-OLS residuals, and enables the user to estimate and test linear and/or non-linear functions of the coefficients, $\boldsymbol{\beta}$. The relevant formulae are in Sections 21.24 and 21.25, with $\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\beta}}_*$, and $\hat{\sigma}^2\widehat{\mathbf{V}}(\widehat{\boldsymbol{\theta}}) = \widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}}_*)$, given by (21.160) and (21.161), respectively.

21.19 Autoregressive distributed lag models

Consider the following augmented autoregressive distributed lag $ARDL(p, q_1, q_2, ..., q_k)$ model:[17]

$$\phi(L,p)y_t = \sum_{i=1}^{k} \beta_i(L,q_i)x_{it} + \boldsymbol{\delta}'\mathbf{w}_t + u_t \tag{21.163}$$

where

$$\phi(L,p) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p \tag{21.164}$$

$$\beta_i(L,q_i) = \beta_{i0} + \beta_{i1}L + ... + \beta_{iq_i}L^{q_i}, \quad i = 1, 2, ..., k \tag{21.165}$$

L is a lag operator such that $Ly_t = y_{t-1}$, and $\mathbf{w}_t$ is a $s \times 1$ vector of deterministic variables such as the intercept term, seasonal dummies or time trends, or exogenous variables with fixed lags.

The $ARDL$ option in *Microfit* (option 6 in the Single Equation Estimation Menu) first estimates (21.163) by the OLS method for all possible values of $p = 0, 1, 2, ..., m$, $q_i = 0, 1, 2, ..., m$, $i = 1, 2, ..., k$; namely a total of $(m+1)^{k+1}$ different $ARDL$ models. The maximum lag, m, is chosen by the user, and all the models are estimated for the same sample period, namely $t = m+1, m+2, ..., n$.

In the second stage the user is given the option of selecting one of the $(m+1)^{k+1}$ estimated models using one of the following four model selection criteria: the $\bar{R}^2$ criterion, Akaike information criterion (AIC), Schwarz Bayesian criterion (SBC), and the Hannan and Quinn criterion (HQC).[18] The program then computes the long-run coefficients and their asymptotic standard errors for the selected $ARDL$ model. It also provides estimates of the error correction model (ECM) that corresponds to the selected $ARDL$ model. The long-run coefficients for the response of y_t to a unit change in x_{it} are estimated by

$$\widehat{\theta}_i = \frac{\widehat{\beta}_i(1,\hat{q}_i)}{\widehat{\phi}(1,\hat{p})} = \frac{\widehat{\beta}_{i0} + \widehat{\beta}_{i1} + ... + \widehat{\beta}_{i\hat{q}_i}}{1 - \widehat{\phi}_1 - \widehat{\phi}_2 - ... - \widehat{\phi}_{\hat{p}}}, \qquad i = 1, 2, ..., k \tag{21.166}$$

where $\hat{p}$ and $\hat{q}_i$, $i = 1, 2, ..., k$ are the selected (estimated) values of p and q_i, $i = 1, 2, ..., k$. Similarly, the long-run coefficients associated with the deterministic/exogenous variables

[17] For a comprehensive early review of $ARDL$ models see Hendry, Pagan, and Sargan (1984).

[18] These model selection criteria are described in Section 21.7.

with fixed lags are estimated by

$$\widehat{\boldsymbol{\psi}} = \frac{\widehat{\boldsymbol{\delta}}(\hat{p}, \hat{q}_1, \hat{q}_2, ..., \hat{q}_k)}{1 - \widehat{\phi}_1 - \widehat{\phi}_2 - ... - \widehat{\phi}_{\hat{p}}} \tag{21.167}$$

where $\widehat{\boldsymbol{\delta}}(\hat{p}, \hat{q}_1, \hat{q}_2, ..., \hat{q}_k)$ denotes the OLS estimate of $\boldsymbol{\delta}$ in (21.163) for the selected $ARDL$ model. The estimates of the asymptotic standard errors of $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$ and $\widehat{\boldsymbol{\psi}}$ are computed using the Bewley (1979) regression approach, which yields the same result as when applying the Δ-method described in Section 21.24, to (21.166) and (21.167).

The error correction model associated with the $ARDL(\hat{p}, \hat{q}_1, \hat{q}_2, ..., \hat{q}_k)$ model can be obtained by writing (21.163) in terms of the lagged levels and the first differences of $y_t, x_{1t}, x_{2t}, ..., x_{kt}$, and $\mathbf{w}_t$. First note that

$$\begin{aligned} y_t &= \Delta y_t + y_{t-1} \\ y_{t-s} &= y_{t-1} - \sum_{j=1}^{s-1} \Delta y_{t-j}, \qquad s = 1, 2, ..., p \end{aligned}$$

and similarly

$$\begin{aligned} \mathbf{w}_t &= \Delta \mathbf{w}_t + \mathbf{w}_{t-1} \\ x_{it} &= \Delta x_{it} + x_{i,t-1} \\ x_{i,t-s} &= x_{i,t-1} - \sum_{j=1}^{s-1} \Delta x_{i,t-j}, \qquad s = 1, 2, ..., q_i \end{aligned}$$

Substituting these relations into (21.163), and after some rearrangements, we have

$$\begin{aligned} \Delta y_t &= -\phi(1, \hat{p}) EC_{t-1} + \sum_{i=1}^{k} \beta_{i0} \Delta x_{it} + \boldsymbol{\delta}' \Delta \mathbf{w}_t \\ &\quad - \sum_{j=1}^{\hat{p}-1} \phi_j^* \Delta y_{t-j} - \sum_{i=1}^{k} \sum_{j=1}^{\hat{q}_i - 1} \beta_{ij}^* \Delta x_{i,t-j} + u_t \end{aligned} \tag{21.168}$$

where EC_t is the correction term defined by

$$EC_t = y_t - \sum_{i=1}^{k} \hat{\theta}_i x_{it} - \widehat{\boldsymbol{\psi}}' \mathbf{w}_t$$

Recall that $\phi(1, \hat{p}) = 1 - \widehat{\phi}_1 - \widehat{\phi}_2 - ... - \widehat{\phi}_{\hat{p}}$, which measures the quantitative importance of the error correction term. The remaining coefficients, ϕ_j^* and β_{ij}^*, relate to the short-run dynamics of the model's convergence to equilibrium. These are given by

$$\begin{aligned} \phi_1^* &= \phi_{\hat{p}} + \phi_{\hat{p}-1} + ... + \phi_3 + \phi_2 \\ \phi_2^* &= \phi_{\hat{p}} + \phi_{\hat{p}-1} + ... + \phi_3 \\ \vdots \quad &\quad \vdots \qquad \vdots \\ \phi_{\hat{p}-1}^* &= \phi_{\hat{p}} \end{aligned}$$

and similarly

$$\begin{array}{lll}\beta^*_{i1} & = & \beta_{i,\hat{q}_i} + \beta_{i,\hat{q}_i-1} + ... + \beta_{i,3} + \beta_{i,2} \\ \beta^*_{i2} & = & \beta_{i,\hat{q}_i} + \beta_{i,\hat{q}_i-1} + ... + \beta_{i,3} \\ \vdots & & \vdots \qquad\qquad \vdots \\ \beta^*_{i,\hat{q}_i-1} & = & \beta_{i,\hat{q}_i}\end{array}$$

The estimates $\hat{\theta}_i$ and $\hat{\boldsymbol{\psi}}$ are already computed using relations (21.166) and (21.167).

The estimates of the parameters of the error correction model (ECM) (21.168) are obtained from the coefficient estimates of the $ARDL$ model using the above relations. The standard errors of these estimates are also obtained using the variance formula (21.196), and allow for possible non-zero covariances between the estimates of the short-run and the long-run coefficients. Notice that the covariances of the short-run and the long-run coefficients are asymptotically uncorrelated only in the case where it is known that the regressors are $I(1)$, and that they are not cointegrated among themselves.

Also see Pesaran, Shin, and Smith (2001) for an analysis of $ARDL$ models when it is not known if the regressors, x_{it}, are $I(1)$ or $I(0)$. In this case the distribution of the F or Wald statistics for testing the existence of the level relations in the $ARDL$ model are non-standard and must be computed by stochastic simulations. In *Microfit* 5 these critical value bounds are computed automatically and reported after the $ARDL$ regression estimates. These are close to the ones provide in Appendix B, but have the advantage that unlike the critical values in Tables B.1 and B.2 they continue to be applicable even if shift dummy variables are included amongst the deterministic variables, $\mathbf{w}_t$.

Dynamic forecasts can also be generated using the error correction equation (21.168). The relevant formulae and other details are the same as those described in Section 21.26.2 on computation of univariate dynamic forecasts.

21.20 Probit and Logit models

Probit and Logit models represent particular formulations of the univariate binary quantitative response models defined by

$$Pr(y_i = 1) = F(\boldsymbol{\beta}'\mathbf{x}_i), \qquad i = 1, 2, ..., n \tag{21.169}$$

where y_i, $i = 1, 2, ..., n$ are independently distributed binary random variables taking the value of 1 or 0, $\mathbf{x}_i$ is a $k \times 1$ vector of explanatory variables, $\boldsymbol{\beta}$ is a $k \times 1$ vector of unknown coefficients, and $F(\cdot)$ is a known function. Under the Probit model $F(\boldsymbol{\beta}'\mathbf{x}_i)$ is specified as

$$F(\boldsymbol{\beta}'\mathbf{x}_i) = \Phi(\boldsymbol{\beta}'\mathbf{x}_i) = \int_{-\infty}^{\boldsymbol{\beta}'\mathbf{x}_i} \frac{1}{\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2}t^2\right\} dt \tag{21.170}$$

which is the cumulative distribution function of the standard normal. Under the Logit model $F\left(\boldsymbol{\beta}'\mathbf{x}_i\right)$ is specified as

$$F(\boldsymbol{\beta}'\mathbf{x}_i) = \Lambda(\boldsymbol{\beta}'\mathbf{x}_i) = \frac{e^{\boldsymbol{\beta}'\mathbf{x}_i}}{1 + e^{\boldsymbol{\beta}'\mathbf{x}_i}} \tag{21.171}$$

The maximum likelihood estimator of $\boldsymbol{\beta}$ is obtained by maximizing the following log-likelihood function

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} y_i \log\left[F(\boldsymbol{\beta}'\mathbf{x}_i)\right] + \sum_{i=1}^{n}(1-y_i)\log\left[1 - F(\boldsymbol{\beta}'\mathbf{x}_i)\right] \tag{21.172}$$

using the Newton-Raphson iterative algorithm. The first and the second derivatives of the log-likelihood function, are given by

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \frac{(y_i - F_i)\, f_i \mathbf{x}_i}{F_i\,(1-F_i)} \tag{21.173}$$

and

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'} &= -\sum_{i=1}^{n} \frac{(y_i - F_i)^2}{F_i\,(1-F_i)} f_i^2 \mathbf{x}_i \mathbf{x}_i' \\ &\quad + \sum_{i=1}^{n} \left[\frac{y_i - F_i}{F_i(1-F_i)}\right] f_i \mathbf{x}_i \mathbf{x}_i' \end{aligned} \tag{21.174}$$

where $f_i = f(\boldsymbol{\beta}'\mathbf{x}_i)$, and $F_i = F(\boldsymbol{\beta}'\mathbf{x}_i)$. These derivative functions simplify when $F(\cdot)$ takes the logistic form and are given by:

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} (y_i - \Lambda_i)\mathbf{x}_i \tag{21.175}$$

and

$$\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'} = -\sum_{i=1}^{n} \Lambda_i(1-\Lambda_i)\mathbf{x}_i \mathbf{x}_i' \tag{21.176}$$

where $\Lambda_i = \Lambda(\boldsymbol{\beta}'\mathbf{x}_i)$.

For the Probit model we have

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \tag{21.177}$$

and

$$\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'} = -\sum_{i=1}^{n} \lambda_i(\lambda_i + \boldsymbol{\beta}'\mathbf{x}_i)\mathbf{x}_i \mathbf{x}_i' \tag{21.178}$$

where

$$\lambda_i = \begin{cases} \dfrac{-\phi_i}{1-\Phi_i}, & \text{if} \quad y_i = 0 \\ \dfrac{\phi_i}{\Phi_i}, & \text{if} \quad y_i = 1 \end{cases} \tag{21.179}$$

$$\phi_i = (2\pi)^{-\frac{1}{2}} \exp\left[-\tfrac{1}{2}(\boldsymbol{\beta}'\mathbf{x}_i)^2\right] \tag{21.180}$$

and

$$\Phi_i = \Phi(\boldsymbol{\beta}'\mathbf{x}_i) \tag{21.181}$$

It is easily seen that the matrix of the second derivatives, $\partial^2\ell(\boldsymbol{\beta})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$, is negative definite under both models, and therefore the ML estimator of $\boldsymbol{\beta}$ (when it exists) is unique.[19]

The numerical computation of $\hat{\boldsymbol{\beta}}$ (the ML estimator of $\boldsymbol{\beta}$) is carried out by the Scoring Method using the following iterations:

$$\boldsymbol{\beta}_{(j)} = \boldsymbol{\beta}_{(j-1)} - \left[E\left(\frac{\partial^2\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\right)\right]^{-1}_{\boldsymbol{\beta}=\boldsymbol{\beta}_{(j-1)}} \left[\frac{\partial\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\right]_{\boldsymbol{\beta}=\boldsymbol{\beta}_{(j-1)}}, \quad j = 0, 1, 2, ... \tag{21.182}$$

where $\boldsymbol{\beta}_{(j-1)}$ is the estimator of $\boldsymbol{\beta}$ at the $(j-1)$ iteration, and

$$E\left(\frac{\partial^2\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\right) = -\sum_{i=1}^{n}\left[\frac{f_i^2}{F_i(1-F_i)}\right]\mathbf{x}_i\mathbf{x}_i' \tag{21.183}$$

Due to the global concavity of the log-likelihood function, in cases where the ML estimator of $\boldsymbol{\beta}$ exists, this iterative procedure is sure to converge, and in practice often converges in less than ten iterations.[20]

The estimator of the variance matrix of $\hat{\boldsymbol{\beta}}$ is computed as

$$\hat{V}(\hat{\boldsymbol{\beta}}) = \left[-E\left(\frac{\partial^2\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\right)\right]^{-1}_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = \sum_{i=1}^{n}\left[\frac{f_i^2}{F_i(1-F_i)}\right]\mathbf{x}_i\mathbf{x}_i' \tag{21.184}$$

where

$$\begin{aligned}
F_i &= \Lambda(\boldsymbol{\beta}'\mathbf{x}_i) = \Lambda_i, \text{and } f_i = \Lambda_i(1-\Lambda_i), \text{ in the case of the Logit model,} \\
F_i &= \Phi(\boldsymbol{\beta}'\mathbf{x}_i) = \Phi_i, \text{and } f_i = \phi_i, \text{ in the case of the Probit model.}
\end{aligned}$$

Comprehensive surveys of the literature on binary choice models and their various extensions can be found in Amemiya (1981), Maddala (1983), andCramer (1991). Also see Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) Chapter 18, and Greene (2002) Chapter 21.

21.20.1 Estimating and testing vector functions of β

To estimate linear/non-linear vector functions of $\boldsymbol{\beta}$, or test linear/non-linear restrictions on elements of β, the variance formula (21.184) can be used in the relations (21.196) and (21.198) given below. The necessary computations can be carried out using options 5 and 6 in the Logit/Probit Post Estimation Menu.

[19] See, for example, Maddala (1983) and Amemiya (1985).

[20] For an example when $\hat{\boldsymbol{\beta}}$ does not exist see the example by Albert and Anderson (1984), discussed in Amemiya (1985) p. 271-272.

21.20.2 Fitted probability and fitted discrete values

Microfit reports fitted probability values, $\Phi(\hat{\boldsymbol{\beta}}'\mathbf{x}_i)$ and $\Lambda(\hat{\boldsymbol{\beta}}'\mathbf{x}_i)$ for the Probit and Logit models, respectively. The estimates under the column 'fitted' values refer to

$$\begin{aligned} \hat{y}_i &= 1, \quad \text{if} \quad F(\hat{\boldsymbol{\beta}}'\mathbf{x}_i) \geq 0.5 \\ &= 0, \quad \text{if} \quad F(\hat{\boldsymbol{\beta}}'\mathbf{x}_i) < 0.5 \end{aligned} \tag{21.185}$$

where $F(\cdot)$ could be either the Probit or the Logit specification.

21.20.3 Measures of goodness of fit and related test statistics

The following statistics are reported after Logit or Probit estimation:

Maximized value of the log-likelihood function	$= \ell(\hat{\boldsymbol{\beta}})$
Akaike information criterion	$= \ell(\hat{\boldsymbol{\beta}}) - k$
Schwarz Bayesian criterion	$= \ell(\hat{\boldsymbol{\beta}}) - \frac{k}{2}\log(n)$
Hannan and Quinn criterion	$= \ell(\hat{\boldsymbol{\beta}}) - k \log\log n$
Mean of y	$= \sum_{i=1}^{n} y_i/n$
Mean of predicted (fitted) y	$= \sum_{i=1}^{n} \hat{y}_i/n$
Goodness of fit	$= \sum_{i=1}^{n} sign(y_i\hat{y}_i)/n$
Pseudo-R^2	$= 1 - (\ell(\hat{\boldsymbol{\beta}})/\ell(\boldsymbol{\beta}_0))$
Chi-squared statistic	$= 2\left(\ell(\hat{\boldsymbol{\beta}}) - \ell(\boldsymbol{\beta}_0)\right)$

where $\hat{y}_i$ is defined by (21.185), and the goodness of fit statistic measures the proportion of observations with correctly predicted (fitted) values of y.

The Pesaran-Timmermann test statistic is computed by applying the $PTTEST$ function to y_i (actual) and $\hat{y}_i$ (fitted) values. Under the null hypothesis that y_i and $\hat{y}_i$ are independently distributed, the $PTTEST$ statistic is asymptotically distributed as a standard normal variate. See Section 21.5 for more details.

Pseudo-R^2 is a popular measure of the model's performance in the binary choice literature and compares the fit of the model (as measured by the maximized log-likelihood value, $\ell(\hat{\boldsymbol{\beta}})$) relative to the maximized value of the log-likelihood function when all the coefficients except the intercept term (if any) in $\boldsymbol{\beta}'\mathbf{x}_i$ are set equal to zero. In the case where $\boldsymbol{\beta}'\mathbf{x}_i$ contains an intercept term

$$\ell(\boldsymbol{\beta}_0) = m\log(\frac{m}{n}) + (n-m)\log(\frac{n-m}{n})$$

where $m = \sum_{i=1}^{n} y_i/n$. See, for example, Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) pp. 766-768. When $\boldsymbol{\beta}'\mathbf{x}_i$ does not contain an intercept term we have $\ell(\boldsymbol{\beta}_0) = n\log(1/2)$.

The chi-squared statistic, $2\left(\ell(\hat{\boldsymbol{\beta}}) - \ell(\boldsymbol{\beta}_0)\right)$, is asymptotically distributed as a χ^2 variate with $k-1$ degrees of freedom when $\boldsymbol{\beta}'\mathbf{x}_i$ contains an intercept term, and it will be asymptotically distributed as a χ^2 variate with k degrees of freedom when $\boldsymbol{\beta}'\mathbf{x}_i$ does not contain an intercept term.

21.20.4 Forecasting with Probit/Logit models

The forecasts of y are obtained by first computing the probability values $\Phi\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_{n+j}\right)$, or $\Lambda\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_{n+j}\right)$, and then setting

$$\begin{aligned} \hat{y}^*_{n+j} &= 1 \qquad \text{if} \quad \Phi\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_{n+j}\right) \geq 0.5 \\ &= 0 \qquad \text{if} \quad \Phi\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_{n+j}\right) < 0.5 \end{aligned}$$

for the Probit model, and

$$\begin{aligned} \hat{y}^*_{n+j} &= 1 \qquad \text{if} \quad \Lambda\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_{n+j}\right) \geq 0.5 \\ &= 0 \qquad \text{if} \quad \Lambda\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_{n+j}\right) < 0.5 \end{aligned}$$

for the Logit model. The index for $j = 1, 2, ..., p$, where p is the forecast horizon. It is assumed that $\mathbf{x}_i$s do not include lagged values of y_i.

The following summary statistics are computed for the estimation and forecast periods:

	Estimation Period	Forecast Period
Mean of y	$\sum_{i=1}^{n} y_i/n$	$\sum_{j=1}^{p} y_{n+j}/p$
Mean of predicted y	$\sum_{i=1}^{n} \hat{y}_i/n$	$\sum_{j=1}^{p} \hat{y}^*_{n+j}/p$
Goodness of fit	$\sum_{i=1}^{n} sign(y_i\hat{y}_i)/n$	$\sum_{j=1} sign(y_{n+j}\hat{y}^*_{j+j})$
Pesaran-Timmermann statistic	$PTTEST(\mathbf{y}, \hat{\mathbf{y}})$	$PTTEST(\mathbf{y}^*, \hat{\mathbf{y}}^*)$

where $\mathbf{y} = (y_1, y_2, ..., y_n)'$, $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, ..., \hat{y}_n)'$, $\mathbf{y}^* = (y_{n+1}, y_{n+2}, ..., y_{n+p})'$, and $\hat{\mathbf{y}}^* = (\hat{y}^*_{n+1}, \hat{y}^*_{n+2}, ..., \hat{y}^*_{n+p})$.

21.21 Non-linear estimation

Consider the non-linear regression equation with additive errors:

$$y_t = f(\mathbf{x}_t, \boldsymbol{\beta}) + u_t, \qquad u_t \sim i.i.d(0, \sigma^2) \tag{21.186}$$

where $\mathbf{x}_t$ is a $k \times 1$ vector of explanatory variables, and $\boldsymbol{\beta}$ is a p-dimensional vector of unknown parameters. It is assumed that the u_t's are serially uncorrelated with mean zero and variance, σ^2. The case where $\{u_t\}$ follows an AR process can be easily dealt with by first transforming the regression equation to remove the residual serial correlation and then applying the non-linear estimation method to the resultant regression. For example, suppose the u_ts have the $AR(1)$ specification

$$u_t = \rho u_{t-1} + \epsilon_t$$

where the ϵ_t's are serially uncorrelated. Then (21.186) may be transformed to yield the following non-linear equation

$$y_t = \psi(\mathbf{x}_t, \mathbf{x}_{t-1}, \mathbf{y}_{t-1}, \boldsymbol{\theta}) + \epsilon_t$$

with serially uncorrelated residuals, where

$$\psi(\mathbf{x}_t, \mathbf{x}_{t-1}, \mathbf{y}_{t-1}, \boldsymbol{\theta}) = f(\mathbf{x}_t, \boldsymbol{\beta}) - \rho f(\mathbf{x}_{t-1}, \boldsymbol{\beta}) + \rho y_{t-1}$$

and $\boldsymbol{\theta} = (\boldsymbol{\beta}',\rho)'$. This method does not work if the error process, u_t, has an MA representation.

21.21.1 The non-linear least squares (NLS) method

The NLS estimates of $\boldsymbol{\beta}$ are computed by finding a $p \times 1$ vector $\hat{\boldsymbol{\beta}}_{NLS}$ that minimizes

$$Q_{LS}(\boldsymbol{\beta}) = \{\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})\}' \{\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})\} \tag{21.187}$$

where $\mathbf{y}$ is the $n \times 1$ vector of observations on the dependent variable y_t, and $\mathbf{X}$ is the $n \times k$ matrix of observations on $\mathbf{x}_t$. The computation of $\hat{\boldsymbol{\beta}}_{NLS}$ is achieved by means of the Gauss-Newton method. Let $\boldsymbol{\beta}_{(j)}$ be the estimate of $\boldsymbol{\beta}$ in the jth iteration, and denote the $n \times p$ matrix of the first derivatives of $\mathbf{f}(\mathbf{X}, \boldsymbol{\beta})$ evaluated at $\boldsymbol{\beta}_{(j)}$ by

$$\mathbf{F}_{(j)} = \partial \mathbf{f}(\mathbf{X}, \boldsymbol{\beta}) / \partial \boldsymbol{\beta} \Big| \boldsymbol{\beta} = \boldsymbol{\beta}_{(j)} \tag{21.188}$$

Then the iterations

$$\boldsymbol{\beta}_{(j+1)} = \boldsymbol{\beta}_{(j)} + (\mathbf{F}'_{(j)}\mathbf{F}_{(j)})^{-1}\mathbf{F}'_{(j)} \left\{\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta}_{(j)})\right\} \tag{21.189}$$

are carried out until convergence is achieved. It is worth noting that the second term on the right-hand side of (21.189) can be computed as the coefficient estimates in the regression of the residual vector $\mathbf{u}_{(j)} = \mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta}_{(j)})$ on $\mathbf{F}_{(j)}$.

The convergence criterion used is

$$\sum_{i=1}^{p} \left| \hat{\beta}_{i(j)} - \hat{\beta}_{i(j-1)} \right| < 0.00001p$$

where $\hat{\beta}_{i(j)}$ is the estimate of the ith element of $\boldsymbol{\beta}$ in the jth iteration, and p is the number of parameters. The derivatives $\mathbf{F}_{(j)}$ are computed numerically.

The estimate of σ^2 is computed as

$$\hat{\sigma}^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}}/(n-p)$$

and the asymptotic variance-covariance matrix of $\hat{\boldsymbol{\beta}}_{OLS}$ as

$$\hat{V}(\hat{\boldsymbol{\beta}}_{NLS}) = \sigma^2(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}$$

where $\hat{\mathbf{F}}$ is the value of $\partial \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})/\partial \boldsymbol{\beta}$ evaluated at the NLS estimates, $\hat{\boldsymbol{\beta}}_{NLS}$. The diagnostic statistics reported for the Non-Linear Least Squares option are computed using the formulae in Section 21.6.2 with

$$e_t = \hat{u}_t = y_t - f(\mathbf{x}_t, \hat{\boldsymbol{\beta}}_{NLS})$$

and $\mathbf{X}$ replaced by $\hat{\mathbf{F}}$.

21.21.2 The non-linear instrumental variables (NL/IV) method

The NL/IV estimates of $\boldsymbol{\beta}$ in (21.186) are computed by finding a $p \times 1$ vector $\hat{\boldsymbol{\beta}}_{IV}$ that minimizes

$$Q_{IV}(\boldsymbol{\beta}) = \{\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})\}' \, \mathbf{P}_z \, \{\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})\} \tag{21.190}$$

where $\mathbf{P}_z$ is the $n \times n$ projection matrix $\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$, and $\mathbf{Z}$ is the $n \times s$ matrix of observations on $s(\geq p)$ instruments.

The numerical procedure followed is similar to the one used for the computation of NLS estimates (see Section 21.21) and utilizes the following iterative algorithm:

$$\boldsymbol{\beta}_{(j+1)} = \boldsymbol{\beta}_{(j)} + \left(\mathbf{F}'_{(j)}\mathbf{P}_z\mathbf{F}_{(j)}\right)^{-1}\mathbf{F}'_{(j)}\mathbf{P}_z\left\{\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta}_{(j)})\right\}$$

where $F_{(j)}$ is defined by (21.188). The convergence criterion and other details are as in Section 21.21.1.

The other estimates and statistics reported by *Microfit* are computed as in Section 21.10, with $\mathbf{X}$ replaced by $\hat{\mathbf{F}} = \partial\mathbf{f}(\mathbf{X}, \boldsymbol{\beta})\partial\boldsymbol{\beta}\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{IV}}$, where $\hat{\boldsymbol{\beta}}_{IV}$ is the NL/IV estimator.

For a comprehensive discussion of non-linear least squares and non-linear instrumental variables methods see Amemiya (1974) and Gallant (1987).

21.22 Heteroscedasticity-consistent variance estimators

In situations where the homoscedasticity assumption, **A2** does not apply (see Section 6.1), the estimator of the covariance matrices of the OLS and the IV estimators given, respectively, by (21.6) and (21.59) are not generally valid, and can result in misleading inferences. Consistent estimators of the covariance matrix of the OLS and the IV estimators when the form of the heteroscedasticity is unknown have been suggested by Eicker (1963), Eicker, LeCam, and Neyman (1967), Rao (1970), White (1980), and White (1982). In the case of the OLS option, the program computes a degrees of freedom corrected version of the White (1980) estimator using the following formula:[21]

$$\widehat{HCV}(\hat{\beta}_{OLS}) = \left(\frac{n}{n-k}\right)(\mathbf{X}'\mathbf{X})^{-1}\left(\sum_{t=1}^{n} e_t^2 \mathbf{x}_t\mathbf{x}_t'\right)(\mathbf{X}'\mathbf{X})^{-1} \tag{21.191}$$

where $e_t = y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}_{OLS}$ are the OLS residuals (see Section 21.6 for the computational details of the OLS option.)

In the case of the IV option, the heteroscedasticity-consistent estimator of the covariance matrix of $\hat{\boldsymbol{\beta}}_{IV}$ is computed according to White (1982) p. 489:

$$\widehat{HCV}(\hat{\boldsymbol{\beta}}_{IV}) = \left(\frac{n}{n-k}\right)\mathbf{Q}_n^{-1}\mathbf{P}'_n\hat{\mathbf{V}}_n\mathbf{P}_n\mathbf{Q}_n^{-1} \tag{21.192}$$

[21] The correction for the degrees of freedom is suggested, amongst others, in MacKinnon and White (1985).

where

$$\begin{aligned}\mathbf{Q}_n &= \mathbf{X}'\mathbf{P}_z\mathbf{X}, \qquad \mathbf{P}_z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}' \\ \mathbf{P}_n &= (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}, \qquad \hat{\mathbf{V}}_n = \sum_{t=1}^{n} e_{t,IV}^2\mathbf{z}_t\mathbf{z}_t'\end{aligned}$$

and, as before, $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n)'$ is the $n \times s$ matrix of instrumental variables, and $e_{t,IV}$, $t = 1, 2, ..., n$ are the IV residuals. (For more details of the computations in the case of the IV option see Section 21.10.)

Notice that (21.192) can also be written as

$$\widehat{HCV}(\hat{\boldsymbol{\beta}}_{IV}) = \left(\frac{n}{n-k}\right)(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1}\left(\sum_{t=1}^{n} e_{t,IV}^2\hat{\mathbf{x}}_t\hat{\mathbf{x}}_t'\right)(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1} \tag{21.193}$$

where

$$\begin{aligned}\hat{\mathbf{X}} &= \mathbf{P}_z\mathbf{X} \\ \hat{\mathbf{x}}_t &= \mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{z}_t\end{aligned}$$

Hence, it readily follows that if $\mathbf{Z}$ is specified to include $\mathbf{X}$, then $\hat{\mathbf{X}} = \mathbf{X}$, $\hat{\mathbf{x}}_t = \mathbf{x}_t$, $e_t = e_{t,IV}$, and $\widehat{HCV}(\hat{\boldsymbol{\beta}}_{OLS}) = \widehat{HCV}(\hat{\boldsymbol{\beta}}_{IV})$.

The relevant expressions for the heteroscedasticity-consistent estimators in the case of the non-linear least squares and the non-linear IV options discussed in Section 21.12 are given by (21.191) and (21.192), respectively, with $\mathbf{X}$ replaced by the matrix of the first derivatives of the non-linear function, namely $\hat{\mathbf{F}}$, defined in Section 21.21.

21.23 Newey-West variance estimators

The Newey and West (1987) heteroscedasticity and autocorrelation consistent variance matrix is a direct generalization of White's estimators described in Section 21.21. In the general case where the non-linear regression model (21.186) is estimated by the IV method, the Newey-West variance matrix is computed according to the following formula:

$$\hat{V}(\hat{\boldsymbol{\beta}}_{IV}) = \left(\frac{n}{n-k}\right)\mathbf{Q}_n^{-1}\mathbf{P}_n'\hat{\mathbf{S}}_n\mathbf{P}_n\mathbf{Q}_n^{-1} \tag{21.194}$$

where

$$\begin{aligned}\mathbf{Q}_n &= \hat{\mathbf{F}}'\mathbf{P}_z\hat{\mathbf{F}}, \qquad \mathbf{P}_z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}' \\ \mathbf{P}_n &= (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\hat{\mathbf{F}}, \qquad \hat{\mathbf{F}} = \frac{\partial\mathbf{f}(\mathbf{X},\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{IV}}\end{aligned}$$

and

$$\hat{\mathbf{S}}_n = \hat{\boldsymbol{\Omega}}_0 + \sum_{j=1}^{m} w(j,m)(\hat{\boldsymbol{\Omega}}_j + \hat{\boldsymbol{\Omega}}_j')$$

in which

$$\begin{aligned}\hat{\mathbf{\Omega}}_j &= \sum_{t=j+1}^{n} \hat{\mathbf{u}}_t \hat{\mathbf{u}}_{t-j} \mathbf{z}_t \mathbf{z}'_{t-j} \\ \hat{u}_t &= y_t - f(\mathbf{x}_t, \hat{\boldsymbol{\beta}}_{IV})\end{aligned}$$

$\mathbf{z}_t$ is the $s \times 1$ vector of instruments, and $w(j,m)$ is the lag window used. *Microfit* allows three choices for the lag window:

Uniform (or rectangular) window

$$w(j,m) = 1, \qquad \text{for} \quad j = 1, 2, ..., m$$

Bartlett window

$$w(j,m) = 1 - \frac{j}{m+1}, \qquad j = 1, 2, ..., m$$

Parzen window

$$\begin{aligned} w(j,m) &= 1 - 6\left(\frac{j}{m+1}\right)^2 + 6\left(\frac{j}{m+1}\right)^3, \qquad 1 \leq j \leq \frac{m+1}{2} \\ &= 2\left(1 - \frac{j}{m+1}\right)^2 \qquad \frac{m+1}{2} < j \leq m \end{aligned}$$

The 'window size' or the 'truncation point', m, is specified by the user.

Newey and West (1987) use the Bartlett window and do not introduce the small sample correction proposed by MacKinnon and White (1985). Users interested in exactly replicating the Newey-West adjusted standard errors should therefore choose the Bartlett window and multiply the standard errors computed by *Microfit* by $\{(n-k)/n\}^{1/2}$. Also note that White's heteroscedasticity-consistent estimators outlined in Section 21.22 can be computed using the Newey-West option by setting the window size, m, equal to zero.

The equal weight (or the uniform) window option is appropriate when estimating a regression model with moving average errors of known order. This type of model arises in testing the market efficiency hypothesis where the forecast horizon exceeds the sampling interval (see, for example, Pesaran (1987b), section 7.6). In other situations a Parzen window is generally preferable to the other two windows. Notice that the positive semi-definiteness of the Newey-West variance matrix is only ensured in the case of the Bartlett and Parzen windows. The choice of the uniform window can result in a negative-definite variance matrix, especially if a large value for m is chosen relative to the number of available observations, n. Also see the discussion in Andrews (1991) and Andrews and Monahan (1992).

21.24 Variance of vector function of estimators

Let $\boldsymbol{\phi} = \boldsymbol{\phi}(\boldsymbol{\theta})$ be an $r \times 1$ first-order differentiable function of the $k \times 1$ parameter vector, $\boldsymbol{\theta}$, of a given econometric model. Suppose also that $\boldsymbol{\Phi}(\boldsymbol{\theta}) = \partial\boldsymbol{\phi}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}'$ is an $r \times k$ matrix of rank $r(\leq k)$. Then the estimator of $\boldsymbol{\phi}$ and the estimator of its asymptotic variance are given by

$$\hat{\boldsymbol{\phi}} = \boldsymbol{\phi}(\hat{\boldsymbol{\theta}}) \tag{21.195}$$

and

$$\hat{V}(\boldsymbol{\phi}) = \hat{\sigma}^2 \left[\frac{\partial\boldsymbol{\phi}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \hat{\mathbf{V}}(\hat{\boldsymbol{\theta}}) \left[\frac{\partial\boldsymbol{\phi}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right]'_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{21.196}$$

where $\hat{\boldsymbol{\theta}}$ represents the estimator of $\boldsymbol{\theta}$, and $\hat{\sigma}^2\hat{V}(\hat{\boldsymbol{\theta}})$ is the estimator of the variance matrix of $\hat{\boldsymbol{\theta}}$. The above procedure for estimation of the variance of $\hat{\boldsymbol{\phi}}$ is also known as the Δ-method. See, for example, Serfling (1980).

21.25 Wald statistic for testing linear and non-linear restrictions

Option 7 in the Hypothesis Testing Menu (see Section 6.23) allows the user to compute Wald statistics for testing r independent linear or non-linear restrictions on the parameters of the regression model $\boldsymbol{\theta}$**.** Let the r restrictions on $\boldsymbol{\theta}$ be given by

$$\boldsymbol{\phi}(\boldsymbol{\theta}) = \mathbf{0} \tag{21.197}$$

where $\boldsymbol{\phi}(\cdot)$ is an $r \times 1$ first-order differentiable function of the unknown parameters of the regression model, and denotes the estimator of the (asymptotic) variance matrix of $\hat{\boldsymbol{\theta}}$ by $\hat{\sigma}^2\hat{\mathbf{V}}(\hat{\boldsymbol{\theta}})$, where $\hat{\boldsymbol{\theta}}$ stands for the estimator of $\boldsymbol{\theta}$. Then the Wald statistic for testing the r restrictions in (21.197) is given by

$$W = \hat{\boldsymbol{\phi}}' \left[\hat{\Phi}\hat{\mathbf{V}}(\theta)\hat{\Phi}'\right]^{-1} \hat{\boldsymbol{\phi}}/\hat{\sigma}^2 \overset{a}{\sim} \chi^2(r) \tag{21.198}$$

where

$$\hat{\boldsymbol{\phi}} = \boldsymbol{\phi}(\hat{\boldsymbol{\theta}}), \qquad \hat{\boldsymbol{\Phi}} = (\partial\boldsymbol{\phi}(\boldsymbol{\theta})/\partial\boldsymbol{\theta})'_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{21.199}$$

Before calculating the W statistic, the program first checks the rank condition on $\hat{\boldsymbol{\Phi}}$, and proceeds with the computations only if $\hat{\boldsymbol{\Phi}}$ is of full rank, namely when $Rank[\hat{\boldsymbol{\Phi}}] = r$.

21.26 Univariate forecasts in regression models

This section considers the problem of forecasting with single-equation linear and non-linear regression models. Forecasting with Probit and Logit models is discussed in Section 21.20.

The following general dynamic regression model underlies the forecasts computed by the program for the linear regression model:

$$y_t = \sum_{i=1}^{\ell} \lambda_i y_{t-i} + \mathbf{g}_t'\boldsymbol{\alpha} + u_t \tag{21.200}$$

$$= \mathbf{x}_t'\boldsymbol{\beta} + u_t, \qquad \mathbf{x}_t = (y_{t-1}, y_{t-2}, ..., y_{t-\ell}, \mathbf{g}_t')' \tag{21.201}$$

where for the AR options

$$u_t = \sum_{i=1}^{m} \rho_i u_{t-i} + \epsilon_t \tag{21.202}$$

and for the MA options

$$u_t = \epsilon_t + \sum_{i=1}^{q} \gamma_i \epsilon_{t-i} \tag{21.203}$$

and ϵ_t are serially uncorrelated random disturbances with zero means. The program computes univariate dynamic forecasts if the regression equation is specified *explicitly* to include lagged values of the dependent variable ($\ell \geq 1$). Otherwise, it will generate static forecasts.

21.26.1 Univariate static forecasts

The static forecasts are computed taking as given the values of the regressors $\mathbf{x}_{t+j}$. The values of y_{t+j}, $j = 1, 2, ..., p$, are forecast by

$$\hat{y}_{t+j}^{*} = \mathbf{x}_{t+j}'\widehat{\boldsymbol{\beta}} + \hat{u}_{t+j}^{*} \tag{21.204}$$

For the OLS and the IV options, $\hat{u}_{t+j}^{*} = 0$ and the estimators of $\hat{y}_{t+j}^{*}$ are given by

$$\hat{y}_{t+j,OLS}^{*} = \mathbf{x}_{t+j}'\widehat{\boldsymbol{\beta}}_{OLS}$$

and

$$\hat{y}_{t+j,IV}^{*} = \mathbf{x}_{t+j}'\widehat{\boldsymbol{\beta}}_{IV}$$

respectively. The forecasts for the non-linear options are given by

$$\hat{y}_{t+j}^{*} = f\left(\mathbf{x}_{t+j}, \hat{\boldsymbol{\beta}}\right)$$

For the AR options, $\hat{u}_{t+j}^{*}$, $j = 1, 2, ..., p$ in (21.204) are computed recursively according to the following relations:

$$\begin{aligned}
\hat{u}_{t+1}^{*} &= \sum_{i=1}^{m} \rho_i \hat{u}_{t+1-i} \\
\hat{u}_{t+2}^{*} &= \hat{\rho}_1 \hat{u}_{t+1}^{*} + \sum_{i=2}^{m} \hat{\rho}_i \hat{u}_{t+2-i} \\
\hat{u}_{t+3}^{*} &= \hat{\rho}_1 \hat{u}_{t+2}^{*} + \hat{\rho}_2 \hat{u}_{t+1}^{*} \sum_{i=3}^{m} \hat{\rho}_i \hat{u}_{t+3-i} \\
&\vdots \\
\hat{u}_{t+m}^{*} &= \sum_{i=1}^{m-1} \hat{\rho}_i \hat{u}_{t+m-i}^{*} + \hat{\rho}_m \hat{u}_t
\end{aligned}$$

and

$$\hat{u}^*_{t+j} = \sum_{i=1}^{m} \hat{\rho}_i \hat{u}^*_{t+j-i}, \qquad \text{for} \quad j = m+1, m+2, ..., p$$

For the MA options, $\hat{u}^*_{t+j}$ in (21.204) are computed as

$$\begin{array}{lll} \hat{u}^*_{t+j} & = \hat{\gamma}_j \hat{\epsilon}_t + \hat{\gamma}_{j+1} \hat{\epsilon}_{t-1} + ... + \hat{\gamma}_q \hat{\epsilon}_{t+j-q} & \text{for } j \leq q \\ & = 0 & \text{for } j > q \end{array}$$

and the $\hat{\epsilon}_t$s are obtained recursively:

$$\hat{\epsilon}_t = -\sum_{i=1}^{q} \hat{\gamma}_i \hat{\epsilon}_{t-i} + \hat{u}_t, \qquad t = 1, 2, ...$$

with the initial values $\hat{\epsilon}_0 = \hat{\epsilon}_{-1} = \hat{\epsilon}_{-2} = ... = \hat{\epsilon}_{q+1} = 0$.

21.26.2 Univariate dynamic forecasts

In computing dynamic forecasts the program takes the values of $\mathbf{g}_{t+j}$, $j = 1, 2, .., p$ in (21.200) as given, and computes y_{t+j}, $j = 1, 2, ..., p$ as j-step ahead forecasts using the following recursive relations:

$$\begin{array}{rcl} y^*_{t+1} & = & \sum_{i=1}^{l} \lambda_i y_{t+1-i} + \mathbf{g}'_{t+1} \boldsymbol{\alpha} + u^*_{t+1} \\ y^*_{t+2} & = & \lambda_1 y^*_{t+1} + \sum_{i=2}^{l} \lambda_i y_{t+1-i} + \mathbf{g}'_{t+2} \boldsymbol{\alpha} + u^*_{t+2} \\ & \vdots & \vdots \\ y^*_{t+l} & = & \sum_{i=1}^{l-1} \lambda_i y^*_{t+\ell-i} + \lambda_l y_t + \mathbf{g}'_{t+l} \boldsymbol{\alpha} + u^*_{t+l} \end{array}$$

and

$$y^*_{t+j} = \sum_{i=1}^{l} \lambda_i y^*_{t+j-i} + \mathbf{g}'_{t+j} \boldsymbol{\alpha} + u^*_{t+j}, \qquad \text{for} \quad j = l+1, l+2, ..., p$$

where, as before, estimates of u^*_{t+j} are obtained by means of the recursive relations given in Section 21.26.1 above. The program computes estimates of y^*_{t+j} by replacing the unknown parameters $\lambda_1, ..., \lambda_q, \boldsymbol{\alpha}, \rho_1, ..., \rho_m, \gamma_1, \gamma_2, ..., \gamma_q$ by their appropriate ML estimators defined in Sections 21.6 to 21.16.

21.26.3 Standard errors of univariate forecast errors: the OLS and IV options

The program computes standard errors of the forecast errors

$$e^*_{t+j} = y_{t+j} - \hat{y}^*_{t+j}, \qquad j = 1, 2, ..., p$$

only for the OLS and the IV options.

Let $\mathbf{e}^* = (e^*_{n+1}, e^*_{n+2}, ..., e^*_{n+p})'$ be the $p \times 1$ vector of forecast errors. For the static forecasts given in Section 21.26.1 we have

$$\text{OLS option:} \qquad \hat{\mathbf{V}}(\mathbf{e}^*) = \hat{\sigma}^2_{OLS} \left\{ \mathbf{I}_p + \mathbf{X}^*(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}^{*'} \right\} \tag{21.205}$$

$$\text{IV option:} \qquad \hat{\mathbf{V}}(\mathbf{e}^*) = \hat{\sigma}^2_{IV} \left\{ \mathbf{I}_p + \mathbf{X}^*(\mathbf{X}'\mathbf{P}_z\mathbf{X})^{-1}\mathbf{X}^{*'} \right\} \tag{21.206}$$

where $\mathbf{I}_p$ is an identity matrix of order p, $\mathbf{X}^*$ is the $p \times k$ matrix of observations on $\mathbf{x}_t$ over the forecast period, $\mathbf{X}$ is the $n \times k$ matrix of observations on $\mathbf{x}_t$ over the estimation period, and $\mathbf{P}_z$ is the projection matrix defined by (21.58).

The variance matrix of the dynamic forecast errors are computed according to the following formula due to Pagan and Nicholls (1984):

$$OLS \text{ option:} \qquad \hat{V}(\mathbf{e}^*) = \hat{\sigma}^2_{OLS}\mathbf{D}^{-1} \left\{ \mathbf{I}_p + \mathbf{X}^*(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}^{*'} \right\} \mathbf{D}'^{-1} \tag{21.207}$$

$$\text{IV option:} \qquad \hat{\mathbf{V}}(\mathbf{e}^*) = \hat{\sigma}^2_{IV}\mathbf{D}^{-1} \left\{ \mathbf{I}_p + \mathbf{X}^*(\mathbf{X}'\mathbf{P}_z\mathbf{X})^{-1}\mathbf{X}^{*'} \right\} \mathbf{D}'^{-1} \tag{21.208}$$

where

$$\mathbf{D} = \begin{bmatrix}
-1 & 0 & . & . & . & 0 & . & . & . & 0 & 0 \\
\hat{\lambda}_1 & -1 & . & . & . & 0 & . & . & . & 0 & 0 \\
\hat{\lambda}_2 & \hat{\lambda}_1 & & -1 & & 0 & . & . & . & 0 & 0 \\
. & . & & . & & . & & & & . & . \\
. & . & & . & & . & & & & . & . \\
. & . & & . & & . & & & & . & . \\
\hat{\lambda}_l & \hat{\lambda}_{l-1} & . & . & . & -1 & & 0 & & 0 & 0 \\
. & . & & & & . & & & & . & . \\
. & . & & & & . & & & & . & . \\
. & . & & & & . & & & & . & 0 \\
0 & \ldots & & 0 & & \hat{\lambda}_l & . & . & \hat{\lambda}_2 & \hat{\lambda}_1 & -1
\end{bmatrix}$$

and $\hat{\lambda}_1, \hat{\lambda}_2, ..., \hat{\lambda}_l$ are the estimates of λ_i in (21.200). The $p \times k$ matrix of observations on $\mathbf{x}_t$ in the case of dynamic forecasts is given by

$$\mathbf{X}^* = \begin{bmatrix}
y_n & y_{n-1} & \cdots & y_{n-l} & \mathbf{g}_{n+1} \\
y_{n+1} & y_n & \cdots & y_{n+1-l} & \mathbf{g}_{n+2} \\
\vdots & \vdots & & \vdots & \vdots \\
\hat{y}_{n+p-1} & \hat{y}_{n+p-2} & \cdots & \hat{y}_{n+p-l} & \mathbf{g}_{n+p}
\end{bmatrix}.$$

21.26.4 Forecasts based on non-linear models

Dynamic forecasts are computed for the non-linear least squares and the non-linear IV options, when the non-linear equation is specified to contain lagged values of y_t. For example, in the case of the non-linear equation:

$$y_t = f(y_{t-1}, \mathbf{g}_t, \boldsymbol{\beta}) + u_t$$

Dynamic forecasts are computed recursively:[22]

$$\hat{y}^*_{t+1} = f(y_t, \mathbf{g}_{t+1}, \hat{\boldsymbol{\beta}})$$

and

$$\hat{y}^*_{t+j} = f(\hat{y}^*_{t+j-1}, \mathbf{g}_{t+j}, \hat{\boldsymbol{\beta}}), \qquad \text{for} \qquad j = 2, 3, ..., p$$

21.26.5 Measures of forecast accuracy

The program also computes the following summary statistics for the forecast values ($\hat{y}^*_{t+j}$, $j = 1, 2, .., p$):

$$\text{Mean prediction errors} = \left(\sum_{j=1}^{p} e^*_{t+j}\right)/p \tag{21.209}$$

where

$$e^*_{t+j} = y_{t+j} - \hat{y}^*_{t+j}, \qquad j = 1, 2, ..., p \tag{21.210}$$

$$\begin{array}{l}\text{Sum of squares of}\\ \text{prediction errors}\end{array} = \sum_{j=1}^{p}(e^*_{t+j})^2 \tag{21.211}$$

$$\begin{array}{l}\text{Root mean sum of}\\ \text{squares of prediction}\\ \text{errors}\end{array} = \sqrt{\left(\sum_{j=1}^{p}(e^*_{t+j})^2/p\right)} \tag{21.212}$$

$$\begin{array}{l}\text{Mean sum of absolute}\\ \text{prediction errors}\end{array} = \sum_{j=1}^{p}\left|e^*_{t+j}\right|/p \tag{21.213}$$

In the table giving the above summary statistics for the *OLS* option, the program also reports the F-statistics for the predictive failure, and the structural stability tests, defined by relations (21.27) and (21.29) respectively. The latter test statistic is reported only if $p > k$.

[22] Notice that the dynamic forecasts in the non-linear case are not necessarily equal to the conditional expectations of y_{t+j}, for $j > 1$, and can therefore be viewed as a 'certainty equivalent' approximation to $E(y_{t+j}\,|\Omega_t)$, for $j > 1$.

Chapter 22

Econometrics of Multiple Equation Models

This chapter complements Chapter 7 and provides the technical details of the econometric methods and algorithms used in *Microfit* for the analysis of multiple time-series models. It covers a number of recent developments in the areas of impulse response analysis and long-run structural modelling. The chapter starts in Section 22.1 with a review of the seemingly unrelated regression equations ($SURE$) originally analyzed by Zellner (1962). Section 22.3 deals with estimation of $SURE$ models subject to general linear restrictions, possibly involving cross-equation restrictions. Section 22.4 reviews the estimation and hypothesis testing in augmented vector autoregressive models. Impulse response analysis, and forecast error variance decomposition of unrestricted VAR models, are set out in Sections 22.5 and 22.6. The remaining sections deal with the long-run structural modelling approach: reviewing the literature on testing for cointegration, identification and maximum likelihood estimation of long-run (or cointegrating) relations, and impulse response and persistence profile analysis in cointegrating VAR models.

Recent detailed treatments of multivariate time-series analysis can be found in Lütkepohl (2005) and Hamilton (1994). More general text-book accounts of $SURE$ estimation and VAR modelling are available in Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) Chapter 16, and Greene (2002) Chapters 17-18. The more recent developments in the areas of impulse response analysis (including generalized impulse response functions and persistence profiles) are covered in Pesaran, Pierce, and Lee (1993), Lee and Pesaran (1993), Koop, Pesaran, and Potter (1996), Pesaran and Shin (1996), and Pesaran and Shin (1998). An excellent survey of the early developments in the literature on cointegration can be found in Banerjee, Dolado, Galbraith, and Hendry (1993), and Watson (1994). For more recent developments and further references to the literature on long-run structural modelling see Pesaran and Shin (1999), Pesaran and Shin (2002), Pesaran, Shin, and Smith (1996), Pesaran, Shin, and Smith (2000), and Pesaran (1997).

22.1 Seemingly unrelated regression equations ($SURE$)

Consider the following m 'seemingly' separate linear regression equations:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{u}_i, \quad i = 1, 2, ..., m \tag{22.1}$$

where $\mathbf{y}_i$ is an $n \times 1$ vector of observations on the dependent variable y_{it}, $i = 1, 2, ..., m$; $t = 1, 2, ..., n$, and $\mathbf{X}_i$ is an $n \times k_i$ matrix of observations on the k_i vector of regressors explaining y_{it}, $\boldsymbol{\beta}_i$ is a $k_i \times 1$ vector of unknown coefficients, and $\mathbf{u}_i$ is an $n \times 1$ vector of disturbances or errors, for $i = 1, 2, ..., m$. It is further assumed that for each i the regressors $\mathbf{X}_i$ and the disturbance $\mathbf{u}_i$ satisfy the classical assumptions **A1** to **A5**, set out in Section 6.1.

In econometric analysis of the system of equations in (22.1), three cases can be distinguished:

1. Contemporaneously uncorrelated disturbances, namely $E\left(\mathbf{u}_i\mathbf{u}_j'\right) = \mathbf{0}$, for $i \neq j$.

2. Contemporaneously correlated disturbances, with identical regressors across all the equations, namely
$$E\left(\mathbf{u}_i\mathbf{u}_j'\right) = \sigma_{ij}\mathbf{I}_n \neq \mathbf{0}$$
where $\mathbf{I}_n$ is an identity matrix of order n, and
$$\mathbf{X}_i = \mathbf{X}_j, \text{ for all } i, j$$

3. Contemporaneously correlated disturbances, with different regressors across the equations, namely
$$E\left(\mathbf{u}_i\mathbf{u}_j'\right) = \sigma_{ij}\mathbf{I}_n \neq \mathbf{0}$$
and
$$\mathbf{X}_i \neq \mathbf{X}_j, \text{ at least for some } i, \text{ and } j$$

In the first case where $E\left(\mathbf{u}_i\mathbf{u}_j'\right) = \mathbf{0}$, for $i \neq j$, there is nothing to be gained by considering the equations in (22.1) as a system, and the application of single equation methods to the individual relations in (22.1) will be valid. There is also no efficiency gain in estimating the equations in (22.1) as a system under case 2 where $\mathbf{X}_i = \mathbf{X}_j$, for all i and j. Once again the application of single-equation methods to each of the equations in the system will be valid. See Zellner (1962), and Section 22.1.1.

It is therefore only under case 3 where there is likely to be some efficiency gains in large samples by estimating the equations in (22.1) as a system in large samples.

22.1.1 Maximum likelihood estimation

In order to compute the maximum likelihood (ML) estimators of the parameters of (22.1), namely
$$\boldsymbol{\theta} = \left(\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', ..., \boldsymbol{\beta}_m', \ \sigma_{11}, \sigma_{12}, ..., \sigma_{1m}; \ \sigma_{22}, \sigma_{23}, ..., \sigma_{2m}; ..., \sigma_{mm}\right)'$$

it is convenient to stack the different equations in the system in the following manner:

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & & & 0 \\ & \mathbf{X}_2 & & \\ & & \ddots & \\ 0 & & & \mathbf{X}_m \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_m \end{pmatrix} + \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_m \end{pmatrix} \tag{22.2}$$

or more compactly

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{22.3}$$

where $\mathbf{y}$, $\mathbf{X}$, $\boldsymbol{\beta}$ and $\mathbf{u}$ have the dimensions $mn \times 1$, $mn \times k$, $k \times 1$ and $mn \times 1$, respectively, where $k = \sum_{i=1}^{m} k_i$. Under the classical assumptions where $E(\mathbf{u}_i) = 0$, $E\left(\mathbf{u}_i\mathbf{u}_j'\right) = \sigma_{ij}\mathbf{I}_n$, we have

$$E\left(\mathbf{u}\mathbf{u}'\right) = \boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \mathbf{I}_n$$

where $\boldsymbol{\Sigma}$ is the $m \times m$ matrix of covariances with its (i, j) elements equal to σ_{ij}, and $\otimes$ stands for Kronecker products.[1] More specifically, we have:

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \mathbf{I}_n = \begin{pmatrix} \sigma_{11}\mathbf{I}_n & \sigma_{12}\mathbf{I}_n & \dots & \sigma_{1m}\mathbf{I}_n \\ \sigma_{21}\mathbf{I}_n & \sigma_{22}\mathbf{I}_n & \dots & \sigma_{2m}\mathbf{I}_n \\ \vdots & & & \\ \sigma_{m1}\mathbf{I}_n & \sigma_{m2}\mathbf{I}_n & \dots & \sigma_{mm}\mathbf{I}_n \end{pmatrix} \tag{22.4}$$

If we now assume that $\mathbf{u}$ has a Gaussian distribution, the log-likelihood function of the stacked system (22.3) can be written as

$$\ell(\boldsymbol{\theta}) = -\frac{nm}{2}\log(2\pi) - \tfrac{1}{2}\log|\Omega| - \tfrac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\,\boldsymbol{\Omega}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

since

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \mathbf{I}_n,\ |\boldsymbol{\Omega}| = |\boldsymbol{\Sigma} \otimes \mathbf{I}_n| = |\boldsymbol{\Sigma}|^n\,|\mathbf{I}_n|^m = |\boldsymbol{\Sigma}|^n$$

and hence

$$\ell(\boldsymbol{\theta}) = -\frac{nm}{2}\log(2\pi) - \frac{n}{2}\log|\boldsymbol{\Sigma}| - \tfrac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\left(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_n\right)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \tag{22.5}$$

Denoting the ML estimators by $\widetilde{\boldsymbol{\theta}} = \left(\widetilde{\boldsymbol{\beta}}_1', \widetilde{\boldsymbol{\beta}}_2', ..., \widetilde{\boldsymbol{\beta}}_m', \ \tilde{\sigma}_{11}, \tilde{\sigma}_{12}, ...\right)'$, it is easily seen that

$$\tilde{\sigma}_{ij} = \frac{\left(\mathbf{y}_i - \mathbf{X}_i\widetilde{\boldsymbol{\beta}}_i\right)'\left(\mathbf{y}_j - \mathbf{X}_j\widetilde{\boldsymbol{\beta}}_j\right)}{n} \tag{22.6}$$

and

$$\widetilde{\boldsymbol{\beta}} = \left(\mathbf{X}'\left(\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_n\right)\mathbf{X}\right)^{-1}\mathbf{X}'\left(\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_n\right)\mathbf{y} \tag{22.7}$$

[1] For the definition of Kronecker products, the $vec(\cdot)$ operators and the rules of their operations see, for example, Magnus and Neudecker (1988).

The computation of the ML estimators $\widetilde{\boldsymbol{\beta}} = \left(\widetilde{\boldsymbol{\beta}}_1', \widetilde{\boldsymbol{\beta}}_2', ..., \widetilde{\boldsymbol{\beta}}_m'\right)'$, and $\tilde{\sigma}_{ij}$, $i, j = 1, 2, ..., m$ are carried out in *Microfit* by iterating between (22.6) and (22.7) starting from the OLS estimators of $\boldsymbol{\beta}_i$, namely $\widehat{\boldsymbol{\beta}}_{i,OLS} = (\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{y}_i$. This iterative procedure is continued until the following convergence criteria are met:

$$\sum_{\ell=1}^{k_i} \left|\boldsymbol{\beta}_{i\ell}^{(r)} - \boldsymbol{\beta}_{i\ell}^{(r-1)}\right| < (0.00001)k_i, \quad i = 1, 2, ..., m \tag{22.8}$$

where $\boldsymbol{\beta}_{i\ell}^{(r)}$ denotes the estimate of the ℓth element of $\boldsymbol{\beta}_i$ at the rth iteration. On convergence, *Microfit* reports $\widehat{\boldsymbol{\beta}}_i, i = 1, 2, ..., m$, and the estimates of their covariances computed as

$$\widehat{\mathrm{Cov}}(\widetilde{\boldsymbol{\beta}}) = \left(\mathbf{X}'\left(\widehat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_n\right)\mathbf{X}\right)^{-1} \tag{22.9}$$

where $\widehat{\boldsymbol{\Sigma}} = (\hat{\sigma}_{ij})$ is a degrees-of-freedom adjusted version of $\widetilde{\boldsymbol{\Sigma}}$ whose (i, j) element is given by

$$\hat{\sigma}_{ij} = \frac{\tilde{\mathbf{u}}_i'\tilde{\mathbf{u}}_j}{\sqrt{(n - k_i)(n - k_j)}}, \quad i, j = 1, 2, ..., m \tag{22.10}$$

For further details see, for example, Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) Chapter 12.

The fitted values, residuals and other statistics such as DW, R^2, $\bar{R}^2$ and maximized log-likelihood values for each equation in the system is computed as in Section (21.6.1). The maximized value of the system log-likelihood function is given by

$$\ell(\widetilde{\boldsymbol{\theta}}) = -\frac{nm}{2}\log(2\pi) - \tfrac{n}{2}\log\left|\widetilde{\boldsymbol{\Sigma}}\right| \tag{22.11}$$

The system's Akaike and Schwarz criteria are computed as

$$\text{System } AIC \quad = \ell(\widetilde{\boldsymbol{\theta}}) - k \tag{22.12}$$

$$\text{System } SBC \quad = \ell(\widetilde{\boldsymbol{\theta}}) - \tfrac{k}{2}\log(n) \tag{22.13}$$

where $k = \sum_{i=1}^{m} k_i$.

22.2 Three-stage least squares

This approach is useful when the dependent variables in some or all equations of system (22.1) appear also as regressors (see Zellner and Theil (1962)). In this case we have the following system of simultaneous equations:

$$\begin{aligned} \mathbf{y}_i &= \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{Y}_i\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i \\ &= \mathbf{W}_i\boldsymbol{\delta}_i + \boldsymbol{\varepsilon}_i \quad i = 1, 2, ..., m \end{aligned} \tag{22.14}$$

where $\mathbf{y}_i$ is a $T \times 1$ vector of observations on the 'normalized' endogenous variable of the i^{th} equation in the system, $\mathbf{X}_i$ are the $T \times k_i$ vector of observations on the exogenous variables in the i^{th} equation, and $\mathbf{Y}_i$ is the $T \times p_i$ vector of endogenous variables in the i^{th} equation whose coefficients are not normalized. Hence equation i has p_i endogenous and k_i exogenous variables. $\mathbf{W}_i$ is $T \times s_i$ where $s_i = k_i + p_i$. Also $\boldsymbol{\delta}_i = \left(\boldsymbol{\beta}_i', \boldsymbol{\gamma}_i'\right)'$.

Choice of the variable to normalize (or equivalently the choice of the left-hand-side variable) can be important in practice and is assumed to be guided by economic theory or other forms of *a priori* information. The order condition for identification of parameters of equation i is given by $k - k_i \geq p_i$, namely the number of excluded exogenous variables in equation i must be at least as large as the number of included endogenous variables minus one (the normalization constant applied to $\mathbf{y}_i$).

The three-stage least squares ($3SLS$) computes the fitted values:

$$\widehat{\mathbf{W}}_i = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}_i \tag{22.15}$$

and estimates the following system of equations by the $SURE$ procedure (see Section 22.1.1):

$$\mathbf{y}_i = \widehat{\mathbf{W}}_i\boldsymbol{\delta}_i + \boldsymbol{\xi}_i \tag{22.16}$$

To obtain explicit expression for the $3SLS$ estimators stack the m equations as

$$\mathbf{y} = \widehat{\mathbf{W}}\boldsymbol{\delta} + \boldsymbol{\xi} \tag{22.17}$$

where

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix}, \ \widehat{\mathbf{W}} = \begin{pmatrix} \widehat{\mathbf{W}}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{W}}_2 & \cdots & \mathbf{0} \\ \mathbf{0} & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \widehat{\mathbf{W}}_m \end{pmatrix}$$

$$\boldsymbol{\delta} = \begin{pmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \vdots \\ \boldsymbol{\delta}_m \end{pmatrix}, \ \boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \vdots \\ \boldsymbol{\xi}_m \end{pmatrix}$$

Then

$$\hat{\boldsymbol{\delta}}_{3SLS} = \left[\widehat{\mathbf{W}}'\left(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_T\right)\widehat{\mathbf{W}}\right]^{-1}\widehat{\mathbf{W}}'\left(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_T\right)\mathbf{y} \tag{22.18}$$

with

$$\hat{\boldsymbol{\Sigma}} = (\hat{\sigma}_{ij}), \ \hat{\sigma}_{ij} = \frac{\left(\mathbf{y}_i - \mathbf{W}_i\hat{\boldsymbol{\delta}}_{i,2SLS}\right)'\left(\mathbf{y}_j - \mathbf{W}_j\hat{\boldsymbol{\delta}}_{j,2SLS}\right)}{T} \tag{22.19}$$

$$\hat{\boldsymbol{\delta}}_{i,2SLS} = \left(\widehat{\mathbf{W}}_i'\widehat{\mathbf{W}}_i\right)^{-1}\widehat{\mathbf{W}}_i'\mathbf{y}_i \tag{22.20}$$

The covariance matrix of the $3SLS$ estimator is given by

$$Var\left(\hat{\boldsymbol{\delta}}_{3SLS}\right) = \left[\widehat{\mathbf{W}}'\left(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_T\right)\widehat{\mathbf{W}}\right]^{-1} \tag{22.21}$$

The estimates $\hat{\Sigma}$ and $\hat{\boldsymbol{\delta}}_{3SLS}$ are updated iteratively until convergence is achieved, as in the $SURE$ estimation (see Section 22.1.1).

22.2.1 Testing linear/non-linear restrictions

Under fairly general conditions the ML estimators, $\widetilde{\boldsymbol{\beta}} = \left(\widetilde{\boldsymbol{\beta}}_1', \widetilde{\boldsymbol{\beta}}_2', ..., \widetilde{\boldsymbol{\beta}}_m'\right)'$, are asymptotically normally distributed with mean $\widetilde{\boldsymbol{\beta}}$ and the covariance matrix given by (22.9). It is therefore possible to test linear or non-linear restrictions on the elements of $\boldsymbol{\beta}$ using the Wald procedure. (see Section 21.25). Notice that the restrictions to be tested could involve coefficients from different equations (there could be cross-equation restrictions). To be more precise, suppose you are interested in testing the following $r \times 1$ general non-linear restrictions on $\boldsymbol{\beta}$:

$$\begin{aligned} H_0 &: \mathbf{h}(\boldsymbol{\beta}) = \mathbf{0} \\ H_1 &: \mathbf{h}(\boldsymbol{\beta}) \neq \mathbf{0} \end{aligned}$$

where $\mathbf{h}(\boldsymbol{\beta})$ is known $r \times 1$ vector function of $\boldsymbol{\beta}$, with continuous partial derivatives. The Wald statistic for testing $\mathbf{h}(\boldsymbol{\beta}) = \mathbf{0}$ against $\mathbf{h}(\boldsymbol{\beta}) \neq \mathbf{0}$ is given by

$$W = \mathbf{h}(\widetilde{\boldsymbol{\beta}})' \left[\mathbf{H}(\widetilde{\boldsymbol{\beta}})\widehat{Cov}(\widetilde{\boldsymbol{\beta}})\mathbf{H}'(\widetilde{\boldsymbol{\beta}})\right]^{-1} \mathbf{h}(\widetilde{\boldsymbol{\beta}}) \tag{22.22}$$

where $\mathbf{H}(\widetilde{\boldsymbol{\beta}})$ is given by $\partial\mathbf{h}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}'$ at $\boldsymbol{\beta} = \widetilde{\boldsymbol{\beta}}$. It will be assumed that $Rank(\mathbf{H}(\boldsymbol{\beta})) = r$.

22.2.2 LR statistic for testing whether $\boldsymbol{\Sigma}$ is diagonal

Suppose it is of interest to test the hypothesis that

$$\begin{array}{lllll} H_0: & \sigma_{12} = & \sigma_{13} = \cdots = & \sigma_{1m} = 0 \\ & & \sigma_{23} = \cdots = & \sigma_{2m} = 0 \\ & & \ddots & \\ & & & \sigma_{mm} = 0 \end{array}$$

against the alternative that one or more of the off-diagonal elements of $\boldsymbol{\Sigma}$ are non-zero. The relevant log-likelihood ratio statistic for testing this hypothesis can be computed in *Microfit* as

$$LR = 2\left[\ell(\tilde{\boldsymbol{\theta}}) - \sum_{i=1}^{m} \ell_i(\hat{\boldsymbol{\theta}}_{i,OLS})\right] \tag{22.23}$$

where $\ell(\tilde{\boldsymbol{\theta}})$ is given by (22.11) and $\ell_i(\hat{\boldsymbol{\theta}}_{i,OLS})$ is the log-likelihood function of the ith equation computed at the OLS estimators. Equivalently, we have

$$LR = T\left[\sum_{i=1}^{m} \log\tilde{\sigma}_i^2 - \log\left|\widetilde{\boldsymbol{\Sigma}}\right|\right] \tag{22.24}$$

where

$$\tilde{\sigma}_{ii} = n^{-1}\left(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_{i,OLS}\right)'\left(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_{i,OLS}\right)$$

Under H_0, LR is asymptotically distributed as a χ^2 with $m(m-1)/2$ degrees of freedom.[2] See Lesson 18.2 for an implementation of this test.

22.3 System estimation subject to linear restrictions

Consider now the problem of estimating the system of equations (22.1) where the coefficient vectors $\boldsymbol{\beta}_i$, $i = 1, 2, ..., m$ are subject to the following $r \times 1$ linear restrictions:

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{b} \tag{22.25}$$

where $\mathbf{R}$ and $\mathbf{b}$ are $r \times k$ matrix and $r \times 1$ vector of known constants, and as in Section 22.1, $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', ..., \boldsymbol{\beta}_m'\right)'$, is a $k \times 1$ vector of unknown coefficients, $k = \sum_{i=1}^m k_i$.

In what follows we distinguish between the cases where the restrictions are applicable to the coefficients $\boldsymbol{\beta}_i$ in each equation separately, and when there are cross-equation restrictions. In the former case the matrix $\mathbf{R}$ is block diagonal, namely

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & & & \mathbf{0} \\ & \mathbf{R}_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{R}_m \end{pmatrix} \tag{22.26}$$

where $\mathbf{R}_i$ is the $r_i \times k_i$ matrix of known constants applicable to $\boldsymbol{\beta}_i$ only. In the more general case where the restrictions involve coefficients from different equations $\mathbf{R}$ is not block-diagonal.

The computations of the ML estimators of $\boldsymbol{\beta}$ in (22.1), subject to the restrictions in (22.25), can be carried out in the following manner. Initially suppose $\boldsymbol{\Sigma}$ is known, and define the $mn \times mn$ matrix $\mathbf{P}$ such that

$$\mathbf{P}\left(\boldsymbol{\Sigma} \otimes \mathbf{I}_n\right)\mathbf{P}' = \mathbf{I}_{mn} \tag{22.27}$$

where $\mathbf{I}_{mn}$ is an identity matrix of order mn. Such a matrix always exists, since $\boldsymbol{\Sigma}$ is a symmetric positive definite matrix. Then compute the transformations

$$\mathbf{X}_* = \mathbf{P}\mathbf{X}, \quad \mathbf{y}_* = \mathbf{P}\mathbf{y} \tag{22.28}$$

Using familiar results from estimation of linear regression models subject to linear restrictions we have (see, for example, Amemiya (1985) Section 1.4)

$$\widetilde{\boldsymbol{\beta}} = \left(\mathbf{X}_*'\mathbf{X}_*\right)^{-1}\mathbf{X}_*'\mathbf{y}_* - \left(\mathbf{X}_*'\mathbf{X}_*\right)^{-1}\mathbf{R}'\tilde{\mathbf{q}} \tag{22.29}$$

[2] An alternative LM test statistic proposed by Breusch and Pagan (1980) is given by

$$LM = n\sum_{i=2}^{m}\sum_{j=1}^{i-1} s_{ij}^2,$$

where $s_{ij}^2 = \tilde{\sigma}_{ij,OLS}/\left\{\tilde{\sigma}_{ii,OLS}\tilde{\sigma}_{jj,OLS}\right\}^{\frac{1}{2}}$. This statistic is also asymptotically distributed as a χ^2 with $m(m-1)/2$ degrees of freedom.

where

$$\tilde{\mathbf{q}} = \left(\mathbf{R}\left(\mathbf{X}_*'\mathbf{X}_*\right)^{-1}\mathbf{R}'\right)\left\{\mathbf{R}\left(\mathbf{X}_*'\mathbf{X}_*\right)^{-1}\mathbf{X}_*'\mathbf{y}_* - \mathbf{b}\right\} \tag{22.30}$$

In practice, since $\boldsymbol{\Sigma}$ is not known we need to estimate it. Starting with unrestricted $SURE$, or other initial estimates of $\boldsymbol{\beta}_i$ (say $\widehat{\boldsymbol{\beta}}_{i,OLS}$), an initial estimate of $\boldsymbol{\Sigma} = (\sigma_{ij})$ can be obtained. Using the OLS estimates of $\boldsymbol{\beta}_i$, the initial estimates of σ_{ij} are given by

$$\hat{\sigma}_{ij,OLS} = \frac{\hat{\mathbf{u}}_{i,OLS}'\hat{\mathbf{u}}_{j,OLS}}{n}, \quad i,j = 1,2,...,m$$

where

$$\hat{\mathbf{u}}_{i,OLS} = \mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_{i,OLS}, \quad i,j = 1,2,...,m$$

With the help of these initial estimates, constrained estimates of $\boldsymbol{\beta}_i$ can be computed using (22.29). Starting from these new estimates of $\boldsymbol{\beta}_i$, another set of estimates for σ_{ij} can be computed. This process can be repeated until the convergence criteria in (22.8) are met.

The covariance matrix of $\widetilde{\boldsymbol{\beta}}$ in this case is given by

$$\widehat{\mathrm{Cov}}(\widetilde{\boldsymbol{\beta}}) = \left(\widehat{\mathbf{X}}_*'\widehat{\mathbf{X}}_*\right)^{-1} - \left(\widehat{\mathbf{X}}_*'\widehat{\mathbf{X}}_*\right)^{-1}\mathbf{R}'\left(\mathbf{R}\left(\widehat{\mathbf{X}}_*'\widehat{\mathbf{X}}_*\right)^{-1}\mathbf{R}'\right)^{-1}\mathbf{R}\left(\widehat{\mathbf{X}}_*'\widehat{\mathbf{X}}_*\right)^{-1}$$

Notice that

$$\widehat{\mathbf{X}}_*'\widehat{\mathbf{X}}_* = \mathbf{X}'\widehat{\mathbf{P}}'\widehat{\mathbf{P}}\mathbf{X} = \mathbf{X}'\left(\widehat{\boldsymbol{\Sigma}}^{-1}\otimes\mathbf{I}_n\right)\mathbf{X}$$

The (i,j) element of $\widehat{\boldsymbol{\Sigma}}$ is computed differently depending on whether or not the matrix $\mathbf{R}$ in (22.26) is block diagonal. When $\mathbf{R}$ is block diagonal, σ_{ij} is estimated by

$$\hat{\sigma}_{ij} = \frac{\mathbf{u}_i'\mathbf{u}_j}{\sqrt{(n-s_i)(n-s_j)}}, \qquad i,j = 1,2,...,m \tag{22.31}$$

where $s_i = k_i - Rank(\mathbf{R}_i) = k_i - r_i$. When $\mathbf{R}$ is not block diagonal σ_{ij} is estimated by

$$\tilde{\sigma}_{ij} = \frac{\mathbf{u}_i'\mathbf{u}_j}{n}, \qquad i,j = 1,2,...,m \tag{22.32}$$

The divisor in (22.31) ensures that the results from the unrestricted and the restricted $SURE$ options in *Microfit* are compatible when there are no cross-equation restrictions. In the case where $\mathbf{R}$ is not block diagonal, an appropriate degrees-of-freedom correction is not available, and hence the ML estimator of σ_{ij} is used in the computation of the covariance matrix of the ML estimators of $\boldsymbol{\beta}$.

The maximum value of the log-likelihood function is computed as in (22.11), and the system AIC and SBC are computed as:

$$AIC = \ell(\widetilde{\boldsymbol{\theta}}) - (k-r)$$

and

$$SBC = \ell(\widetilde{\boldsymbol{\theta}}) - \tfrac{1}{2}(k-r)\log n$$

where $k = \sum_{i=1}^{m} k_i$, and $r = Rank(\mathbf{R})$. When $\mathbf{R}$ is block, $r = \sum_{i=1}^{m} r_i$.

Wald statistics for testing linear and/or non-linear restrictions on the elements of $\boldsymbol{\beta}$ can also be computed after the restricted $SURE$ option. The relevant formula is given by (22.22).

22.4 Augmented vector autoregressive models

Microfit allows estimation of the following augmented vector autoregressive model:[3]

$$\begin{aligned}\mathbf{z}_t &= \mathbf{a}_0 + \mathbf{a}_1 t + \sum_{i=1}^{p} \mathbf{\Phi}_i \mathbf{z}_{t-i} + \mathbf{\Psi}\mathbf{w}_t + \mathbf{u}_t, \qquad t = 1, 2, ..., n \\ &= \mathbf{A}'\mathbf{g}_t + \mathbf{u}_t\end{aligned} \tag{22.33}$$

where $\mathbf{z}_t$ is an $m \times 1$ vector of jointly determined dependent variables, and $\mathbf{w}_t$ is a $q \times 1$ vector of deterministic or exogenous variables. For example, $\mathbf{w}_t$ could include seasonal dummies, or exogenously given variables such as oil prices, foreign interest rates and prices in the case of small open economies. The $m \times 1$ vector of disturbances satisfy the following assumptions:

B1 $E(\mathbf{u}_t) = \mathbf{0}$.
B2 $E(\mathbf{u}_t \mathbf{u}_t') = \mathbf{\Sigma}$ for all t,
B3 $E(\mathbf{u}_t \mathbf{u}_{t'}') = \mathbf{0}$ for all $t \neq t'$,

where $\mathbf{\Sigma}$ is an $m \times m$ positive definite matrix.

B4 $E(\mathbf{u}_t \,|\mathbf{w}_t\,) = 0$.
B5 The augmented $VAR(p)$ model, (22.33), is stable; that is, all the roots of the determinantal equation

$$\left| \mathbf{I}_m - \mathbf{\Phi}_1 \lambda - \mathbf{\Phi}_2 \lambda^2 - \cdots - \mathbf{\Phi}_p \lambda^p \right| = 0 \tag{22.34}$$

fall outside the unit circle[4]

B6 The $m \times 1$ vector of disturbances have a multivariate normal distribution
B7 The observations $\mathbf{g}_t = (1, t, \mathbf{z}_{t-1}, \mathbf{z}_{t-2}, ..., \mathbf{z}_{t-p}, \mathbf{w}_t)$, for $t = 1, 2, ..., n$ are not perfectly collinear

Since the system of equations (22.33) is in the form of a $SURE$ model with all the equations having the same set of regressors $\mathbf{g}_t = (1, t, \mathbf{z}_{t-1}, \mathbf{z}_{t-2}, ..., \mathbf{z}_{t-p}, \mathbf{w}_t)$ in common, it then follows that when $\mathbf{u}_t$s are Gaussian the ML estimators of the unknown coefficients can be computed by OLS regressions of $\mathbf{z}_t$ on $\mathbf{g}_t$. Writing (22.33) in matrix notation we have

$$\underset{n \times m}{\mathbf{Z}} = \underset{n \times s}{\mathbf{G}} \; \underset{s \times m}{\mathbf{A}} + \underset{n \times m}{\mathbf{U}} \tag{22.35}$$

where $s = mp + q + 2$,

$$\underset{n \times m}{\mathbf{Z}} = (\mathbf{z}_1, \mathbf{z}_2, \ldots \mathbf{z}_n)'$$

[3] In the analysis of trend-stationary VAR models, without any loss of generality, the intercept and the trend terms $\boldsymbol{\iota}_n$ and $\mathbf{t}_n$ can be subsumed in $\mathbf{w}_t$. However, as it should become clear later, an explicit modelling of intercepts and trends is required in the case of cointegrating VAR models. See section 22.7.

[4] The case where one or more roots of (22.34) fall on the unit circle will be discussed in Section 22.7.

$$\underset{m\times s}{\mathbf{A}'} = (\mathbf{a}_0, \mathbf{a}_1, \mathbf{\Phi}_1, \mathbf{\Phi}_2, ..., \mathbf{\Phi}_p, \mathbf{\Psi})$$

$$\underset{n\times(mp+q+2)}{\mathbf{G}} = (\boldsymbol{\iota}_n, \mathbf{t}_n, \mathbf{Z}_{-1}, \mathbf{Z}_{-2}, ..., \mathbf{Z}_{-p}, \mathbf{W})$$

where $\boldsymbol{\iota}_n$ and $\mathbf{t}_n$ are the n-dimensional vectors $(1, 1, ..., 1)'$ and $(1, 2, ..., n)'$, respectively, and

$$\underset{n\times q}{\mathbf{W}} = (\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n)'$$

The ML estimators of $\mathbf{A}$ and $\mathbf{\Sigma}$ are given by

$$\widehat{\mathbf{A}} = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{Z} \tag{22.36}$$

and

$$\widetilde{\mathbf{\Sigma}} = n^{-1}(\mathbf{Z} - \mathbf{G}\widehat{\mathbf{A}})'(\mathbf{Z} - \mathbf{G}\widehat{\mathbf{A}}) \tag{22.37}$$

The maximized value of the system's log-likelihood function is given by

$$\ell\left(\widehat{\mathbf{A}}, \widetilde{\mathbf{\Sigma}}\right) = -\frac{nm}{2}(1 + \log 2\pi) - \frac{n}{2}\log\left|\widetilde{\mathbf{\Sigma}}\right| \tag{22.38}$$

The covariance matrix of the coefficients of the individual equations in the VAR model are computed using the standard least squares formula given in Section 21.6.1, making the usual degrees of freedom corrections, namely $\mathbf{\Sigma}$ is estimated by

$$\widehat{\mathbf{\Sigma}} = (n-s)^{-1}(\mathbf{Z} - \mathbf{G}\widehat{\mathbf{A}})'(\mathbf{Z} - \mathbf{G}\widehat{\mathbf{A}}) \tag{22.39}$$

The individual equation log-likelihood function, R^2, $\bar{R}^2$ and other summary and diagnostic statistics for individual equations are also computed using the formulae in Section 21.6.1.

22.4.1 VAR order selection

The order of the augmented VAR model (22.33), p, can be selected either with the help of model selection criteria such as the Akaike information criterion (AIC) and the Schwarz Bayesian criterion (SBC), or by means of a sequence of log-likelihood ratio tests. The values of the AIC and SBC for model (22.33) are given by

$$AIC_p = \frac{-nm}{2}(1 + \log 2\pi) - \frac{n}{2}\log\left|\widetilde{\mathbf{\Sigma}}_p\right| - ms \tag{22.40}$$

and

$$SBC_p = \frac{-nm}{2}(1 + \log 2\pi) - \frac{n}{2}\log\left|\widetilde{\mathbf{\Sigma}}_p\right| - \frac{ms}{2}\log(n) \tag{22.41}$$

where $s = mp + q + 2$, and $\widetilde{\mathbf{\Sigma}}_p$ is defined by (22.37). *Microfit* reports AIC_p and SBC_p for values of $p = 0, 1, 2, ..., P$, where P is the maximum order of the VAR model chosen by the user. The same augmenting set of variables, $\boldsymbol{\iota}_n$, $\mathbf{t}_n$and $\mathbf{w}_t$ (if any) are used in the computations as the order of VAR is changed.

The log-likelihood ratio statistic for testing the hypothesis that the order of the VAR is p against the alternative that it is $P\,(P > p)$ are given by

$$LR_{P,p} = n\left(\log\left|\widetilde{\mathbf{\Sigma}}_p\right| - \log\left|\widetilde{\mathbf{\Sigma}}_P\right|\right) \tag{22.42}$$

For $p = 0, 1, 2, ..., P-1$, where P is the maximum order for the VAR model selected by the user, $\widetilde{\mathbf{\Sigma}}_p$ is defined by (22.37), and $\widetilde{\mathbf{\Sigma}}_0$ refers to the ML estimator of the system covariance matrix in the regression of $\mathbf{z}_t$ on $\boldsymbol{\iota}_n$, $\mathbf{t}_n$and $\mathbf{w}_t$.

Under the null hypothesis, the LR statistic in (22.42) is asymptotically distributed as a chi-squared variate with $m^2(P-p)$ degrees of freedom.

In small samples the use of the LR statistic (22.42) tends to result in over-rejection of the null hypothesis. In an attempt to take some account of this small sample problem, in practice the following degrees of freedom adjusted LR statistics are also computed:.

$$LR^*_{P,p} = (n-q-2-mP)\left(\log\left|\widetilde{\mathbf{\Sigma}}_p\right| - \log\left|\widetilde{\mathbf{\Sigma}}_P\right|\right) \tag{22.43}$$

for $p = 0, 1, 2, ..., P-1$. These adjusted LR statistics have the same asymptotic distribution as the unadjusted statistics given by (22.42).

22.4.2 Testing the deletion of deterministic/exogenous variables

Microfit computes log-likelihood ratio statistics for testing the deletion of $\boldsymbol{\iota}_n$, $\mathbf{t}_n$and $\mathbf{w}_t$ or a sub-set of these variables from the $VAR(p)$ model (22.33). For notational convenience, from here onwards until notice to the contrary we shall be subsuming the intercept and the trend terms, $\boldsymbol{\iota}_n$ and $\mathbf{t}_n$, in $\mathbf{w}_t$. Let

$$\underset{q\times 1}{\mathbf{w}_t} = \begin{pmatrix} \mathbf{w}_{1t}\,, & q_1\times 1 \\ \mathbf{w}_{2t} & q_2\times 1 \end{pmatrix} \quad \text{and} \quad \underset{m\times q}{\mathbf{\Psi}} = \begin{pmatrix} \underset{m\times q_1}{\mathbf{\Psi}_1}\,, & \underset{m\times q_2}{\mathbf{\Psi}_2} \end{pmatrix} \tag{22.44}$$

where $q = q_1 - q_2$. The log-likelihood ratio statistic for testing the null hypothesis

$$H_0 : \mathbf{\Psi}_1 = \mathbf{0}, \quad \text{against} \quad H_1 : \mathbf{\Psi}_1 \neq \mathbf{0}$$

is computed as

$$LR(\mathbf{\Psi}_1 = \mathbf{0}) = 2\left\{LL_U - LL\left(\mathbf{\Psi}_1 = \mathbf{0}\right)\right\} \tag{22.45}$$

where LL_U is the unrestricted maximized value of the log-likelihood function given by (22.38) and $LL(\mathbf{\Psi}_1 = \mathbf{0})$ is the maximized value of the log-likelihood function obtained under $\mathbf{\Psi}_1 = 0$. Asymptotically $LR(\mathbf{\Psi}_1 = \mathbf{0})$ is distributed as a chi-squared variate with mq_1 degrees of freedom.

22.4.3 Testing for block Granger non-causality

Let $\mathbf{z}_t = (\mathbf{z}'_{1t}, \mathbf{z}'_{2t})'$, where $\mathbf{z}_{1t}$ and $\mathbf{z}_{2t}$ are $m_1 \times 1$ and $m_2 \times 1$, $(m_1 + m_2 = m)$ variables, and partition the system of equations (22.33) (or equivalently (22.35)) into the two sub-systems

$$\mathbf{Z}_1 = \mathbf{Y}_1\mathbf{A}_{11} + \mathbf{Y}_2\mathbf{A}_{12} + \mathbf{W}\mathbf{A}_{13} + \mathbf{U}_1 \tag{22.46}$$

$$\mathbf{Z}_2 = \mathbf{Y}_1\mathbf{A}_{21} + \mathbf{Y}_2\mathbf{A}_{22} + \mathbf{W}\mathbf{A}_{23} + \mathbf{U}_2 \tag{22.47}$$

where $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are $n \times m_1$ and $n \times m_2$ matrices of observations on $\mathbf{z}_{1t}$ and $\mathbf{z}_{2t}$ respectively; $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are $n \times pm_1$ and $n \times pm_2$ matrices of observations on the p lagged values of $\mathbf{z}_{1,t-\ell}$, and $\mathbf{z}_{2,t-\ell}$, for $t = 1, 2, ..., n$, $\ell = 1, 2, ..., p$, respectively. The hypothesis that '$\mathbf{z}_{2t}$ do not Granger cause $\mathbf{z}_{1t}$' is defined by the m_1m_2p restrictions $\mathbf{A}_{12} = \mathbf{0}$.[5]

The log-likelihood ratio statistic for the test of these restrictions is computed as

$$LR_G\left(\mathbf{A}_{12} = \mathbf{0}\right) = 2\left(\log\left|\widetilde{\mathbf{\Sigma}}_R\right| - \log\left|\widetilde{\mathbf{\Sigma}}\right|\right)$$

where $\widetilde{\mathbf{\Sigma}}$ is ML estimator of $\mathbf{\Sigma}$ for the unrestricted (full) system (22.37), and $\widetilde{\mathbf{\Sigma}}_R$ is the ML estimator of $\mathbf{\Sigma}$ when the restrictions $\mathbf{A}_{12} = \mathbf{0}$ are imposed. Under the null hypothesis that $\mathbf{A}_{12} = \mathbf{0}$, LR_G is asymptotically distributed as a chi-squared variate with m_1m_2p degrees of freedom.

Since under $\mathbf{A}_{12} = \mathbf{0}$, the system of equations (22.46) and (22.47) are block recursive, $\widetilde{\mathbf{\Sigma}}_R$ can be computed in the following manner:

1. Run OLS regressions of $\mathbf{Z}_1$ on $\mathbf{Y}_1$ and $\mathbf{W}$, and compute the $n \times m_1$ matrix of residuals, $\widehat{\mathbf{U}}_1$.

2. Run the OLS regressions

$$\mathbf{Z}_2 = \mathbf{Y}_1\mathbf{A}^*_{21} + \mathbf{Y}_2\mathbf{A}^*_{22} + \mathbf{W}\mathbf{A}^*_{23} + \widehat{\mathbf{U}}_1\mathbf{A}^*_{24} + \mathbf{V}_2 \tag{22.48}$$

and compute the $n \times m_2$ matrix of residuals:

$$\widehat{\mathbf{U}}_2 = \mathbf{Z}_2 - \mathbf{Y}_1\widehat{\mathbf{A}}^*_{21} - \mathbf{Y}_2\widehat{\mathbf{A}}^*_{22} - \mathbf{W}\widehat{\mathbf{A}}^*_{23}$$

where $\widehat{\mathbf{A}}^*_{21}$, $\widehat{\mathbf{A}}^*_{22}$ and $\widehat{\mathbf{A}}^*_{23}$ are the OLS estimators of $\mathbf{A}^*_{21}$, $\mathbf{A}^*_{22}$ and $\mathbf{A}^*_{23}$, in (22.48). Define

$$\widehat{\mathbf{U}} = \left(\widehat{\mathbf{U}}_1 : \widehat{\mathbf{U}}_2\right)$$

Then

$$\widetilde{\mathbf{\Sigma}}_R = n^{-1}\left(\widehat{\mathbf{U}}'\widehat{\mathbf{U}}\right) \tag{22.49}$$

22.5 Impulse response analysis

The impulse response function measures the time profile of the effect of shocks on the future states of a dynamical system. In the case of the $VAR(p)$ model (22.33), two different impulse response functions can be computed using *Microfit*:

1. The Orthogonalized Impulse Response Function (IRF) advanced by Sims (1980) and Sims (1981).

[5] See Engle, Hendry, and Richard (1983) for a discussion on the notions of predeterminedness, strict exogeneity and causality.

2. The Generalized IR Function (GIRF) proposed by Koop, Pesaran, and Potter (1996), and Pesaran and Shin (1998).

Both impulse response functions work with the $m \times m$ coefficient matrices $\mathbf{A}_i$, in the infinite moving average representation of (22.33):[6]

$$\mathbf{z}_t = \sum_{j=0}^{\infty} \mathbf{A}_j \mathbf{u}_{t-j} + \sum_{j=0}^{\infty} \mathbf{B}_j \mathbf{w}_{t-j}, \tag{22.50}$$

where the matrices $\mathbf{A}_j$ are computed using the recursive relations

$$\mathbf{A}_j = \boldsymbol{\Phi}_1 \mathbf{A}_{j-1} + \boldsymbol{\Phi}_2 \mathbf{A}_{j-2} + ... + \boldsymbol{\Phi}_p \mathbf{A}_{j-p}, \quad j = 1, 2, ..., \tag{22.51}$$

with $\mathbf{A}_0 = \mathbf{I}_m$, and $\mathbf{A}_j = \mathbf{0}$, for $j < 0$, and $\mathbf{B}_j = \mathbf{A}_j \boldsymbol{\Psi}$, for $j = 1, 2, ..$

22.5.1 Orthogonalized impulse responses

Sims' approach employs the following Cholesky decomposition of $\boldsymbol{\Sigma}$ (the covariance matrix of the shocks, $\mathbf{u}_t$):

$$\boldsymbol{\Sigma} = \mathbf{T}\mathbf{T}' \tag{22.52}$$

where $\mathbf{T}$ is a lower triangular matrix. Sims then rewrites the moving average representation (22.50) as

$$\begin{aligned} \mathbf{z}_t &= \sum_{j=0}^{\infty} (\mathbf{A}_j \mathbf{T}) \left(\mathbf{T}^{-1} \mathbf{u}_{t-j}\right) + \sum_{j=0}^{\infty} \mathbf{B}_j \mathbf{w}_{t-j} \\ &= \sum_{j=0}^{\infty} \mathbf{A}_j^* \boldsymbol{\epsilon}_{t-j} + \sum_{j=0}^{\infty} \mathbf{B}_j \mathbf{w}_{t-j} \end{aligned} \tag{22.53}$$

where

$$\mathbf{A}_j^* = \mathbf{A}_j \mathbf{T}, \text{ and } \ \boldsymbol{\epsilon}_t = \mathbf{T}^{-1} \mathbf{u}_t$$

It is now easily seen that

$$E\left(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t'\right) = \mathbf{T}^{-1} E\left(\mathbf{u}_t \mathbf{u}_t'\right) \mathbf{T}'^{-1} = \mathbf{T}^{-1} \boldsymbol{\Sigma} \mathbf{T}'^{-1} = \mathbf{I}_m$$

and the new errors $\boldsymbol{\epsilon}_t$ obtained using the transformation matrix, $\mathbf{T}$, are now contemporaneously uncorrelated and have unit standard errors. In other words the shocks $\boldsymbol{\epsilon}_t = (\boldsymbol{\epsilon}_{1t}, \boldsymbol{\epsilon}_{2t}, ..., \boldsymbol{\epsilon}_{mt})'$ are orthogonal to each other.

The Orthogonalized IR function of a 'unit shock' (equal to one standard error) at time t to the ith orthogonalized error, namely $\boldsymbol{\epsilon}_{it}$, on the jth variable at time $t + N$, is given by the jth element of:

$$\begin{matrix}\text{Orthogonalized IR function to} \\ \text{the } i\text{th variable (equation)}\end{matrix} = \mathbf{A}_N^* \mathbf{e}_i = \mathbf{A}_N \mathbf{T} \mathbf{e}_i \tag{22.54}$$

[6] Notice that the existence of such an infinite MA representation is ensured by condition (**B**5).

where $\mathbf{e}_i$ is the $m \times 1$ selection vector,

$$\mathbf{e}_i = (0, \;\; 0, \;\; ... \;\; 0, \;\; \underset{\substack{\uparrow \\ i\text{th element}}}{1}, \;\; 0, \;\; ... \;\; 0)' \tag{22.55}$$

or, written more compactly

$$OI_{ij,N} = \mathbf{e}_j' \mathbf{A}_N \mathbf{T} \mathbf{e}_i, \quad i, j, = 1, 2, ..., m \tag{22.56}$$

These orthogonalized impulse responses are not unique and in general depend on the particular ordering of the variables in the VAR. The orthogonalized responses are invariant to the ordering of the variables only if $\mathbf{\Sigma}$ is diagonal (or almost diagonal).[7] The non-uniqueness of the orthogonalized impulse responses is also related to the non-uniqueness of the matrix $\mathbf{T}$ in the Cholesky decomposition of $\mathbf{\Sigma}$ in (22.52). For more details see Lütkepohl (2005) Section 2.3.2.

22.5.2 Generalized impulse responses

The main idea behind the Generalized IR function is to circumvent the problem of the dependence of the orthogonalized impulse responses to the ordering of the variables in the VAR. The concept of the Generalized Impulse Response function, advanced in Koop, Pesaran, and Potter (1996) was originally intended to deal with the problem of impulse response analysis in the case of non-linear dynamical systems, but can also be readily applied to multivariate time-series models such as VAR, as set out in Pesaran and Shin (1998).

The Generalized IR analysis deals explicitly with three main issues that arise in impulse response analysis:

1. How was the dynamical system hit by shocks at time t? Was it hit by a variable-specific shock or system-wide shocks?

2. What was the state of the system at time $t-1$, before the system was hit by shock(s)? Was the trajectory of the system in an upward or in a downward phase?

3. How would one expect the system to be shocked in the future, namely over the interim period from $t+1$, to $t+N$?

In the context of the VAR model (22.33), the Generalized Impulse Response function for a system-wide shock, $\mathbf{u}_t^0$, is defined by

$$GI_{\mathbf{z}}\left(N, \mathbf{u}_t^0, \Omega_{t-1}^0\right) = E\left(\mathbf{z}_{t+N} \left| \mathbf{u}_t = \mathbf{u}_t^0, \Omega_{t-1}^0\right.\right) - E\left(\mathbf{z}_{t+N} \left| \Omega_{t-1}^0\right.\right) \tag{22.57}$$

where $E(\cdot\,|\cdot)$ is the conditional mathematical expectation taken with respect to the VAR model (22.33), and Ω_{t-1}^0 is a particular historical realization of the process at time $t-1$.

[7] Tests of the diagonality of $\mathbf{\Sigma}$ are discussed in Section 22.2.2.

In the case of the VAR model having the infinite moving-average representation (22.50) we have

$$GI_{\mathbf{z}}\left(N, \mathbf{u}_t^0, \Omega_{t-1}^0\right) = \mathbf{A}_N \mathbf{u}_t^0 \tag{22.58}$$

which is independent of the 'history' of the process. This history invariance property of the impulse response function (also shared by the traditional methods of impulse response analysis) is, however, specific to linear systems, and does not carry over to non-linear dynamic models.

In practice, the choice of the vector of shocks, $\mathbf{u}_t^0$, is arbitrary; one possibility would be to consider a large number of likely shocks and then examine the empirical distribution function of $\mathbf{A}_N \mathbf{u}_t^0$ for all these shocks. In the case where $\mathbf{u}_t^0$ is drawn from the same distribution as $\mathbf{u}_t$, namely a multivariate normal with zero means and a constant covariance matrix $\mathbf{\Sigma}$, we have the analytical result that

$$GI_{\mathbf{z}}\left(N, \mathbf{u}_t^0, \Omega_{t-1}^0\right) \sim N\left(0, \mathbf{A}_N \mathbf{\Sigma} \mathbf{A}_N'\right) \tag{22.59}$$

The diagonal elements of $\mathbf{A}_N \mathbf{\Sigma} \mathbf{A}_N'$, when appropriately scaled, are the 'persistence profiles' proposed in Lee and Pesaran (1993) and applied in Pesaran and Shin (1996) to analyze the speed of convergence to equilibrium in cointegrated systems (see Section 22.9.5). It is also worth noting that when the underlying VAR model is stable (i.e. condition **B5** is met), the limit of the persistence profile as $N \to \infty$ tends to the spectral density function of $\mathbf{z}_t$ (without the $\mathbf{w}_t$s) at zero frequency (apart from a multiple of π).

Consider now the effect of a variable-specific shock on the evolution of $\mathbf{z}_{t+1}, \mathbf{z}_{t+2}, ..., \mathbf{z}_{t+N}$, and suppose that for a given $\mathbf{w}_t$, the VAR model is perturbed by a shock of size $\delta_i = \sqrt{\sigma_{ii}}$ to its ith equation at time t. By the definition of the Generalized IR function we have:

$$GI_{\mathbf{z}}\left(N, \delta_i, \Omega_{t-1}^0\right) = E\left(\mathbf{z}_t \left| u_{it} = \delta_i, \Omega_{t-1}^0\right.\right) - E\left(\mathbf{z}_t \left| \Omega_{t-1}^0\right.\right) \tag{22.60}$$

Once again using the infinite moving-average representation (22.50), we obtain

$$GI_{\mathbf{z}}\left(N, \delta_i, \Omega_{t-1}^0\right) \sim \mathbf{A}_N E\left(\mathbf{u}_t \left| \mathbf{u}_{it} = \delta_i\right.\right) \tag{22.61}$$

which is history invariant (it does not depend on Ω_{t-1}^0). The computation of the conditional expectations $E\left(\mathbf{u}_t \left| \mathbf{u}_{it} = \delta_i\right.\right)$ depends on the nature of the multivariate distribution assumed for the disturbances, $\mathbf{u}_t$. In the case where $\mathbf{u}_t \sim N\left(0, \mathbf{\Sigma}\right)$, we have

$$E\left(\mathbf{u}_t \left| \mathbf{u}_{it} = \delta_i\right.\right) = \begin{pmatrix} \sigma_{1i}/\sigma_{ii} \\ \sigma_{2i}/\sigma_{ii} \\ \vdots \\ \sigma_{mi}/\sigma_{ii} \end{pmatrix} \delta_i \tag{22.62}$$

where as before, $\mathbf{\Sigma} = (\boldsymbol{\sigma}_{ij})$. Hence for a 'unit shock' defined by $\delta_i = \sqrt{\sigma_{ii}}$, we have

$$GI_{\mathbf{z}}\left(N, \delta_i = \sqrt{\sigma_{ii}}, \Omega_{t-1}^0\right) = \frac{\mathbf{A}_N \mathbf{\Sigma} \mathbf{e}_i}{\sqrt{\sigma_{ii}}}, \quad i, j, = 1, 2, ..., m \tag{22.63}$$

where $\mathbf{e}_i$ is a selection vector given by (22.55). The Generalized Impulse Response Function (GIRF) of a unit shock to the ith equation in the VAR model (22.33) on the jth variable at horizon N is given by the jth element of (22.63), or expressed more compactly

$$GI_{ij,N} = \frac{\mathbf{e}_j' \mathbf{A}_N \mathbf{\Sigma} \mathbf{e}_i}{\sqrt{\sigma_{ii}}}, \quad i, j, = 1, 2, ..., m \tag{22.64}$$

Unlike the Orthogonalized Impulse Responses in (22.54), the Generalized Impulse Responses in (22.63) are invariant to the ordering of the variables in the VAR. It is also interesting to note that the two impulse responses coincide only for the first variable in the VAR, or when $\mathbf{\Sigma}$ is a diagonal matrix. See Pesaran and Shin (1998).

22.6 Forecast error variance decompositions

The forecast error variance decomposition provides a decomposition of the variance of the forecast errors of the variables in the VAR at different horizons.

22.6.1 Orthogonalized forecast error variance decomposition

In the context of the orthogonalized moving-average representation of the VAR model given by (22.53), the forecast error variance decomposition for the ith variable in the VAR is given by

$$\theta_{ij,N} = \frac{\sum_{\ell=0}^{N} \left(\mathbf{e}_i' \mathbf{A}_\ell \mathbf{T} \mathbf{e}_j\right)^2}{\sum_{\ell=0}^{N} \mathbf{e}_i' \mathbf{A}_\ell \mathbf{\Sigma} \mathbf{A}_\ell' \mathbf{e}_i}, \quad i, j = 1, 2, ..., m \tag{22.65}$$

where $\mathbf{T}$ is defined by the Cholesky decomposition of $\mathbf{\Sigma}$, (22.52), $\mathbf{e}_i$ is the selection vector defined by (22.55), and $\mathbf{A}_\ell$, $\ell = 0, 1, 2, ...$ are the coefficient matrices in the moving-average representation, (22.50). Notice that $\mathbf{e}_i' \mathbf{A}_\ell \mathbf{\Sigma} \mathbf{A}_\ell' \mathbf{e}_i$ is simply the ith diagonal element of the matrix $\mathbf{A}_\ell \mathbf{\Sigma} \mathbf{A}_\ell'$, which also enters the persistence profile analysis (see Lee and Pesaran (1993)).

$\theta_{ij,N}$ measures the proportion of the N-step ahead forecast error variance of variable i, which is accounted for by the orthogonalized innovations in variable j. For further details, see, for example, Lütkepohl (2005) Section 2.3.3.[8] As with the Orthogonalized Impulse Response function, the orthogonalized forecast error variance decompositions in (22.65) are not invariant to the ordering of the variables in the VAR.

22.6.2 Generalized forecast error variance decomposition

An alternative procedure to the orthogonalized forecast error variance decomposition would be to consider the proportion of the variance of the N-step forecast errors of $\mathbf{z}_t$ that are

[8] Notice also that $\sum_{j=1}^{m} \theta_{ij,N} = 1$.

explained by conditioning on the non-orthogonalized shocks, $u_{it}, u_{i,t+1}, ..., u_{i,t+N}$, but explicitly allowing for the contemporaneous correlations between these shocks and the shocks to the other equations in the system.

Using the MA representation (22.50),[9] the forecast error of predicting $\mathbf{z}_{t+N}$ conditional on the information at time $t-1$ is given by

$$\underset{m\times 1}{\boldsymbol{\xi}_t(N)} = \sum_{\ell=0}^{N} \mathbf{A}_\ell \mathbf{u}_{t+N-\ell} \tag{22.66}$$

with the *total* forecast error covariance matrix

$$\mathrm{Cov}\left(\boldsymbol{\xi}_t(N)\right) = \sum_{\ell=0}^{N} \mathbf{A}_\ell \boldsymbol{\Sigma} \mathbf{A}'_\ell \tag{22.67}$$

Consider now the forecast error covariance matrix of predicting $\mathbf{z}_{t+N}$ conditional on the information at time $t-1$, *and* given values of the shocks to the ith equation, u_{it}, $u_{i,t+1}$, ..., $u_{i,t+N}$. Using (22.50) we have[10]

$$\underset{m\times 1}{\boldsymbol{\xi}_t^{(i)}(N)} = \sum_{\ell=0}^{N} \mathbf{A}_\ell \left(\mathbf{u}_{t+N-\ell} - E\left(\mathbf{u}_{t+N-\ell} \left| u_{i,t+N-\ell}\right.\right)\right) \tag{22.68}$$

As in the case of the Generalized Impulse Responses, assuming $\mathbf{u}_t \sim N\left(0, \boldsymbol{\Sigma}\right)$ we have

$$E\left(\mathbf{u}_{t+N-\ell} \left| u_{i,t+N-\ell}\right.\right) = \left(\sigma_{ii}^{-1}\boldsymbol{\Sigma}\mathbf{e}_i\right) u_{i,t+N-\ell} \qquad \text{for} \quad \begin{array}{l} \ell = 0,1,2,...,N \\ i = 1,2,...,m \end{array}$$

Substituting this result back in (22.68)

$$\boldsymbol{\xi}_t^{(i)}(N) = \sum_{\ell=0}^{N} \mathbf{A}_\ell \left(\mathbf{u}_{t+N-\ell} - \sigma_{ii}^{-1}\boldsymbol{\Sigma}\mathbf{e}_i u_{i,t+N-\ell}\right)$$

and taking unconditional expectations, yields

$$\mathrm{Cov}\left(\boldsymbol{\xi}_t^{(i)}(N)\right) = \sum_{\ell=0}^{N} \mathbf{A}_\ell \boldsymbol{\Sigma} \mathbf{A}'_\ell - \sigma_{ii}^{-1}\left(\sum_{\ell=0}^{N} \mathbf{A}_\ell \boldsymbol{\Sigma}\mathbf{e}_i\mathbf{e}'_i\boldsymbol{\Sigma}\mathbf{A}'_\ell\right) \tag{22.69}$$

Therefore, using (22.67) and (22.69) it follows that the decline in the N-step forecast error variance of $\mathbf{z}_t$ obtained as a result of conditioning on the future shocks to the ith equation is given by

$$\begin{aligned} \boldsymbol{\Delta}_{iN} &= \mathrm{Cov}\left[\boldsymbol{\xi}_t(N)\right] - \mathrm{Cov}\left[\boldsymbol{\xi}_t^{(i)}(N)\right] \\ &= \sigma_{ii}^{-1}\sum_{\ell=0}^{N} \mathbf{A}_\ell \boldsymbol{\Sigma}\mathbf{e}_i\mathbf{e}'_i\boldsymbol{\Sigma}\mathbf{A}'_\ell. \end{aligned} \tag{22.70}$$

[9] We continue to assume that $\mathbf{w}_t$s are given.

[10] Notice that since $\mathbf{u}_t$s are serially uncorrelated, $E\left(\mathbf{u}_{t+n-\ell} \left| u_{it},\ u_{i,t+1}, ..., u_{i,t+n}\right.\right) = E\left(\mathbf{u}_{t+N-\ell} \left| u_{i,t+n-\ell}\right.\right), \quad \ell = 0,1,2,...,N.$

Scaling the jth diagonal element of $\mathbf{\Delta}_{iN}$, namely $\mathbf{e}_j'\mathbf{\Delta}_{iN}\mathbf{e}_j$, by the N-step ahead forecast error variance of the ith variable in $\mathbf{z}_t$, we have the following generalized forecast error variance decomposition:

$$\Psi_{ij,N} = \frac{\sigma_{ii}^{-1}\sum_{\ell=0}^{N}(\mathbf{e}_j'\mathbf{A}_\ell\mathbf{\Sigma}\mathbf{e}_i)^2}{\sum_{\ell=0}^{N}\mathbf{e}_j'\mathbf{A}_\ell\mathbf{\Sigma}\mathbf{A}_\ell'\mathbf{e}_j} \tag{22.71}$$

Note that the denominator of this measure is the ith diagonal element of the total forecast error variance formula in (22.67), and is the same as the denominator of the orthogonalized forecast error variance decomposition formula (22.65). Also $\boldsymbol{\theta}_{ij,N} = \mathbf{\Psi}_{ij,N}$ when z_{it} is the first variable in the VAR, and/or $\mathbf{\Sigma}$ is diagonal. However, in general the two decompositions differ.

For computational purposes it is worth noting that the numerator of (22.71) can also be written as the sum of squares of the generalized responses of the shocks to the ith equation on the jth variable in the model, namely $\sum_{\ell=0}^{N}(GI_{ij,\ell})^2$, where $GI_{ij,\ell}$ is given by (22.64).

22.7 Cointegrating VAR

The statistical framework for the cointegrating VAR options in *Microfit* is the following general vector error correction model ($VECM$):

$$\Delta\mathbf{y}_t = \mathbf{a}_{0y} + \mathbf{a}_{1y}t - \mathbf{\Pi}_y\mathbf{z}_{t-1} + \sum_{i=1}^{p-1}\mathbf{\Gamma}_{iy}\Delta\mathbf{z}_{t-i} + \mathbf{\Psi}_y\mathbf{w}_t + \boldsymbol{\epsilon}_t,\ t = 1, 2, ..., n \tag{22.72}$$

where

- $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$, $\mathbf{y}_t$ is an $m_y \times 1$ vector of jointly determined (endogenous) $I(1)$ variables
- $\mathbf{x}_t$ is an $m_x \times 1$ vector of exogenous $I(1)$ variables[11]

$$\Delta\mathbf{x}_t = \mathbf{a}_{0x} + \sum_{i=1}^{p-1}\mathbf{\Gamma}_{ix}\Delta\mathbf{z}_{t-i} + \mathbf{\Psi}_x\mathbf{w}_t + \mathbf{v}_t \tag{22.73}$$

- $\mathbf{w}_t$ is a $q \times 1$ vector of exogenous/deterministic $I(0)$ variables, excluding the intercepts and/or trends
- The disturbance vectors $\boldsymbol{\epsilon}_t$ and $\mathbf{v}_t$ satisfy the following assumptions:

$$\mathbf{u}_t = \begin{pmatrix}\boldsymbol{\epsilon}_t \\ \mathbf{v}_t\end{pmatrix} \sim IID\left(\mathbf{0}, \mathbf{\Sigma}\right), \tag{22.74}$$

  where $\mathbf{\Sigma}$ is a symmetric positive-definite matrix

[11] Notice that (22.73) allows for feedbacks from $\Delta\mathbf{y}$ to $\Delta\mathbf{x}$, but does not allow for level feedbacks, and hence assumes that $\mathbf{x}_t$s are not themselves cointegrated.

- The disturbances in the combined model, $\mathbf{u}_t$, are distributed independently of $\mathbf{w}_t$

$$E\left(\mathbf{u}_t \,|\mathbf{w}_t\right) = \mathbf{0} \tag{22.75}$$

The intercept and the trend coefficients, $\mathbf{a}_{0y}$ and $\mathbf{a}_{1y}$ are $m_y \times 1$ vectors, $\mathbf{\Pi}_y$ is the long-run multiplier matrix of order $m_y \times m$, where $m = m_x + m_y$, $\mathbf{\Gamma}_{1y}, \mathbf{\Gamma}_{2y}, ..., \mathbf{\Gamma}_{p-1,y}$ are $m_y \times m$ coefficient matrices capturing the short-run dynamic effects, and $\mathbf{\Psi}_y$ is the $m_y \times q$ matrix of coefficients on the $I(0)$ exogenous variables.

The $VECM$ in (22.72) differs in a number of important respects from the usual VAR formulation for the $VECM$ analyzed *inter alia* by Johansen (1991). Firstly, (22.72) allows for a sub-system approach in which the m_x-vector of random variables $\mathbf{x}_t$ are the forcing variables, or common 'stochastic trends', in the sense that the error correction terms do not enter in the sub-system for $\mathbf{x}_t$ (given by (22.73)). Therefore, cointegrating analysis in *Microfit* allows for contemporaneous and short-term feedbacks from y_t to $\mathbf{x}_t$, but requires that no such feedbacks are possible in the long-run. We refer to $\mathbf{x}_t$ as the 'long-run forcing' variables of the system. Secondly, the cointegration analysis critically depends on whether the underlying $VECM$ contains intercepts and/or time trends, and whether the intercepts, $\mathbf{a}_{0y}$, and the trend coefficients, $\mathbf{a}_{1y}$, are restricted. Accordingly, the cointegration analysis in *Microfit* distinguishes between five cases of interest ordered according to the importance of the trends:

Case I: $\mathbf{a}_{0y} = \mathbf{a}_{1y} = \mathbf{0}$ (no intercepts and no trends)

Case II: $\mathbf{a}_{1y} = \mathbf{0}$, and $\mathbf{a}_{0y} = \mathbf{\Pi}_y \boldsymbol{\mu}_y$ (restricted intercepts and no trends)

Case III: $\mathbf{a}_{1y} = \mathbf{0}$, and $\mathbf{a}_{0y} \neq \mathbf{0}$ (unrestricted intercepts and no trends)

Case IV: $\mathbf{a}_{0y} \neq \mathbf{0}$ and $\mathbf{a}_{1y} = \mathbf{\Pi}_y \gamma_y$ (unrestricted intercepts and restricted trends)

Case V: $\mathbf{a}_{0y} \neq \mathbf{0}$, and $\mathbf{a}_{1y} \neq \mathbf{0}$ (unrestricted intercepts and trends)

The rationale behind the restricted intercepts and the restricted trend cases are discussed below.

22.7.1 Cointegrating relations

The cointegrating VAR analysis is concerned with the estimation of (22.72) when the rank of the long-run multiplier matrix, $\mathbf{\Pi}$, could at most be equal to m_y. Therefore, rank deficiency of $\mathbf{\Pi}$ can be represented as

$$H_r : \text{Rank}\left(\mathbf{\Pi}_y\right) = r < m_y.$$

In this case we can write

$$\mathbf{\Pi}_y = \boldsymbol{\alpha}_y \boldsymbol{\beta}'$$

where $\boldsymbol{\alpha}_y$ and $\boldsymbol{\beta}$ are $m_y \times r$ and $m \times r$ matrices, each with full column rank, r. In the case where $\mathbf{\Pi}_y$ is rank deficient we have $\mathbf{y}_t \sim I(1)$, $\Delta\mathbf{y}_t \sim I(0)$, and $\boldsymbol{\beta}'\mathbf{z}_t \sim I(0)$. The $r \times 1$ trend-stationary relations, $\boldsymbol{\beta}'\mathbf{z}_t$, are referred to as the cointegrating relations, and characterize the long-run equilibrium (steady state) of the $VECM$ (22.72).

It is, however, important to recognize that in the case where the $VECM$ (22.72) contains deterministic trends ($\mathbf{a}_{1y} \neq \mathbf{0}$), in general there will also be a linear trend in the cointegrating relations. To see this, combining the equation systems (22.72) and (22.73) we have

$$\Delta \mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t - \mathbf{\Pi} \mathbf{z}_{t-1} + \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta \mathbf{z}_{t-i} + \mathbf{\Psi} \mathbf{w}_t + \mathbf{u}_t \tag{22.76}$$

for $t = 1, 2, ..., n$, where

$$\mathbf{z}_t = \begin{pmatrix} \mathbf{y}_t \\ \mathbf{x}_t \end{pmatrix}, \quad \mathbf{u}_t = \begin{pmatrix} \mathbf{u}_{yt} \\ \mathbf{v}_t \end{pmatrix}, \quad \mathbf{a}_0 = \begin{pmatrix} \mathbf{a}_{0y} \\ \mathbf{a}_{0x} \end{pmatrix}, \quad \mathbf{a}_1 = \begin{pmatrix} \mathbf{a}_{1y} \\ \mathbf{0} \end{pmatrix}$$

$$\mathbf{\Pi} = \begin{pmatrix} \mathbf{\Pi}_y \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{\Gamma}_i = \begin{pmatrix} \mathbf{\Gamma}_{iy} \\ \mathbf{\Gamma}_{ix} \end{pmatrix}, \quad \mathbf{\Psi} = \begin{pmatrix} \mathbf{\Psi}_y \\ \mathbf{\Psi}_x \end{pmatrix}$$

which is the vector error correction form of (22.33).

In the case where $\mathbf{\Pi}$ is rank deficient, the solution of (22.76) involves common stochastic trends, and is given by,[12]

$$\mathbf{z}_t = \mathbf{z}_0 + \mathbf{b}_0 t + \mathbf{b}_1 \left\{ \frac{t(t+1)}{2} \right\} + \mathbf{C}(1) \mathbf{S}_t + \mathbf{C}^*(L) (\mathbf{h}_t - \mathbf{h}_0) \tag{22.77}$$

where

$$\mathbf{h}_t = \mathbf{\Psi} \mathbf{w}_t + \mathbf{u}_t \tag{22.78}$$

$$\mathbf{S}_t = \sum_{i=1}^{t} \mathbf{u}_i, \quad t = 1, 2, ... \tag{22.79}$$

$$\mathbf{b}_0 = \mathbf{C}(1) \mathbf{a}_0 + \mathbf{C}^*(1) \mathbf{a}_1 \tag{22.80}$$

$$\mathbf{b}_1 = \mathbf{C}(1) \mathbf{a}_1 \tag{22.81}$$

$$\mathbf{C}(L) = \mathbf{C}(1) + (1 - L) \mathbf{C}^*(L), \tag{22.82}$$

$$\mathbf{C}^*(L) = \sum_{i=0}^{\infty} \mathbf{C}_i^* L^i$$

where L is the one-period lag operator and the $m \times m$ matrices $\mathbf{C}_i^*$ are obtained recursively from

$$\mathbf{C}_i^* = \mathbf{C}_{i-1}^* \mathbf{\Phi}_1 + \cdots + \mathbf{C}_{i-p}^* \mathbf{\Phi}_p \tag{22.83}$$

$i = 1, 2, ..,$ with $\mathbf{C}_0^* = \mathbf{I}_m - \mathbf{C}(1)$, $\mathbf{C}_i^* = \mathbf{0}$, $i < 0$, and

$$\mathbf{\Pi} \mathbf{C}(1) = \mathbf{0} = \mathbf{C}(1) \mathbf{\Pi} \tag{22.84}$$

[12] See, for example, Pesaran and Shin (2002), and Pesaran, Shin, and Smith (1996).

The matrices $\mathbf{\Phi}_1, \mathbf{\Phi}_2, ..., \mathbf{\Phi}_p$ are the coefficient matrices in the VAR form of (22.76), and in terms of $\mathbf{\Pi}$, $\mathbf{\Gamma}_1$, $\mathbf{\Gamma}_2, ...,$ and $\mathbf{\Gamma}_{p-1}$ are given by

$$\begin{aligned} \mathbf{\Phi}_1 &= \mathbf{I}_m - \mathbf{\Pi} + \mathbf{\Gamma}_1 \\ \mathbf{\Phi}_i &= \mathbf{\Gamma}_i - \mathbf{\Gamma}_{i-1}, \qquad i = 2, 3, ..., p-1 \\ \mathbf{\Phi}_p &= -\mathbf{\Gamma}_{p-1} \end{aligned}$$

From solution (22.77) it is clear that in general $\mathbf{z}_t$ will contain a quadratic trend. When $\mathbf{a}_1 \neq \mathbf{0}$, the quadratic trend disappears only if $\mathbf{C}(1)\mathbf{a}_1 = \mathbf{0}$, otherwise the number of independent quadratic trend terms in the solution of $\mathbf{z}_t$ will be equal to the Rank of $\mathbf{C}(1)$ and hence depends on the number of cointegrating relations. Note that Rank$[\mathbf{C}(1)] = m - r$. Therefore, without some restrictions on the trend coefficients, $\mathbf{a}_1$, the solution (22.77) has the unsatisfactory property that the nature of the trend in z_t varies with the assumed number of cointegrating relations. This outcome can be avoided by restricting the trend coefficients as in Case IV, namely by setting $\mathbf{a}_1 = \mathbf{\Pi}\boldsymbol{\gamma}$. Under these restrictions using (22.81) and (22.84) we have

$$\mathbf{b}_1 = \mathbf{C}(1)\mathbf{a}_1 = \mathbf{C}(1)\mathbf{\Pi}\boldsymbol{\gamma} = \mathbf{0}$$

and the $VECM$ in (22.76) becomes

$$\Delta\mathbf{z}_t = \mathbf{a}_0 - \mathbf{\Pi}\left(\mathbf{z}_{t-1} - \boldsymbol{\gamma}t\right) + \sum_{i=1}^{p-1}\mathbf{\Gamma}_i\Delta\mathbf{z}_{t-i} + \mathbf{\Psi}\mathbf{w}_t + \mathbf{u}_t \tag{22.85}$$

A similar consideration also applies where the $VECM$ contains intercepts, but no trends. In this case, unless the intercepts are appropriately restricted (as in Case II) the nature of the trend in $\mathbf{z}_t$ will vary with the number of the cointegrating relations.

Using (22.77) the cointegrating relations, $\boldsymbol{\beta}'\mathbf{z}_t$, can also be derived in terms of the shocks $\mathbf{u}_{t-i}$, $i = 0, 1, 2, ...$, and the current and past values of the $I(0)$ exogenous values. Pre-multiplying (22.77) by $\boldsymbol{\beta}'$, and bearing in mind the cointegration restrictions $\boldsymbol{\beta}'\mathbf{C}(1) = \mathbf{0}$, we obtain[13]

$$\boldsymbol{\beta}'\mathbf{z}_t = \boldsymbol{\beta}'\mathbf{z}_0 + \left(\boldsymbol{\beta}'\mathbf{b}_0\right)t + \boldsymbol{\beta}'\mathbf{C}^*(L)\left(\mathbf{h}_t - \mathbf{h}_0\right) \tag{22.86}$$

Using (22.80) we also have

$$\boldsymbol{\beta}'\mathbf{b}_0 = \boldsymbol{\beta}'\mathbf{C}^*(1)\mathbf{a}_1 \tag{22.87}$$

and hence when $\mathbf{a}_1 \neq \mathbf{0}$, the cointegrating relations $\boldsymbol{\beta}'\mathbf{z}_t$, in general, contain deterministic trends, which do not disappear even if a_1 is restricted as in Case IV. When $a_1 = \Pi\gamma$, the coefficients of the deterministic trend in the cointegrating relations are given by

$$\boldsymbol{\beta}'\mathbf{b}_0 = \boldsymbol{\beta}'\mathbf{C}^*(1)\mathbf{\Pi}\boldsymbol{\gamma}.$$

But as shown in Pesaran and Shin (2002), $\mathbf{C}^*(1)\mathbf{\Pi} = \mathbf{I}_m$, and $\boldsymbol{\beta}'\mathbf{b}_0 = \boldsymbol{\beta}'\boldsymbol{\gamma} \neq \mathbf{0}$. Using this result in (22.86) we have

$$\boldsymbol{\beta}'\mathbf{z}_t = \boldsymbol{\beta}'\mathbf{z}_0 + \left(\boldsymbol{\beta}'\boldsymbol{\gamma}\right)t + \boldsymbol{\beta}'\mathbf{C}^*(L)\left(\mathbf{h}_t - \mathbf{h}_0\right) \tag{22.88}$$

[13] Notice that $\boldsymbol{\beta}'\mathbf{b}_1 = \boldsymbol{\beta}'\mathbf{C}(1)\mathbf{a}_1 = 0$, irrespective of whether or not the trend coefficients $\mathbf{a}_1$ are restricted.

A test of whether the cointegrating relations are trended can be carried out by testing the following r restrictions:

$$\boldsymbol{\beta}'\boldsymbol{\gamma} = \mathbf{0} \tag{22.89}$$

We shall refer to these as the 'cotrending' restrictions.[14]

22.8 ML estimation and tests of cointegration

Suppose that n observations $\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n$ and $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n$ are available on the variables $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$, and $\mathbf{w}_t$. Then, stacking the $VECM$ in (22.72) we have

$$\Delta\mathbf{Y} = \boldsymbol{\iota}_n\mathbf{a}_{0y}' + \mathbf{t}_n\mathbf{a}_{1y}' - \mathbf{Z}_{-1}\boldsymbol{\Pi}_y' + \Delta\mathbf{Z}_p\boldsymbol{\Gamma}_y' + \mathbf{W}\boldsymbol{\Psi}_y' + \mathbf{E} \tag{22.90}$$

where

$$\Delta\mathbf{Y} = (\Delta\mathbf{y}_1, \Delta\mathbf{y}_2, ..., \Delta\mathbf{y}_n)'$$

$$\mathbf{E} = (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, ..., \boldsymbol{\epsilon}_n)'$$

$$\boldsymbol{\tau}_n = (1, 1, ..., 1)', \mathbf{t}_n = (1, 2, ..., n)'$$

$$\boldsymbol{\Gamma}_y = (\boldsymbol{\Gamma}_{1y}, \boldsymbol{\Gamma}_{2y}, ..., \boldsymbol{\Gamma}_{p-1,y})$$

$$\Delta\mathbf{Z}_p = (\Delta\mathbf{Z}_{-1}, \Delta\mathbf{Z}_{-2}, ..., \Delta\mathbf{Z}_{1-p})$$

$$\Delta\mathbf{Z}_{-i} = (\Delta\mathbf{z}_{1-i}, \Delta\mathbf{z}_{2-i}, ..., \Delta\mathbf{z}_{n-i})', i = 1, 2, ..., p-1$$

The log-likelihood function of (22.90) is given by

$$\ell_n(\boldsymbol{\varphi}; r) = \frac{-nm_y}{2}\log 2\pi - \frac{n}{2}\log|\boldsymbol{\Sigma}_y| - \tfrac{1}{2}Tr\left(\boldsymbol{\Sigma}_y^{-1}\mathbf{E}'\mathbf{E}\right) \tag{22.91}$$

where $\boldsymbol{\Sigma}_y = E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t')$, $\boldsymbol{\varphi}$ denotes for the vector of the unknown parameters of the model, and r is the assumed rank of $\boldsymbol{\Pi}_y$. Writing $\boldsymbol{\Pi}_y = \boldsymbol{\alpha}_y\boldsymbol{\beta}'$, and maximizing the log-likelihood function with respect to the elements of $\boldsymbol{\Sigma}_y$, $\mathbf{a}_{0y}$, $\mathbf{a}_{1y}$, $\boldsymbol{\Gamma}_{iy}$, $i = 1, 2, ..., p-1$, $\boldsymbol{\alpha}_y$ and $\boldsymbol{\Psi}_y$, we have the following concentrated log-likelihood function:

$$\ell_n^c(\boldsymbol{\beta}; r) = \frac{-nm_y}{2}(1 + \log 2\pi) - \frac{n}{2}\log\left|\widetilde{\boldsymbol{\Sigma}}_y(\boldsymbol{\beta})\right| \tag{22.92}$$

where

$$\left|\widetilde{\boldsymbol{\Sigma}}_y(\boldsymbol{\beta})\right| = \frac{|\mathbf{S}_{00}|\left|\boldsymbol{\beta}'\mathbf{A}_n\boldsymbol{\beta}\right|}{\left|\boldsymbol{\beta}'\mathbf{B}_n\boldsymbol{\beta}\right|} \tag{22.93}$$

$$\mathbf{A}_n = \mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}, \text{ and } \mathbf{B}_n = \mathbf{S}_{11} \tag{22.94}$$

$$\mathbf{S}_{ij} = n^{-1}\sum_{t=1}^{n}\mathbf{r}_{it}\mathbf{r}_{jt}', \quad i, j = 0, 1 \tag{22.95}$$

and $\mathbf{r}_{0t}$ and $\mathbf{r}_{1t}$ for $t = 1, 2, ..., n$ are the residual vectors obtainable from the following regressions:

[14] Also see Park (1992) and Ogaki (1992).

Case I: $(\mathbf{a}_{0y} = \mathbf{a}_{1y} = \mathbf{0})$

$\mathbf{r}_{0t}$ is the residual vector from the OLS regressions of $\Delta\mathbf{y}_t$ on $(\Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$, and $\mathbf{r}_{1t}$ is the residual vector from the OLS regressions of $\mathbf{z}_{t-1}$ on $(\Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$.

Case II: $\left(\mathbf{a}_{1y} = \mathbf{0}, \mathbf{a}_{0y} = \mathbf{\Pi}_y \boldsymbol{\mu}_y\right)$

$\mathbf{r}_{0t}$ is the residual vector from the OLS regressions of $\Delta\mathbf{y}_t$ on $(\Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$, and $\mathbf{r}_{1t}$ is the residual vector from the OLS regressions of $\begin{pmatrix} 1 \\ \mathbf{z}_{t-1} \end{pmatrix}$ on $(\Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$.

Case III: $(\mathbf{a}_{1y} = \mathbf{0}, \mathbf{a}_{0y} \neq \mathbf{0})$

$\mathbf{r}_{0t}$ is the residual vector from the OLS regressions of $\Delta\mathbf{y}_t$ on $(1, \Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$, and $\mathbf{r}_{1t}$ is the residual vector from the OLS regressions of $\mathbf{z}_{t-1}$ on $(1, \Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$.

Case IV: $\left(\mathbf{a}_{0y} \neq \mathbf{0}, \mathbf{a}_{1y} = \mathbf{\Pi}_y \boldsymbol{\gamma}_y\right)$

$\mathbf{r}_{0t}$ is the residual vector from the OLS regressions of $\Delta\mathbf{y}_t$ on $(1, \Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$, and $\mathbf{r}_{1t}$ is the residual vector from the OLS regressions of $\begin{pmatrix} t \\ \mathbf{z}_{t-1} \end{pmatrix}$ on $(1, \Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$.

Case V: $(\mathbf{a}_{0y} \neq \mathbf{0}, \mathbf{a}_{1y} \neq \mathbf{0})$

$\mathbf{r}_{0t}$ is the residual vector from the OLS regressions of $\Delta\mathbf{y}_t$ on $(1, t, \Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$, and $\mathbf{r}_{1t}$ is the residual vector from the OLS regressions of $\mathbf{z}_{t-1}$ on $(1, t, \Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ..., \Delta\mathbf{z}_{t-p+1}, \mathbf{w}_t)$.

Substituting (22.93) in (22.92) yields

$$\begin{aligned} \ell_n^c(\boldsymbol{\beta}; r) &= \frac{-nm_y}{2}(1 + \log 2\pi) - \frac{n}{2}\log|\mathbf{S}_{00}| \\ &\quad -\frac{n}{2}\left\{\log\left|\boldsymbol{\beta}'\mathbf{A}_n\boldsymbol{\beta}\right| - \log\left|\boldsymbol{\beta}'\mathbf{B}_n\boldsymbol{\beta}\right|\right\} \end{aligned} \tag{22.96}$$

The dimension of $\boldsymbol{\beta}$ depends on whether or not the intercepts, a_{0y}, and/or the trend coefficients, a_{1y}, are restricted. For example, in Case IV where $a_{0y} \neq 0$, and $a_{1y} = \Pi_y \gamma_y$, the term $t_n a'_{1y} - Z_{-1}\Pi'_y$in (22.90) can be written as

$$\mathbf{t}_n \mathbf{a}'_{1y} - \mathbf{Z}_{-1}\mathbf{\Pi}'_y = -\,\mathbf{Z}^*_{-1}\mathbf{\Pi}^{*\prime}_y$$

where

$$\mathbf{\Pi}^*_y = \mathbf{\Pi}_y\left(-\boldsymbol{\gamma}_y, \mathbf{I}_m\right)$$
$$\mathbf{Z}^*_{-1} = \left(\mathbf{z}^*_0, \mathbf{z}^*_1, ..., \mathbf{z}^*_{n-1}\right)$$

and $\mathbf{z}^*_t = \begin{pmatrix} t \\ \mathbf{z}_{t-1} \end{pmatrix}$. In this case the cointegrating vectors are defined by

$$\mathbf{\Pi}^*_y = \boldsymbol{\alpha}_y \boldsymbol{\beta}'_*$$

and $\boldsymbol{\beta}$ in (22.96) should be replaced by the $(m+1)\times r$ matrix $\boldsymbol{\beta}_*$.

The unconstrained maximization of $\ell_n^c(\boldsymbol{\beta};r)$ will not lead to unique estimates of $\boldsymbol{\beta}$ (or $\boldsymbol{\beta}_*$), and $\boldsymbol{\beta}$ can only be identified up to post-multiplication by an $r\times r$ non-singular matrix. It is easily seen that

$$\ell_n^c(\boldsymbol{\beta};r) = \ell_n^c(\boldsymbol{\beta}\mathbf{Q};r)$$

where $\mathbf{Q}$ is any non-singular $r\times r$ matrix. Therefore, r^2 just-identifying restriction on $\boldsymbol{\beta}$ (or $\boldsymbol{\beta}_*$) are required for exact identification. The resultant maximized concentrated log-likelihood function $\ell_n^c(\boldsymbol{\beta};r)$ at the ML estimator of $\boldsymbol{\beta}$ does not, however, depend on $\mathbf{Q}$, and is given by

$$\ell_n^c(r) = \frac{-nm_y}{2}(1+\log 2\pi) - \frac{n}{2}\log|\mathbf{S}_{00}| - \frac{n}{2}\sum_{i=1}^{r}\log\left(1-\widehat{\lambda}_i\right) \tag{22.97}$$

for all exactly identified choices of $\boldsymbol{\beta}$, where $\widehat{\lambda}_1 > \widehat{\lambda}_2 > \cdots \widehat{\lambda}_r > 0$ are the r largest eigenvalues of $\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\mathbf{S}_{11}^{-1}\mathbf{S}_{10}$ (or equivalently the eigenvalues of $\mathbf{S}_{11}^{-1}\mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}$).

22.8.1 Maximum eigenvalue statistic

Suppose the interest is in testing the null hypothesis of r cointegrating relations

$$H_r : \text{Rank}(\boldsymbol{\Pi}_y) = r \tag{22.98}$$

against the alternative hypothesis that

$$H_{r+1} : \text{Rank}(\boldsymbol{\Pi}_y) = r+1 \tag{22.99}$$

$r = 0, 1, 2, ..., m_y - 1$, in the $VECM$ (22.72). Then the appropriate test statistic is given by the log-likelihood ratio statistic

$$\mathfrak{LR}\left(H_r\,|H_{r+1}\right) = -n\log\left(1-\widehat{\lambda}_{r+1}\right) \tag{22.100}$$

where $\widehat{\lambda}_r$ is the rth largest eigenvalue of $\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\mathbf{S}_{11}^{-1}\mathbf{S}_{10}$, and the matrices $\mathbf{S}_{00}$, $\mathbf{S}_{01}$, and $\mathbf{S}_{11}$ are defined by (22.95).

The bootstrapped critical values are also provide in *Microfit*.

22.8.2 Trace statistic

Suppose the interest is in testing the null hypothesis $H(r)$ defined in (22.98) against the alternative of trend-stationarity, that is

$$H_{m_y} : \text{Rank}(\boldsymbol{\Pi}_y) = m_y \tag{22.101}$$

for $r = 0, 1, 2, ..., m_y - 1$. The log-likelihood ratio statistic for this test is given by

$$\mathfrak{LR}\left(H_r\,\left|H_{m_y}\right.\right) = -n\sum_{i=r+1}^{m_y}\log\left(1-\widehat{\lambda}_{r+1}\right) \tag{22.102}$$

where $\widehat{\lambda}_{r+1}$, $\widehat{\lambda}_{r+1}$, ..., $\widehat{\lambda}_{m_y}$ are the largest eigenvalues of $\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\mathbf{S}_{11}^{-1}\mathbf{S}_{10}$, where the matrices $\mathbf{S}_{00}$, $\mathbf{S}_{01}$, and $\mathbf{S}_{11}$ are defined by (22.95).

The critical values for the maximum eigenvalue and the trace statistics defined by (22.100) and (22.102), respectively, depend on $m_y - r$, m_x, whether the $VECM$ (22.72) contains intercepts and/or trends, and whether these are restricted; namely, which one of the five cases set out above is applicable. They are computed in Pesaran, Shin, and Smith (1996), using stochastic simulation techniques. They cover all the five cases and allow for up to twelve endogenous $I(1)$ variables and up to five exogenous $I(1)$ variables in the $VECM$ (22.72). The critical values do not, however, depend on the order of the VAR, p, and the stochastic properties of the $I(0)$ exogenous variables, $\mathbf{w}_t$ at least in large samples.

Following the Johansen (1991) approach, Osterwald-Lenum (1992) also provided the critical values of both trace and the maximum eigenvalue statistic only in the case where up to eleven endogenous variables and no $I(1)$ exogenous variables are considered in the system. There are also differences between his critical values and the critical values tabulated by Pesaran, Shin, and Smith (1996) for Cases III and V, because Johansen does not restrict the intercept coefficient in Case III or the trend coefficients in Case V.

Monte Carlo simulation results indicate that these cointegrating rank test statistics generally tend to under-reject in small samples see Pesaran, Shin, and Smith (2000)). Bootstrapped critical values can be obtained in *Microfit*.

22.8.3 Model selection criteria for choosing the number of cointegrating relations

The model selection criteria AIC, SBC and HQC, defined in Section 21.7, are also computed for different values of r, the rank of the long-run matrix $\mathbf{\Pi}_y$ in (22.72). We have

$$
\begin{aligned}
AIC &= \ell_n^c(r) - \mathfrak{s} && (22.103)\\
SBC &= \ell_n^c(r) - \left(\frac{\mathfrak{s}}{2}\right)\log n && (22.104)\\
HQC &= \ell_n^c(r) - \left(\frac{\mathfrak{s}}{2}\right)\log\log n && (22.105)
\end{aligned}
$$

where $\ell_n^c(r)$ is given by (22.97), and $\mathfrak{s}$ is the total number of coefficients estimated. The value of $\mathfrak{s}$ depends on whether the intercepts and the trend coefficients in (22.72) are restricted. The value of $\mathfrak{s}$ for the five cases distinguished in *Microfit* are as follows:[15]

Case I: $(\mathbf{a}_{0y} = \mathbf{a}_{1y} = \mathbf{0})$

$$\mathfrak{s} = mm_y(p-1) + (m + m_y)\,r - r^2 + qm_y$$

[15] The number of free parameters in the $m_y \times m$ long-run matrix $\mathbf{\Pi}_y$ depends on its rank. When Rank$(\mathbf{\Pi}_y) = r$, $\mathbf{\Pi}_y$ contains $(m_y + m)\,r - r^2$ free parameters (or equivalently $\mathbf{\Pi}_y$ will be subject to $(m-r)\,(m_y - r)$ restrictions). This result follows from the so-called 'UDV' decomposition of $\mathbf{\Pi}_y = \mathbf{UDV}$, where $\mathbf{U}$, $\mathbf{D}$, and $\mathbf{V}$ are $m_y \times r$, $r \times r$ and $r \times m$ matrices such that $\mathbf{U}'\mathbf{U} = \mathbf{I}_r$, $\mathbf{V}'\mathbf{V} = \mathbf{I}_r$, and $\mathbf{D}$ is a diagonal matrix of rank r.

Case II: $\left(\mathbf{a}_{1y} = \mathbf{0}, \mathbf{a}_{0y} = \mathbf{\Pi}_y \boldsymbol{\mu}_y\right)$

$$\mathfrak{s} = mm_y(p-1) + (m + m_y + 1)\, r - r^2 + qm_y$$

Case III: $(\mathbf{a}_{1y} = \mathbf{0}, \mathbf{a}_{0y} \neq \mathbf{0})$

$$\mathfrak{s} = mm_y(p-1) + (m + m_y)\, r - r^2 + (q+1)\, m_y$$

Case IV: $\left(\mathbf{a}_{0y} \neq \mathbf{0}, \mathbf{a}_{1y} = \mathbf{\Pi}_y \boldsymbol{\gamma}_y\right)$

$$\mathfrak{s} = mm_y(p-1) + (m + m_y + 1)\, r - r^2 + (q+1) m_y$$

Case V: $(\mathbf{a}_{0y} \neq \mathbf{0}, \mathbf{a}_{1y} \neq \mathbf{0})$

$$\mathfrak{s} = mm_y(p-1) + (m + m_y)\, r - r^2 + (q+2)\, m_y$$

Recall also that $m = m_x + m_y$.

22.9 Long-run structural modelling

As we have seen already, the estimation of the $VECM$ (22.72) subject to deficient rank restrictions on the long-run multiplier matrix, $\mathbf{\Pi}_y$, does not generally lead to a unique choice for the cointegrating relations. The identification of $\boldsymbol{\beta}$ (in $\mathbf{\Pi}_y = \boldsymbol{\alpha}_y \boldsymbol{\beta}'$) requires at least r restrictions per each of the r cointegrating relations.[16] In the simple case where $r = 1$, the one restriction needed to identify the cointegrating relation can be viewed as a 'normalizing' restriction which could be applied to the coefficient of any one of the integrated variables that enter the cointegrating relation. However, in the more general case where $r > 1$, the number of such 'normalizing' restrictions is just equal to r, which needs to be supplemented with further $r^2 - r$ *a priori* restrictions, preferably derived from a suitable economic theory.[17]

22.9.1 Identification of the cointegrating relations

The structural estimation of the cointegrating relations requires maximization of the concentrated log-likelihood function (22.96) subject to appropriate just-identifying or over-identifying restrictions on $\boldsymbol{\beta}$. The just-identifying restrictions utilized in Johansen (1988), and Johansen (1991) estimation procedure involve the observation matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ defined by (22.94), and are often referred to as 'empirical' or 'statistical' identifying restrictions, as compared to *a priori* restrictions on $\boldsymbol{\beta}$ which are independent of particular values of $\mathbf{A}_n$ and $\mathbf{B}_n$. Johansen's estimates of $\boldsymbol{\beta}$, which we denote by $\widehat{\boldsymbol{\beta}}_J$, are obtained as the first r eigenvectors of $\mathbf{B}_n - \mathbf{A}_n$ with respect to $\mathbf{B}_n$, satisfying the following 'normalization' and 'orthogonalization' restrictions:

$$\widehat{\boldsymbol{\beta}}_J' \mathbf{B}_n \widehat{\boldsymbol{\beta}}_J = \mathbf{I}_r \tag{22.106}$$

[16] Readers interested in more details should consult Pesaran and Shin (2002).

[17] The role of economic theory in providing suitable identifying restrictions on the cointegrating vectors is discussed in Pesaran (1997).

and

$$\widehat{\boldsymbol{\beta}}'_{iJ}\left(\mathbf{B}_n - \mathbf{A}_n\right)\widehat{\boldsymbol{\beta}}_{jJ} = 0, \qquad i \neq j, \quad i, j = 1, 2, ..., r \tag{22.107}$$

where $\widehat{\boldsymbol{\beta}}_{iJ}$ represents the ith column of $\widehat{\boldsymbol{\beta}}_J$. The conditions (22.106) and (22.107) together exactly impose r^2 just-identifying restrictions on $\boldsymbol{\beta}$. It is, however, clear that the r^2 restrictions in (22.106) and (22.107) are adopted for their mathematical convenience, and not because they are meaningful from the perspective of any long-run economic theory.

A more satisfactory procedure would be to directly estimate the concentrated log-likelihood function (22.96) subject to exact or over-identifying *a priori* restrictions obtained from the long-run equilibrium properties of a suitable underlying economic model (see Pesaran (1997)). *Microfit* enables you to compute ML estimates of $\boldsymbol{\beta}$ (and hence the other parameters in the $VECM$ (22.72)), when the elements of $\boldsymbol{\beta}$ are subject to the following general linear restrictions:

$$\mathbf{R}\,\text{vec}(\boldsymbol{\beta}) = \mathbf{b} \tag{22.108}$$

where $\mathbf{R}$ and $\mathbf{b}$ are $k \times rm$ matrix and $k \times 1$ vector of known constants (with Rank$(\mathbf{R}) = k$), and vec$(\boldsymbol{\beta})$ is $rm \times 1$ vector of long-run coefficients, which stacks the r columns of $\boldsymbol{\beta}$ into a vector. As in Section 22.3, we can distinguish between the cases where the restrictions are applicable to columns of $\boldsymbol{\beta}$ separately, and when they involve parameters from two or more cointegrating vectors. In the former case the matrix $\mathbf{R}$ is block-diagonal and (22.108) can be written as

$$\mathbf{R}_i\boldsymbol{\beta}_i = \mathbf{b}_i\,, \; i = 1, 2, ..., r \tag{22.109}$$

where $\boldsymbol{\beta}_i$ is the ith cointegrating vector, and $\mathbf{R}_i$ is the ith block in matrix $\mathbf{R}$, and $\mathbf{b}_i$ is defined by $\mathbf{b}' = (\mathbf{b}'_1, \mathbf{b}'_2, ..., \mathbf{b}'_r)$. In this case the necessary and sufficient conditions for identification of the cointegrating vectors are given by

$$\text{Rank}\left(\mathbf{R}_i\boldsymbol{\beta}\right) = r, \qquad i = 1, 2, ..., r \tag{22.110}$$

This result also implies that *there must be at least r independent restrictions on each of the r cointegrating vectors.*

The identification condition in the case where $\mathbf{R}$ is not block diagonal is given by

$$\text{Rank}\left\{\mathbf{R}\left(\mathbf{I}_r \otimes \boldsymbol{\beta}\right)\right\} = r^2 \tag{22.111}$$

A necessary condition for (22.111) to hold is given by the order condition $k \geq r^2$. As with the Cowles Commission approach, three cases of interest can be distinguished:

1. $k < r^2$, the under-identified case
2. $k = r^2$, the exactly identified case
3. $k > r^2$, the over-identified case.

22.9.2 Estimation of the cointegrating relations under general linear restrictions

Here we distinguish between two cases: when the long-run restrictions are exactly identified ($k = r^2$), and when there are over-identifying restriction on the cointegrating vectors ($k > r^2$).

Exactly identified case $\left(k = r^2\right)$

In this case the ML estimator of $\boldsymbol{\beta}$ that satisfy the restrictions (22.108) are readily computed using Johansen's estimates, $\widehat{\boldsymbol{\beta}}_J$. We have:

$$\operatorname{vec}(\widehat{\boldsymbol{\beta}}) = \left(\mathbf{I}_r \otimes \widehat{\boldsymbol{\beta}}_J\right)\left[\mathbf{R}\left(I_r \otimes \widehat{\boldsymbol{\beta}}_J\right)\right]^{-1}\mathbf{b} \tag{22.112}$$

where $\otimes$ denotes the Kronecker product. It is easily verified that this estimator satisfies the restriction (22.108), and is invariant to non-singular transformations of the cointegrating space spanned by columns of $\widehat{\boldsymbol{\beta}}$.[18]

Over-identified case $\left(k > r^2\right)$

In this case there are $k-r^2$ additional restrictions that need to be taken into account at the estimation stage. This can be done by maximization of the concentrated log-likelihood function given by (22.96), subject to the restrictions given by (22.108). We assume that the normalization restrictions on each of the r cointegrating vectors is also included in $\mathbf{R}\operatorname{vec}(\boldsymbol{\beta}) = \mathbf{b}$. The advantage of working with (22.96) lies in the fact that the data matrices $\mathbf{A}_n$ and $\mathbf{B}_n$,defined by (22.94), need to be computed only once, and the speed of convergence of the proposed algorithm does not depend on the sample size, T. The Lagrangian function for this problem is given by

$$\begin{aligned}\Lambda(\boldsymbol{\theta},\ \boldsymbol{\lambda}) &= \frac{1}{n}\ell_n^c(\boldsymbol{\theta}; r) - \frac{1}{2}\boldsymbol{\lambda}'(\mathbf{R}\boldsymbol{\theta} - \mathbf{b}) \\ &= \text{constant} - \frac{1}{2}\left\{\log\left|\boldsymbol{\beta}'\mathbf{A}_n\boldsymbol{\beta}\right| - \log\left|\boldsymbol{\beta}'\mathbf{B}_n\boldsymbol{\beta}\right|\right\} - \frac{1}{2}\boldsymbol{\lambda}'(\mathbf{R}\boldsymbol{\theta} - \mathbf{b})\end{aligned}$$

where $\boldsymbol{\theta} = \operatorname{vec}(\boldsymbol{\beta})$, $\boldsymbol{\lambda}$ is a $k \times 1$ vector of Lagrange multipliers, and $\mathbf{A}_n$ and $\mathbf{B}_n$ are defined in (22.94). Then, first order conditions are given by

$$\mathbf{d}_n\left(\widetilde{\boldsymbol{\theta}}\right) = \mathbf{R}'\widetilde{\boldsymbol{\lambda}} \tag{22.113}$$

$$\mathbf{R}\widetilde{\boldsymbol{\theta}} = \mathbf{b} \tag{22.114}$$

where $\widetilde{\boldsymbol{\theta}}$ and $\widetilde{\boldsymbol{\lambda}}$ denote the restricted ML estimators, and $\mathbf{d}_n(\widetilde{\boldsymbol{\theta}})$ is the score function defined by

$$\mathbf{d}_n(\widetilde{\boldsymbol{\theta}}) = \left\{\left[\left(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n\widetilde{\boldsymbol{\beta}}\right)^{-1} \otimes \mathbf{A}_n\right] - \left[\left(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n\widetilde{\boldsymbol{\beta}}\right)^{-1} \otimes \mathbf{B}_n\right]\right\}\widetilde{\boldsymbol{\theta}} \tag{22.115}$$

Here we propose two different but related numerical procedures for the computation of $\widetilde{\boldsymbol{\theta}}$. The *first procedure* is a 'back-substitution' algorithm and uses only the information on the first derivatives. It involves solving the system of equations, (22.113) and (22.114) numerically for $\widetilde{\boldsymbol{\theta}}$ $\left(= \operatorname{vec}(\widetilde{\boldsymbol{\beta}})\right)$, after eliminating $\widetilde{\boldsymbol{\lambda}}$. Define

$$\mathbf{P}_n = \left(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n\widetilde{\boldsymbol{\beta}}\right) \otimes \mathbf{A}_n^{-1}, \quad \text{and} \quad \mathbf{F}_n = \left(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n\widetilde{\boldsymbol{\beta}}\right)\left(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n\widetilde{\boldsymbol{\beta}}\right)^{-1} \otimes \mathbf{A}_n^{-1}\mathbf{B}_n \tag{22.116}$$

[18] For a derivation of this result see Pesaran and Shin (2002)

and pre-multiply (22.113) by $\mathbf{P}_n$ to obtain[19]

$$\widetilde{\boldsymbol{\theta}} = \mathbf{F}_n\widetilde{\boldsymbol{\theta}} + \mathbf{P}_n\mathbf{R}'\widetilde{\boldsymbol{\lambda}} \tag{22.117}$$

Now multiplying both sides of this relation by $\mathbf{R}$ we have

$$\mathbf{R}\widetilde{\boldsymbol{\theta}} = \mathbf{R}\mathbf{F}_n\widetilde{\boldsymbol{\theta}} + (\mathbf{R}\mathbf{P}_n\mathbf{R}')\widetilde{\boldsymbol{\lambda}} = \mathbf{b} \tag{22.118}$$

Since by assumption $\mathbf{P}_n$ is non-singular, then Rank$(\mathbf{R}\mathbf{P}_n\mathbf{R}') = \text{Rank}\,(\mathbf{R}) = k$, which means that $\mathbf{R}\mathbf{P}_n\mathbf{R}'$ is also non-singular, and $\widetilde{\boldsymbol{\lambda}}$ is given by

$$\widetilde{\boldsymbol{\lambda}} = (\mathbf{R}\mathbf{P}_n\mathbf{R}')^{-1}(\mathbf{b} - \mathbf{R}\mathbf{F}_\mathbf{n}\widetilde{\boldsymbol{\theta}}) \tag{22.119}$$

Next, eliminating $\widetilde{\boldsymbol{\lambda}}$ from (22.117) using (22.119), we have

$$\widetilde{\boldsymbol{\theta}} = \mathbf{f}\left(\widetilde{\boldsymbol{\theta}}\right) \tag{22.120}$$

where

$$\mathbf{f}\left(\widetilde{\boldsymbol{\theta}}\right) = \mathbf{S}_n^{-1}\mathbf{P}_n\mathbf{R}'(\mathbf{R}\mathbf{P}_n\mathbf{R}')^{-1}\mathbf{b} \tag{22.121}$$

$$\mathbf{S}_n = \mathbf{I}_{mr} - \mathbf{F}_n + \mathbf{P}_n\mathbf{R}'(\mathbf{R}\mathbf{P}_n\mathbf{R}')^{-1}\mathbf{R}\mathbf{F}_n \tag{22.122}$$

The $mr \times 1$ vector function $\mathbf{f}(\cdot)$ depends on $\widetilde{\boldsymbol{\theta}}$ through the positive definite matrices $\widetilde{\boldsymbol{\beta}}'\mathbf{A}_T\widetilde{\boldsymbol{\beta}}$ and $\widetilde{\boldsymbol{\beta}}'\mathbf{B}_T\widetilde{\boldsymbol{\beta}}$. The numerical problem to be solved is to find the fixed point of $\widetilde{\boldsymbol{\theta}} = \mathbf{f}\left(\widetilde{\boldsymbol{\theta}}\right)$. This can be achieved by starting with an initial estimate of $\boldsymbol{\theta}$, say $\widetilde{\boldsymbol{\theta}}^{(0)}$, and using (22.120) to compute a new estimate of $\boldsymbol{\theta}$, namely $\widetilde{\boldsymbol{\theta}}^{(1)} = \mathbf{f}\left(\widetilde{\boldsymbol{\theta}}^{(0)}\right)$, and so on until convergence.

The *second procedure* (which we shall refer to as the generalized Newton-Raphson procedure) makes use of both the first and second derivatives of the concentrated log-likelihood function to solve numerically for $\widetilde{\boldsymbol{\theta}}$.

Let $\tilde{\boldsymbol{\theta}}^{(0)}$ and $\tilde{\boldsymbol{\lambda}}^{(0)}$ be the initial estimates of the ML estimators of $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$. Taking the Taylor series expansion of (22.113) around $\widetilde{\boldsymbol{\theta}}^{(0)}$ and $\widetilde{\boldsymbol{\lambda}}^{(0)}$, and using (22.114) we obtain

$$\begin{bmatrix} \mathbf{G}_n(\widetilde{\boldsymbol{\theta}}^{(0)}) & \mathbf{R}' \\ \mathbf{R} & 0 \end{bmatrix} \begin{bmatrix} n(\widetilde{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}^{(0)}) \\ \widetilde{\boldsymbol{\lambda}} - \widetilde{\boldsymbol{\lambda}}^{(0)} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_n(\widetilde{\boldsymbol{\theta}}^{(0)}) - \mathbf{R}'\widetilde{\boldsymbol{\lambda}}^{(0)} \\ -n(\mathbf{R}\widetilde{\boldsymbol{\theta}}^{(0)} - \mathbf{b}) \end{bmatrix} + o_p(1) \tag{22.123}$$

where $\mathbf{G}_n(\widetilde{\boldsymbol{\theta}})$ is given by

$$\begin{aligned} \mathbf{G}_n(\widetilde{\boldsymbol{\theta}}) \quad = \quad & -n^{-1}(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n\widetilde{\boldsymbol{\beta}})^{-1} \otimes [(\mathbf{A}_n\widetilde{\boldsymbol{\beta}})(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n\widetilde{\boldsymbol{\beta}})^{-1}(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n)] \\ & +n^{-1}(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n\widetilde{\boldsymbol{\beta}})^{-1} \otimes [(\mathbf{B}_n\widetilde{\boldsymbol{\beta}})(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n\widetilde{\boldsymbol{\beta}})^{-1}(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n)] \\ & -n^{-1}\mathbf{C}_{rm}\{[(\mathbf{A}_n\widetilde{\boldsymbol{\beta}})(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n\widetilde{\boldsymbol{\beta}})^{-1}] \otimes [(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n\widetilde{\boldsymbol{\beta}})^{-1}(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n)] \\ & +n^{-1}\mathbf{C}_{rm}\{[(\mathbf{B}_n\widetilde{\boldsymbol{\beta}})(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n\widetilde{\boldsymbol{\beta}})^{-1}] \otimes [(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n\widetilde{\boldsymbol{\beta}})^{-1}(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n)] \\ & +n^{-1}(\widetilde{\boldsymbol{\beta}}'\mathbf{A}_n\widetilde{\boldsymbol{\beta}})^{-1} \otimes \mathbf{A}_n - n^{-1}(\widetilde{\boldsymbol{\beta}}'\mathbf{B}_n\widetilde{\boldsymbol{\beta}})^{-1} \otimes \mathbf{B}_n \end{aligned}$$

[19]The computations are carried out assuming that r, the number of cointegrating vectors, is known, and hence $\boldsymbol{\beta}'\mathbf{A}_T\boldsymbol{\beta}$ and $\boldsymbol{\beta}'\mathbf{B}_T\boldsymbol{\beta}$ are of full rank. Notice that the data matrices $\mathbf{A}_T$ and $\mathbf{B}_T$ are assumed to be non-singular.

and $\mathbf{C}_{rm}$ is the $rm \times rm$ commutation matrix (see section 3.7 in Magnus and Neudecker (1988)). To deal with the singularity of $\mathbf{G}_n(\boldsymbol{\theta})$, partition $\mathbf{R} = (\mathbf{R}'_A, \mathbf{R}_B)'$, where $\mathbf{R}_A$ and $\mathbf{R}_B$ are matrices of order $r^2 \times rm$, and $(k - r^2) \times rm$; representing the r^2 just-identifying restrictions, and the $(k - r^2)$, over-identifying restrictions, respectively, and let $\mathbf{J}_n(\widetilde{\boldsymbol{\theta}}) = \mathbf{G}_n(\widetilde{\boldsymbol{\theta}}) + \mathbf{R}'_A\mathbf{R}_A$. Then, the solution of (22.123) using a generalized inverse based on $\mathbf{J}_n(\widetilde{\boldsymbol{\theta}})$ is given by[20]

$$\begin{bmatrix} n(\widetilde{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}^{(0)}) \\ \widetilde{\boldsymbol{\lambda}} - \widetilde{\boldsymbol{\lambda}}^{(0)} \end{bmatrix} = \begin{bmatrix} \mathbf{V}_{\theta\theta}(\widetilde{\boldsymbol{\theta}}^{(0)}) & \mathbf{V}_{\theta\lambda}(\widetilde{\boldsymbol{\theta}}^{(0)}) \\ \mathbf{V}'_{\theta\lambda}(\widetilde{\boldsymbol{\theta}}^{(0)}) & \mathbf{V}_{\lambda\lambda}(\widetilde{\boldsymbol{\theta}}^{(0)}) \end{bmatrix} \begin{bmatrix} \mathbf{d}_n(\widetilde{\boldsymbol{\theta}}^{(0)}) - \mathbf{R}'\widetilde{\boldsymbol{\lambda}}^{(0)} \\ -n(\mathbf{R}\widetilde{\boldsymbol{\theta}}^{(0)} - \mathbf{b}) \end{bmatrix} + o_p(1) \tag{22.124}$$

where

$$\begin{aligned} \mathbf{V}_{\theta\theta}(\widetilde{\boldsymbol{\theta}}) &= \mathbf{J}_n^{-1}(\widetilde{\boldsymbol{\theta}}) - \mathbf{J}_n^{-1}(\widetilde{\boldsymbol{\theta}})\mathbf{R}'(\mathbf{R}\mathbf{J}_n^{-1}(\widetilde{\boldsymbol{\theta}})\mathbf{R}')^{-1}\mathbf{R}\mathbf{J}_n^{-1}(\widetilde{\boldsymbol{\theta}}) \\ \mathbf{V}_{\theta\lambda}(\widetilde{\boldsymbol{\theta}}) &= \mathbf{J}_n^{-1}(\widetilde{\boldsymbol{\theta}})\mathbf{R}'(\mathbf{R}\mathbf{J}_n^{-1}(\widetilde{\boldsymbol{\theta}})\mathbf{R}')^{-1} \\ \mathbf{V}_{\lambda\lambda}(\widetilde{\boldsymbol{\theta}}) &= (\mathbf{R}\mathbf{J}_n^{-1}(\widetilde{\boldsymbol{\theta}})\mathbf{R}')^{-1} \end{aligned} \tag{22.125}$$

Hence, we obtain the following generalized version of the Newton-Raphson algorithm:

$$\begin{bmatrix} \widetilde{\boldsymbol{\theta}}^{(i)} \\ \widetilde{\boldsymbol{\lambda}}^{(i)} \end{bmatrix} = \begin{bmatrix} \widetilde{\boldsymbol{\theta}}^{(i-1)} \\ \widetilde{\boldsymbol{\lambda}}^{(i-1)} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{\theta\theta}(\widetilde{\boldsymbol{\theta}}^{(i-1)}) & \mathbf{V}_{\theta\lambda}(\widetilde{\boldsymbol{\theta}}^{(i-1)}) \\ \mathbf{V}'_{\theta\lambda}(\widetilde{\boldsymbol{\theta}}^{(i-1)}) & \mathbf{V}_{\lambda\lambda}(\widetilde{\boldsymbol{\theta}}^{(i-1)}) \end{bmatrix} \begin{bmatrix} \mathbf{n}^{-1}\{\mathbf{d}_n(\widetilde{\boldsymbol{\theta}}^{(i-1)}) - \mathbf{R}'\widetilde{\boldsymbol{\lambda}}^{(i-1)}\} \\ -n(\mathbf{R}\widetilde{\boldsymbol{\theta}}^{(i-1)} - \mathbf{b}) \end{bmatrix} \tag{22.126}$$

For the initial estimates, $\widetilde{\boldsymbol{\theta}}^{(0)}$, we use the linearized exactly identified estimators given by (22.112), and for $\widetilde{\boldsymbol{\lambda}}^{(0)}$ we start from zero. Our experience with using this algorithm in a number of applications suggests that the generalized Newton-Raphson algorithm based on (22.109) has good convergence properties, and converges reasonably fast. Finally, a consistent estimator of the asymptotic variance of $\widetilde{\boldsymbol{\theta}}$ is given by (22.125).

22.9.3 Log-likelihood ratio statistics for tests of over-identifying restrictions on the cointegrating relations

Consider now the problem of testing over-identifying restrictions on the coefficients of the cointegrating (or long-run) relations. Suppose there are r cointegrating relations and the interest is to test the restrictions

$$\mathbf{R}\,\mathrm{vec}(\boldsymbol{\beta}) = \mathbf{b} \tag{22.127}$$

where $\mathbf{R}$ is a $k \times mr$ matrix, and b is a $k \times 1$ vector of known constants such that $\mathrm{Rank}(\mathbf{R}) = k > r^2$. As before, let $\boldsymbol{\theta} = \mathrm{vec}(\boldsymbol{\beta})$ and decompose the k restriction defined by (22.127) into r^2 and $k - r^2$ set of restrictions

$$\underset{r^2 \times rm}{\mathbf{R}_A}\ \underset{rm \times 1}{\boldsymbol{\theta}} = \underset{r^2 \times 1}{\mathbf{b}_A} \tag{22.128}$$

[20] In general, the Newton-Raphson algorithm gives the same solution when $\mathbf{R}$ instead of $\mathbf{R}_A$ is used in construction of $\mathbf{J}_n(\widetilde{\boldsymbol{\theta}})$. See Pesaran and Shin (2002) for more details.

$$\underset{(k-r^2)\times rm}{\mathbf{R}_B}\;\underset{rm\times 1}{\boldsymbol{\theta}} = \underset{(k-r^2)\times 1}{\mathbf{b}_B} \tag{22.129}$$

where $\mathbf{R} = (\mathbf{R}_A', \mathbf{R}_B')'$, and $\mathbf{b} = (\mathbf{b}_A, \mathbf{b}_B)$, such that $\text{Rank}(\mathbf{R}_A) = r^2$, $\text{Rank}(\mathbf{R}_B) = k - r^2$, and $\mathbf{b}_A \neq 0$. Without loss of generality, the restriction, characterized by (22.128), can be viewed as the just-identifying restrictions, and the remaining restriction defined by (22.129) will then constitute the $k - r^2$ over-identifying restrictions. Let $\widehat{\boldsymbol{\theta}}$ be the ML estimators of $\boldsymbol{\theta}$ obtained subject to the r^2 exactly-identifying restrictions, and $\widetilde{\boldsymbol{\theta}}$ be the ML estimators of $\boldsymbol{\theta}$ obtained under all the k restriction in (22.127). Then, the log-likelihood ratio statistic for testing the over-identifying restrictions is simply given by

$$\mathfrak{LR}\left(\mathbf{R}\,|\mathbf{R}_A\right) = 2\left\{\ell_n^c\left(\widehat{\boldsymbol{\theta}};r\right) - \ell_n^c\left(\widetilde{\boldsymbol{\theta}};r\right)\right\} \tag{22.130}$$

where $\ell_n^c\left(\widehat{\boldsymbol{\theta}};r\right)$ is given by (22.97) and represents the maximized value of the log-likelihood function under the just-identifying restriction, (say $\mathbf{R}_A\boldsymbol{\theta} = \mathbf{b}_A$), and $\ell_n^c\left(\widetilde{\boldsymbol{\theta}};r\right)$ is the maximized value of the log-likelihood function under the k just- and over-identifying restrictions given by (22.127).

Under the null hypothesis that the restrictions (22.127) hold, the log-likelihood ratio statistic $(\mathfrak{LR})(\mathbf{R}\,|\mathbf{R}_A)$ defined by (22.130) is asymptotically distributed as a χ^2 variate with degrees of freedom equal to the number of the over-identifying restrictions, namely $k-r^2 > 0$.

The above testing procedure is also applicable when interest is on testing restrictions in a single cointegrating vector of a sub-set of cointegrating vectors. For this purpose, one simply needs to impose just-identifying restrictions on all the vectors except for the vector(s) that are to be subject to the over-identifying restrictions. The resultant test statistic will be invariant to the nature of the just-identifying restrictions. Notice that this test of the over-identifying restrictions on the cointegrating relations pre-assumes that the variables $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$, are $I(1)$, and that the number of cointegrating relations, r, is correctly chosen.

Another application of the above log-likelihood ratio procedure is to the problem of testing the 'co-trending' restriction (22.89), discussed in Section 22.7.1. The relevant test statistic is given by

$$\mathfrak{LR}\left(\boldsymbol{\beta}'\boldsymbol{\gamma} = \mathbf{0}\right) = 2\left\{\ell_n^c\left(\widehat{\boldsymbol{\theta}};r\right) - \ell_n^c\left(\widetilde{\boldsymbol{\theta}};r\right)\right\} \tag{22.131}$$

where, as before, $\ell_n^c\left(\widehat{\boldsymbol{\theta}};r\right)$ is the maximized value of the log-likelihood function when the cointegrating relations are just-identified, and $\ell_n^c\left(\widetilde{\boldsymbol{\theta}};r\right)$ is the maximized value of the log-likelihood function obtained subject to the just-identified restrictions plus the additional r co-trending restrictions, $\boldsymbol{\beta}'\boldsymbol{\gamma} = \mathbf{0}$. Under the co-trending null hypothesis, $\mathfrak{LR}\left(\boldsymbol{\beta}'\boldsymbol{\gamma} = \mathbf{0}\right)$ is asymptotically distributed as a χ^2 with r degrees of freedom.

22.9.4 Impulse response analysis in cointegrating VAR models

The impulse response analysis of the cointegrating model given by the equation systems (22.72) and (22.73) can be carried out along the lines set out in Section 22.5. In the present application it is important that the parametric restrictions implied by the deficiency in the rank of the long-run multiplier matrix, $\mathbf{\Pi}$, is taken into account. It is also important to note that due to the rank deficiency of the long-run multiplier matrix, shocks (whether equation-specific or system-wide shocks) will have persistence effects on the individual variables in the model, and their effects do not generally die out.

The computation of the impulse response function for the cointegrating VAR model can be based on the $VECM$ (22.76), which combines the equation systems for $\mathbf{y}_t$ and $\mathbf{x}_t$ given by (22.72) and (22.73), respectively. The solution of the combined model is given by (22.77), and the orthogonalized impulse response function of the effect of a unit shock to the ith variable at time t in (22.76) on the jth variable at time $t+N$ is given by

$$OI_{ij,N} = \mathbf{e}_j' \left(\mathbf{C}(1) + \mathbf{C}_N^*\right) \mathbf{T}\mathbf{e}_i \tag{22.132}$$

where, as before, $\mathbf{T}$ is a lower triangular matrix such that $\mathbf{\Sigma} = \mathbf{T}\mathbf{T}'$, $\mathbf{e}_i$ is the selection vector defined by (22.55), and $\mathbf{C}(1)$ and $\mathbf{C}_N^*$ are defined by relations (22.82)-(22.84). Alternatively, let

$$\mathbf{A}_i = \mathbf{C}(1) + \mathbf{C}_i^* \tag{22.133}$$

Then substituting $\mathbf{C}_i^* = \mathbf{A}_i - \mathbf{C}(1)$ in (22.83) and using (22.84) it also follows that

$$\mathbf{A}_i = \mathbf{A}_{i-1}\mathbf{\Phi}_i + \cdots + \mathbf{A}_{i-p}\mathbf{\Phi}_p, \qquad i = 1, 2, ... \tag{22.134}$$

where $\mathbf{A}_0 = \mathbf{I}_m$, and $\mathbf{A}_i = \mathbf{0}$, for $i < 0$.[21] However, from (22.133) it is clear that

$$\lim_{i\to\infty} \mathbf{A}_i = \mathbf{C}(1) \tag{22.135}$$

which is a non-zero matrix with rank $m - r$.[22] Therefore, the orthogonalized impulse responses for the cointegrating VAR model can be computed in exactly the same way as in the case of stationary VAR models; the main difference being that the matrices $\mathbf{A}_i$, in the moving-average representation of the $\mathbf{z}_t$-process tend to zero when the underlying VAR model is trend-stationary, and tends to a non-zero rank deficient matrix $\mathbf{C}(1)$ when the underlying VAR model is first-difference stationary.[23]

The generalized impulse response function, the orthogonalized and the generalized forecast error variance decomposition can also be computed for the cointegrating VAR models, along similar lines as in Sections 22.5.2 and 22.6.

22.9.5 Impulse response functions of cointegrating relations

We saw in the previous section that effects of shocks on individual variables in a cointegrating VAR model do not die out and persist for ever! An alternative approach would be to consider the effect of system-wide shocks or variable-specific shocks on the cointegrating relations, $\boldsymbol{\beta}'\mathbf{z}_t$, rather than on the individual variables in the model. The effect of shocks on

[21] For an alternative derivation of this result see Appendix A in Pesaran and Shin (1996). Also, note that the recursive relations defined by (22.134) and (22.51) produce the same results.

[22] The matrices $\mathbf{C}_n^*$, $n = 0, 1, 2, ...$ belong to the stationary component of $\mathbf{z}_t$ and tend to zero as $n \to \infty$.

[23] See Lütkepohl and Reimers (1992) and Mellander, Vredin, and Warne (1992) for more details and a derivation of the asymptotic distribution of the estimators of the orthogonalized impulse responses.

cointegrating relations is bound to die out, and their time profile contains useful information on the speed of convergence of the model to its cointegrating (or 'equilibrium') relations. See Lee and Pesaran (1993) and Pesaran and Shin (1996).

Consider first the time profile of the effect of a unit shock to the variable in $\mathbf{z}_t$ on the jth cointegrating relation, namely $\boldsymbol{\beta}_i'\mathbf{z}_t$. Once again we can obtain such a time profile both using Sims' orthogonalization method or the generalized impulse response approach. Using (22.86), we have:

$$OI_i\left(\boldsymbol{\beta}_j'\mathbf{z}_t, N\right) = \boldsymbol{\beta}_j'\mathbf{A}_N\mathbf{T}\mathbf{e}_i \tag{22.136}$$

for $i = 1, 2, ..., m$, $j = 1, 2, ..., r$, $N = 0, 1, 2, ...$ which give the responses of a unit change in the ith orthogonalized shock (equal to $\sqrt{\sigma_{ii}}$) on the jth cointegrating relation $\boldsymbol{\beta}_j'\mathbf{z}_t$.[24] The corresponding generalized impulse responses are given by

$$GI_i\left(\boldsymbol{\beta}_j'\mathbf{z}_t, N\right) = \frac{\boldsymbol{\beta}_j'\mathbf{A}_N\boldsymbol{\Sigma}\mathbf{e}_i}{\sqrt{\sigma_{ii}}} \tag{22.137}$$

for $i = 1, 2, ..., m$, $j = 1, 2, ..., r$, and $N = 0, 1, 2, ...$ Once again the two impulse response functions coincide either if $\boldsymbol{\Sigma}$ is diagonal, or if $i = 1$.

22.9.6 Persistence profiles for cointegrating relations and speed of convergence to equilibrium

Given the ambiguities that surround the impulse response analysis with respect to variable-specific shocks, it is of some interest to consider the effect of system-wide shocks on cointegrating relations. Such a time profile, referred to as the persistence profile, has been proposed in Pesaran and Shin (1996). The (scaled) persistence profile of the effect of system-wide shocks on the jth cointegrating relationship is given by

$$h\left(\boldsymbol{\beta}_j'\mathbf{z}_t, N\right) = \frac{\boldsymbol{\beta}_j'\mathbf{A}_n\boldsymbol{\Sigma}\mathbf{A}_n'\boldsymbol{\beta}_j}{\boldsymbol{\beta}_j'\boldsymbol{\Sigma}\boldsymbol{\beta}_j} \tag{22.138}$$

for $j = 1, 2, ..., r$, and $N = 0, 1, 2, ...$ The value of this profile is equal to unity on impact, but should tend to zero as $N \to \infty$, if $\boldsymbol{\beta}_j$ is indeed a cointegrating vector. The persistence profile, $h\left(\boldsymbol{\beta}_j'\mathbf{z}_t, N\right)$, viewed as a function of N, provides important information on the speed with which the effect of system-wide shocks on the cointegrating relation, $\boldsymbol{\beta}_j'\mathbf{z}_t$, disappears, even though these shocks generally have lasting impacts on the individual variables in $\mathbf{z}_t$. This is a useful addition to the long-run structural modelling techniques advanced in *Microfit*, and provides the users with estimates of the speed with which the economy or the markets under consideration return to their equilibrium states.

The persistence profiles are also useful in the case of time-series that are close to being $I(1)$, or 'near integrated'. The persistence profiles of near integrated variables eventually converge to zero, but can be substantially different from zero for protracted periods.

[24] Notice that $\boldsymbol{\beta}_j'\mathbf{C}_N^* = \boldsymbol{\beta}_j'(\mathbf{A}_N - \mathbf{C}(1)) = \boldsymbol{\beta}_j'\mathbf{A}_N$. See (22.133) and recall that $\boldsymbol{\beta}_j'\mathbf{C}(1) = 0$.

22.10 VARX Models

In what follows, we provide a brief account of the econometric issues involved in the modelling approach advanced by Pesaran, Shin, and Smith (2000). We start by describing a general structural $VARX$ model, which allows for the possibility of distinguishing between endogenous and weakly exogenous variables. We then turn our attention to the analysis of cointegrating $VARX$ models, and discuss forecasting and impulse response analysis in this framework.

For further details see Harbo, Johansen, Nielsen, and Rahbek (1998), Pesaran, Shin, and Smith (2000), and Garratt, Lee, Pesaran, and Shin (2006).

22.10.1 The structural VARX model

The general structural $VARX$ model for an $m_y \times 1$ vector of endogenous variables $\mathbf{y}_t$, is given by

$$\mathbf{A}\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{B}_0\mathbf{x}_t + \mathbf{B}_1\mathbf{x}_{t-1} + \cdots + \mathbf{B}_q\mathbf{x}_{t-q} + \mathbf{D}\mathbf{d}_t + \boldsymbol{\varepsilon}_t \tag{22.139}$$

for $t = 1, 2, ..., T$, where $\mathbf{d}_t$ is a $s \times 1$ vector of deterministic variables (for example, intercept, trend and seasonal variables), $\mathbf{x}_t$ is an $m_x \times 1$ vector of exogenous variables, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{m_y t})'$ is an $m_y \times 1$ vector of serially uncorrelated errors distributed independently of $\mathbf{x}_t$ with a zero mean and a constant positive definite variance-covariance matrix, $\boldsymbol{\Omega} = (\omega_{ij})$, where ω_{ij} is the $(i, j)$the element of $\boldsymbol{\Omega}$.

For given values of $\mathbf{d}_t$ and $\mathbf{x}_t$, the above dynamic system is stable if all the roots of the determinantal equation

$$\left|\mathbf{A} - \mathbf{A}_1\lambda - \mathbf{A}_2\lambda^2 - \cdots - \mathbf{A}_p\lambda^p\right| = 0 \tag{22.140}$$

lie strictly outside the unit circle. This stability condition ensures the existence of long-run relationships between $\mathbf{y}_t$ and $\mathbf{x}_t$, which will be cointegrating when one or more elements of $\mathbf{x}_t$ contain unit roots.

Model (22.139) is structural in the sense that it explicitly allows for instantaneous interactions between the endogenous variables through the contemporaneous coefficient matrix, $\mathbf{A}$. It can also be written as

$$\mathbf{A}(L)\mathbf{y}_t = \mathbf{B}(L)\mathbf{x}_t + \mathbf{D}\mathbf{d}_t + \boldsymbol{\varepsilon}_t \tag{22.141}$$

where

$$\mathbf{A}(L) = \mathbf{A} - \mathbf{A}_1 L - \cdots - \mathbf{A}_p L^p;\ \mathbf{B}(L) = \mathbf{B}_0 + \mathbf{B}_1 L + \cdots + \mathbf{B}_q L^q$$

Of particular interest are the system long-run effects of the exogenous variables which are given by

$$\mathbf{A}(1)^{-1}\mathbf{B}(1) = \left(\mathbf{A} - \sum_{i=1}^{p}\mathbf{A}_i\right)^{-1}\sum_{i=0}^{q}\mathbf{B}_i$$

Notice that since all the roots of (22.140) fall outside the unit circle by assumption, the inverse of $\mathbf{A}(1)$, which we denote by $\mathbf{A}(1)^{-1}$, exists.

The decision to work with a model of the type (22.139) presents the applied econometrician with a number of important choices, namely:

1. The number and list of the endogenous variables to be included, $(m_y, \mathbf{y}_t)$
2. The number and list of the exogenous variables (if any) to be included, $(m_x, \mathbf{x}_t)$
3. The nature of the deterministic variables (intercepts, trends, seasonals) and whether the intercepts and/or the trend coefficients need to be restricted
4. The lag orders p and q in the $VARX$ (the lag order of the $\mathbf{y}_t$ and $\mathbf{x}_t$ components of the $VARX$ need not be the same)
5. The order of integration of the variables.

These choices change the maximised value of the log-likelihood so that, in principle, they could be made on the basis of hypothesis testing exercises or by means of model selection criteria such as the Akaike information criterion (AIC), or the Schwarz Bayesian criterion (SBC) (see Section 22.4.1).

22.10.2 The reduced form VARX model

The reduced form of the structural model (22.139), which expresses the endogenous variables in terms of the predetermined and exogenous variables, is given by

$$\mathbf{y}_t = \mathbf{\Phi}_1\mathbf{y}_{t-1} + \cdots + \mathbf{\Phi}_p\mathbf{y}_{t-p} + \mathbf{\Psi}_0\mathbf{x}_t + \mathbf{\Psi}_1\mathbf{x}_{t-1} + \cdots + \mathbf{\Psi}_p\mathbf{x}_{t-q} + \mathbf{\Upsilon}\mathbf{d}_t + \mathbf{u}_t \tag{22.142}$$

where $\mathbf{\Phi}_i = \mathbf{A}^{-1}\mathbf{A}_i$, $\mathbf{\Psi}_i = \mathbf{A}^{-1}\mathbf{B}_i$, $\mathbf{\Upsilon} = \mathbf{A}^{-1}\mathbf{D}$, $\mathbf{u}_t = \mathbf{A}^{-1}\boldsymbol{\varepsilon}_t$ is $IID(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = \mathbf{A}^{-1}\mathbf{\Omega}\mathbf{A}^{'-1} = (\sigma_{ij})$. The classical identification problem is how to recover the structural form parameters

$$\left(\mathbf{A}, \mathbf{A}_{i+1}, \mathbf{B}_j, i = 1, .., p; j = 0, 1, ..., q;\ \mathbf{D} \text{ and } \mathbf{\Omega}\right)$$

from the reduced form parameters

$$(\mathbf{\Phi}_i, \mathbf{\Psi}_i, i = 1, .., p; j = 0, 1, ..., q; \mathbf{\Upsilon}, \text{ and } \mathbf{\Sigma})$$

Exact identification of the structural parameters requires m_y^2 *a priori* restrictions, of which m_y restrictions would be provided by normalisation conditions.

22.10.3 The cointegrated VAR model with $I(1)$ exogenous variables

Let $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$, set $n = max\{p, q\}$, and consider the following *extended* vector error correction model ($VECM$) in $\mathbf{z}_t$:

$$\Delta\mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t - \mathbf{\Pi}\mathbf{z}_{t-1} + \sum_{i=1}^{n-1}\mathbf{\Gamma}_i\Delta\mathbf{z}_{t-i} + \mathbf{u}_t \tag{22.143}$$

where the matrices $\{\boldsymbol{\Gamma}_i\}_{i=1}^{n-1}$ are the short-run responses, and $\boldsymbol{\Pi}$ is the long-run multiplier matrix.

By partitioning the error term $\mathbf{u}_t$ conformably with $\mathbf{z}_t = (\mathbf{y}'_t, \mathbf{x}'_t)'$ as $\mathbf{u}_t = \left(\mathbf{u}'_{yt}, \mathbf{u}'_{xt}\right)'$ and its variance matrix as

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{pmatrix}$$

we are able to express $\mathbf{u}_{yt}$ conditionally in terms of $\mathbf{u}_{xt}$ as

$$\mathbf{u}_{yt} = \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\mathbf{u}_{xt} + \boldsymbol{\upsilon}_t \tag{22.144}$$

where $\boldsymbol{\upsilon}_t \sim iid\,(\mathbf{0}, \boldsymbol{\Sigma}_{\upsilon\upsilon})$, $\boldsymbol{\Sigma}_{\upsilon\upsilon} \equiv \boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xy}$ and $\boldsymbol{\upsilon}_t$ is uncorrelated with $\mathbf{u}_{xt}$ by construction. Substitution of (22.144) into (22.143) together with a similar partitioning of the parameter vectors and matrices $\mathbf{a}_0 = \left(\mathbf{a}'_{y0}, \mathbf{a}'_{x0}\right)'$, $\mathbf{a}_1 = \left(\mathbf{a}'_{y1}, \mathbf{a}'_{x1}\right)'$, $\boldsymbol{\Pi} = \left(\boldsymbol{\Pi}'_y, \boldsymbol{\Pi}'_x\right)'$, $\boldsymbol{\Gamma}_i = \left(\boldsymbol{\Gamma}'_{yi}, \boldsymbol{\Gamma}'_{xi}\right)'$, $i = 1, ..., n-1$, provides a conditional model for $\Delta\mathbf{y}_t$ in terms of $\mathbf{z}_{t-1}$, $\Delta\mathbf{x}_t$, $\Delta\mathbf{z}_{t-1}$, $\Delta\mathbf{z}_{t-2}, ...$

$$\Delta\mathbf{y}_t = \mathbf{c}_0 + \mathbf{c}_1 t - \boldsymbol{\Pi}_{yy.x}\mathbf{z}_{t-1} + \boldsymbol{\Lambda}\Delta\mathbf{x}_t + \sum_{i=1}^{n-1} \boldsymbol{\Psi}_i\Delta\mathbf{z}_{t-i} + \boldsymbol{\upsilon}_t \tag{22.145}$$

where $\boldsymbol{\Pi}_{yy.x} \equiv \boldsymbol{\Pi}_y - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Pi}_x$, $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}$, $\boldsymbol{\Psi}_i \equiv \boldsymbol{\Gamma}_{yi} - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Gamma}_{xi}$, $i = 1, ..., n-1$, $\mathbf{c}_0 \equiv \mathbf{a}_{y0} - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\mathbf{a}_{x0}$ and $\mathbf{c}_1 \equiv \mathbf{a}_{y1} - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\mathbf{a}_{x1}$,

Following Johansen (1995), we assume that the process $\{\mathbf{x}_t\}_{t=1}^{\infty}$ is *weakly exogenous* with respect to the matrix of long-run multiplier parameters $\boldsymbol{\Pi}$, namely

$$\boldsymbol{\Pi}_x = \mathbf{0} \tag{22.146}$$

Therefore,

$$\boldsymbol{\Pi}_{yy.x} = \boldsymbol{\Pi}_y \tag{22.147}$$

Consequently, from (22.143) and (22.145) we obtain the following system of equations:

$$\Delta\mathbf{y}_t = -\boldsymbol{\Pi}_y\mathbf{z}_{t-1} + \boldsymbol{\Lambda}\Delta\mathbf{x}_t + \sum_{i=1}^{n-1} \boldsymbol{\Psi}_i\Delta\mathbf{z}_{t-i} + \mathbf{c}_0 + \mathbf{c}_1 t + \boldsymbol{\upsilon}_t \tag{22.148}$$

$$\Delta\mathbf{x}_t = \sum_{i=1}^{n-1} \boldsymbol{\Gamma}_{xi}\Delta\mathbf{z}_{t-i} + \mathbf{a}_{x0} + \mathbf{u}_{xt} \tag{22.149}$$

where now the restrictions on trend coefficients are modified to

$$\mathbf{c}_1 = \boldsymbol{\Pi}_y\boldsymbol{\gamma} \tag{22.150}$$

The restriction $\boldsymbol{\Pi}_x = \mathbf{0}$ in (22.146) implies that the elements of the vector process $\{\mathbf{x}_t\}_{t=1}^{\infty}$ are not cointegrated among themselves. Moreover, the information available from the differenced $VAR(p-1)$ model (22.149) for $\{\mathbf{x}_t\}_{t=1}^{\infty}$ is redundant for efficient conditional estimation and inference concerning the long-run parameters $\boldsymbol{\Pi}_y$ as well as the deterministic and short-run parameters $\mathbf{c}_0$, $\mathbf{c}_1$, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}_i$, $i = 1, ..., p-1$, of (22.148). $\{\mathbf{x}_t\}_{t=1}^{\infty}$ may be regarded as *long-run forcing* for $\{\mathbf{y}_t\}_{t=1}^{\infty}$. Note that this restriction does not preclude $\{\mathbf{y}_t\}_{t=1}^{\infty}$ being *Granger-causal* for $\{\mathbf{x}_t\}_{t=1}^{\infty}$ in the *short run*.

When there are r cointegrating relations among $\mathbf{z}_t$, we may express

$$\mathbf{\Pi}_y = \boldsymbol{\alpha}_y \boldsymbol{\beta}' \tag{22.151}$$

where $\boldsymbol{\alpha}_y$ $(m_y \times r)$ and $\boldsymbol{\beta}$ $(m \times r)$ are matrices of error correction coefficients and of the long-run (or cointegrating) coefficients, both of which are of full column rank, r. For the purpose of empirical analysis, we assume that the lag order p is large enough so that $\mathbf{u}_t$ and $\boldsymbol{\upsilon}_t$ are serially uncorrelated, and have zero mean and positive definite covariance matrices, $\mathbf{\Sigma}$ and $\mathbf{\Sigma}_{\upsilon\upsilon}$, respectively. For the purpose of the ML estimation, we also assume that $\mathbf{u}_t$ and $\boldsymbol{\upsilon}_t$ are normally distributed, although this is not binding if the number of the time-series observations available is large enough.

The analysis of a cointegrated VAR model containing exogenous $I(1)$ variables follows similar lines to that described in Section 22.7 above. Again, to avoid the unsatisfactory possibility that there exist quadratic trends in the level solution of the data generating process for $\mathbf{z}_t$ when there is no cointegration, we can assume that there are restrictions on the intercepts and/or time trends, as in the following five cases:

Case I ($\mathbf{c}_0 = \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{0}$). The structural $VECM$ (22.148) becomes

$$\Delta \mathbf{y}_t = -\mathbf{\Pi}_y \mathbf{z}_{t-1} + \mathbf{\Lambda} \Delta \mathbf{x}_t + \sum_{i=1}^{n-1} \mathbf{\Psi}_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\upsilon}_t \tag{22.152}$$

Case II ($\mathbf{c}_0 = \mathbf{\Pi}_y \boldsymbol{\mu}$ and $\mathbf{c}_1 = \mathbf{0}$). The structural $VECM$ (22.148) is

$$\Delta \mathbf{y}_t = \mathbf{\Pi}_y \boldsymbol{\mu} - \mathbf{\Pi}_y \mathbf{z}_{t-1} + \mathbf{\Lambda} \Delta \mathbf{x}_t + \sum_{i=1}^{n-1} \mathbf{\Psi}_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\upsilon}_t \tag{22.153}$$

Case III ($\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{0}$). In this case, the intercept restriction $\mathbf{c}_0 = \mathbf{\Pi}_y \boldsymbol{\mu}$ is ignored, and the structural $VECM$ estimated is

$$\Delta \mathbf{y}_t = \mathbf{c}_0 - \mathbf{\Pi}_y \mathbf{z}_{t-1} + \mathbf{\Lambda} \Delta \mathbf{x}_t + \sum_{i=1}^{n-1} \mathbf{\Psi}_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\upsilon}_t \tag{22.154}$$

Case IV ($\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 = \mathbf{\Pi}_y \boldsymbol{\gamma}$). Thus,

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + (\mathbf{\Pi}_y \boldsymbol{\gamma})\, t - \mathbf{\Pi}_y \mathbf{z}_{t-1} + \mathbf{\Lambda} \Delta \mathbf{x}_t + \sum_{i=1}^{n-1} \mathbf{\Psi}_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\upsilon}_t. \tag{22.155}$$

Case V ($\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 \neq \mathbf{0}$). Here, the deterministic trend restriction $\mathbf{c}_1 = \mathbf{\Pi}_y \boldsymbol{\gamma}$ is ignored and the structural $VECM$ estimated is

$$\Delta \mathbf{y}_t = \mathbf{c}_0 + \mathbf{c}_1 t - \mathbf{\Pi}_y \mathbf{z}_{t-1} + \mathbf{\Lambda} \Delta \mathbf{x}_t + \sum_{i=1}^{n-1} \mathbf{\Psi}_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\upsilon}_t \tag{22.156}$$

Tests of the cointegrating rank are obtained along exactly the same lines as those in Sections 22.8.1 and 22.8.2.[25] Estimation of the $VECM$ subject to exact- and over-identifying long-run restrictions can be carried out by maximum likelihood methods as in Section 22.8, applied to (22.148) subject to the appropriate restrictions on the intercepts and trends, subject to $\text{Rank}(\mathbf{\Pi}_y) = r$, and subject to k general linear restrictions. Having computed ML estimates of the cointegrating vectors, estimation of the short-run parameters of the conditional $VECM$ can be computed by OLS regressions.

The investigation of the dynamic properties of the system including exogenous $I(1)$ variables does require a little care, however. For this, we require the full-system $VECM$, obtained by augmenting the conditional model for $\Delta\mathbf{y}_t$, (22.148), with the marginal model for $\Delta\mathbf{x}_t$, (22.149). This is written as

$$\Delta\mathbf{z}_t = -\boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{z}_{t-1} + \sum_{i=1}^{n-1}\mathbf{\Gamma}_i\Delta\mathbf{z}_{t-i} + \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{H}\boldsymbol{\zeta}_t \tag{22.157}$$

where $\boldsymbol{\beta}$ is defined by (22.151),

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\alpha}_y \\ \mathbf{0} \end{pmatrix},\ \mathbf{\Gamma}_i = \begin{pmatrix} \mathbf{\Psi}_i + \mathbf{\Lambda}\mathbf{\Gamma}_{xi} \\ \mathbf{\Gamma}_{xi} \end{pmatrix},\ \mathbf{a}_0 = \begin{pmatrix} \mathbf{c}_0 + \mathbf{\Lambda}\mathbf{a}_{x0} \\ \mathbf{a}_{x0} \end{pmatrix},\ \mathbf{a}_1 = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{0} \end{pmatrix}, \tag{22.158}$$

$$\boldsymbol{\zeta}_t = \begin{pmatrix} \boldsymbol{\upsilon}_t \\ \mathbf{u}_{xt} \end{pmatrix},\ \mathbf{H} = \begin{pmatrix} \mathbf{I}_{m_y} & \mathbf{\Lambda} \\ \mathbf{0} & \mathbf{I}_{m_x} \end{pmatrix},\ Cov(\boldsymbol{\zeta}_t) = \mathbf{\Sigma}_{\zeta\zeta} = \begin{pmatrix} \mathbf{\Sigma}_{\upsilon\upsilon} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{xx} \end{pmatrix}. \tag{22.159}$$

While estimation and inference on the parameters of (22.148) can be conducted without reference to the marginal model (22.149), for forecasting and impulse response analysis the processes driving the weakly exogenous variables must be specified. In other words, one needs to take into account the possibility that changes in one variable may have an impact on the weakly exogenous variables, and that these effects will continue and interact over time. Hence, impulse response analysis can be conducted following the lines of the arguments set out in Sections 22.9.4-22.9.6, but *applied to the full system in* (22.157) (for more details see Section 22.10.4).

This last point is worth emphasising, and applies to any analysis involving counterfactuals, including impulse response analysis and forecasting exercises. Macro-modellers frequently consider the dynamic response of a system to a change in an exogenous variable by considering the effects of a once-and-for-all increase in the variable. This (implicitly) imposes restrictions on the processes generating the exogenous variable, assuming that there is no serial correlation in the variable, and that a shock to one exogenous variable can be considered without having to take into account changes in other exogenous variables. These counterfactual exercises might be of interest. But, generally speaking, one needs to take into account the possibility that changes in one exogenous variable will have an impact on other exogenous variables, and that these effects might continue and interact over time. This

[25] Asymptotic distributions of the trace and maximum eigenvalue statistics are again non-standard, and depend on whether the intercepts and/or the coefficients on the deterministic trends are restricted or unrestricted.

requires an explicit analysis of the dynamic processes driving the exogenous variables, as captured by the marginal model in (22.149). The whole point of the approach to investigating model dynamics reflected in the model of (22.157) and incorporated in the idea of generalised impulse response analysis is to explicitly allow for the conditional correlation structure in errors and the interactions between endogenous and exogenous variables to provide a 'realistic' counterfactual exercise based on the contemporaneous covariances and interactions between all the variables ($\mathbf{y}_t$ and $\mathbf{x}_t$) observed historically in the data.

22.10.4 Forecasting and impulse response analysis in VARX models

Consider the full system (22.157), and rewrite it as follows:

$$\mathbf{z}_t = \sum_{i=1}^{n} \mathbf{\Phi}_i \mathbf{z}_{t-i} + \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{H}\boldsymbol{\zeta}_t \tag{22.160}$$

where $\mathbf{\Phi}_1 = \mathbf{I}_m - \boldsymbol{\alpha}\boldsymbol{\beta}' + \mathbf{\Gamma}_1$, $\mathbf{\Phi}_i = \mathbf{\Gamma}_i - \mathbf{\Gamma}_{i-1}$, $i = 2, ..., n-1$, $\mathbf{\Phi}_n = -\mathbf{\Gamma}_{n-1}$.

Equation (22.160) can be used for forecasting purposes and in impulse response analysis.

The generalized impulse responses are derived from the moving-average representation of equation (22.160). Consider

$$\Delta \mathbf{z}_t = \mathbf{C}\left(L\right)\left(a_0 + a_1 t + \mathbf{H}\boldsymbol{\zeta}_t\right)$$

where

$$\mathbf{C}(L) = \sum_{j=0}^{\infty} \mathbf{C}_j L^j = \mathbf{C}(1) + (1-L)\mathbf{C}^*(L)$$

$$\mathbf{C}^*(L) = \sum_{j=0}^{\infty} \mathbf{C}_j^* L^j, \text{ and } \mathbf{C}_j^* = -\sum_{i=j+1}^{\infty} \mathbf{C}_i$$

$$\mathbf{C}_i = \mathbf{\Phi}_1 \mathbf{C}_{i-1} + \mathbf{\Phi}_2 \mathbf{C}_{i-2} + ... + \mathbf{\Phi}_p \mathbf{C}_{i-p}, \text{ for } i = 2, 3, ... \tag{22.161}$$

and $\mathbf{C}_0 = \mathbf{I}_m$, $\mathbf{C}_1 = \mathbf{\Phi}_1 - \mathbf{I}_m$ and $\mathbf{C}_i = 0$, for $i < 0$. Cumulating forward one obtains the level moving average representation,

$$\mathbf{z}_t = \mathbf{z}_0 + \mathbf{b}_0 t + \mathbf{C}(1)\sum_{j=1}^{t} \mathbf{H}\boldsymbol{\zeta}_j + \mathbf{C}^*(L)\mathbf{H}(\boldsymbol{\zeta}_t - \boldsymbol{\zeta}_0)$$

where $\mathbf{b}_0 = \mathbf{C}(1)\mathbf{a}_0 + \mathbf{C}^*(1)\mathbf{a}_1$ and $\mathbf{C}(1)\mathbf{\Pi}\boldsymbol{\gamma} = 0$, with $\boldsymbol{\gamma}$ being an arbitrary $m \times 1$ vector of fixed constants. The latter relation applies because the trend coefficients are restricted to lie in the cointegrating space.

The generalized and orthogonalized impulse response functions of individual variables $\mathbf{z}_{t+N} = (\mathbf{y}'_{t+N}, \mathbf{x}'_{t+N})'$ at horizon N to a unit change in the error, ζ_{it}, measured by one standard deviation, $\sqrt{\sigma_{\zeta,ii}}$ are

$$GI\left(N, \mathbf{z} : \zeta_i\right) = \frac{1}{\sqrt{\sigma_{\zeta,ii}}} \tilde{\mathbf{C}}_n \mathbf{H} \mathbf{\Sigma}_{\zeta\zeta} \mathbf{e}_i, \; N = 0, 1, ..., \; i = 1, ..., m \tag{22.162}$$

$$OI\left(N, \mathbf{z} : \zeta_i^*\right) = \tilde{\mathbf{C}}_n \mathbf{H} \mathbf{P}_\zeta \mathbf{e}_i,\ N = 0, 1, ...,\ i = 1, ..., m \tag{22.163}$$

where $\boldsymbol{\zeta}_t$ is $iid\left(\mathbf{0}, \boldsymbol{\Sigma}_{\zeta\zeta}\right)$, ζ_i^* is an orthogonalized residual, $\sigma_{\zeta,ij}$ is $(i,j)^{th}$ element of $\boldsymbol{\Sigma}_{\zeta\zeta}$, $\tilde{\mathbf{C}}_n = \sum_{j=0}^h \mathbf{C}_j$, with $\mathbf{C}_j$'s given by the recursive relations (22.161), $\mathbf{H}$ and $\boldsymbol{\Sigma}_{\zeta\zeta}$ are given in (22.159), $\mathbf{e}_i$ is a selection vector of zeros with unity as its ith element, $\mathbf{P}_\zeta$ is a lower triangular matrix obtained by the Cholesky decomposition of $\boldsymbol{\Sigma}_{\zeta\zeta} = \mathbf{P}_\zeta \mathbf{P}_\zeta'$, and $m = m_x + m_y$. Also see Pesaran and Shin (1998).

Similarly, the generalized and orthogonalized impulse response functions for the cointegrating relations with respect to a unit change in the error, ζ_{it} are given by

$$GI\left(N, \boldsymbol{\xi} : \zeta_i\right) = \frac{1}{\sqrt{\sigma_{\zeta,ii}}} \boldsymbol{\beta}' \tilde{\mathbf{C}}_n \mathbf{H} \boldsymbol{\Sigma}_{\zeta\zeta} \mathbf{e}_i,\ N = 0, 1, ...,\ i = 1, ..., m \tag{22.164}$$

$$OI\left(N, \boldsymbol{\xi} : \zeta_i^*\right) = \boldsymbol{\beta}' \tilde{\mathbf{C}}_n \mathbf{H} \mathbf{P}_\zeta \mathbf{e}_i,\ N = 0, 1, ...,\ i = 1, ..., m \tag{22.165}$$

where $\boldsymbol{\xi}_t = \boldsymbol{\beta}' \mathbf{z}_{t-1}$.

While the impulse responses show the effect of a shock to a particular variable, the persistence profile, as developed by Lee and Pesaran (1993) and Pesaran and Shin (1996), show the effects of system-wide shocks on the cointegrating relations. In the case of the cointegrating relations the effects of the shocks (irrespective of their sources) will eventually disappear. Therefore, the shape of the persistence profiles provide valuable information on the speed of convergence of the cointegrating relations towards equilibrium. The persistence profile for a given cointegrating relation defined by the cointegrating vector $\boldsymbol{\beta}_j$, in the case of a $VARX$ model, is given by

$$h(\boldsymbol{\beta}_j' \mathbf{z}, n) = \frac{\boldsymbol{\beta}_j' \widetilde{\mathbf{C}}_n \mathbf{H} \boldsymbol{\Sigma}_{\zeta\zeta} \mathbf{H}' \widetilde{\mathbf{C}}_n' \boldsymbol{\beta}_j}{\boldsymbol{\beta}_j' \mathbf{H} \boldsymbol{\Sigma}_{\zeta\zeta} \mathbf{H}' \boldsymbol{\beta}_j},\ N = 0, 1, ...,\ j = 1, ...r$$

where $\boldsymbol{\beta}$, $\widetilde{\mathbf{C}}_n$, $\mathbf{H}$ and $\boldsymbol{\Sigma}_{\zeta\zeta}$ are as defined above.

22.11 Trend/cycle decomposition in VARs

The trend/cycle decomposition available in *Microfit* allows partitioning of a vector of random variables in the sum of a stationary process, called the transitory or cyclical component, and a permanent component, which may be further sub-divided into a deterministic (trend) and a stochastic part (Garratt, Lee, Pesaran, and Shin (2006), Mills (2003), Robertson, Garratt, and Wright (2006), Evans and Reichlin (1994)). Such decomposition can be considered as a modification of the classic multivariate Beveridge-Nelson decomposition (see Beveridge and Nelson (1981) and Engle and Granger (1987)), extended to include possible restrictions in the intercept and/or trend, as well as the existence of long-run relationships.

Consider a $m \times 1$ vector of random variables $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$. For any arbitrary partitioning of $\mathbf{z}_t$ into a permanent and cyclical component $\mathbf{z}_t^P$ and $\mathbf{z}_t^C$, the cyclical part, since stationary, must satisfy

$$\lim_{h \to \infty} E_t\left(\mathbf{z}_{t+h}^C\right) = \mathbf{0} \tag{22.166}$$

where $E_t(.)$ denotes the expectation operator conditional on the information at time t, taken to be $\{\mathbf{z}_t, \mathbf{z}_{t-1}, ..., \mathbf{z}_0\}$. Hence, denote

$$\mathbf{z}_{dt}^P = \mathbf{g}_0 + \mathbf{g}t$$

as the deterministic part of $\mathbf{z}_t$, where $\mathbf{g}_0$ is an $m \times 1$ vectors of fixed intercepts, $\mathbf{g}$ is an $m \times 1$ vector of (restricted) trend growth rates, and t is a deterministic trend term. From (22.166) it follows that (see Garratt, Robertson, and Wright (2005))

$$\mathbf{z}_{st}^P = \lim_{h \to \infty} E_t\left(\mathbf{z}_{t+h} - (\mathbf{g}_0 + \mathbf{g}h)\right)$$

The above result is at the basis of the trend/cycle decomposition of $\mathbf{z}_t$ available in *Microfit*. Suppose that $\mathbf{z}_t$ has the following vector error correction representation with unrestricted intercept and restricted trend (which in *Microfit* corresponds to option 4 from the cointegrating VAR option in the System Estimation Menu), then

$$\Delta \mathbf{z}_t = \mathbf{a} - \boldsymbol{\alpha}\boldsymbol{\beta}^{'}\left[\mathbf{z}_{t-1} - \boldsymbol{\gamma}(t-1)\right] + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{z}_{t-i} + \mathbf{u}_t \tag{22.167}$$

where $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$ is an $m \times 1$ vector of random variables. Denote the deviation of the variables in $\mathbf{z}_t$ from their deterministic components as $\widetilde{\mathbf{z}}_t$, namely

$$\widetilde{\mathbf{z}}_t = \mathbf{z}_t - \mathbf{g}_0 - \mathbf{g}t$$

Then, in terms of $\widetilde{\mathbf{z}}_t$ we have

$$\Delta \widetilde{\mathbf{z}}_t = \mathbf{a} - \boldsymbol{\alpha}\boldsymbol{\beta}^{'}\mathbf{g}_0 - \left(\mathbf{I}_m - \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i\right)\mathbf{g} - \boldsymbol{\alpha}\boldsymbol{\beta}'\left(\mathbf{g} - \boldsymbol{\gamma}\right)(t-1) - \boldsymbol{\alpha}\boldsymbol{\beta}^{'}\widetilde{\mathbf{z}}_t + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \widetilde{\mathbf{z}}_{t-i} + \mathbf{u}_t$$

Since $\widetilde{\mathbf{z}}_t$ has no deterministic components by construction, it must be that

$$\mathbf{a} = \boldsymbol{\alpha}\boldsymbol{\beta}^{'}\mathbf{g}_0 + \left(\mathbf{I}_m - \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i\right)\mathbf{g} \tag{22.168}$$

and

$$\boldsymbol{\beta}'\mathbf{g} = \boldsymbol{\beta}'\boldsymbol{\gamma} \tag{22.169}$$

Hence

$$\Delta \widetilde{\mathbf{z}}_t = -\boldsymbol{\alpha}\boldsymbol{\beta}'\widetilde{\mathbf{z}}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \widetilde{\mathbf{z}}_{t-i} + \mathbf{u}_t \tag{22.170}$$

or, equivalently,

$$\widetilde{\mathbf{z}}_t = \sum_{i=1}^{p} \boldsymbol{\Phi}_i \widetilde{\mathbf{z}}_{t-i} + \mathbf{u}_t \tag{22.171}$$

where

$$\mathbf{\Phi}_1 = \mathbf{I}_m + \mathbf{\Gamma}_1 - \boldsymbol{\alpha\beta}', \ \mathbf{\Phi}_i = \mathbf{\Gamma}_i - \mathbf{\Gamma}_{i-1}, \ i = 2, ..., p-1, \ \mathbf{\Phi}_p = -\mathbf{\Gamma}_{p-1}$$

If we then apply the classic multivariate Beveridge-Nelson decomposition to (22.171) (see Stock and Watson (1988) and Evans and Reichlin (1994)), we determine that $\mathbf{z}_t$ can be written as

$$\mathbf{z}_t = \mathbf{z}_0 + \mathbf{g}t + \mathbf{C}(1)\sum_{i=1}^{t}\mathbf{u}_i + \mathbf{C}^*(L)(\mathbf{u}_t - \mathbf{u}_0) \tag{22.172}$$

where

$$\mathbf{C}(1) = \sum_{i=0}^{\infty}\mathbf{C}_i$$

$$\mathbf{C}^*(L) = \sum_{i=0}^{\infty}\mathbf{C}_i^* L^i$$

with

$$\begin{aligned} \mathbf{C}_i &= \mathbf{C}_{i-1}\mathbf{\Phi}_1 + \mathbf{C}_{i-2}\mathbf{\Phi}_2 + \cdots + \mathbf{C}_{i-p}\mathbf{\Phi}_p \quad \text{for } i = 1, 2, ... \\ \mathbf{C}_i^* &= \mathbf{C}_{i-1}^* + \mathbf{C}_i \end{aligned}$$

and $\mathbf{C}_0 = \mathbf{I}_m$, $\mathbf{C}_1 = -(\mathbf{I}_m - \mathbf{\Phi}_1)$, $\mathbf{C}_i = \mathbf{0}$ for $i < 0$, $\mathbf{C}_0^* = \mathbf{C}_0 - \mathbf{C}(1)$.

Hence, the stochastic term and the cyclical component are defined respectively as

$$\mathbf{z}_{st}^P = \mathbf{C}(1)\sum_{i=1}^{t}\mathbf{u}_i \tag{22.173}$$

$$\mathbf{z}_t^C = \mathbf{C}^*(L)(\mathbf{u}_t - \mathbf{u}_0) \tag{22.174}$$

As for the estimation of the various components, note that $\mathbf{z}_{st}^P$ can be easily estimated, since the coefficients for $\mathbf{C}_i$ can be derived recursively in terms of $\mathbf{\Phi}_i$, which in turn can be obtained from the $\mathbf{\Gamma}_i$. Once $\mathbf{z}_{st}^P$ has been estimated, consider the difference

$$\hat{\mathbf{w}}_t = \mathbf{z}_t - \hat{\mathbf{C}}(1)\sum_{i=1}^{t}\hat{\mathbf{u}}_i$$

and notice that this is also equal to

$$\hat{\mathbf{w}}_t = \mathbf{z}_0 + \hat{\mathbf{g}}t + \hat{\mathbf{z}}_t^C$$

Hence, to obtain $\hat{\mathbf{g}}$ and $\hat{\mathbf{z}}_t^C$, one can perform a seemingly unrelated ($SURE$) regression of $\hat{\mathbf{w}}_t$ on an intercept and a time trend t, subject to the restrictions

$$\hat{\boldsymbol{\beta}}'\hat{\mathbf{g}} = \hat{\boldsymbol{\beta}}'\hat{\boldsymbol{\gamma}} \tag{22.175}$$

where $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\beta}}$ have already been estimated, under the assumption that the cointegrating vectors are exactly identified. Residuals obtained from such regression will be an estimate of the cyclical component $\mathbf{z}_t^C$.

In the case of a cointegrating VAR with no intercept and no trends (option 1 from the cointegrating VAR option in the System Estimation Menu), we have

$$\mathbf{w}_t = \mathbf{z}_t - \hat{\mathbf{C}}(1) \sum_{i=1}^{t} \hat{\mathbf{u}}_i = \mathbf{z}_0 + \hat{\mathbf{z}}_t^C$$

while the deterministic component is given by $\mathbf{z}_0$. In the case of a cointegrating VAR with restricted intercepts and no trends (option 2 in the cointegrating VAR option in the System Estimation Menu), consistent estimates of $\mathbf{g}$ and $\mathbf{z}_t^C$ can be obtained by running the $SURE$ regressions of $\mathbf{w}_t$ on an intercept, subject to the restrictions

$$\hat{\boldsymbol{\beta}}'\mathbf{g}_0 = \hat{\boldsymbol{\beta}}'\hat{\mathbf{a}}$$

where, once again, $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{a}}$ have already been estimated from the $VECM$ model.

In the case of a cointegrating VAR with unrestricted intercepts and no trends (option 3 from the cointegrating VAR option in the System Estimation Menu), $\mathbf{g} = \mathbf{0}$, and $\mathbf{g}_0$ can be consistently estimated by computing the sample mean of $\mathbf{w}_t$ (or running OLS regressions of $\mathbf{w}_t$ on the intercepts).

Finally, for a cointegrating VAR with unrestricted intercepts and trends (option 5 in the cointegrating VAR option in the System Estimation Menu), consistent estimates of $\mathbf{g}$ can be obtained by running OLS regressions of $\mathbf{w}_t$ on an intercept and a linear trend. The cyclical component $\hat{\mathbf{z}}_t^C$ is in all cases the residual from the above regressions.

22.12 Principal components

Principal components (PC) are linear combinations of a given set of variables which have special properties in terms of their variances. The first PC has the largest variance, the second PC has the second largest variance and so on. The linear combinations are orthogonal to each other and summarize important characteristics of the original variables. The PCs with largest variances are often used in statistics and econometrics as proxies for the original set of high-dimensional variables, thus reducing the dimensionality of the data set under consideration. For comprehensive treatments of the PC literature see Chapter 11 of Anderson (2003) and Jolliffe (2004).

Let $\mathbf{x}_{.t} = (x_{1t}, x_{2t}, ..., x_{nt})'$ be an $n \times 1$ vector of observations at time t on n variables. Suppose that the elements of $\mathbf{x}_{.t}$ have zero means, or are converted to zero mean variables by de-meaning, and that there are T observations available on $\mathbf{x}_{.t}$, with $T > n$. Consider the sample covariance matrix

$$\underset{n\times n}{\mathbf{S}} = \frac{\mathbf{X}'\mathbf{X}}{T}$$

where $\mathbf{X} = (\mathbf{x}_{.1}, ..., \mathbf{x}_{.T})'$ is a $T \times n$ matrix of observations. Let

$$\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_n \geq 0$$

be the n eigenvalue of $\mathbf{S}$, in descending order. Then the first k principal components are given by

$$\hat{\mathbf{f}}_t^{(k)} = \hat{\mathbf{B}}_k'\mathbf{x}_{.t}, \text{for } t = 1, 2, ..., T,$$

where $\hat{\mathbf{B}}'_k = \left(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, ..., \hat{\boldsymbol{\beta}}_k\right)$ is $n \times k$, with $\hat{\boldsymbol{\beta}}_i$ is the $n \times 1$ eigenvector of $\mathbf{S}$ corresponding to λ_i, and is such that

$$\begin{aligned} \hat{\boldsymbol{\beta}}_i' \hat{\boldsymbol{\beta}}_i &= 1, \quad i = 1, ..., n \\ \hat{\boldsymbol{\beta}}_i' \hat{\boldsymbol{\beta}}_j &= 0, \quad i \neq j \end{aligned}$$

In the context of a static factor model with $k < n$ factors

$$x_{it} = \mathbf{a}_i' \mathbf{z}_t + \boldsymbol{\beta}_i^{(k)\prime} \mathbf{f}_t^{(k)} + \varepsilon_{it}, \; i = 1, 2, ..., n, \; t = 1, 2, ..., T < n$$

$\boldsymbol{\beta}_i^{(k)}$ is the $k \times 1$ vector of factor loadings, $\mathbf{f}_t^{(k)}$ is the $k \times 1$ vector of unobserved common factors, and $\mathbf{z}_t$ is a $s \times 1$ vector of observed common factors. For example, $\mathbf{z}_t$ could include an intercept, time trend or other common observed variables, such as oil prices. To identify the unobserved factors we assume they have mean zero, unit variance and are mutually uncorrelated. To ensure that factors have mean zero, one of the elements of $\mathbf{z}_t$ must be specified to be equal to 1, unless $x_{it}'s$ are already demeaned. When n is sufficiently large the common factors can be estimated by the first k PCs of $T^{-1}\mathbf{X}'\mathbf{X}$. Note also that by construction the estimated factors are pair-wise uncorrelated, namely $\widehat{Corr}\left(\mathbf{f}_t^{(k)}, \mathbf{f}_t^{(s)}\right) = 0$ for all $s \neq k$ and $\widehat{Var}\left(\mathbf{f}_t^{(k)}\right) = \lambda_k$, for $k = 1, 2, ..., n$.

If $n \leq T$, *Microfit* reports λ_i, $i = 1, 2, ..., n$, the associated eigenvectors (or factor loadings), $\hat{\boldsymbol{\beta}}_i$, and the principal components $\hat{\mathbf{f}}_t^{(k)}$ for $k = 1, 2, ..., n$; $t = 1, 2, ..., T$, and allows these estimates to be saved as CSV or FIT files. The PCs can also be saved on the workspace, but not the eigenvectors or factor loadings.

Microfit also allows estimating principal components for $\mathbf{X}$, once these have been filtered by a set of variables, contained in a $T \times s$ matrix $\mathbf{Z}$, that might influence $\mathbf{X}$. In this case, principal components are computed from eigenvectors and eigenvalues of the matrix

$$\mathbf{S} = \frac{\mathbf{X}'\mathbf{M}_Z\mathbf{X}}{T}$$

where $\mathbf{M}_Z = \mathbf{I}_T - \mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$, and $\mathbf{I}_T$ is a $T \times T$ identity matrix.

If $n > T$, eigenvalues and principal components are computed using the $T \times T$ matrix $n^{-1}\mathbf{X}\mathbf{X}'$. In this case factor loadings will have the same dimension as the observations on the workspace and hence can be saved on the workspace as well.

22.12.1 Selecting the number of PCs or factors

There are a number of methods that can be used to select, $k < n$, the number of PC's or factors. The simplest and most popular procedures are the Kaiser criterion and the scree test. To use the Kaiser (1960) criterion the observations are standardized so that the variables, x_{it}, have unit variances (in sample), and hence $\sum_{i=1}^{n} \lambda_i = Trace(\mathbf{S}) = n$.

According to Kaiser (1960) criterion one would then retain only factors with eigenvalues greater than 1. In effect only factors that explain as much as the equivalent of one original variable are retained.

The scree test is based on a graphical method, first proposed by Cattell (1966). A simple line plot of the eigenvalues is used to identify a visual break in this time plot. There is no formal method for identifying the threshold, and certain degree of personal judgment is required. More formal procedures are suggested by Bai and Ng (2002), when n and T are large.

22.13 Canonical correlations

Canonical correlations (CC) measure the degree of correlations between two sets of variables. Let $\underset{T\times n_y}{\mathbf{Y}}$ be a matrix of T observations on n_y random variables, and $\underset{T\times n_x}{\mathbf{X}}$ be a matrix of T observations on n_x random variables, and suppose that $T > \max\{n_y, n_x\}$. CC is concerned with finding linear combinations of the $\mathbf{Y}$ variables and linear combinations of the $\mathbf{X}$ variables that are most highly correlated. In particular, let

$$u_{it} = \boldsymbol{\alpha}'_{(i)}\mathbf{y}_t \text{ and } v_{it} = \boldsymbol{\gamma}'_{(i)}\mathbf{x}_t,\ i = 1, 2, ...n = \min\{n_y, n_x\}$$

where $\mathbf{y}_t = (y_{1t}, y_{2t}, ..., y_{n_y,t})'$, $\mathbf{x}_t = (x_{1t}, x_{2t}, ..., x_{n_x,t})'$, and $\boldsymbol{\alpha}_i$ and $\boldsymbol{\gamma}_j$ are the associated $n_y \times 1$ and $n_x \times 1$ loading vectors, respectively. The first canonical correlation of $\mathbf{y}_t$ and $\mathbf{x}_t$ is given by those values of $\boldsymbol{\alpha}_{(1)}$ and $\boldsymbol{\gamma}_{(1)}$ that maximize the correlation of u_{1t} and v_{1t}. These variables are known as canonical variates. The second canonical correlation refers to $\boldsymbol{\alpha}_{(2)}$ and $\boldsymbol{\gamma}_{(2)}$ such that u_{2t} and v_{2t} have maximum correlation subject to the restriction that they are uncorrelated with u_{1t} and v_{1t}. The loadings are typically normalized so that the canonical variates have unit variances. As shown in Chapter 12 of Anderson (2003), the loading vectors should satisfy the following set of equations

$$\begin{pmatrix} -\rho\mathbf{S}_{yy} & \mathbf{S}_{yx} \\ \mathbf{S}_{xy} & -\rho\mathbf{S}_{xx} \end{pmatrix}\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\gamma} \end{pmatrix} = \mathbf{0}$$

where $\mathbf{S}_{yy} = T^{-1}(\mathbf{Y}'\mathbf{Y})$, $\mathbf{S}_{xx} = T^{-1}(\mathbf{X}'\mathbf{X})$ and $\mathbf{S}_{yx} = T^{-1}(\mathbf{Y}'\mathbf{X})$. Now set

$$\mathbf{S}_{yxy} = \mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}, \text{ if } n_y \leq n_x$$

and

$$\mathbf{S}_{xyx} = \mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}, \text{ if } n_x < n_y$$

and let $\rho_1^2 \geq \rho_2^2 \geq .... \geq \rho_{n_y}^2 \geq 0$ be the eigenvalues of $\mathbf{S}_{yxy}$. Then the k^{th} squared canonical correlation of $\mathbf{Y}$ and $\mathbf{X}$ is given by the k^{th} largest eigenvalue associated to the matrix $\mathbf{S}_{yxy}$, ρ_k^2. These coefficients measure the strength of the overall relationships between the two canonical variates, or weighted sums of $\mathbf{Y}$ and $\mathbf{X}$.

The canonical variates, u_{kt} and v_{kt}, associated with the k^{th} squared canonical correlation, ρ_k^2 is given by

$$u_{kt} = \boldsymbol{\alpha}'_{(k)}\mathbf{y}_t \text{ and } v_{kt} = \boldsymbol{\gamma}'_{(k)}\mathbf{x}_t$$

where
$$\begin{pmatrix} -\rho_k \mathbf{S}_{yy} & \mathbf{S}_{yx} \\ \mathbf{S}_{xy} & -\rho_k \mathbf{S}_{xx} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_{(k)} \\ \boldsymbol{\gamma}_{(k)} \end{pmatrix} = \mathbf{0}$$

But it is easily seen that

$$\begin{aligned} \left(\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy} - \rho_k^2 \mathbf{S}_{yy}\right)\boldsymbol{\alpha}_{(k)} &= 0 \\ \left(\mathbf{S}_{xy}\mathbf{S}_{yy}^{-1}\mathbf{S}_{yx} - \rho_k^2 \mathbf{S}_{xx}\right)\boldsymbol{\gamma}_{(k)} &= 0 \end{aligned}$$

and hence $\boldsymbol{\alpha}_{(k)}$ can be computed as the eigenvector associated with the k^{th} largest root of $\mathbf{S}_{yxy} = \mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}$, and $\boldsymbol{\gamma}_{(k)}$ can be computed as the eigen vector associated with the k^{th} largest root of $\mathbf{S}_{xyx} = \mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}$. These eigenvectors are normalized such that

$$\boldsymbol{\alpha}'_{(k)}\mathbf{S}_{yy}\boldsymbol{\alpha}_{(k)} = 1, \ \boldsymbol{\gamma}'_{(k)}\mathbf{S}_{xx}\boldsymbol{\gamma}_{(k)} = 1, \text{ and } \boldsymbol{\alpha}'_{(k)}\mathbf{S}_{yx}\boldsymbol{\gamma}_{(k)} = \rho_k$$

Under the null hypothesis $H_0 : Cov(\mathbf{X}, \mathbf{Y}) = \mathbf{0}$, the statistic

$$T \cdot Trace\left(\mathbf{S}_{yxy}\right) \overset{a}{\sim} \chi^2_{(n_y-1)(n_x-1)}$$

The above analysis can be extended to control for a third set of variables that might influence $\mathbf{Y}$ and $\mathbf{X}$. Consider the $T \times n_z$ observation matrix $\mathbf{Z}$, and suppose that $T > \max\{n_y, n_x, n_z\}$. Let $\mathbf{M}_z = \mathbf{I}_T - \mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$. Compute

$$\hat{\mathbf{Y}} = \mathbf{M}_z\mathbf{Y}, \quad \hat{\mathbf{X}} = \mathbf{M}_z\mathbf{X}$$

Then in this case the $\mathbf{S}$ matrix is given by

$$\mathbf{S}_{\hat{y}\hat{x}\hat{y}} = \left(\frac{\hat{\mathbf{Y}}'\hat{\mathbf{Y}}}{T}\right)^{-1}\left(\frac{\hat{\mathbf{Y}}'\hat{\mathbf{X}}}{T}\right)\left(\frac{\hat{\mathbf{X}}'\hat{\mathbf{X}}}{T}\right)^{-1}\left(\frac{\hat{\mathbf{X}}'\hat{\mathbf{Y}}}{T}\right)$$

if $n_y \leq n_x$ and

$$\mathbf{S}_{\hat{x}\hat{y}\hat{x}} = \left(\frac{\hat{\mathbf{X}}'\hat{\mathbf{X}}}{T}\right)^{-1}\left(\frac{\hat{\mathbf{X}}'\hat{\mathbf{Y}}}{T}\right)\left(\frac{\hat{\mathbf{Y}}'\hat{\mathbf{Y}}}{T}\right)^{-1}\left(\frac{\hat{\mathbf{X}}'\hat{\mathbf{Y}}}{T}\right)$$

if $n_y > n_x$.

Similarly, the covariates in this case are defined by

$$u_{kt} = \boldsymbol{\alpha}'_{(k)}\hat{\mathbf{y}}_t \text{ and } v_{kt} = \boldsymbol{\gamma}'_{(k)}\hat{\mathbf{x}}_t$$

where $\boldsymbol{\alpha}_{(k)}$ is the eigen vector of $\mathbf{S}_{\hat{y}\hat{x}\hat{y}}$ associated with its k^{th} largest eigenvalue, and $\boldsymbol{\gamma}_{(k)}$ is the eigenvector of $\mathbf{S}_{\hat{x}\hat{y}\hat{x}}$ associated with its k^{th} largest eigenvalue.

Note that by construction $\widehat{Corr}(u_{kt}, v_{kt}) = \rho_k$, $\widehat{Var}(u_{kt}) = \widehat{Var}(v_{kt}) = 1$, for $k = 1, 2, .., \min(n_y, n_x)$.

See also 4.4.7 and 10.17.

Chapter 23

Econometrics of Volatility Models

This chapter provides the technical details of the econometric methods and the algorithms that underlie the computation of the various estimators and test statistics in the case of univariate and multivariate conditionally heteroscedastic models. It complements Chapter 19, which describes the estimation options in *Microfit 5.0* for volatility modelling. Textbook treatments of some of the topics covered here can be found in Hamilton (1994), Satchell and Knight (2002), Campbell, Lo, and MacKinlay (1997), Bollerslev, Chou, and Kroner (1992) and Engle (1995). For extensions to the multivariate case, see Bollerslev, Engle, and Wooldridge (1988), Engle (2002), and Pesaran and Pesaran (2007). An application to risk management can be found in Pesaran, Schleicherc, and Zaffaroni (2009), where a large number of alternative multivariate volatility models are considered.

23.1 Univariate conditionally heteroscedastic models

Consider the following linear regression model:

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + u_t, \quad t = 1, ..., n \tag{23.1}$$

where y_t is the dependent variable, $\boldsymbol{\beta}$ is a $k \times 1$ vector of unknown coefficients, $\mathbf{x}_t$ is the $k \times 1$ vector of regressors, and u_t is a disturbance term. Under the classical normal assumptions (A1 to A5) set out in Section 6.1, the disturbances u_t, in the regression model (23.1) have a constant variance both *unconditionally* and *conditionally*. However, in many applications in macroeconomics and finance, the assumption that the conditional variance of u_t is constant over time is not valid. Regression models that allow the conditional variance of u_t to vary over time as a function of past errors are known as Autoregressive Conditional Heteroscedastic ($ARCH$) models.

23.1.1 GARCH-in-mean models

The $ARCH$ model was introduced into the econometric literature by Engle (1982), and was subsequently generalized by Bollerslev (1986), who proposed the Generalized $ARCH$ (or $GARCH$) models. Other related models where the conditional variance of u_t is used as one of the regressors explaining the conditional mean of y_t have also been suggested in the literature, and are known as $ARCH$-in-Mean and $GARCH$-in-Mean (or $GARCH$-M) models (see, for example, Engle, Lillien, and Robins (1987)).

The various members of the $GARCH$ and $GARCH$-M class of models can be written compactly as

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + \gamma h_t^2 + u_t \tag{23.2}$$

where

$$\begin{aligned} h_t^2 &= V(u_t \,|\Omega_{t-1}) = E(u_t^2 \,|\Omega_{t-1}) \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \phi_i h_{t-i}^2 \end{aligned} \tag{23.3}$$

and Ω_{t-1} is the information set at time $t-1$, containing at least observations on $\mathbf{x}_t$ and on lagged values of y_t and $\mathbf{x}_t$; namely, $\Omega_{t-1} = (\mathbf{x}_t, \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, ..., y_{t-1}, y_{t-2}, ..)$. The unconditional variance of u_t is constant and is given by[1]

$$V(u_t) = \sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{p} \phi_i} > 0 \tag{23.4}$$

and the necessary condition for (23.2) to be covariance stationary is given by

$$\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \phi_i < 1 \tag{23.5}$$

In addition to the restrictions (23.4) and (23.5), Bollerslev (1986) also assumes that $\alpha_i \geq 0$, $i = 1, 2, ..., q$, and $\phi_i \geq 0$, $i = 1, 2, ..., q$. Although these additional restrictions are sufficient for the conditional variance to be positive, they are not necessary (see Nelson and Cao (1992)).

Microfit computes approximate maximum likelihood estimates of the parameters of a generalization of the $GARCH$-M model, where in addition to past disturbances, other variables could also influence h_t^2:

$$h_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \phi_i h_{t-i}^2 + \boldsymbol{\delta}'\mathbf{w}_t \tag{23.6}$$

where $\mathbf{w}_t$ is a vector of covariance stationary variables in Ω_{t-1}. The unconditional variance of u_t in this case is given by

$$\sigma^2 = \frac{\alpha_0 + \boldsymbol{\delta}'\boldsymbol{\mu}_w}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{p} \phi_i} > 0 \tag{23.7}$$

[1] Notice that $V(u_t) = \lim_{s \to \infty} E\left(u_{t+s}^2 \,|\Omega_{t-1}\right)$.

where $\boldsymbol{\mu}_\omega = E(\mathbf{w}_t)$.

The ML estimation of the above augmented $GARCH$-M model can be carried out in *Microfit* under two different assumptions concerning the conditional distribution of the disturbances, namely Gaussian and Student's t-distribution. In both cases the *exact* log-likelihood function depends on the joint density function of the initial observations, $f(y_1, y_2, ..., y_q)$, which is non-Gaussian and intractable analytically. In most applications where the sample size is large (as is the case with most financial time-series) the effect of the distribution of the initial observations is relatively small and can be ignored.[2]

23.1.2 ML estimation with Gaussian errors

The log-likelihood function used in computation of the ML estimators for the Gaussian case is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= -\frac{(n-q)}{2}\log(2\pi) - \tfrac{1}{2}\sum_{t=q+1}^{n}\log h_t^2 \\ &\quad -\tfrac{1}{2}\sum_{t=q+1}^{n} h_t^{-2}u_t^2 \end{aligned} \tag{23.8}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}', \gamma, \alpha_0, \alpha_1, \alpha_2, ..., \alpha_q, \phi_1, \phi_2, ..., \phi_p, \boldsymbol{\delta}')'$, and u_t and h_t^2 are given by (23.2) and (23.6), respectively.

23.1.3 ML estimation with Student's t-distributed errors

Under the assumption that conditional on Ω_{t-1}, the disturbances are distributed as a Student's t-distribution with v degrees of freedom ($v > 2$), the log-likelihood function is given by

$$\ell(\boldsymbol{\theta}, v) = \sum_{t=q+1}^{n} \ell_t(\boldsymbol{\theta}, v) \tag{23.9}$$

where

$$\begin{aligned}\ell_t(\boldsymbol{\theta}, v) &= -\log\left\{B\left(\tfrac{v}{2}, \tfrac{1}{2}\right)\right\} - \tfrac{1}{2}\log(v-2) \\ &\quad -\tfrac{1}{2}\log h_t^2 - \left(\frac{v+1}{2}\right)\log\left(1 + \frac{u_t^2}{h_t^2(v-2)}\right)\end{aligned} \tag{23.10}$$

and $B\left(\frac{v}{2}, \frac{1}{2}\right)$ is the complete Beta function.[3]

The degrees of freedom of the underlying t-distribution, v, is then estimated along with the other parameters. The Gaussian log-likelihood function (23.8) is a special case of (23.10), and can be obtained from it for large values of v. In most applications two log-likelihood

[2] Diebold and Schuermann (1992) examine the quantitative importance of the distribution of the initial observations in the case of simple ARCH models, and find their effect to be negligible.

[3] Notice that $B(\frac{v}{2}, \frac{1}{2}) = \Gamma\left(\frac{v+1}{2}\right)/\Gamma\left(\frac{v}{2}\right)\Gamma\left(\frac{1}{2}\right)$. The constant term $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ is omitted from the expression used by Bollerslev (1987). See his equation (1).

functions provide very similar results for values of v around 20. The t -distribution is particularly appropriate for the analysis of stock returns, where the distribution of the standardized residuals, $\hat{u}_t/\hat{h}_t$, are often found to have fatter tails than the normal distribution.

23.1.4 Exponential GARCH-in-Mean models

It is often the case that the conditional variance, h_t^2 is not an even function of the past disturbances, $u_{t-1}, u_{t-2}, \ldots$. The Exponential *GARCH* (or *EGARCH*) model proposed by Nelson (1991) aims at capturing this important feature, often observed when analysing stock market returns. *Microfit* provides *ML* estimates of the following augmented version of the *EGARCH-M* model:

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + \gamma h_t^2 + u_t \tag{23.11}$$

where as before $h_t^2 = V(u_t\,|\Omega_{t-1}) = E(u_t^2\,|\Omega_{t-1})$. However, for h_t^2 Nelson uses an exponential functional form, which can be written as

$$\begin{aligned}\log h_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \left(\frac{u_{t-i}}{h_{t-i}}\right) + \sum_{i=1}^{q} \alpha_i^* \left(\left|\frac{u_{t-i}}{h_{t-i}}\right| - \mu\right) \\ &\quad + \sum_{i=1}^{p} \phi_i \log h_{t-i}^2 + \boldsymbol{\delta}'\mathbf{w}_t \end{aligned} \tag{23.12}$$

where $\mu = E\left(\left|\frac{u_t}{h_t}\right|\right)$. The value of μ depends on the density function assumed for the standardized disturbances, $\varepsilon_t = u_t/h_t$. We have

$$\mu = \sqrt{\frac{2}{\pi}}, \qquad \text{if} \qquad \varepsilon_t \sim N(0,1) \tag{23.13}$$

and

$$\mu = \frac{2(v-2)^{\frac{1}{2}}}{(v-1)B\left(\frac{v}{2},\frac{1}{2}\right)} \tag{23.14}$$

if ε_t has a standardized t-distribution with v degrees of freedom.[4]

The (approximate) log-likelihood function for the *EGARCH* model has the same form as in (23.8) and (23.9) for the Gaussian and Student t-distributions, respectively. Unlike the *GARCH-M* class of models, the *EGARCH-M* model always yields a positive conditional variance, h_t^2, for any choice of the unknown parameters; it is only required that the roots of $1 - \sum_{i=1}^{p} \phi_i z^i = 0$ should all fall outside the unit circle. The unconditional variance of u_t in the case of the *EGARCH* model does not have a simple analytical form.

[4] Notice that in this case $E\left(|u_t/h_t|\right) = \frac{2}{\sqrt{v}B\left(\frac{v}{2},\frac{1}{2}\right)} \int_0^\infty a\left(1+\frac{a^2}{v}\right)^{-(v+1)/2} da.$

23.1.5 Absolute GARCH-in-Mean models

This is the third class of conditionally heteroscedastic models that can be estimated in *Microfit,* and has the following specification:[5]

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + \gamma h_t^2 + u_t \tag{23.15}$$

where h_t is given by

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i |u_{t-i}| + \sum_{i=1}^{p} \phi_i h_{t-i} + \boldsymbol{\delta}'\mathbf{w}_t \tag{23.16}$$

The $AGARCH$ model can also be estimated for different error distributions. The log-likelihood functions for the cases where $\varepsilon_t = u_t/h_t$ has a standard normal distribution, and when it has a standardized Student-t distribution, these are given by (23.8) and (23.9), where u_t and h_t are now specified by (23.15) and (23.16), respectively.

23.1.6 Computational considerations

The computation of the ML estimators for the above models are carried out by the Newton-Raphson algorithm using numerical derivatives, after appropriate scaling of the parameters. We also use a 'damping factor' to control for the step-size in the iterations. The value of this damping factor can be changed by the user in the range $[0.01, 2.0]$.[6] The convergence of the iterations often depends on the nature of the conditional heteroscedasticity in the data and the choice of the initial values for the parameters.

Once the ML estimates are computed, *Microfit* then computes their asymptotic standard errors using the inverse of the Hessian matrix (the second partial derivatives of the log-likelihood function).

23.1.7 Testing for ARCH (or GARCH) effects

The simplest way to test for $ARCH(p)$ effects is to use the Lagrange Multiplier procedure proposed by Engle (1982) (see, in particular, page 1000). The test involves two steps. In the first step the OLS residuals, $\hat{u}_{t,OLS}$, from the regression of y_t on $\mathbf{x}_t$, are obtained, and in the second step $\hat{u}_{t,OLS}^2$ is regressed on a constant and p of its own lagged values

$$\hat{u}_{t,OLS}^2 = a_0 + b_1\hat{u}_{t-1,OLS}^2 + ... + b_q\hat{u}_{t-q,OLS}^2 + e_t$$

for $t = q+1, q+2, ..., n$. A test of the $ARCH(q)$ effect can now be carried out by testing the statistical significance of the slope coefficients $b_1 = b_2 = ... = b_q = 0$. This statistic is automatically computed by *Microfit* using option 2 in the Hypothesis Testing Menu (after the OLS regression).

[5] The $AGARCH$ model has been proposed in the literature by Heutschel (1991).

[6] The unobserved initial values of h_t are set equal to $\tilde{\sigma}$, the OLS estimator of the unconditional variance of u_t.

23.1.8 Residuals, DW, $\mathbf{R}^2$ and other statistics

Microfit reports the values of the unscaled residuals $\hat{u}_t$ computed as

$$\hat{u}_t = y_t - \widehat{\boldsymbol{\beta}}'\mathbf{x}_t - \hat{\gamma}\hat{h}_t^2,\ t = 1, 2, ..., n$$

and the conditional standard errors, $\hat{h}_{t,}$ (namely the ML estimates of $h_{t,}$) using the relations (23.6), (23.12) and (23.16). The scaled residuals, $\hat{u}_t/\hat{h}_{t,}$ are used in the histogram plots. The program also reports the maximized value of the log-likelihood function, AIC, and SBC, DW,R^2 using formulae similar to those for the other estimation procedures.

The summary statistics are computed using the unscaled residuals. Those interested in computing these statistics using the scaled residuals can do so by first saving the values of $\hat{u}_t$ and $\hat{h}_t$ in the Post Regression Menu of the $GARCH$ option, and then carrying out the necessary computations at the data processing stage.

23.1.9 Forecasting with conditionally heteroscedastic models

There are two components in $GARCH$ or $GARCH$-M models that require forecasting: the conditional mean and the conditional variance. The forecasts of the former are given by

$$\hat{y}^*_{t+j} = \hat{\boldsymbol{\beta}}'\ {}_t\hat{\mathbf{x}}_{t+j} + \hat{\gamma}\ {}_t\hat{h}^2_{t+j},\quad j\ = 1, 2, ..., p$$

where $\hat{\boldsymbol{\beta}}$, and $\hat{\gamma}$ are ML estimators of the regression coefficients, ${}_t\hat{\mathbf{x}}_{t+j}$ is the j-step ahead forecast of $\mathbf{x}_t$, and ${}_t\hat{h}^2_{t+j}$ is the j-step ahead forecast of the conditional variance, namely $E(u^2_{t+j} \mid \Omega_t)$, $j = 1, 2, ..., p$. The computation of ${}_t\hat{\mathbf{x}}_{t+j}$ is carried out along the lines set out in Section 21.26.2, and depends on whether $\mathbf{x}_t$ contains lagged values of the dependent variable or not.

The computation of ${}_t\hat{h}^2_{t+j}$ varies depending on the conditional heteroscedasticity model under consideration. For the $GARCH$ specification defined by (23.6), the one-step ahead forecast of the conditional volatility is given by

$${}_t\hat{h}^2_{t+1} = \hat{\alpha}_0 + \sum_{i=1}^{q}\hat{\alpha}_i\hat{u}^2_{t+1-i} + \sum_{i=1}^{p}\hat{\phi}_i\hat{h}^2_{t+1-i} + \hat{\boldsymbol{\delta}}'\ {}_t\hat{\mathbf{w}}_{t+1}$$

where ${}_t\hat{\mathbf{w}}_{t+1}$ is the one-step ahead predictor of $\mathbf{w}_t$,

$$\hat{u}_{t-i} = y_{t-i} - \hat{\boldsymbol{\beta}}'\mathbf{x}_{t-i} - \gamma\hat{h}^2_{t-i}, i = 0, 1, 2, .. \tag{23.17}$$

and

$$\hat{h}^2_{t-j} = \hat{\alpha}_0 + \sum_{i=1}^{q}\hat{\alpha}_i\hat{u}^2_{t-j-i} + \sum_{i=1}^{p}\hat{\phi}_i\hat{h}^2_{t-j-i} + \hat{\boldsymbol{\delta}}'\mathbf{w}_{t-j},\ \text{for } j = 0, 1, 2, ...$$

with the unobserved initial values of $\hat{h}^2$ set equal to the estimate of the unconditional variance of u_t. Similarly, two- and three-step ahead forecasts are given by

$$\begin{aligned}
{}_t\hat{h}^2_{t+2} &= \hat{\alpha}_0 + (\hat{\alpha}_1 + \hat{\phi}_1)\ {}_t\hat{h}^2_{t+1} + \sum_{i=2}^{q} \hat{\alpha}_i \hat{u}^2_{t+2-i} + \sum_{i=2}^{p} \hat{\phi}_i \hat{h}^2_{t+2-i} + \hat{\boldsymbol{\delta}}'\ {}_t\hat{\mathbf{w}}_{t+2} \\
{}_t\hat{h}^2_{t+3} &= \hat{\alpha}_0 + (\hat{\alpha}_1 + \hat{\phi}_1)\ {}_t\hat{h}^2_{t+2} + (\hat{\alpha}_2 + \hat{\phi}_2)\ {}_t\hat{h}^2_{t+1} + \sum_{i=3}^{q} \hat{\alpha}_i \hat{u}^2_{t+3-i} \\
&\quad + \sum_{i=3}^{p} \hat{\phi}_i \hat{h}^2_{t+3-i} + \hat{\boldsymbol{\delta}}'\ {}_t\hat{\mathbf{w}}_{t+3}
\end{aligned}$$

and so on.

For the *EGARCH* model defined by (23.12), the one-step ahead forecast of conditional volatility is given by

$$\widehat{\log\ {}_t h^2_{t+1}} = \hat{\alpha}_0 + \sum_{i=1}^{p} \hat{\phi}_i \log \hat{h}^2_{t+1-i} + \hat{\boldsymbol{\delta}}'\ {}_t\hat{\mathbf{w}}_{t+1} \tag{23.18}$$

where

$$\begin{aligned}
\log \hat{h}^2_{t-j} &= \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i \left(\frac{\hat{u}_{t-i-j}}{\hat{h}_{t-i-j}} \right) + \sum_{i=1}^{q} \hat{\alpha}^*_i \left(\left| \frac{\hat{u}_{t-i-j}}{\hat{h}_{t-i-j}} \right| - \hat{\mu} \right) \\
&\quad + \sum_{i=1}^{p} \hat{\phi}_i \log \hat{h}^2_{t-j-i} + \hat{\boldsymbol{\delta}}' \mathbf{w}_{t-j}, \ \text{for } j = 0, 1, 2, ..
\end{aligned}$$

and $\hat{u}_t$ is already defined by (23.17). For higher-step ahead forecasts we have[7]

$$\begin{aligned}
\widehat{\log\ {}_t h^2_{t+2}} &= \hat{\alpha}_0 + \hat{\phi}_1 \widehat{\log\ {}_t h^2_{t+1}} + \sum_{i=2}^{p} \hat{\phi}_i \log \hat{h}^2_{t+2-i} + \hat{\boldsymbol{\delta}}'\ {}_t\hat{\mathbf{w}}_{t+2}, \\
\widehat{\log\ {}_t h^2_{t+3}} &= \hat{\alpha}_0 + \hat{\phi}_1 \widehat{\log\ {}_t h^2_{t+2}} + \hat{\phi}_2 \widehat{\log\ {}_t h^2_{t+1}} + \sum_{i=3}^{p} \hat{\phi}_i \log \hat{h}^2_{t+3-i} + \hat{\boldsymbol{\delta}}'\ {}_t\hat{\mathbf{w}}_{t+3},
\end{aligned}$$

and so on.

For the Absolute *GARCH* model defined by (23.16) we have

$${}_t\hat{h}_{t+1} = \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i \left| \hat{u}_{t+1-i} \right| + \sum_{i=1}^{p} \hat{\phi}_i \hat{h}_{t+1-i} + \hat{\boldsymbol{\delta}}'\ {}_t\hat{\mathbf{w}}_{t+1}$$

where

$$\hat{h}_{t-j} = \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i \left| \hat{u}_{t-j-i} \right| + \sum_{i=1}^{p} \hat{\phi}_i \hat{h}_{t-j-i} + \hat{\boldsymbol{\delta}}'\ \mathbf{w}'_{t-j}{}' \ \text{for } j = 0, 1, 2, ..$$

[7]The volatility forecasts for more than one-step ahead are computed assuming that $\exp(\widehat{\log\ {}_t h^2_{t+1}})$ is approximately equal to ${}_t\hat{h}^2_{t+1}$. Exact computation of ${}_t\hat{h}^2_{t+1}$ involves computing integrals by stochastic simulation techniques. The same point also applies to volatility forecasts based on the Absolute *GARCH* model.

and $\hat{u}_{t-j}$, $j = 0, 1, 2, ..$ are defined by (23.17). For higher-step ahead forecasts we have:[8]

$$
\begin{aligned}
{}_t\hat{h}_{t+2} &= \hat{\alpha}_0 + (\hat{\alpha}_1 + \hat{\phi}_1)\, {}_t\hat{h}_{t+1} + \sum_{i=2}^{q} \hat{\alpha}_i \left|\hat{u}_{t+1-i}\right| + \sum_{i=2}^{p} \hat{\phi}_i \hat{h}_{t+1-i} + \hat{\boldsymbol{\delta}}'\, {}_t\hat{\mathbf{w}}_{t+2} \\
{}_t\hat{h}_{t+3} &= \hat{\alpha}_0 + (\hat{\alpha}_1 + \hat{\phi}_1)\, {}_t\hat{h}_{t+2} + (\hat{\alpha}_2 + \hat{\phi}_2)\, {}_t\hat{h}_{t+1} + \sum_{i=3}^{q} \hat{\alpha}_i \left|\hat{u}_{t+1-i}\right| \\
&\quad + \sum_{i=3}^{p} \hat{\phi}_i \hat{h}_{t+1-i} + \hat{\boldsymbol{\delta}}'\, {}_t\hat{\mathbf{w}}_{t+3}
\end{aligned}
$$

and so on.

In all the above formulae the forecasts of $\mathbf{w}_{t+j}$ are obtained recursively if $\mathbf{w}_t$ contains lagged values of y_t. Otherwise, actual values of $\mathbf{w}_{t+j}$ will be used. Forecasts will not be computed if future values of $\mathbf{w}_t$ are not available or cannot be computed using recursive forecasts of y_t.

23.2 Multivariate conditionally heteroscedastic models

Let $\mathbf{r}_t = (r_{1t}, ..., r_{mt})'$ be an $m \times 1$ vector of asset returns at close day t, with conditional mean and variance

$$
\begin{aligned}
\boldsymbol{\mu}_{t-1} &= E\left(\mathbf{r}_t \left|\Omega_{t-1}\right.\right) \\
\boldsymbol{\Sigma}_{t-1} &= Cov\left(\mathbf{r}_t \left|\Omega_{t-1}\right.\right)
\end{aligned}
$$

where Ω_{t-1} is the information set available at close of day $t-1$, and $\boldsymbol{\Sigma}_{t-1}$ is assumed to be non-singular. Here we are not concerned with how mean returns are predicted, and take $\boldsymbol{\mu}_{t-1}$ as given.[9] The conditional covariance matrix $\boldsymbol{\Sigma}_{t-1}$ may be uniquely expressed in terms of the decomposition

$$
\boldsymbol{\Sigma}_{t-1} = \mathbf{D}_{t-1}\mathbf{R}_{t-1}\mathbf{D}_{t-1} \tag{23.19}
$$

where

$$
\mathbf{D}_{t-1} = \begin{bmatrix} \sigma_{1,t-1} & 0 & \ldots & 0 \\ 0 & \sigma_{2,t-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & \ldots & \sigma_{m,t-1} \end{bmatrix}
$$

[8] Here we are making use of the approximation $E\left(|u_{t+j}| \mid \Omega_t\right) \simeq {}_t h_{t+j}$.

[9] Although the estimation of $\boldsymbol{\mu}_{t-1}$ and $\boldsymbol{\Sigma}_{t-1}$ are interrelated, in practice mean returns are predicted by least squares techniques (such as recursive estimation or recursive modelling) which do not take account of the conditional volatility. This might involve some loss in efficiency of estimating $\boldsymbol{\mu}_{t-1}$, but considerably simplifies the estimation of the return distribution needed in portfolio decisions and risk management.

$$\mathbf{R}_{t-1} = \begin{bmatrix} 1 & \rho_{12,t-1} & \rho_{13,t-1} & \cdots & \rho_{1m,t-1} \\ \rho_{21,t-1} & 1 & \rho_{23,t-1} & \cdots & \rho_{2m,t-1} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & & \rho_{m-1,m,t-1} \\ \rho_{m1,t-1} & \cdots & \cdots & \rho_{m,m-1,t-1} & 1 \end{bmatrix}$$

$\mathbf{D}_{t-1}$ is an $m \times m$ diagonal matrix with elements $\sigma_{i,t-1}$, $i = 1, 2, \ldots, m$ denoting the conditional volatilities of assets returns, and $\mathbf{R}_{t-1}$ is the symmetric $m \times m$ matrix of pairwise conditional correlations. More specifically, the conditional volatility for the ith asset return is defined as

$$\sigma^2_{i,t-1} = Var\left(r_{it} \mid \Omega_{t-1}\right)$$

and the conditional pair-wise return correlation between the ith and the jth asset is

$$\rho_{ij,t-1} = \rho_{ji,t-1} = \frac{Cov\left(r_{it}, r_{jt} \mid \Omega_{t-1}\right)}{\sigma_{i,t-1}\sigma_{j,t-1}}$$

Clearly, $-1 \le \rho_{ij,t-1} \le 1$, and $\rho_{ij,t-1} = 1$, for $i = j$.

The decomposition of $\mathbf{\Sigma}_{t-1}$ in (23.19) allows separate specification for the conditional volatilities and conditional cross-asset returns correlations. *Microfit* allows estimating the following $GARCH(1{,}1)$ model for $\sigma^2_{i,t-1}$

$$\sigma^2_{i,t-1} = \bar{\sigma}^2_i\left(1 - \lambda_{1i} - \lambda_{2i}\right) + \lambda_{1i}\sigma^2_{i,t-2} + \lambda_{2i}r^2_{i,t-1} \tag{23.20}$$

where $\bar{\sigma}^2_i$ is the unconditional variance of the ith asset return. Notice that in (23.20) we allow the parameters $\lambda_{1i}, \lambda_{2i}$ to differ across assets. Under the restriction $\lambda_{1i} + \lambda_{2i} = 1$, the unconditional variance does not exist. In this case we have the integrated $GARCH$ ($IGARCH$) model used extensively in the professional financial community:[10]

$$\sigma^2_{i,t-1}\left(\lambda_i\right) = \left(1 - \lambda_i\right)\sum_{s=1}^{\infty}\lambda_i^{s-1}r^2_{i,t-s} \qquad 0 < \lambda_i < 1 \tag{23.21}$$

or, written recursively,

$$\sigma^2_{i,t-1}\left(\lambda_i\right) = \lambda_i\sigma^2_{i,t-2} + \left(1 - \lambda_i\right)r^2_{i,t-1}$$

As for cross-asset correlations, *Microfit* estimates the (i, j)th conditional correlation as

$$\tilde{\rho}_{ij,t-1}\left(\phi\right) = \frac{q_{ij,t-1}}{\sqrt{q_{ii,t-1}q_{jj,t-1}}}$$

where $q_{ij,t-1}$ are given by

$$q_{ij,t-1} = \bar{\rho}_{ij}(1 - \phi_1 - \phi_2) + \phi_1 q_{ij,t-2} + \phi_2\tilde{r}_{i,t-1}\tilde{r}_{j,t-1} \tag{23.22}$$

In (23.22), $\bar{\rho}_{ij}$ is the (i, j)th unconditional correlation, ϕ_1, ϕ_2 are parameters such that $\phi_1 + \phi_2 < 1$, and $\tilde{r}_{i,t-1}$ are the standardized assets returns. Returns are standardized to achieve normality.

[10] See, for example, Litterman and Winkelmann (1998).

Engle (2002) proposes the following standardization for returns:

$$\tilde{r}_{i,t-1} = \tilde{r}^{exp}_{i,t-1} = \frac{r_{it}}{\sigma_{i,t-1}} \tag{23.23}$$

where $\sigma_{i,t-1}$ is given either by (23.20) or, in the case of non-mean reverting volatilities, by (23.21). We refer to (23.23) as the 'exponentially weighted returns'.

An alternative way of standardizing returns is by using a measure of the actual or realized volatility (Pesaran and Pesaran (2007)):

$$\tilde{r}_{i,t-1} = \tilde{r}^{devol}_{i,t-1} = \frac{r_{it}}{\sigma^{realized}_{i,t-1}} \tag{23.24}$$

where $\sigma^{realized}_{i,t-1}$ is a proxy of the realized volatility of the i^{th} return in day t. Pesaran and Pesaran (2007) have suggested the following approximation for the realized volatility:

$$\tilde{\sigma}^2_{it}(p) = \frac{\sum_{s=0}^{p-1} r^2_{i,t-s}}{p} \tag{23.25}$$

The lag order, p, needs to be chosen carefully; its default value for p in *Microfit* is 20. We refer to returns (23.24), where the realized volatility is estimated using (23.25) as 'devolatized returns'.

In a series of papers, Andersen, Bollerslev and Diebold show that daily returns on foreign exchange and stock returns standardized by realized volatility are approximately Gaussian (see, for example, Andersen, Bollerslev, Diebold, and Ebens (2001), and Andersen, Bollerslev, Diebold, and Labys (2001)). The transformation of returns to Gaussianity is important, since the use of correlation as a measure of dependence can be misleading in the case of (conditionally) non-Gaussian returns (see Embrechts, Hoing, and Juri (2003)). In contrast, estimation of correlations based on devolatized returns that are nearly Gaussian is likely to be more generally meaningful.

In (23.22) it is required that $\phi_1 + \phi_2 < 1$, that is the process is mean reverting. In the case $\phi_1 + \phi_2 = 1$, we have

$$q_{ij,t-1} = \phi q_{ij,t-2} + (1-\phi)\, \tilde{r}_{i,t-1}\tilde{r}_{j,t-1}$$

In practice, the hypothesis that $\phi_1 + \phi_2 < 1$ needs to be tested.

23.2.1 Initialization, estimation and evaluation samples

Estimation and evaluation of the dynamic conditional correlation (DCC) model given by (23.20) and (23.22) is carried out in a recursive manner.

Suppose daily observations are available on m daily returns in the $m \times 1$ vector $\mathbf{r}_t$ over the period $t = 1, 2, ..., T, T+1, ..., T+N$. The sample period can be divided into three sub-periods, choosing s, T_0 and T such that $p < T_0 < s < T$. We call:

– Initialization sample: $\mathcal{S}_0 = \{\mathbf{r}_t, \, t = 1, 2, ..., T_0\}$. The first T_0 observations are used for initialization of the recursions in (23.20) and (23.22).

– Estimation sample: $\mathcal{S}_{est} = \{\mathbf{r}_t,\, t = s, s+1, ..., T\}$. A total of $T - s + 1$ observations are used for estimation of (23.20) and (23.22) (see Section 23.2.2).

– Evaluation sample: $\mathcal{S}_{eval} = \{\mathbf{r}_t,\, t = T+1, T+2, ..., T+N\}$. The last N observations are used for testing the validity of the model (see Section 23.2.3). In *Microfit* $N \geq 35$.

This decomposition allows variations of the size of the estimation window by moving the index s along the time axis in order to accommodate estimation of the unknown parameters using expanding or rolling observation windows, with different estimation update frequencies. For example, for an expanding estimation window we set $s = T_0 + 1$, and for a rolling window of size W we need to set $s = T + 1 - W$. The whole estimation process can then be rolled into the future with an update frequency of h by carrying the estimations at $T+h, T+2h$, ..., using either expanding or rolling estimation samples from $t = s$.

23.2.2 Maximum likelihood estimation

ML estimation of the DCC model can be carried out in *Microfit* under two different assumptions concerning the conditional distribution of assets returns: the multivariate Gaussian distribution and the multivariate Student's t-distribution.

In its most general formulation (the non-mean reverting specifications given by (23.20) and (23.22)) the DCC model contains $2m + 2$ unknown parameters: $2m$ coefficients $\boldsymbol{\lambda}_1 = (\lambda_{11}, \lambda_{12}, \ldots, \lambda_{1m})'$ and $\boldsymbol{\lambda}_2 = (\lambda_{21}, \lambda_{22}, \ldots, \lambda_{2m})'$, which enter the individual asset returns volatilities, and the coefficients ϕ_1 and ϕ_2 that enter the conditional correlations. In the case of t-distributed returns, a further parameter, the degrees of freedom of the multivariate Student t-distribution, v, need to be estimated.

The intercepts $\bar{\sigma}_i^2$ and $\bar{\rho}_{ij}$ in (23.20) and (23.22) refer to the unconditional volatilities and return correlations and can be estimated as

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^{T} r_{it}^2}{T} \tag{23.26}$$

$$\bar{\rho}_{ij} = \frac{\sum_{t=1}^{T} r_{it} r_{jt}}{\sqrt{\sum_{t=1}^{T} r_{it}^2}\sqrt{\sum_{t=1}^{T} r_{jt}^2}} \tag{23.27}$$

In the non-mean reverting case these intercept coefficients disappear, but for initialization of the recursive relations (23.20) and (23.22) it is still advisable to use unconditional estimates of the correlation matrix and asset returns volatilities.

ML estimation with Gaussian returns

Denote the unknown coefficients by

$$\boldsymbol{\theta} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \phi_1, \phi_2)'$$

Based on a sample of observations on returns, $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_t$, available at time t, the time t log-likelihood function based on the decomposition (23.19) is given by

$$l_t(\boldsymbol{\theta}) = \sum_{\tau=s}^{t} f_\tau(\boldsymbol{\theta})$$

where $s < t$ is the start date of the estimation window and

$$\begin{aligned} f_\tau(\boldsymbol{\theta}) &= -\frac{m}{2}\ln(\pi) - \frac{1}{2}\ln|R_{\tau-1}(\boldsymbol{\theta})| - \ln|D_{\tau-1}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)| \\ &\quad - \ln\left[\mathbf{e}_\tau' D_{\tau-1}^{-1}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) R_{\tau-1}^{-1}(\boldsymbol{\theta}) D_{\tau-1}^{-1}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)\mathbf{e}_\tau\right] \\ &= f_{\tau,R}(\boldsymbol{\theta}) + f_{\tau,C}(\boldsymbol{\theta}) \end{aligned}$$

with

$$\mathbf{e}_\tau = \mathbf{r}_\tau - \boldsymbol{\mu}_{\tau-1}$$

For estimation of the unknown parameters, $\lambda_1, \lambda_2, ..., \lambda_m$, and ϕ, Engle (2002) proposes a two-step procedure whereby in the first step individual $GARCH(1,1)$ models are estimated for each of the m asset returns separately, and then the coefficient of the conditional correlations, ϕ, is estimated by the Maximum Likelihood method assuming that asset returns are conditionally Gaussian.

Note that under Engle's specification, $\mathbf{R}_{t-1}$ depends on $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ as well as on ϕ_1 and ϕ_2.

This procedure has two main drawbacks. First, the Gaussianity assumption in general does not hold for daily returns, and its use can under-estimate the portfolio risk. Second, the two-stage approach is likely to be inefficient even under Gaussianity.

For further details on ML estimation using Gaussian returns, see Engle (2002).

ML estimation with Student's t-distributed returns

Denote the unknown coefficients by

$$\boldsymbol{\theta} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \phi_1, \phi_2, v)'$$

where v are the (unknown) degrees of freedom of the t-distribution. The time t log-likelihood function based on the decomposition (23.19) is given by

$$l_t(\boldsymbol{\theta}) = \sum_{\tau=s}^{t} f_\tau(\boldsymbol{\theta}) \tag{23.28}$$

where

$$\begin{aligned} f_\tau(\boldsymbol{\theta}) &= -\frac{m}{2}\ln(\pi) - \frac{1}{2}\ln|R_{\tau-1}(\boldsymbol{\theta})| - \ln|D_{\tau-1}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)| \\ &\quad + \ln\left[\Gamma\left(\frac{m+v}{2}\right)/\Gamma\left(\frac{v}{2}\right)\right] - \frac{m}{2}\ln(v-2) \\ &\quad - \left(\frac{m+v}{2}\right)\ln\left[1 + \frac{\mathbf{e}_\tau' D_{\tau-1}^{-1}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) R_{\tau-1}^{-1}(\boldsymbol{\theta}) D_{\tau-1}^{-1}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)\mathbf{e}_\tau}{v-2}\right] \end{aligned} \tag{23.29}$$

and $\mathbf{e}_\tau = \mathbf{r}_\tau - \boldsymbol{\mu}_{\tau-1}$. Note that

$$\ln \mid D_{\tau-1}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \mid = \sum_{i=1}^{m} \ln\left[\sigma_{i,\tau-1}\left(\lambda_{1i}, \lambda_{2i}\right)\right]$$

Under the specification based on devolatized returns, $\mathbf{R}_{t-1}$ does not depend on $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$, but depends on ϕ_1 and ϕ_2, and p, the lag order used in the devolatization process. Under the specification based on exponentially weighted returns, $\mathbf{R}_{t-1}$ depends on $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ as well as on ϕ_1 and ϕ_2.

The ML estimate of $\boldsymbol{\theta}$ based on the sample observations, $\mathbf{r}_s, \mathbf{r}_2, ..., \mathbf{r}_T$, can now be computed by maximization of $l_t\left(\boldsymbol{\theta}\right)$ with respect to $\boldsymbol{\theta}$, which we denote by $\hat{\boldsymbol{\theta}}_t$. More specifically

$$\hat{\boldsymbol{\theta}}_t = Arg\max_{\boldsymbol{\theta}}\left\{l_t\left(\boldsymbol{\theta}\right)\right\} \tag{23.30}$$

for $t = T, T+h, T+2h, ..., T+N$,where h is the (estimation) update frequency, and N is the length of the evaluation sample (see Section 23.2.1). The standard errors of the ML estimates are computed using the asymptotic formula

$$\widehat{Cov}(\hat{\boldsymbol{\theta}}_t) = \left\{\sum_{\tau=s}^{t}\left[\frac{-\partial^2 f_\tau\left(\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_t}\right\}^{-1}$$

In practice the simultaneous estimation of all the parameters of the DDC model could be problematic, since it can encounter convergence problems, or could lead to a local maxima of the likelihood function. When the returns are conditionally Gaussian one could simplify (at the expense of some loss of estimation efficiency) the computations by adopting Engle's two-stage estimation procedure. But in the case of t-distributed returns the use of such a two-stage procedure could lead to contradictions. For example, estimation of separate t-$GARCH(1,1)$ models for individual asset returns can lead to different estimates of v, while the multivariate t-distribution requires v to be the same across all assets.[11]

23.2.3 Simple diagnostic tests of the DCC model

In the following, we assume that the $m \times 1$ vector of returns $\mathbf{r}_t$ follows a multivariate Student's t-distribution, though the same line of reasoning applies in the case of Gaussian returns. Consider a portfolio based on m assets with returns $\mathbf{r}_t$, using a $m \times 1$ vector of predetermined weights, $\mathbf{w}_{t-1}$. The return on this portfolio is given by

$$\rho_t = \mathbf{w}_{t-1}'\mathbf{r}_t \tag{23.31}$$

Suppose that we are interested in computing the capital Value at Risk (VaR) of this portfolio expected at the close of business on day $t-1$ with probability $1-\alpha$, which we denote by $VaR(\mathbf{w}_{t-1},\alpha)$. For this purpose we require that

$$\Pr\left[\mathbf{w}_{t-1}'\mathbf{r}_t < -VaR(\mathbf{w}_{t-1},\alpha)\ |\Omega_{t-1}\right] \leq \alpha$$

[11] Marginal distributions associated with a multivariate t-distribution with v degrees of freedom are also t-distributed with the same degrees of freedom.

Under our assumptions, conditional on Ω_{t-1}, $\mathbf{w}'_{t-1}\mathbf{r}_t$ has a Student t-distribution with mean $\mathbf{w}'_{t-1}\boldsymbol{\mu}_{t-1}$, variance $\mathbf{w}'_{t-1}\boldsymbol{\Sigma}_{t-1}\mathbf{w}_{t-1}$, and degrees of freedom v. Hence

$$z_t = \sqrt{\frac{v}{v-2}}\left(\frac{\mathbf{w}'_{t-1}\mathbf{r}_t - \mathbf{w}'_{t-1}\boldsymbol{\mu}_{t-1}}{\sqrt{\mathbf{w}'_{t-1}\boldsymbol{\Sigma}_{t-1}\mathbf{w}_{t-1}}}\right)$$

conditional on Ω_{t-1}, will also have a Student t-distribution with v degrees of freedom. It is easily verified that $E(z_t|\Omega_{t-1}) = 0$, and $V(z_t|\Omega_{t-1}) = v/(v-2)$. Denoting the cumulative distribution function of a Student's t with v degrees of freedom by $F_v(z)$, $VaR(\mathbf{w}_{t-1},\alpha)$ will be given a solution to

$$F_v\left(\frac{-VaR(\mathbf{w}_{t-1},\alpha) - \mathbf{w}'_{t-1}\boldsymbol{\mu}_{t-1}}{\sqrt{\frac{v-2}{v}\left(\mathbf{w}'_{t-1}\boldsymbol{\Sigma}_{t-1}\mathbf{w}_{t-1}\right)}}\right) \leq \alpha$$

But since $F_v(z)$ is a continuous and monotonic function of z, we have

$$\frac{-VaR(\mathbf{w}_{t-1},\alpha) - \mathbf{w}'_{t-1}\boldsymbol{\mu}_{t-1}}{\sqrt{\frac{v-2}{v}\left(\mathbf{w}'_{t-1}\boldsymbol{\Sigma}_{t-1}\mathbf{w}_{t-1}\right)}} = F_v^{-1}(\alpha) = -c_\alpha$$

where c_α is the α per cent critical value of a Student t-distribution with v degrees of freedom. Therefore,

$$VaR(\mathbf{w}_{t-1},\alpha) = \tilde{c}_\alpha\sqrt{\left(\mathbf{w}'_{t-1}\boldsymbol{\Sigma}_{t-1}\mathbf{w}_{t-1}\right)} - \mathbf{w}'_{t-1}\boldsymbol{\mu}_{t-1} \tag{23.32}$$

where $\tilde{c}_\alpha = c_\alpha\sqrt{\frac{v-2}{v}}$.

Following Engle and Manganelli (2004), a simple test of the validity of t-DCC model can be computed recursively using the indicator statistics

$$d_t = I\left(\mathbf{w}'_{t-1}\mathbf{r}_t + VaR(\mathbf{w}_{t-1},\alpha)\right) \tag{23.33}$$

where $I(A)$ is an indicator function, equal to unity if $A > 0$ and zero otherwise. These indicator statistics can be computed in-sample, or preferably can be based on recursive out-of-sample one-step ahead forecast of $\boldsymbol{\Sigma}_{t-1}$ and $\boldsymbol{\mu}_{t-1}$, for a given (pre-determined set of portfolio weights, $\mathbf{w}_{t-1}$). In such an out–of-sample exercise the parameters of the mean returns and the volatility variables ($\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, respectively) could be either kept fixed at the start of the evaluation sample, or changed with an update frequency of h periods (for example, with $h = 5$ for weekly updates, or $h = 20$ for monthly updates). For the evaluation sample, $\mathcal{S}_{\text{eval}} = \{\mathbf{r}_t,\, t = T+1, T+2, ..., T+N\}$, the mean hit rate is given by

$$\hat{\pi}_N = \frac{1}{N}\sum_{t=T+1}^{T+N} d_t \tag{23.34}$$

Under the t-DCC specification, $\hat{\pi}_N$ will have mean $1-\alpha$ and variance $\alpha(1-\alpha)/N$, and the standardized statistic,

$$z_\pi = \frac{\sqrt{N}\left[\hat{\pi}_N - (1-\alpha)\right]}{\sqrt{\alpha(1-\alpha)}} \tag{23.35}$$

will have a standard normal distribution for a sufficiently large evaluation sample size, N. This result holds irrespective of whether the unknown parameters are estimated recursively or fixed at the start of the evaluation sample. In such cases the validity of the test procedure requires that $N/T \to 0$ as $(N,T) \to \infty$. For further details on this statistic, see Pesaran, Schleicherc, and Zaffaroni (2009).

The z_π statistic provides evidence on the performance of $\mathbf{\Sigma}_{t-1}$ and $\boldsymbol{\mu}_{t-1}$ in an average (unconditional) sense. An alternative conditional evaluation procedure can be based on probability integral transforms:

$$\hat{U}_t = F_v\left(\frac{\mathbf{w}'_{t-1}\mathbf{r}_t - \mathbf{w}'_{t-1}\hat{\boldsymbol{\mu}}_{t-1}}{\sqrt{\frac{v-2}{v}\mathbf{w}'_{t-1}\hat{\mathbf{\Sigma}}_{t-1}\mathbf{w}_{t-1}}}\right), \; t = T+1, T+2, ..., T+N \tag{23.36}$$

Under the null hypothesis of correct specification of the t-DCC model, the probability transform estimates, $\hat{U}_t$, are serially uncorrelated and uniformly distributed over the range $(0,1)$. Both of these properties can be readily tested. The serial correlation property of $\hat{U}_t$ can be tested by Lagrange Multiplier tests using OLS regressions of Z_t on an intercept, and the lagged values $\hat{U}_{t-1}, \hat{U}_{t-2}, ..., \hat{U}_{t-s}$, where the maximum lag length, s, can be selected by using the AIC criterion. The uniformity of the distribution of $\hat{U}_t$ over t can be tested using the Kolmogorov-Smirnov statistic defined by $KS_N = \sup_x \left|F_{\hat{U}}(x) - U(x)\right|$, where $F_{\hat{U}}(x)$ is the empirical cumulative distribution function (CDF) of the $\hat{U}_t$, for $t = T+1, T+2, ..., T+N$, and $U(x) = x$ is the CDF of $IIDU[0,1]$. Large values of the Kolmogorov-Smirnov statistic, KS_N, indicate that the sample CDF is not similar to the hypothesized uniform CDF.[12]

23.2.4 Forecasting volatilities and conditional correlations

Having obtained the recursive ML estimates, $\hat{\boldsymbol{\theta}}_t$, given by (23.30), the following one step-ahead forecasts can be obtained. For volatilities we have

$$V\widehat{(r_{i,T+1} \mid \Omega_T)} = \hat{\sigma}^2_{i,T} = \bar{\sigma}^2_{i,T}\left(1 - \hat{\lambda}_{1i,T} - \hat{\lambda}_{2i,T}\right) + \hat{\lambda}_{1i,T}\hat{\sigma}^2_{i,T-1} + \hat{\lambda}_{2i,T}r^2_{iT}$$

where $\bar{\sigma}^2_{i,T}$ is the estimate of the unconditional mean of r^2_{it}, computed as

$$\bar{\sigma}^2_{i,T} = T^{-1}\sum_{\tau=1}^{T} r^2_{i\tau}$$

$\hat{\lambda}_{1i,T}$ and $\hat{\lambda}_{2i,T}$ are the ML estimates of λ_{1i} and λ_{2i} computed using the observations over the estimation sample $\mathcal{S}_{est} = \{\mathbf{r}_t, \, t = s, s+1, ..., T\}$, and $\hat{\sigma}^2_{i,T-1}$ is the ML estimate of $\sigma^2_{i,T-1}$, based on the estimates $\bar{\sigma}^2_{i,T-1}$, $\hat{\lambda}_{1i,T-1}$ and $\hat{\lambda}_{2i,T-1}$.

Similarly, the one step-ahead forecast of $\rho_{ij,T}$ (using either exponentially weighted returns (23.23) or devolatilized returns (23.24)) is given by

$$\hat{\rho}_{ij,T}\left(\phi\right) = \frac{\hat{q}_{ij,T}}{\sqrt{\hat{q}_{ii,T}\hat{q}_{jj,T}}}$$

[12] For details of the Kolmogorov-Smirnov test and its critical values see, for example, Neave and Worthington (1992), pp.89-93.

where

$$\hat{q}_{ij,T} = \bar{\rho}_{ij,T}(1 - \hat{\phi}_{1T} - \hat{\phi}_{2T}) + \hat{\phi}_{1T}\hat{q}_{ij,T-1} + \hat{\phi}_{2,T}\tilde{r}_{i,T}\tilde{r}_{j,T}$$

As before, $\hat{\phi}_{1T}$ and $\hat{\phi}_{2T}$ are the ML estimates of ϕ_{1T} and ϕ_{2T} computed using the estimation sample, and $\hat{q}_{ij,T-1}$ is the ML estimate of $q_{ij,T-1}$, based on the estimates $\bar{\rho}_{ij,T-1}$,$\hat{\phi}_{1T-1}$ and $\hat{\phi}_{2T-1}$.

Part VI

Appendices

Appendix A

Size Limitations

This version of *Microfit* is subject to the following limits:

(a) **At the data processing stage**

NV	=	Total number of variables ≤ 200
NO	=	Total number of observations $\leq$ 5,000,000

(b) **At the linear and non-linear estimation and hypothesis testing stage**

IK	=	Total number of regressors (including an intercept if one is included) ≤ 102
IN	=	Total number of observations in the estimation period $\leq 3,000$
IP	=	Total number of parameters of the autoregressive (AR) error process ≤ 12
$Stack$	=	Size of the stack for the specification of the non-linear restrictions or equations ≤ 600. This allows typing 600 'items' in the screen editor provided. An item could be a bracket, a variable name, an arithmetic operator or a function name.
MAR	=	Maximum order of the AR error process ≤ 50
MMA	=	Maximum order of the MA error process ≤ 12
INP	=	Number of observations used in the predictive failure/structural stability tests $IN + INP \leq 3,000$; $INP \leq 500$ $(IP + IK) * (IN - MAR) \leq 150,000$

(c) **At the forecasting stage**

For non-linear least squares, non-linear two-stage least squares, $ARDL$, Unrestricted VAR, Cointegrating VAR and $SURE$ estimation:

$IQ =$ Number of observations in the forecast period ≤ 500
$IN + IQ \leq 3,100$

For all other options:

$IQ \leq Min(500, 3,000/IK)$
$IN + IQ \leq 3,100$

(d) **At the cointegrating VAR estimation stage**

q = Number of $I(0)$ variables ≤ 18
m = Number of endogeneous $I(1)$ variables ≤ 12
k = Number of exogeneous $I(1)$ variables ≤ 5
MP = Maximum lag order of the VAR model

$$MP \leq Min\left(\frac{100-s-m}{m+k}, 24\right)$$

(e) **Unrestricted VAR**

q = Number of deterministic/exogeneous variables ≤ 18
m = Number of variables in the VAR ≤ 12

$$MP \leq Min\left(\frac{102-s}{m}, 24\right)$$

(f) $ARDL$ **Estimation**

s = Number of deterministic/fixed lag regressors ≤ 18
m = Number of variables in the $ARDL$ model ≤ 10

$$MP \leq Min\left(\frac{102-s}{m}, 24\right)$$

(g) $SURE$ **Estimation**

NEQ = Number of equations in the $SURE$ model ≤ 10
$MPARM$ = Maximum number of total parameters in the $SURE$ model ≤ 200
$MPEQ$ = Maximum number of regressors in each equation ≤ 102

(h) $GARCH$, $EGARCH$ and $AGARCH$ **Estimation**

MAR = Maximum order of AR ≤ 99
MMA = Maximum order of MA ≤ 99
Maximum number of parameters in the AR part ≤ 12
Maximum number of parameters in the MA part ≤ 12
Maximum number of parameters included in the $GARCH$ model (inclusive of the number of parameters in regression) ≤ 102

Appendix B

Statistical Tables

B.1 Upper and lower bound F-test and W-test critical values of Pesaran, Shin and Smith single-equation cointegration test

The critical value bounds reported in Tables B.1 and B.2 below are computed using stochastic simulation for $T = 500$ and $20,000$ replications in the case of Wald- and F-statistic for testing the joint null hypothesis of $\phi = \gamma_1 = \gamma_2 = ... = \gamma_k = 0$ in the following models:

Case I: No trend and no intercept

$$\Delta y_t = \phi y_{t-1} + \sum_{i=1}^{k} \gamma_i x_{i,t-1} + u_t,$$

Case II: With intercept, but without a trend

$$\Delta y_t = a_0 + \phi y_{t-1} + \sum_{i=1}^{k} \gamma_i x_{i,t-1} + u_t,$$

Case III: With an intercept and a linear trend

$$\Delta y_t = a_0 + a_1 t + \phi y_{t-1} + \sum_{i=1}^{k} \gamma_i x_{i,t-1} + u_t,$$

where $t = 1, ..., T$, and k is the number of the forcing variables.

The critical values for $k = 0$ are the same as the square of the critical values of the Dickey-Fuller unit root t-statistic. The columns headed '$I(0)$' refer to the lower-bound critical values, computed when all the k regressors are $I(0)$, and the figures in the columns headed '$I(1)$' refer to the upper-bound critical values, and are computed assuming all the k regressors are $I(1)$.

When using ARDL option in Microfit 5 the program automatically computes the critical value bounds using stochastic simulations following a procedure similar to the above. These simulated critical values are close to the ones provide in the following tables, but have the advantage that unlike the tabulated values they continue to be applicable even if shift dummy variables are included amongst the deterministic variables.

See Pesaran, Shin, and Smith (2001) for further details.

Table B.1: Testing the existence of a long-run relationship: critical value bonds of the F-statistic

Case I: no intercept and no trend								
	90%		95%		97.5%		99%	
k	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0	3.016	3.016	4.136	4.136	5.347	5.347	7.381	7.381
1	2.458	3.342	3.145	4.153	3.893	4.927	5.020	6.006
2	2.180	3.211	2.695	3.837	3.258	4.458	3.939	5.341
3	2.022	3.112	2.459	3.625	2.901	4.161	3.372	4.797
4	1.919	3.016	2.282	3.474	2.618	3.924	3.061	4.486
5	1.825	2.943	2.157	3.340	2.481	3.722	2.903	4.261
6	1.760	2.862	2.082	3.247	2.367	3.626	2.744	4.124
7	1.718	2.827	2.003	3.199	2.288	3.536	2.595	3.909
8	1.678	2.789	1.938	3.133	2.198	3.445	2.481	3.826
9	1.640	2.774	1.873	3.072	2.122	3.351	2.396	3.725
10	1.606	2.738	1.849	3.026	2.076	3.291	2.319	3.610
Case II: intercept and no trend								
	90%		95%		97.5%		99%	
k	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0	6.597	6.597	8.199	8.199	9.679	9.679	11.935	11.935
1	4.042	4.788	4.934	5.764	5.776	6.732	7.057	7.815
2	3.182	4.126	3.793	4.855	4.404	5.524	5.288	6.309
3	2.711	3.800	3.219	4.378	3.727	4.898	4.385	5.615
4	2.425	3.574	2.850	4.049	3.292	4.518	3.817	5.122
5	2.262	3.367	2.649	3.805	3.056	4.267	3.516	4.781
6	2.141	3.250	2.476	3.646	2.823	4.069	3.267	4.540
7	2.035	3.153	2.365	3.553	2.665	3.871	3.027	4.296
8	1.956	3.085	2.272	3.447	2.533	3.753	2.848	4.126
9	1.899	3.047	2.163	3.349	2.437	3.657	2.716	3.989
10	1.840	2.964	2.099	3.270	2.331	3.569	2.607	3.888
Case III: intercept and trend								
	90%		95%		97.5%		99%	
k	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0	9.830	9.830	11.722	11.722	13.503	13.503	16.133	16.133
1	5.649	6.335	6.606	7.423	7.643	8.451	9.063	9.786
2	4.205	5.109	4.903	5.872	5.672	6.554	6.520	7.584
3	3.484	4.458	4.066	5.119	4.606	5.747	5.315	6.414
4	3.063	4.084	3.539	4.667	4.004	5.172	4.617	5.786
5	2.782	3.827	3.189	4.329	3.573	4.782	4.011	5.331
6	2.578	3.646	2.945	4.088	3.277	4.492	3.668	4.978
7	2.410	3.492	2.752	3.883	3.044	4.248	3.418	4.694
8	2.290	3.383	2.604	3.746	2.882	4.081	3.220	4.411
9	2.192	3.285	2.467	3.614	2.723	3.898	3.028	4.305
10	2.115	3.193	2.385	3.524	2.607	3.812	2.885	4.135

Table B.2: Testing the existence of a long-run relationship: critical value bonds of the W-statistic

Case I: no intercept and no trend

	90%		95%		97.5%		99%	
k	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0	3.016	3.016	4.136	4.136	5.347	5.347	7.381	7.381
1	4.916	6.684	6.291	8.307	7.786	9.853	10.040	12.011
2	6.541	9.632	8.086	11.512	9.774	13.374	11.816	16.023
3	8.086	12.449	9.836	14.501	11.603	16.645	13.489	19.189
4	9.593	15.078	11.412	17.370	13.092	19.622	15.305	22.429
5	10.949	17.657	12.940	20.042	14.888	22.330	17.417	25.565
6	12.323	20.036	14.575	22.729	16.566	25.385	19.207	28.866
7	13.742	22.616	16.025	25.590	18.301	28.290	20.759	31.272
8	15.100	25.105	17.444	28.196	19.779	31.003	22.325	34.434
9	16.405	27.738	18.730	30.724	21.215	33.509	23.958	37.245
10	17.671	30.116	20.339	33.289	22.839	36.203	25.507	39.715

Case II: intercept and no trend

	90%		95%		97.5%		99%	
k	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0	6.597	6.597	8.199	8.199	9.679	9.679	11.935	11.935
1	8.085	9.576	9.867	11.528	11.552	13.463	14.114	15.630
2	9.546	12.378	11.380	14.566	13.211	16.571	15.864	18.926
3	10.844	15.199	12.875	17.512	14.907	19.591	17.540	22.460
4	12.124	17.868	14.252	20.247	16.460	22.591	19.085	25.612
5	13.569	20.205	15.896	22.831	18.339	25.601	21.097	28.689
6	14.989	22.751	17.330	25.520	19.760	28.486	22.868	31.783
7	16.279	25.223	18.920	28.421	21.322	30.965	24.215	34.367
8	17.601	27.766	20.448	31.021	22.797	33.774	25.634	37.136
9	18.993	30.466	21.634	33.488	24.368	36.574	27.158	39.891
10	20.238	32.609	23.087	35.967	25.640	39.262	28.673	42.766

Case III: intercept and trend

	90%		95%		97.5%		99%	
k	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0	9.830	9.830	11.722	11.722	13.503	13.503	16.133	16.133
1	11.299	12.670	13.212	14.847	15.286	16.902	18.126	19.571
2	12.616	15.326	14.710	17.617	17.017	19.661	19.561	22.752
3	13.936	17.831	16.264	20.477	18.423	22.989	21.259	25.655
4	15.316	20.420	17.694	23.335	20.022	25.861	23.085	28.932
5	16.690	22.963	19.135	25.971	21.441	28.692	24.066	31.984
6	18.047	25.521	20.614	28.617	22.942	31.443	25.678	34.844
7	19.282	27.936	22.013	31.065	24.354	33.984	27.347	37.553
8	20.611	30.443	23.432	33.715	25.940	36.727	28.979	39.697
9	21.924	32.846	24.666	36.138	27.225	38.985	30.280	43.050
10	23.262	35.126	26.240	38.760	28.682	41.928	31.738	45.482

Part VII

References

Acton, F. S. (1970). *Numerical Methods That Work*. New York: Harper and Row.

Affandi, J. (2007). *A Small Monetary System for Indonesia: A Long Run Structural Approach*. Ph. D. thesis, Unpublished Ph.D. Thesis, University of Cambridge.

Akaike, H. (1973). Information Theory and the Extension of the Maximum Likelihood Principle. In B. Petrov and B. F. Csaki (Eds.), *Proceeding of the Second International Symposium on Information Theory*, pp. 267–281.

Akaike, H. (1974). A New Look at the Statistical Identification Model. *IEEE: Transactions on Automatic Control 19*, 716–723.

Albert, A. and J. A. Anderson (1984). On the Existence of Maximum Likelihood Estimates in Logistic Regression Models. *Biometrika 71*, 1–10.

Almon, S. (1965). The Distributed Lag Between Capital Appropriations and Expenditures. *Econometrica 33*, 178–196.

Alogoskoufis, G. S. and R. Smith (1991a). On Error Correction Models: Specification, Interpretation, Estimation. *Journal of Economic Surveys 5*, 97–128.

Alogoskoufis, G. S. and R. Smith (1991b). The Phillips Curve, The Persistence of Inflation, and the Lucas Critique: Evidence from Exchange-Rate Regimes. *American Economic Review 81*, 1254–1275.

Amemiya, T. (1974). The Nonlinear Two-Stage Least Squares Estimator. *Journal of Econometrics 2*, 105–110.

Amemiya, T. (1980). Selection of Regressors. *International Economic Review 21*, 331–354.

Amemiya, T. (1981). Qualitative Response Models: A Survey. *Journal of Economic Literature 19*, 1483–1536.

Amemiya, T. (1985). *Advanced Econometrics*. Oxford: Basil Blackwell.

Andersen, T., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The Distribution of Realized Stock Return Volatility. *Journal of Financial Economics 61*, 43–76.

Andersen, T., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association 96*, 42–55.

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis* (3rd ed.). Wiley.

Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica 59*, 817–858.

Andrews, D. W. K. and J. C. Monahan (1992). An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator. *Econometrica 60*, 953–966.

Anscombe, F. J. (1961). Examination of Residuals. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability 4*, Berkeley, pp. 1–36. University of California Press.

Attfield, C. L. F. and J. R. W. Temple (2003). Measuring trend output: how useful are the great ratios? Unpublished Manuscript, University of Bristol.

Bai, J. and S. Ng (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica 70*, 191–221.

Banerjee, A., J. J. Dolado, J. W. Galbraith, and D. Hendry (1993). *Cointegration, Error Correction and the Econometric Analysis of Non-stationary Data*. Oxford: Oxford University Press.

Bartlett, M. S. (1946). On the Theoretical Specification and Sampling Properties of Autocorrelated Time-Series. *Journal of the Royal Statistical Society Supplement, 8*, 27–41.

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006). Multivariate GARCH Models: a Survey. *Journal of Applied Econometrics 21*, 79–109.

Beach, C. M. and J. G. MacKinnon (1978). A Maximum Likelihood Procedure for Regression with Autocorrelated Errors. *Econometrica 46*, 51–58.

Belsley, D. A., E. Kuh, and R. E. Welsch (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley.

Bera, A. K. and C. M. Jarque (1981). An Efficient Large-Sample Test for Normality of Observations and Regression Residuals. Working Papers in Econometrics, Australian National University.

Bera, A. K. and M. McAleer (1989). Nested and Non-Nested procedures for Testing Linear and Log-Linear Regression Models. *Sankhya B: The Indian Journal of Statistics 21*, 212–224.

Beveridge, S. and C. R. Nelson (1981). A New Approach to the Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'. *Journal of Monetary Economics 7*, 151–174.

Bewley, R. (1979). The Direct Estimation of the Equilibrium Response in a Linear Dynamic Model. *Economics Letters 3*, 251–276.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics 69*, 542–547.

Bollerslev, T. (1990). Modelling the Coherence in Short Run Nominal Exchange Rates: a Multivariate Generalized ARCH Model. *Review of Economics and Statistics 72*, 498–505.

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics 52*, 5–59.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988). A Capital Asset Pricing Model with Time Varying Covariances. *Journal of Political Economy 96*, 116–131.

Bowman, A. W. and A. Azzalini (1997). *Applied smoothing techniques for data analysis : the kernel approach with S-Plus illustrations*. Clarendon.

Box, G. E. P. and G. M. Jenkins (1970). *Time Series Analysis: Forecasting and Control* (Revised edition, 1976 ed.). San Francisco: Holden-Day.

Box, G. E. P. and D. A. Pierce (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated-Moving Average Time Series Models. *Journal of American Statistical Association 65*, 1509–1526.

Breen, W., L. R. Glosten, and R. Jagannathan (1989). Economic Significance of Predictable Variations in Stock Index Returns. *Journal of Finance 44*, 1177–1189.

Brent, R. P. (1973). *Algorithms for Minimization Without Derivatives*. New Jersey: Prentice-Hall, Englewood Cliffs.

Breusch, T. S. and L. G. Godfrey (1981). A Review of Recent Work on Testing for Autocorrelation in Dynamic Simultaneous Models. In D. Currie, R. Nobay, and D. Peel (Eds.), *Macroeconomic Analysis: Essays in Macroeconomics and Econometrics*. Croom Helm.

Breusch, T. S. and A. R. Pagan (1980). The Lagrange Multiplier Test and its Application to Model Specifications in Econometrics. *Review of Economic Studies 47*, 239–253.

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods* (2nd Edition ed.). New York: Springer.

Brown, R. L., J. Durbin, and J. M. Evans (1975). Techniques for Testing the Constancy of Regression Relations Over Time with Discussion. *Journal of the Royal Statistical Society, Series B 37*, 149–192.

Campbell, J., A. Lo, and A. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton: Princeton University Press.

Campbell, J. Y. (1987). Stock Returns and the Term Structure. *Journal of Financial Economics 18*, 373–399.

Campbell, J. Y. and N. G. Mankiw (1987). Are Output Fluctuations Transitory? *Quarterly Journal of Economics 102*, 857–880.

Canova, F. (1995). Vector autoregressive models: Specification, estimation, inference and forecasting. In M. Pesaran and M. Wickens (Eds.), *Handbook of Applied Econometrics*. Basil Blackwell.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research 1*, 245–276.

Champernowne, D. G. (1960). An Experimental Investigation of the Robustness of Certain Procedures for Estimating Means and Regression coefficients. *Journal of the Royal Statistical Society, Series A 123*, 398–412.

Chatfield, C. (2003). *The Analysis of Time Series: an Introduction* (6th Edition ed.). London: Chapman and Hall.

Chesher, A. and I. Jewitt (1987). The Bias of a Heteroskedasticity Consistent Covariance Matrix Estimator. *Econometrica 55*, 1217–1222.

Cheung, Y. and K. S. Lai (1993). Finite-Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration. *Oxford Bulletin of Economics and Statistics 55*, 315–328.

Chow, G. C. (1960). Test of Equality Between Sets of Coefficients in Two Linear Regression. *Econometrica 28*, 591–605.

Cobb, C. W. and P. H. Douglas (1928). A Theory of Production. *American Economic Review 18*, 139–165.

Cochrane, D. and G. H. Orcutt (1949). Application of Least Squares Regression to Relationship Containing Autocorrelated Error Terms. *Journal of the American Statistical Association 44*, 32–61.

Cochrane, J. H. (1988). How Big is the Random Walk Component in GNP? *Journal of Political Economy 96*, 893–920.

Cogley, J. (1995). Effects of Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research. *Journal of Economic Dynamics and Control 19*, 253–278.

Cook, R. D. and S. Weisberg (1982). *Residuals and Influence in Regression*. New York: Chapman and Hall.

Cox, D. R. (1961). Tests of separate families of hypotheses. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, Berkeley. University of California Press.

Cox, D. R. (1962). Further Results on Tests of Separate Families of Hypotheses. *Journal of the Royal Statistical Society, Series B 24*, 406–424.

Cramer, J. S. (1991). *The Logit Model: An Introduction for Economists*. London: Edward Arnold.

D'Agostino, R. B., A. Belanger, and R. B. D'Agostino Jr. (1990). A Suggestion for Using Powerful and Informative Tests of Normality. *The American Statistician 44*, 316–321.

Dastoor, N. K. (1983). Some Aspects of Testing Non-Nested Hypotheses. *Journal of Econometrics 21*, 213–228.

Davidson, R. and J. G. MacKinnon (1981). Several Tests for Model Specification in the Presence of Alternative Hypothesis. *Econometrica 49*, 781–793.

Davidson, R. and J. G. MacKinnon (1984). Model Specification Tests Based on Artificial Linear Regressions. *International Economic Review 25*, 485–502.

Davidson, R. and J. G. MacKinnon (1993). *Estimation and Inference in Econometrics*. Oxford University Press.

Deaton, A. S. (1977). Involuntary Saving Through Unanticipated Inflation. *American Economic Review 67*, 899–910.

Deaton, A. S. (1982). Model selection procedures, or does the consumption function exist? In G. Chow and P. Corsi (Eds.), *Evaluating the Reliability of Macroeconometric Models*, New York. John Wiley.

Deaton, A. S. and J. Muellbauer (1980). An Almost Ideal Demand System. *American Economic Review 70*, 312–336.

Dees, S., F. di Mauro, M. H. Pesaran, and L. V. Smith (2007). Exploring The International Linkages Of The Euro Area: A Global Var Analysis. *Journal of Applied Econometrics 22*, 1–38.

Dhrymes, P. J. (1971). *Distributed Lags: Problems of Estimation and Formulation*. San Francisco: Holden Day.

Dickey, D. and W. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association 74*, 427–431.

Diebold, F. X. and T. Schuermann (1992). Exact Maximum Likelihood Estimation of ARCH Models. Unpublished Manuscript, Department of Economics, University of Pennsylvania.

Dufour, J. M. (1980). Dummy Variables and Predictive Tests for Structural Change. *Economics Letters 6*, 241–247.

Durbin, J. (1970). Testing for Serial Correlation in Least-Squares Regression when Some of the Regressors are Lagged Dependent Variables. *Econometrica 38*, 410–421.

Durbin, J. and G. S. Watson (1950). Testing for Serial Correlation in Least Squares Regression I. *Biometrika 37*, 409–428.

Durbin, J. and G. S. Watson (1951). Testing for Serial Correlation in Least Squares Regression II. *Biometrika 38*, 159–178.

Eicker, F. (1963). Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions. *The Annals of Mathematical Statistics 34*, 447–456.

Eicker, F., L. M. LeCam, and J. Neyman (1967). Limit theorems for regressions with unequal and dependent errors. In *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, University of California, Berkeley, pp. 59–82.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica 64*, 813–836.

Embrechts, P., A. Hoing, and A. Juri (2003). Using Copulas to Bound VaR for Functions of Dependent Risks. *Finance and Stochastics 7*, 145–167.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica 50*, 987–1007.

Engle, R. F. (1995). *ARCH: Selected Readings*. Oxford: Oxford University Press.

Engle, R. F. (2002). Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. *Journal of Business Economics & Statistics 20*, 339–350.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica 55*, 251–276.

Engle, R. F. and C. W. J. Granger (1991). *Long-Run Economic Relationships: Readings in Cointegration*. Oxford: Oxford University Press.

Engle, R. F., D. F. Hendry, and J. F. Richard (1983). Exogeneity. *Econometrica 51*, 277–304.

Engle, R. F. and K. F. Kroner (1995). Multivariate Simultaneous Generalized Arch. *Econometric Theory 11*, 122–150.

Engle, R. F., D. M. Lillien, and R. P. Robins (1987). Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica 55*, 391–407.

Engle, R. F. and S. Manganelli (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics 22*, 367–381.

Engle, R. F. and B. S. Yoo (1987). Forecasting and Testing in Co-integrated Systems. *Journal of Econometrics 35*, 143–159.

Evans, G. (1989). Output and Unemployment Dynamics in the United States: 1950-1985. *Journal of Applied Econometrics 4*, 213–237.

Evans, G. and L. Reichlin (1994). Information, forecasts, and measurement of the business cycle. *Journal of Monetary Economics 33*, 233–254.

Fabiani, S. (1995). Technological Change and Output Fluctuations: An Empirical analysis for the G7 Countries. Unpublished Ph.D. Thesis, University of Cambridge.

Fama, E. F. and K. R. French (1989). Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics 25*, 23–49.

Fisher, G. R. and M. McAleer (1981). Alternative Procedures and Associated Tests of Significance for Non-Nested Hypotheses. *Journal of Econometrics 16*, 103–119.

Fuller, W. A. (1996). *Introduction to Statistical Time Series* (2nd Edition ed.). New York: Wiley.

Gallant, A. R. (1987). *Nonlinear Statistical Models*. New York: John Wiley.

Garratt, A., D. Robertson, and S. Wright (2005). Permanent vs transitory components and economic fundamentals. *Journal of Applied Economics 21*, 521–542.

Garratt, T., K. Lee, M. H. Pesaran, and Y. Shin (2006). *Global and National Macroeconometric Modelling: A Long Run Structural Approach*. Oxford: Oxford University Press.

Glosten, C. R., R. Jagannathan, and D. E. Runkle (1993). On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance 48*, 1779–1801.

Godfrey, L. G. (1978a). A Note on the Use of Durbin's h Test When the Equation is Estimated by Instrumental Variables. *Econometrica 46*, 225–228.

Godfrey, L. G. (1978b). Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables. *Econometrica 46*, 1293–1301.

Godfrey, L. G. (1978c). Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables. *Econometrica 46*, 1303–1310.

Godfrey, L. G. (1988). *Misspecification Test in Econometrics: The Lagrange Multiplier Principle and Other Approaches*. Cambridge: Cambridge University Press.

Godfrey, L. G. and M. H. Pesaran (1983). Test of Non-Nested Regression Models: Small Sample Adjustments and Monte Carlo Evidence. *Journal of Econometrics 21*, 133–154.

Gourierous, C., A. Holly, and A. Monfort (1982). Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters. *Econometrica 50*, 63–80.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica 37*, 424–438.

Granger, C. W. J. and P. Newbold (1974). Spurious Regressions in Econometrics. *Journal of Econometrics 2*, 111–120.

Greene, W. (2002). *Econometric Analysis* (Fifth ed.). Prentice Hall.

Gregory, A. W. and M. R. Veall (1985). Formulating Wald Tests of Nonlinear Restrictions. *Econometrica 53*, 1465–1468.

Gregory, A. W. and M. R. Veall (1987). Formulating Wald Tests of the Restrictions Implied by the Rational Expectations Hypothesis. *Journal of Applied Econometrics 2*, 61–68.

Grossman, S. J. and R. J. Shiller (1981). The Determinants of the Variability of Stock Market Prices. *American Economic Review 71*, 222–227.

Grunfeld, Y. (1960). The determinants of corporate investment. In A. Harberger (Ed.), *The Demand for Durable Goods*, pp. 211–266. Chicago: University of Chicago Press.

Grunfeld, Y. and Z. Griliches (1960). Is Aggregation Necessarily Bad? *Review of Economics and Statistics 42*, 1–13.

Hall, A. D., H. M. Anderson, and C. W. J. Granger (1992). A Cointegration Analysis of Treasury Bill Yields. *Review of Economics and Statistics 74*, 116–126.

Hall, S. R. (1978). Stochastic Implications of the Life Cycle Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy 86*, 971–987.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.

Hannan, E. J. and B. G. Quinn (1979). The Determination of the Order of an Autoregression. *Journal of the Royal Statistical Society, Series B 41*, 190–195.

Hansen, B. (1992). Efficient Estimation and Testing of Cointegrating Vectors in the Presence of Deterministic Trends. *Journal of Econometrics 53*, 87–121.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica 50*, 1029–1054.

Hansen, L. P. and K. J. Singleton (1983). Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns. *Journal of Political Economy 91*, 249–265.

Harbo, I., S. Johansen, B. Nielsen, and A. Rahbek (1998). Asymptotic Inference on Cointegrating Rank in Partial Systems. *Journal of Business & Economic Statistics 16*, 388–399.

Harvey, A. C. (1981). *The Econometric Analysis of Time Series*. London: Philip Allan.

Harvey, A. C. (1989). *Forecasting Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Harvey, A. C. and A. Jaeger (1993). Detrending, Stylized Facts and the Business Cycle. *Journal of Applied Econometrics 8*, 231–247.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica 46*, 1251–1272.

Hayashi, F. and C. Sims (1983). Nearly Efficient Estimation of Time Series Models with Predetermined, but not Exogenous, Instruments. *Econometrica 51*, 783–798.

Hendry, D. F., A. R. Pagan, and J. D. Sargan (1984). Dynamic specification. In Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics, Vol. II*, Amsterdam, pp. 1023–1100. Elsevier.

Heutschel, L. (1991). The Absolute Value GARCH Model and the Volatility of U.S. Stock Returns. Unpublished Manuscript, Princeton University.

Hildreth, C. and W. Dent (1974). An adjusted maximum likelihood estimator. In W. Sellekaert (Ed.), *Econometrics and Economic Theory: Essays in Honour of Jan Tinbergen*. London: Macmillan.

Hildreth, C. and J. Y. Lu (1960). Demand relations with autocorrelated disturbances. Technical Bulletin 276, Michigan State University Agricultural Experiment Station, East Lansing, Michigan.

Hodrick, R. and E. Prescott (1997). Post-War U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking 29*, 1–16.

Jarque, C. M. and A. K. Bera (1980). Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. *Economics Letters 6*, 255–259.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control 12*, 231–254.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica 59*, 1551–1580.

Johansen, S. (1995). *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model*. Oxford: Oxford University Press.

Johansen, S. and K. Juselius (1992). Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and UIP for UK. *Journal of Econometrics 53*, 211–244.

Jolliffe, I. T. (2004). *Principal components analysis* (2nd ed.). Springer.

Jones, M. C., J. S. Marron, and S. J. Sheather (1996). A Brief Survey of Bandwidth Selection for density estimation. *Journal of the American Statistical Association 91*, 401–407.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lütkepohl, and T. C. Lee (1985). *The Theory and Practice of Econometrics* (2nd ed.). New York: John Wiley.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement 20*, 141–151.

Kendall, M., A. Stuart, and J. K. Ord (1983). *The Advanced Theory of Statistics*, Volume 3. London: Charles Griffin & Co.

King, R. G., C. I. Plosser, S. H. Stock, and M. W. Watson (1991). Stochastic Trends and Economic Fluctuations. *American Economic Review 81*, 819–840.

Kiviet, J. F. (1986). On the Rigour of Some Misspecification Tests for Modelling Dynamic Relationships. *Review of Economic Studies 53*, 241–261.

Kmenta, J. and R. F. Gilbert (1970). Estimation of Seemingly Unrelated Regressions with Autoregressive Disturbances. *Journal of the American Statistical Association 65*, 186–197.

Koenker, R. (1981). A Note on Studentizing a Test for Heteroskedasticity. *Journal of Econometrics 17*, 107–112.

Koop, G., M. H. Pesaran, and S. M. Potter (1996). Impulse Response Analysis in Nonlinear Multivariate Models. *Journal of Econometrics 74*, 119–147.

Kullback, S. and R. A. Leibler (1951). On Information and Sufficiency. *Annals of Mathematical Statistics 22*, 79–86.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root. *Journal of Econometrics 54*, 159–178.

Lee, K. and M. H. Pesaran (1993). Persistence Profiles and Business Cycle Fluctuations in a Disaggregated Model of UK Output Growth. *Ricerche Economiche 47*, 293–322.

Leybourne, S. J. (1995). Testing for Unit Roots Using Forward and Reverse Dickey-Fuller Regressions. *Oxford Bulletin of Economics and Statistics 57*, 559–571.

Litterman, R. and K. Winkelmann (1998). *Estimating Covariance Matrices.* Risk Management Series. Goldman Sachs.

Ljung, G. M. and G. E. P. Box (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika 65*, 297–303.

Lucas, R. E. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy 1*, 19–46. Supplementary Series.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis.* Springer-Verlag.

Lütkepohl, H. and H. E. Reimers (1992). Impulse Response Analysis of Cointegrated Systems. *Journal of Economic Dynamics and Control 16*, 53–78.

MacKinnon, J. and H. White (1985). Some Heteroskedasticity-Consistent Matrix Estimators with Improved Finite Sample Properties. *Journal of Econometrics 29*, 305–325.

MacKinnon, J. G. (1991). Critical values for cointegration tests. In C. G. R.F Engle (Ed.), *Run Economic Relationships: Readings in Cointegration*, Chapter 13, pp. 267–276. Oxford University Press.

MacKinnon, J. G., H. White, and R. Davidson (1983). Tests for Model Specification in the Presence of Alternative Hypothesis: Some Further Results. *Journal of Econometrics 21*, 53–70.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Maddala, G. S. (1988). *Introduction to Econometrics.* New York: MacMillan.

Magnus, J. R. and H. Neudecker (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* New York: John Wiley and Sons.

Marcellino, M., J. Stock, and M. Watson (2006). A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series. *Journal of Econometrics 135*, 499–526.

Massey, F. J. (1951). The Kolmogorov-Smirnov test of goodness of fit. *Journal of the American Statistical Association 46*, 68–78.

McAleer, M. and M. H. Pesaran (1986). Statistical Inference in Non-Nested Econometric Models. *Applied Mathematics and Computation 20*, 271–311.

McCallum, B. T. (1976). Rational Expectations and the Natural Rate Hypothesis: Some Consistent Estimates. *Econometrica 44*, 43–52.

McFadden, D. (1976). Quantal Choice Analysis: A Survey. *Annals of Economic and Social Measurement 5*, 363–390.

Mellander, E., A. Vredin, and A. Warne (1992). Stochastic Trends and Economic Fluctuations in a Small Open Economy. *Journal of Applied Econometrics 7*, 369–394.

Merton, R. C. (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics 8*, 323–361.

Mills, T. C. (2003). *Modelling Trends and Cycles in Economic Series.* Palgrave Texts in Econometrics.

Mizon, G. E. and J. F. Richard (1986). The Encompassing Principle and its Application to Testing Non-Nested Hypotheses. *Econometrica 54*, 657–678.

Muellbauer, J. (1983). Surprises in the Consumption Function. *Economic Journal, Supplement 93*, 34–50.

Nakamura, A. and M. Nakamura (1981). On the Relationships Among Several Specification Error Tests Presented by Durbin, Wu, and Hausman. *Econometrica 49*, 1583–1588.

Neave, H. R. and P. L. Worthington (1992). *Distribution-free Tests*. London: Routledge.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: a New Approach. *Econometrica 59*, 347–370.

Nelson, D. B. and C. Q. Cao (1992). Inequality Constraints in the Univariate GARCH Model. *Journal of Business and Economic Statistics 10*, 229–235.

Newey, W. K. and K. D. West (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica 55*, 703–708.

Ogaki, M. (1992). Engle's Law and Cointegration. *Journal of Political Economy 100*, 1027–1046.

Osterwald-Lenum, M. (1992). A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics. *Oxford Bulletin of Economics and Statistics 54*, 461–472.

Pagan, A. and A. Ullah (1999). *Nonparametric Econometrics*. Cambridge: Cambridge University Press.

Pagan, A. R. (1984). Econometric Issues in the Analysis of Regression with Generated Regressors. *International Economic Review 25*, 221–247.

Pagan, A. R. (1996). The Econometrics of Financial Markets. *Journal of Empirical Finance 3*, 15–102.

Pagan, A. R. and D. F. Nicholls (1984). Estimating Predictions, Prediction Errors and their Standard Deviations Using Constructed Variables. *Journal of Econometrics 24*, 293–310.

Pantula, S. G., G. Gonzalez-Farias, and W. A. Fuller (1994). A Comparison of Unit-Root Test Criteria. *Journal of Business and Economic Statistics 12*, 449–459.

Park, H. and W. Fuller (1995). Alternative Estimators and Unit Root Tests for the Autoregressive Process. *Journal of Time Series Analysis 16*, 415–429.

Park, J. Y. (1992). Canonical Cointegrating Regressions. *Econometrica 60*, 119–143.

Parks, R. W. (1967). Efficient Estimation of a System of Regression Equations when Disturbances are Both Serially and Contemporaneously Correlated. *Journal of the American Statistical Association 62*, 500–509.

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica 57*, 1361–1401.

Pesaran, B. (1988a). Exact maximum likelihood estimation of regression models with invertible general order moving disturbances. National Institute of Economic and Social Research, Discussion Paper No.136.

Pesaran, B. (1990). Instrumental Variable Estimation of Regression Models with Invertible General order Moving Average Disturbances. Unpublished Manuscript.

Pesaran, B. and M. H. Pesaran (2007). Modelling Volatilities and Conditional Correlations in Futures Markets with a Multivariate t Distribution. CESifo Working Paper No. 2056.

Pesaran, B. and G. Wright (1995). The Use of Spreads in Forecasting Medium Term U.K. Interest Rates. Working Paper, University of East London.

Pesaran, M. H. (1972). *Small Sample Estimation of Dynamic Economic Models*. Ph. D. thesis, Cambridge University.

Pesaran, M. H. (1974). On the General Problem of Model Selection. *Review of Economic Studies 41*, 153–171.

Pesaran, M. H. (1981). Diagnostic Testing and Exact Maximum Likelihood Estimation of Dynamic Models. In E. Charatsis (Ed.), *Proceedings of the Econometric Society European Meeting, 1979, Selected Econometric Papers in memory of Stefan Valvanis*, Amsterdam, pp. 63–87. North-Holland.

Pesaran, M. H. (1982a). Comparison of Local Power of Alternative Tests of Non-Nested Regression Models. *Econometrica 50*, 1287–1305.

Pesaran, M. H. (1982b). A Critique of the Proposed Tests of the Natural Rate-Rational Expectations Hypothesis. *Economic Journal 92*, 529–554.

Pesaran, M. H. (1987a). Global and Partial Non-Nested Hypotheses and Asymptotic Local Power. *Econometric Theory 3*, 69–97.

Pesaran, M. H. (1987b). *The Limits to Rational Expectations*. Oxford: Basil Blackwell. Reprinted with corrections 1989.

Pesaran, M. H. (1988b). On the Policy Ineffectiveness Proposition and a Keynesian Alternative: A Rejoinder. *Economic Journal 98*, 504–508.

Pesaran, M. H. (1991). On the Volatility and Efficiency of Stock Prices. *Cuadernos Economicos 49*, 121–155. In Spanish.

Pesaran, M. H. (1997). The Role of Economic Theory in Modelling the Long Run. *Economic Journal 107*, 178–191.

Pesaran, M. H. and R. A. Evans (1984). Inflation, Capital Gains and UK Personal Savings: 1953-81. *Economic Journal 94*, 237–257.

Pesaran, M. H. and B. Pesaran (1991). *Microfit 3.0: An Interactive Econometric Software Package*. Oxford: Oxford University Press.

Pesaran, M. H. and B. Pesaran (1993). A Simulation Approach to the Problem of Computing Cox's Statistic for Testing Non-nested Models. *Journal of Econometrics 57*, 377–392.

Pesaran, M. H. and B. Pesaran (1995). A Non-nested Test of Level-Differenced Versus Log-Differenced Stationary Models. *Econometric Reviews 14*, 213–227.

Pesaran, M. H., R. G. Pierce, and K. C. Lee (1993). Persistence, Cointegration and Aggregation: A Disaggregated Analysis of Output Fluctuations in the U.S. Economy. *Journal of Econometrics 56*, 57–88.

Pesaran, M. H., C. Schleicherc, and P. Zaffaroni (2009). Model Averaging in Risk Management with an Application to Futures Markets. *Journal of Empirical Finance 16*, 280–305.

Pesaran, M. H., T. Schuermann, and S. Weiner (2004). Modelling Regional Interdependencies using a Global Error-Correcting Macroeconometric Model. *Journal of Business and Economics Statistics 22*, 129–162.

Pesaran, M. H. and Y. Shin (1996). Cointegration and Speed of Convergence to Equilibrium. *Journal of Econometrics 71*, 117–143.

Pesaran, M. H. and Y. Shin (1998). Generalised Impulse Response Analysis in Linear Multivariate Models. *Economics Letters 58*, 17–29.

Pesaran, M. H. and Y. Shin (1999). An autoregressive distributed lag modelling approach to cointegration analysis. In S. Strom (Ed.), *Econometrics and Economic Theory in 20th Century: The Ragnar Frisch Centennial Symposium*, Chapter 11, pp. 371–413. Cambridge: Cambridge University Press.

Pesaran, M. H. and Y. Shin (2002). Long Run Structural Modelling. *Econometrics Reviews 21*, 49–87.

Pesaran, M. H., Y. Shin, and R. J. Smith (1996). Testing for the existence of a long run relationship. DAE working paper, no. 9622, University of Cambridge.

Pesaran, M. H., Y. Shin, and R. J. Smith (2000). Structural Analysis of Vector Error Correction Models with Exogenous I(1) Variables. *Journal of Econometrics 97*, 293–343.

Pesaran, M. H., Y. Shin, and R. J. Smith (2001). Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics 16*, 289–326. Special issue in honour of J D Sargan on the theme 'Studies in Empirical Macroeconometrics'.

Pesaran, M. H. and L. J. Slater (1980). *Dynamic Regression: Theory and Algorithms.* Chichester: Ellis Horwood.

Pesaran, M. H. and R. J. Smith (1990). A Unified Approach to Estimation and Orthogonality Tests in Linear Single Equation Econometric Models. *Journal of Econometrics 44*, 41–66.

Pesaran, M. H. and R. J. Smith (1994). A Generalized R2 Criterion for Regression Models Estimated by the Instrumental Variables Method. *Econometrica 62*, 705–710.

Pesaran, M. H. and R. P. Smith (1985). Evaluation of Macroeconometric Models. *Economic Modelling 2*, 125–134.

Pesaran, M. H., R. P. Smith, and S. Yeo (1985). Testing for Structural Stability and Predictive Failure: a Review. *Manchester School 53*, 280–295.

Pesaran, M. H. and A. Timmermann (1992). A Simple Nonparametric Test of Predictive Performance. *Journal of Business and Economic Statistics 10*, 461–465.

Pesaran, M. H. and A. Timmermann (1994). Forecasting Stock Returns: An Examination of Stock Market Trading in the Presence of Transaction Costs. *Journal of Forecasting 13*, 335–367.

Phillips, A. W. (1958). The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom 1861-1957. *Economica 25*, 283–299.

Phillips, P. C. B. (1986). Understanding Spurious Regressions in Econometrics. *Journal of Econometrics 33*, 311–340.

Phillips, P. C. B. and B. E. Hansen (1990). Statistical Inference in Instrumental Variables Regression with I(1) Processes. *Review of Economic Studies 57*, 99–125.

Phillips, P. C. B. and S. Ouliaris (1990). Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica 58*, 165–193.

Phillips, P. C. B. and P. Perron (1988). Testing for a Unit Root in Time Series Regression. *Biometrika 75*, 335–346.

Potter, S. M. (1995). A Nonlinear Approach to US GNP. *Journal of Applied Econometrics 10*, 109–125.

Powell, M. J. D. (1964). An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives. *The Computer Journal 7*, 155–162.

Power, S. (1990). An Exact Transformation Matrix for Use with Rational Expectations Models with MA(1) Composite Disturbances. *Oxford Bulletin of Economics and Statistics 52*, 89–94.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1989). *Numerical Recipes: The Art of Scientific Computing FORTRAN version.* Cambridge: Cambridge University Press.

Priestley, M. B. (1981). *Spectral Analysis and Time Series.* London: Academic Press.

Ramsey, J. B. (1969). Tests for Specification Errors in Classical Linear Least Squares Regression Analysis. *Journal of the Royal Statistical Society, Series B 31*, 350–371.

Ramsey, J. B. (1970). Models, Specification Error and Inference: A Discussion of Some Problems in Econometric Methodology. *Bulletin of the Oxford Institute of Economics and Statistics 32*, 301–318.

Rao, C. (1948). Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Applications to Problems of Estimation. In *Proceedings of the Cambridge Philosophical Society*, pp. 50–57.

Rao, C. R. (1970). Estimation of Heteroscedastic Variances in Linear Models. *Journal of the American Statistical Association 65*, 161–172.

Robertson, D., A. Garratt, and S. Wright (2006). Permanent vs Transitory Components and Economic Fundamentals. *Journal of Applied Econometrics 21*, 521–542.

Rush, M. and D. Waldo (1988). On the Policy Ineffectiveness Proposition and a Keynesian Alternative. *Economic Journal 98*, 498–503.

Said, E. S. and D. Dickey (1984). Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order. *Biometrika 71*, 599–607.

Salkever, D. S. (1976). The Use of Dummy Variables to Compute Predictions, Prediction Errors and Confidence Intervals. *Journal of Econometrics 4*, 393–397.

Sargan, J. D. (1958). The Estimation of Economic Relationships Using Instrumental Variables. *Econometrica 26*, 393–415.

Sargan, J. D. (1959). The Estimation of Relationships with Autocorrelated Residuals by the Use of Instrumental Variables. *Journal of the Royal Statistical Society, Series B 21*, 91–105.

Sargan, J. D. (1964). Econometrics analysis for national economic planning. In P. Hart, G. Mills, and J. Whitaker (Eds.), *Wages and Prices in the United Kingdom: A Study in Econometric Methodology*. London: Butterworths.

Sargan, J. D. (1976). Testing for Misspecification after Estimation Using Instrumental Variables. Unpublished manuscript.

Satchell, S. and J. L. Knight (Eds.) (2002). *Forecasting Volatility in the Financial Markets*. Butterworth-Heinemann Finance.

Scheffe, H. (1959). *The Analysis of Variance*. New York: John Wiley.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics 6*, 461–464.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley.

Sheather, S. J. (2004). Density estimation. *Statistical Science 19*, 588–597.

Shephard, N. (2005). *Stochastic Volatility: Selected Readings*. Oxford University Press.

Silverman, B. (1986). *Density estimation for statistics and data analysis*. London: Chapman and Hall.

Sims, C. (1980). Macroeconomics and Reality. *Econometrica 48*, 1–48.

Sims, C. (1981). An autoregressive index model for the us 1948-1975. In J. Ramsey (Ed.), *Large-Scale Econometric Models*. Netherlands: North-Holland.

Spanos, A. (1989). *Statistical Foundations of Econometric Modelling*. Cambridge: Cambridge University Press.

Stock, J. M. (1987). Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. *Econometrica 55*, 1035–1056.

Stock, J. M. (1994). Unit roots, structural breaks and trends. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, pp. 2738–2841. North Holland.

Stock, J. M. and M. W. Watson (1988). Testing for Common Trends. *Journal of the American Statistical Association 83*, 1097–1107.

Söderlind, P. (1994). Cyclical Properties of a Real Business Cycle Model. *Journal of Applied Econometrics 9*, 113–122.

Tong, H. (1990). *Non-Linear Time Series*. Oxford: Oxford University Press.

Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica 57*, 307–333.

Watson, M. W. (1994). Vector autoregression and cointegration. In D. MacFadden and R. Engle (Eds.), *Handbook of Econometrics*, pp. 843–915. North Holland.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica 48*, 817–838.

White, H. (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica 50*, 1–25.

Wickens, M. R. (1982). The Efficient Estimation of Econometric Models with Rational Expectations. *Review of Economic Studies 49*, 55–67.

Wu, D. M. (1973). Alternative Tests of Independence Between Stochastic Regressors and Disturbances. *Econometrica 41*, 733–750.

Yule, G. U. (1926). Why Do We Sometimes Get Nonsense-Correlations Between Time Series? A Study in Sampling and the Nature of Time-Series. *Journal of the Royal Statistical Society 89*, 1–63.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association 57*, 348–368.

Zellner, A. and H. Theil (1962). Three Stage Least Squares: Simultaneous Estimation of Simultaneous Equations. *Econometrica 30*, 54–78.