# Time Series Modelling Version 4.52
# User's Manual

## James Davidson

*This document is also accessible through the program Help pages.*

*There is no conventional index since the text is revised regularly. The PDF file with the Search option in Acrobat Reader provide an effective indexing capability.*

# Contents

# 0. Introduction

## *0.1 TSM Basics*

TIME SERIES MODELLING Version 4 is an interactive package for modelling and forecasting time series. It is designed primarily for nonlinear dynamic model estimation, including ARMA and ARFIMA models, conditional variance models (ARCH/GARCH and several variants), regime-switching and smooth transition. It also functions well as a user-friendly, general purpose regression package.

To enhance its power and ease of use, the program has various special features not found in comparable packages. It is strongly recommended to read through this introduction before starting work with TSM. It does not aim to describe all the capabilities of the package. These can best be discovered by browsing the menus in conjunction with the Help pages provided for each. Its purpose is to outline TSM's unique design and operating conventions to first-time users.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
Working with TSM
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

1. ENTERING DATA
Regression packages often have their own proprietary database formats, and getting data into them from different sources can be troublesome. By contrast, TSM reads either ASCII files or spreadsheet files directly. It works smoothly in conjunction with popular spreadsheet programs such as Microsoft Excel, and also OxMetrics. It can itself merge datasets with different start and finish dates and use and display date information, including daily dates in Excel format. It offers flexible single-observation editing, series transformations, and dummy creation capabilities.

2. MENUS AND DIALOGS.
TSM is operated in Graphical User Interface (GUI) mode by setting the options in a number of dialogs. These can be opened from the menu bar or by shortcut buttons on the tool bar. Some operations can be performed in several different ways. For example, an estimation run can be launched from the menus, from the "Running Man" toolbar button, and also from the "Go" buttons in several dialogs. The action of the toolbar button can vary depending on the dialog currently open.

3. COMPUTING ESTIMATES, TESTS AND FORECASTS
The basic method of operation is as follows:
1) Specify the calculations and settings desired, using the Setup, Model, Values and Options menus.
2) Launch the estimation module using one of the commands on the Actions menu, a dialog button, or a tool bar button.
Both the "Running Man" and the "Calculator" buttons on the tool bar launch closed form (non-iterative) estimations, such as OLS and IV, and associated tests and forecasts. For nonlinear estimation (if enabled) the "Running Man" button launches the optimization algorithm, while the "Calculator" button just performs post-estimation computations (forecasts or tests) using the currently stored parameter values, either obtained on the latest run or entered by the user.

NOTE: forecasts and tests are not computed by merely specifying them in the

Options dialog! Use the menu items in the Actions menu for this purpose. The Calculate button performs all the currently specified calculations and generates the complete estimation output for the current parameter values.

Each run has a unique ID number to identify the outputs associated with it, such as graphics files, spreadsheet files, and settings files.

## 4. SAVING PROGRAM SETTINGS

TSM has a large number of optional settings that most users will want to change only occasionally. By default, the program automatically saves all current settings in a special file called (by default) "settings.tsm". The current data set, listings and charts are stored (by default) in "settings.tsd". When the program is restarted the working environment, with all selected options, is then exactly as it was in the last session, even if the original data files have been moved or deleted in the meantime. Named settings files (with .tsm and .tsd extensions, and 'red TSM' and 'blue TSD' Windows  icons) can be saved and re-loaded manually at any time.

## 5. ORGANIZING YOUR WORK

The File / Settings / Export... command saves a complete image of the current session, including options, model specifications, data, generated series, tables and graphics. 'Exported' .tsm files do not contain local path information and are fully portable between installations. When they are opened, the data file and temporary storage (.tsd) files are recreated. This provides an ideal way to share work with collaborators, move between home/office installations and distribute classroom exercises. Double-clicking on a .tsm file icon in Windows Explorer starts the program and loads the file contents automatically.

## 6. SIMULATION

Part of the TSM philosophy is that any model that can be estimated by the program can also be simulated, using randomly generated disturbances or bootstrapped (randomly resampled) residuals. The former can be Gaussian, or generated from the distribution specified by the selected likelihood function. This feature can be used for one-off simulations whose output is graphed. Comparing the simulation of the fitted model with the original data can be a useful informal diagnostic tool. However, the main application for the simulation module is to running bootstrap tests and Monte Carlo experiments. A flexible interactive Monte Carlo module is provided.

## 7. MODELS

A "model" is a complete set of specifications and values to estimate, simulate or forecast an equation or system of equations. Any number of  these specifications can be stored and recalled during a session, as well as saved permanently in the settings (.tsm) file. For example, this option allows the user to run an exploratory regression on the fly, while working on a complex multiple-equation model, without losing any settings and values. Just use the Model Manager to store the current settings and values to a named model (optionally including the data set). Load the model to restore them again. The generated series, graphics and (optionally) data set associated with a model are stored in a file with '.tsd' extension and 'blue TSM' Windows icon

Models are also used for running Monte Carlo experiments. Select one model

to generate the data using the simulation module, and another (or the same) model for estimating, allowing a very flexible approach to misspecification analysis.

## 8. PROGRAMMING WITH TSM

TSM can be called as a module in your own Ox program. It is easy to write out the commands and options set by TSM dialogs as lines of Ox code. Your program can call the main program functions such as "Run_Estimation" and "Run_Simulation", and perform further operations on the output.

The special scripting language is fully detailed in the programming manual. To see what your current set of (non-default) options looks like in coded form, give the command File / Settings / Display/Save Text... and  inspect the listing created. The reverse operation is also possible. Write out a set of coded commands in the TSM launch file, and these options are then set when the GUI program is started. This provides an alternative way to set and maintain your list of favourite program options.

****************
Modelling with TSM
****************

## 9. TYPES OF MODEL

There are two main model specification dialogs, called respectively Linear Regression and Dynamic Equation. The "regression scatter" and "space shuttle" toolbar buttons give direct access to these. (The latter can be optionally hidden to simplify the interface, if these features are not being used.)
 Linear Regression is used to specify linear models for estimation by a closed-form expression (or, at most, a fixed, finite sequence of calculations). These include OLS, 2SLS, SUR and 3SLS.
Although linear regressions can also be specified in the Dynamic Equation dialog, its special role is to set up nonlinear dynamic models. All estimation in this dialog, even of linear models, is done by optimizing a log-likelihood or other criterion function numerically, using the BFGS and/or simulated annealing algorithms. Model choices include Gaussian, Student t and GED continuous distributions, binary logit and probit models, and count data models. All these models can feature conditional variances, and Markov switching or smooth transition nonlinearity.

## 10. SELECTING VARIABLES

Model specification dialogs include a list of variable names, corresponding to the currently loaded data file. Selecting a variable is a two-step procedure. First select a radio button specifying the variable Type - dependent variable(s), one of several types of explanatory variables, instruments, and so forth. Next, click on the desired name in the list, to highlight it. The list can be scrolled, if it is too long to display complete in the dialog. Note that when one or more variables are selected in a category, the corresponding radio button is highlighted with a lighter-grey panel. To see which variables are  currently selected in a particular category, click on the radio button in question. A little practice helps to get variable selection smooth and rapid, and it is a good idea to click on each highlighted button in turn to check the specification is as desired, before launching the estimation run. To deselect a variable, simply click on it again to remove the highlight. There is also a "Clear"

button to remove all the current selections.
NOTE: to display time plots of one or more variables, highlight them in the list
and click the 'Data Graphics' button on the toolbar

## 11. VARIABLE "TYPES"

Explanatory variables in an equation can be of two, or three, different "Types", with
a different radio button assigned to each. In some models "typing" is irrelevant
(in which case, just choose Type 1), but it has a number of common uses. In linear
regressions, it is used to allow lags to be treated differently. Lags up to a specified
order can be included automatically, so that lagged values do not need to be created
and stored individually in the database. The number of lags is selected with the scroll
bar, for all variables of the given Type. For example, assign non-lagged variables
such as dummies to Type 1, and distributed lag regressors to Type 2. To allow easy
inclusion of the lagged dependent variable, the dependent variable in a regression
can be assigned as a Type 2 regressor with lags specified. The current value is
omitted from the regressor set automatically. This also works in system models such
as VARs.

In the Dynamic Equation dialog, "typing" regressors in combination with specifying
autoregressive and moving average components has a special additional role,
allowing an equation to feature "structural dynamics" and/or "error dynamics".  See
the relevant Help pages and the Models and Methods document for details on this.

## 12. PARAMETER VALUES

A special feature of TSM is the Values dialogs, where values and conditions for
model parameters can be set. Among the uses of these dialogs are
*        Setting starting values for numerical optimization
*        Fixing parameters at chosen values during estimation
*        Setting inequality constraints on parameters during estimation
*        Setting zero or linear restrictions on parameters, either for constrained
         estimation, or calculation of Wald tests.
The Values dialogs can be accessed from the menu bar, or by the "Values" button in
the relevant model specification dialog.

Parameter values obtained in the latest estimation run can be viewed in the Values
dialogs, and will form the starting values for the next run unless edited or cleared.
Alternatively, the menu item Evaluate at Current Values (or the "Calculator"
button) generates the program  outputs at these values, without re-optimizing.

## 13. CODED FUNCTIONS

General nonlinear models can be estimated by creating coded formulae. The
program features a formula parser which can evaluate functions of data and/or
parameters typed using standard notations. This feature is used to create
functions for estimation and parameter restrictions for testing, as well as data
transformations.

Alternatively, TSM can estimate a model programmed by the user in the Ox
language, while making use of all the estimation, testing and forecasting features
of the package. (This feature is distinct from, though compatible with, calling
TSM from within an Ox program.)

© James Davidson 2024

## 14. SYSTEMS OF EQUATIONS

Setting up a system of equations is greatly simplified by requiring that the right-hand side of every equation has the same specification. In this way, only one specification has to be created and stored. This is the natural approach for an unrestricted VAR, for example. To have the equations different from each other (e.g. with identifying restrictions imposed) the method is to create an inclusive specification of which all the actual equations are special cases, and then "fix" the surplus parameters at zero in the Values dialogs - enter 0 in the value fields, and check the 'Fixed' checkboxes. While editing values, one can switch easily from one equation to another, using the 'Next Equation' button or the choice widget on the tool bar. Parameters fixed at 0 are not reported in the output, although those fixed at non-zero values are listed as such.

Systems of coded nonlinear functions can be estimated in the same way. While their specifications can differ, each has the common set of named parameters assigned to it - "fix" the surplus ones in Values / Equation, so that the search algorithm ignores them.

Simultaneous equation systems can be specified by including variables as both dependent variables and Type 1 regressors. When their presence in both sets is detected by the program, the system is estimated by FIML. It is the user's responsibility to ensure that identifying restrictions are imposed on the equations, in this case.

In vector ARMA and GARCH models specified in the Dynamic Equation dialog, the lags of all the variables are included by default in each equation. For example, a VAR(2) system of three equations has 6 regressors in each equation. Of course, some of these regressors can be suppressed if desired, by fixing their coefficients at 0 in the Values dialogs.

Note that a VAR can be set up in either the Linear Regression dialog or the Dynamic Equation dialog. In the first case it will be estimated in two steps, by SUR (equivalently OLS, if unrestricted). In the second case it will be estimated numerically, by Least Generalized Variance or ML.

## 15. OUTPUT CONVENTIONS

The outputs from TSM are not always the same as those of other packages. t-ratios and p-values are not reported in those cases where "zero" is not the natural null hypothesis for the parameter (variances for example). The Durbin Watson statistic is not reported by default, since it is often not valid in dynamic regressions. See the Tests and Diagnostics Options dialog to select it. A valid LM statistic or M-statistic for residual autocorrelation, or neglected ARCH, can always be computed if desired, and provides an equivalent test to the DW in those cases where it is valid. Autocorrelation Q statistics for residuals and squared residuals are reported by default, as are standard model selection criteria, although these outputs can be optionally suppressed.

Robust standard errors are reported by default. These will not match the "naive" (conventional) standard error formulae reported by most packages, and note that the latter are often based on incorrect assumptions, although they can always be computed as an option. Heteroscedasticity and autocorrelation consistent (HAC)

© James Davidson 2024

standard errors are also optionally available. Test p-values and confidence intervals can also be computed by the parametric bootstrap, based on resampling model residuals and using the fitted model to simulate data under the null hypothesis.

Asymptotic chi-squared statistics are reported by default for the standard tests of restrictions and mis-specification, although reporting in 'F-statistic' form is a selectable option. (Note that in most time series applications, 'F statistics' are not truly F distributed in finite samples.) Bootstrap p-values can be computed if desired, for improved test performance in small samples. Another approach related to the bootstrap is to tabulate test statistics by simulation and use the EDF tabulations to generate p-values.

## 0.2  How To ....

For the first-time user who doesn't want to spend too much time with the user's manual, here are some simple step-by-step instructions to get you started.

1. HOW TO LOAD A DATA SET
*      EITHER: click the "Open File" button on the tool bar. In the file dialog, navigate to the Windows folder containing your data file and click on it. OR: simply drag the data file icon from the Windows Explorer window and drop it onto the TSM window.
*      Several different file formats are supported, but an Excel worksheet is a popular choice. The series should be stored in columns, starting in column 2, with the first row of the spreadsheet containing the variable names.

2. HOW TO PLOT DATA SERIES
*      Click with the mouse on the variable list in any open dialog.
*      Click the "Chart" button on the toolbar, to display plots of all the highlighted series on the list.
*      Clicking the "Chart" button without first clicking a variable list opens the Graphics / Show Data Graphic dialog.

3. HOW TO RUN A SIMPLE REGRESSION.
*      Click on the "regression scatter" button to open the Linear Regression dialog.
*      In the Select Estimator box choose Ordinary Least Squares (the default).
*      Choose the dependent variable from the list, and highlight the name by clicking on it with the mouse. At most one name can be selected at once. Note that the radio button is highlighted to show you have made a selection
*      In the "Select Regressors" panel, click on the "Type 1" radio button. Then select the regressors from the list, as for the dependent variable. Any number of names can be selected. The radio button is highlighted.
*      Click the check-box for Intercept, and, if appropriate, for Trend.
*      To ensure your selection of variables is as you intend, it's a good idea the click alternately on the highlighted radio buttons. Note how your selections on the data list are highlighted in turn.
*      Click the "Go" button in the dialog - or the "Running Man" or "Calculator" buttons on the tool bar. All have the same action, in this case.
*      The results appear in the window. To view the Actual-Fitted and Residual plots, click the "Twin-Graph" button on the tool bar.

*     To use less than the complete sample for estimation, click the "Select Sample" button to open the sample-setting dialog. Use the scroll bars to select the first and last observation.
*     When you launch an estimation run, open dialogs are closed automatically to show the results window. Click the "Windows" button on the tool bar to restore them to their previous locations on the screen.

## 4. HOW TO GENERATE QUARTERLY DUMMIES
*     Click the "f(x)" button on the tool bar.
*     Click the Edit button, then scroll down the choice widget (pull-down menu) until you find "Make Seasonals".
*     A scroll bar to choose the frequency appears. Select "4", and press Go.
*     Four dummy variables are added to the data set. Add only three of them to your regression if you have an intercept!

## 5. HOW TO TAKE LOGARITHMS OF YOUR DATA
*     Click the "f(x)" button on the tool bar.
*     Click the Transform button, then scroll down the choice widget until you find "Logarithm".
*     Highlight all the variables on the list you wish to transform.
*     Click Go. The transformed variables are added to the data set, identified with the prefix "Log-" added to the name.
*     To give a more convenient name, if desired, select Edit and Rename with the choice widget, and enter the new name in the field provided.

## 6. HOW TO TEST THE SIGNIFICANCE OF THE REGRESSION
*     The following procedure provides a valid time series implementation of the "F test of the Regression" reported routinely by many packages. It tests all exogenous regressors, but automatically excludes the trend, seasonal dummies, and lagged dependent  variables from the test set. Thus, the null hypothesis can be a valid univariate representation of the dependent variable.
*     In the Linear Regression dialog, check  "Wald Test of Constraints".
*     Use the button beside the checkbox to open the Constraints dialog, and check the box "Test Joint Significance of  Regressors".
*     The test is computed by running the regression, or by choosing Actions / Compute Test Statistics / Wald Test of Set Restrictions.

## 7. HOW TO TEST FOR A UNIT ROOT
*     Open the Setup / Compute Summary Statistics dialog.
*     Check the "Report I(1) Tests" checkbox.
*     Choose the variable you want to test and press Go.
*     The Augmented Dickey-Fuller test and Phillips-Perron statistics are among the results reported. The order of lags in the ADF test is chosen automatically to optimize the Schwarz model selection criterion over the range up to $M = O(T^{1/3})$. The bandwidth settings for the Phillips-Perron test can be changed in Options / General...

## 8. HOW TO ESTIMATE AN ARMA(p,q) MODEL
*     Click the "Space Shuttle" button on the tool bar to open the Dynamic

Equation dialog. (If this is not shown, open Options / General, check "Enable Optimization Estimators", and restart TSM.)
* In the Select Estimator box, choose Least Squares.
* Select your dependent variable from the list.
* Use the scroll bars to select the desired AR and MA orders (p and q).
* Select a Type 2 intercept (this is identified even if you have a unit root). You cannot have both types at once, so deselect the Type 1 intercept first, if necessary.
* Press "Go", or the "Running Man" button.
* If your run has been successful, you should see "Strong Convergence" in the results window.
* If you aren't sure what p and q to choose, you can have the program try each pair in succession up to a chosen maximum. Click Setup / Multiple ARMA Models... and select the maximum values you want to try. Click the Running Man with this dialog open, and see the estimates computed successively in the results window.
* You can choose a preferred specification by comparing the Akaike, Schwarz or Hannan-Quinn selection criteria for each model.
* To estimate an ARIMA(p,1,q) model check the "Impose Unit Root" box.

### 9. HOW TO FORECAST WITH AN ARMA/ARIMA MODEL
* Having selected your model, choose Options / Forecasting..., and use the scroll bar to select the number of post-sample periods to forecast. Note that you cannot forecast beyond the end of the data set if your model contains exogenous variables.
* Select the options "Ex-ante Multi-Step" and "Analytic".
* Now open the Options / Output and Retrieval dialog, and in the panel Print in Results Window, check the option "Forecasts & MA Coeffs".
* If the model has already been estimated, click the "Calculator" button on the tool bar. This will evaluate the model and forecasts. Otherwise, click "Running Man" to run the estimation and forecast calculation in one step.
* To see the point forecasts and confidence intervals graphically, select Graphics / Ex-Ante Forecasts.
* The forecasts can also be exported to a spreadsheet file. Click File / Listings / Save Forecasts.

### 10. HOW TO ESTIMATE A SIMPLE VAR MODEL
* Open the Linear Regression dialog and check the box "System of Equations". You are now allowed to select two or more dependent variables.
* In the Select Regressors panel, click on the Type 2 radio button. Make the *same* selections as you made for the dependent variables.
* Use the Lags scroll bar to choose the VAR order.
* Select an intercept, if desired. Any exogenous variables can be added to the model as regressors of Type 1.
* Press "Go", or the "Running Man" or "Calculator" buttons to estimate the model.
* Note that when endogenous variables are entered as Type 2 Regressors, the current values are automatically suppressed, making it easy to specify lagged dependent variables.

\*       A VAR or VARMA can also be estimated in the Dynamic Equation dialog.

11. HOW TO ESTIMATE A GARCH MODEL
 \*       Open the Dynamic Equations dialog.
 \*       Specify the 'mean equation' as appropriate. This could be an ARMA, or a regression model, for example.
 \*       In the Select Estimator panel, choose Gaussian ML.
 \*       Click on Conditional Variance Model. The Conditional Variance dialog box opens.
 \*       Select the GAR and GMA orders.
 \*       To see the model in the form usually reported (the 'Bollerslev form') open the Options / ML and Dynamics dialog box and, under "GARCH Settings", uncheck the first two checkboxes. (Note, the model you fit is identical in either case. Only the interpretation of the coefficients is affected by these options.)
 \*       Press the Go or Running Man buttons to estimate.
 \*       Estimation of GARCH models can sometimes be tricky. Poor starting values can cause convergence failure. See Help / Hints and Tips for advice on optimization in case of difficulty.

# 1. User Interface

## *1.1 The Menus*

Clicking on most menu items opens a dialog in which program settings can be changed. As many dialogs as desired can be open simultaneously.

* The File menu controls loading and saving of settings, data and results.

* The Setup menu controls basic operations including choice of data set and sample size, data editing and transformations, special estimation routines, preliminary data analyses, Monte Carlo experiments, a test tabulation viewer and a calculator.

* The Model menu gives access to dialogs for setting model specifications. The "Load a Model..." and "Save Current Model" menu items provide a quick link to the basic Model Manager operations.

* The Values menu is for setting starting values and selecting options for model parameters.

* The Actions menu does things. Some of these menu items open dialogs, but most perform actions directly.

* The Graphics menu is to display graphs. An estimation must be run before the equation graphics options are available.

* The Options menu is to change program settings.

* The Help menu gives access to the complete user's manual (also available in PDF format) and the list of data names and descriptions, if these have been created. It also contains commands to open the Registration dialog and "About" window.

Notes:
1. Close a dialog by clicking the "Close" button ([X]) on the menu bar.

2. Positions on the screen of closed dialogs are remembered, even between sessions. This allows the user to create a personalized work layout on the desktop. Use the Reset Dialog Positions command to centre all dialog positions.

3. Submenus and menu items are "greyed out" when the command is currently unavailable. For example: a data file must be loaded before any action apart from changing options is possible.

4. Model options (e.g. Conditional Variances, Regime Switching) can be switched on and off without altering the settings. To activate a model option either check the "Active" checkbox in the dialog, or select the option in the Model / Dynamic Equation dialog. Just setting the required options in the

© James Davidson 2024

dialog does not activate the option automatically!

5.    Values dialogs will not reflect a change in the model specification until either the dialog is refreshed or the Values button is pressed in the relevant dialog.

6.    If a dialog gets hidden behind others, bring it to the top by reselecting it from the menu bar.

## *1.2 Status Bar and Tool Bar*

The status bar at the bottom of the results window shows:
*       the name of the current data set,
*       the currently selected estimator and currently selected sample,
*       the current run ID number.
The symbol "(B)" also appears if bootstrap inference is selected, and the symbol "(S)" if subsampling inference is selected.

The tool bar buttons provide quick access to the most frequently used actions and dialogs.

*       "Open Folder" button: shows the File / Data / Open... popup menu by default.
        When the "Load Multiple Data Sets" option is enabled (see Options / General), shows the Setup / Data Sets / Switch To... popup menu.

*       "f(x)" button: opens the Setup / Data Transformation and Editing dialog

*       "MM" button: opens the Setup / Model Manager dialog.

*       "Regression line" button: opens the Model / Linear Regression dialog.

*       "Space Shuttle" button (optional): opens the Model / Dynamic Equation dialog

*       "Running Man" button (optional): normally, launches Actions / Run Estimation. Also duplicates the function of the "Go" button in an open dialog.

*       "Calculator" button: launches Actions / Evaluate at Current Values. This is equivalent to Actions / Run Estimation in the case of linear regression.

*       "Chart" button: Dual action.
        o       Displays the plot of the variable(s) selected in the variable list in any dialog (first click the list).
        o       Opens the Graphics / Show Data Graphic dialog if no variable list has the focus.

*       "Double Chart" button: displays combined chart of actual and fitted values, and residuals. Only available following an estimation run. Optionally, can also show the actual-fitted scatter and residual histogram and kernel density (see Options / Graphics).

© James Davidson 2024

\*       "Dialog 1" button: can be assigned by the user to open any dialog lacking a dedicated toolbar button. By default, it shows the last dialog opened.
Up to 6 assignable buttons are available, but only the first one is visible by default. To assign a button, go to the Options / General dialog. See Section 8.8 for details.

\*       "+|x|/|=" button: displayed optionally (see Options/General/Interface Settings) for direct access to the manual calculator.

\*       "Windows" button: Toggle action.
o          Restores dialogs in previous positions (duplicating Actions  / Restore Dialogs) if dialogs closed,
o          Closes all dialogs (duplicating Actions / Close All Dialogs) if dialogs open.

\*       "Stop" button: only displayed when an iterative calculation is in progress. Pauses the calculation and opens a choice box, to Abort or Continue.

Notes:
1.       The "Space Shuttle" and  "Running Man" buttons can be optionally omitted from the button bar (deselect "Enable Optimization Estimators" in Options / General). This option can be used to simplify the interface for classroom use.

2.       When TSM is running, two "TSM" icons appear on the Windows task bar. The icon enclosed in a white square represents Ox, running in a minimized console window. Click it to view the window. The icon without the white square represents the GUI "front end" running under the Java Runtime Environment.
Note: if  TSM is started by clicking on a settings file in Windows explorer, the Windows "DOS box" icon appears instead of the first TSM icon.

3.       The "Windows" button remembers which dialogs were open between sessions. If the "Remember Dialog Positions" checkbox is checked in Options / General. Pressing it at start-up restores the window configuration existing at a recent close-down.  To avoid an excessive clutter of dialogs getting re-opened the configuration is reset if the time lag between sessions exceeds 11 hours.

4.       When lengthy computations such as optimizations, recursive estimations, grid plots, and Monte Carlo experiments are running, a "Stop" button appears on the tool bar. Clicking the button opens a choice box, with the choice of aborting the run or continuing.
Note: there may be some delay before the box opens.

5.       If a panel data set is loaded, the program switches to 'Panel Data' mode and the "Space Shuttle" and "Running Man" are automatically omitted from the button bar. Modelling options are selected from the Linear Regression and Panel Data dialogs in this case.

*1.3 The Results Window*

The text area is editable. Comments and aide-memoires can be typed in and saved
with the output.

*	Text can be highlighted with the mouse, hold down the mouse button and
	drag the "I-beam" cursor.

*	Use the right mouse button to open a context menu, with commands to cut
	or copy the text to the Windows clipboard.

*	The TSM command File / Save Selected Text allows highlighted text to be
	written to a named file.

*	Highlighted text can also be "dragged and dropped" into another application
	such as Word or Notepad. Hold down the Control key while dropping
	to drop a copy, while leaving the original behind.

The results for an estimation run are printed with a heading showing the version
number of the program, a run ID number and the date and time.

*	The run ID number is incremented each time an estimation is run, and also
	each time a stored model is loaded.

*	The ID number is used to identify retrieved series, and also listings, graphics
	and settings files optionally generated in the run. This makes it easy to
	verify which piece of output goes with which run, at the end of a session.

*	The current ID is stored in the file settings.tsm. It is preserved when settings
	are restored to defaults (File / Settings / New). It  is reset to 0 only when
	settings.tsm is deleted.

*	It can be assigned a new value (zero or otherwise) by going to Options /
	General / Special Settings and selecting "Initialize Run ID".

The text size can be changed from the default of 12-point in the Options / General
dialog.

The window can be resized by dragging the corner with the mouse, and also
maximized to fill the screen.

Dialogs are displayed on top of the window. In maximized display mode,
they can conveniently be arranged on the right-hand side of the screen, with the
results printed on the left-hand side.

The command File / Load Text File allows the window to be used as a text
editor. To save text, highlight it and use the File / Save Selected Text command.

*	This feature allows a user's code to be conveniently edited "on the fly",
	without needing to start up OxEdit. The command File / Restart closes
	and restarts the program in exactly its previous state, including window

© James Davidson 2024

contents, allowing the amended code to be compiled.

## *1.4 Dialogs*

Dialogs are where most interactions between user and program take place.
They contain objects of the following kinds

| | |
|---|---|
| Buttons: | to initiate actions, including opening/closing other dialogs. |
| Checkboxes: | to turn program options on and off. |
| Radio Buttons: | to choose from a set of mutually exclusive options. |
| Choice Widgets: | also known as pull-down menus, to make a choice from a list of mutually exclusive options. |
| Scrollbars: | to input numerical values, usually integers, e.g. sample size. |
| Text fields: | to input strings of characters from the keyboard. These may represent names to label the output, or real numerical values. |
| Lists: | columns of names, usually of data series, that can be selected and deselected with the mouse. |

Notes:
1.	In model specification dialogs, the choice of radio button may determine the interpretation of list selections. The correct sequence of actions is:
	*	Click the radio button to select the type of variable.
	*	Click list items to select/deselect them.
	To avoid unintended choices it is a good practice, before launching an estimation run, to click each radio button in turn, and note which variables are highlighted in the list for each case.

2.	Choice widgets are similar in function to radio buttons, convenient for larger numbers of alternatives. The widget shows the current selection in the window. Click the down-arrow to display the list, which can be scrolled if necessary, then click the desired selection.

3.	Lists are equipped with vertical scroll bars when there are more items than can be displayed at one time. Drag the bar to bring them into view. Horizontal scrolling allows long names to be viewed. On occasion Java can fail to display the scrollbars correctly. If this happens try closing and re-opening the dialog, or just drag on the list to scroll it.

4.	In some cases (e.g. selecting the dependent variable) only one list selection is allowed at a time. In others, multiple selections are allowed. Clicking a name has a toggle action, select when unselected, deselect when selected. In some dialogs, double-clicking a list item simultaneously selects the item and launches the appropriate action on it, e.g. displaying a plot.

5.	There are two ways to quickly select a block of adjacent names on a list. One way is to select the first name of the block and then select the last name using the *right* mouse button. This action highlights all the names in between. Another way is to use 'drag-select'. In other words: click the first name, then click the last name and drag the mouse briefly with the button held down, Again, all names between are highlighted.

De-selection of a selected block works in the same way - click the first highlighted name, then either right-click, or drag, the last highlighted name. Use the "Clear" button to deselect the whole list.

* Don't try the Windows Explorer method of holding down "Shift" while selecting! This can crash Java, unfortunately.

6. The dialog object most recently clicked on is said to 'have the focus', and is outlined by a broken rectangle. If the variable list of any open dialog has the focus, clicking the "Chart" button on the toolbar displays the plot of the highlighted variable(s).

   Note: The current settings in the Graphics / Show Data Graphic... dialog control what items are plotted.

7. Options may be "greyed-out" (disabled) if they are not compatible with other selected options. For example, GARCH and Markov switching options are not available unless a maximum likelihood estimator is selected. In some cases, a reminder message is displayed if an illegal selection is attempted.

8. Make a selection with a horizontal scrollbar by dragging the slider, or by clicking the arrow buttons at each end, for step-by-step selection. Some scrollbars (e.g. for numbers of Monte Carlo replications) are incremented in jumps (10s, 100s) depending on the range of the selection.

## *1.5 Entering Formulae*

Algebraic formulae can be submitted to the program in seven different contexts.
A. Data transformation
B. Nonlinear parameter constraints
C. Nonlinear equations for estimation or simulation (including equilibrium relations, nonlinear MA terms and nonlinear ECMs)
D. Random shock generation for simulations.
E: Automatic ("on the fly") data transformations.
F: Calculator (evaluates numerical formulae).
G: Matrix calculator.

Formulae are typed into a text box which is opened by the following commands:
Case A: Setup / Data Transformation and Editing / Formula ...
Setup / Make Data Formula...
Case B: Model / Constraints / Enter Code (select "Coded Restrictions")
Case C: Model / Coded Function.
Case D: Options / Simulation and Resampling / Enter Formula (Select "Formula")
Case E: The formula is entered in the variable's description field, preceded by #FN (case sensitive).
Case F: Setup / Calculator...
Case G: Setup / Matrix Calculator / Matrix Calculator...

TYPING FORMULAE

The general syntax in cases A, B, C "Equation" and G is of the form "name = formula",

© James Davidson 2024

| Case A: | "name" is a new variable name |
|---|---|
| Case B: | "name" is an existing parameter name. |
| Case C "Equation": | " = " is required. "name" is optional and if given, must match the selected dependent variable(s) in Dynamic Equation dialog. |
| Case G: | "name" is a new matrix name. This is optional, if omitted a text field opens for entry of a name. |

The general syntax in cases C "Residuals", D, E and F is of the form "formula".

A formula can be constructed from the following components depending
on the formula type, as indicated.

ABCDEFG: Real numeric constants, containing digits 0-9, optionally including
a decimal point (.) and optionally preceded by a minus sign (-).

ABCDEFG: The operators + (add), - (subtract), * (multiply), / (divide),
^(raise to a power).

ABCDEFG: Parentheses (...) enclosing sub-formulae.

ABCDEFG: The functions log( ), exp( ), cos( ), sin( ), tan( ), atn( ), abs( ),
sqrt( ), fct( ), lgm(), erf(), int( ), sgn( ), pos( ), neg( ), flr( ), ips( ),
ing( ), izo( ) and imi(), where the parentheses can contain a
variable, parameter, constant or  sub-formula.

ABCDEFG: The names Pi, pi, PI and pI are reserved, and are replaced by
the constant 3.14159..., to machine accuracy.

ACDE: Variables, by name or number (see Notes 1, 4) optionally
followed by an integer lag (negative) value or lead (positive) value,
enclosed in braces {}.

BCD: Parameters, by name or number (see Notes 1, 5).

F: Algebraic variables, by name.

G: Matrices, by name.
The operators ' (transpose), ** (Kronecker product),
@ (Hadamard product).
The functions inv(), gnv(), cho(), eig(), evc(), vec(), vcr(),
dia(), det(), tra(), nrm().

Notes:
1. The first character of a variable or parameter name MUST be a letter
(from the sets a-z, A-Z).

2. The following characters can also appear in variable or parameter names:
!#$%&0123456789<>?\_~

3. Names are case-sensitive! E.g. 'Alpha', 'ALPHA' and 'alpha' are all

treated as different names.

4. An alternative to typing a variable name in full is to type x[i] or X[i] where "i" denotes the number of the variable in the data list. When the text field is opened, variable numbers are pre-pended to names in the appropriate data list , i.e.:
 * In case A, Setup / Data Transformation and Editing
 * In case C, Model / Dynamic Equation.
 * In case G Setup / Matrix Calculator (Matrix list)

5. An alternative to typing a parameter name in full is to type a[i] or p[i] or A[i] or P[i] (all forms are equivalent) where "i" is the parameter number. Parameter numbers are shown in the Values / Equation dialog(s).

6. The characters +-*/^.'()[]{}, and also spaces, must NOT appear in names of variables or parameters. Otherwise spaces are ignored, use these optionally to improve readability. The character ; is used, in case F only, to separate formulae for plotting. Other characters are ignored.
 WARNING: Variable names generated automatically in Setup / Data Transformation and Editing may contain spaces and other illegal characters.
 * Illegal characters can be removed automatically, or replaced with underscores, by entering special strings in the text field in Edit / Rename. See Section 3.3.

7. The * sign to denote multiplication is generally mandatory, omitting it will cause a syntax error. The exception is where it would follow ' (transposition symbol).

8. Numeric constants can contain the symbols .0123456789
 The symbol "." (decimal point) may appear ONLY as part of a numeric constant. An isolated "." is parsed as the number 0.

9. Constants are nonnegative. A leading "-" (minus) in any formula is parsed as a unary operator, and changes the sign of the operand.

10. Definitions of special functions:
 $atn(x) = arctan(x)$
 $fct(x) = x!$
 $lgm(x) = log(Gamma(x))$
 $erf(x) = $ error function
 $sgn(x) = -1$ if $x < 0$, 1 otherwise
 $pos(x) = x$ if $x > 0$, 0 otherwise
 $neg(x) = x$ if $x < 0$, 0 otherwise
 $int(x) = $ integer nearest to x
 $flr(x) = $ largest integer not exceeding x
 Indicators:
 $ips(x) = 1$ if $x > 0$, 0 otherwise
 $ing(x) = 1$ if $x < 0$, 0 otherwise
 $izo(x) = 1$ if $x = 0$, 0 otherwise
 $imi(x) = 1$ if $x = $ NaN (missing observation), 0 otherwise.

These functions are not differentiable, so may not be used in parameter constraints. This results in a parsing error.

Also note:

pos(x - y) + y = max(x, y)

neg(x - y) + y = min(x, y)

x - y*flr(x/y) = x mod y (x,y integers)

(1 + erf(x/sqrt(2)))/2 = Phi(x) (normal c.d.f)

1 - ips(y - x) = 1 if y <= x, 0 otherwise

etc, etc. Other uses include creating period dummies. If t is an integer,

dummy = neg(pos(Trend - t) - 1) + 1

= 0 up to date t, 1 from date t onwards.

11.    Reminder: the following uses are mandatory!
*     Parentheses ( ) for function arguments and sub-formulae;
*     Square brackets [ ] for parameter/variable numbers;
*     Braces { } for lags and leads.

12.    If the formula cannot be parsed (e.g. non-matching parentheses, missing operators, illegal characters) an error message is printed and the calculations are terminated.

If an operation is illegal (e.g. square root of negative number) the output of the formula will be NaN.

RESERVED NAMES

13.    The reserved names U#, u#, and N#, n# (upper and lower case equivalent) are treated differently depending on the function call.

Cases ABE, C in an estimation run:

Assigned the value 0.

Case C in a simulation run:

U# and u# are replaced by a uniform[0,1] random drawing, and N# and n# are replaced by a standard normal random drawing, playing the role of a randomly selected parameter. This feature allows the generation of randomized models for (e.g.) test power evaluation.

Case D: U# and u# are replaced by a sample-dimension vector of uniform[0,1] random drawings, and N# and n# are replaced a sample-dimension vector of standard normal random drawings.

If the symbols appear more than once in a formula, a different drawing (vector of drawings) replaces each.

14.    (Cases AC)

The variable names E#{-j}, e#{-j}, where j represents a positive integer, are reserved for use in recursive functions. They are replaced by the function value itself lagged j periods. E# without a lag is replaced by 0. In systems of equations, the names E#1, E#2, ..., and E#1{-j}, E#2{-j}, ..., denote the values for each equation, current and lagged respectively. Current values from lower-numbered equations can appear in higher-numbered equations, otherwise are replaced by 0.

15.    (Case C)

The variable names Z# and z# are reserved for use in nonlinear ECMs,

and are replaced by the values of the equilibrium relation. In models with two or more equilibrium relations, use Z#1, Z#2, ..., to denote the values.

16. (Case C)
The variable names W#, w#, and also W#{-j}, w#{-j} where j represents a positive integer, are reserved for use in simulation models. These represent the random shock used to generate the series, current and lagged j periods, respectively. In systems of equations, the names W#1, W#2, ..., etc., denote the shock values for each equation.
NOTE: Attempting to estimate a model with this option generates an error. Equations in which the current shock enters nonlinearly cannot be inverted automatically. To estimate a model of this type, create the specification in a separate model.

17. (Case C)
The variable names X#[i] and x#[i],  where i >= 0 denotes the position of a variable in the data set, as in 4 above, cause the indicated variable to be over-written by a column of standard normal random numbers before a simulation is run. Leads and lags are accessed in the usual way, by entering X#[i]{-1}, etc. When estimating a model with this option selected, the "#" is ignored and X#[i] is equivalent to X[i].

18. (Case C)
The variable names H# and h# are reserved to represent the conditional variance or conditional standard deviation in models with conditional heteroscedasticity. In multi-equation models the names H#1, H#2,.. etc, denote the conditional variances/s.d.s for each equation.

19. (Case E)
The variable names S# and s# denote the variable being transformed.

MATRIX CALCULATIONS

20. The available functions are
> inv() = inverse matrix
> gnv() = generalized inverse
> cho() = Cholesky decomposition
> eig() = eigenvalues, returned as a diagonal matrix
> evc() = eigenvector matrix  (columns normalized to unit
>     length)
> eim() = moduli of eigenvalues, returned as a column
> vec() = vectorize a matrix (stacked columns)
> vcr() = vectorize a matrix (stacked transposed rows)
> dia() = diagonal of a matrix, returned as a column
>     vector
> det() = determinant
> tra() = trace
> nrm() = Frobenius norm
> len() = Euclidean lengths of columns (returned as row)

21. All standard functions can be used in expressions, and will be

applied elementwise.

22.   The operators +, - and @ (Hadamard product) require the
      arguments to be conformable.

23.   The operator * denotes either scalar or matrix multiplication. Either
      the arguments must be conformable matrices, or one or other
      must be a scalar (1 x 1 matrix).

24.   The operators ^ and / require the exponent or divisor to be scalar,
      and are applied elementwise.
      Example: A*A yields the matrix product (A square), but A^2 yields
      elementwise squaring (any matrix).

25.   The transpose ' can be used as a unary operator, or to form
      a product.
      Example: A'B yields the same result as (A')*B.

26.   For matrices A and B, A**B yields the Kronecker product and
      and A@B the Hadamard product, respectively.

27.   Complex-valued matrices are not handled, and result in errors.
      Use eim() to get eigenvalue moduli.

## 1.6 File Drag-and-Drop

Data and settings files can be loaded by dragging the icon(s) from Windows Explorer
and dropping over the program window.
*       The action taken depends on the file extension.
*       See 2.2 File / Data for a list of the supported data file types.
*       Multiple file drops are supported. Cases where only single file drops allowed
        are indicated by (S).
*       Special actions are performed on files with the indicated name prefixes.

| File Type | Equivalent Menu Command |
|-----------|-------------------------|
| settings file (.tsm) | File / Settings / Open... (S) |
| model file (.tsd) | File / Settings / Import Model... |
| data file | File / Data / Open... |
| text file (.txt, .ox, .h) | File / Load Text File |
| data file, "EDF_"… | File / Data / Tabulations / Load EDF Table. (S) |
| data file, "DNS_"… | File / Data / Tabulations / Load Density Plot. (S) |
| .tsd file, "Batch_Run"… | File / Data / Tabulations / MC Batch Results. (S) |
| matrix file, "MAT_"... | Setup / Matrix Calculator / Load File (S) |

© James Davidson 2024

# 2. File

## *2.0 File: General Information*

This menu gives access to submenus for:

| | | |
|---|---|---|
| Settings | >> | Save and load ".tsm" files. |
| Data | >> | Save and load data files, including empirical distributions for test statistics. |
| Results | >> | Control the saving to a file of output from the results window. |
| Listings | >> | Save items of output from estimation or simulation runs to spreadsheet files. |
| Graphics | >> | Save a copy of the most recently displayed graphic. |

Load Text File         Allows the results window to be used as a text editor. The contents of  the chosen file will be appended to the existing window contents.

Save Selected Text     Allows any part of the text in the results window to be saved to a file. Give this command after highlighting the desired text by dragging with the mouse. The options presented are to append the text to the current results file, and to write to (overwrite) a new (existing) named file.

Restart         >>     These options close the program and automatically restart it with all run settings intact, including results window contents. These are convenient shortcuts for loading user's code.
In some situations a manual restart is necessary, see Notes.

Restart/Load User Code
                       A choice box gives the options of loading the currently specified user Ox code, and selecting a different code file in the file dialog.

Restart/Load Text Input
                       A choice box gives the option of selecting a text file containing a set of TSM text commands, for loading as program settings and/or model specifications.

Restart/Load Text Defaults
                       Loads default settings which (if different from the programmed defaults) will be set with the File / Settings / Clear All command, and also if the settings file is renewed.

Quit                   Closes the program. The prompts to save specification and data before closing can be optionally switched off; go to Options / General.

Notes:
1       Selected text can be cut, copied or deleted. After highlighting, click the right

© James Davidson 2024

mouse button and select from the context menu.

2.    The contents of the Windows clipboard can be pasted into the results window at the cursor position. Right-click with the mouse, and select "Paste".

3.    Commands for loading, saving and exporting settings and data files open a popup menu listing the most recently used (MRU) files. The leading item, 'File Dialog...', opens the operating system's file dialog for selecting other files.

4.    If a list is empty, the popup menu is skipped and the file dialog opens directly.

5.    The maximum number of files on the MRU lists is an installation option - see Appendix B. The default setting is 20.

6.    MRU lists are stored in the settings file. When settings are loaded the current lists are appended to the new ones, as space permits.

FILE DIALOG

7.    The file dialog is sensitive to the type of file being loaded/saved, and presents only files of the appropriate type for selection, specified by extension. (See Section 11: Directories and Files for details.) Use the "Files of Type:" choice widget to present different categories of file types, where appropriate.

RESTART

8.    The restart options work automatically only if the program is launched from the Windows Start Menu. If started in Windows Explorer (by double-clicking on a .tsm file) the restart must be performed manually after the program closes, by double-clicking "settings.tsm". However, the main feature of the restart mode, the preservation of the results window contents, is operative in this case.

9.    The Restart/Load User Code command causes the file "usercode.ox" to be created if necessary, then edited to contain the line
        #include "[chosen file]".
The chosen file must have extension '.ox'! If necessary, the executable Ox file tsmod_run.ox is also edited to include and #define USER_CODE '
' #include "usercode.ox" ' compiler directives.
Notes:
*    Only one code file at a time can be compiled. Any other '#include...' statements are deleted from usercode.ox. Other lines are commented out.
*    The currently stored usercode file name is used by default. Open the file dialog at the prompt to select a different one.
*    To remove the currently loaded code file from memory, open the file dialog at the prompt and press "Cancel".

10.    The Restart/Load Text Input command causes tsmod_run.ox to be edited to include the directive '#define TEXT_INPUT ' and the statement

' Text_Input(){ #include "[chosen file]" } ', which include the contents of the chosen text file as lines in the Text_Input function. After the restart, these lines are deleted to restore tsmod_run.ox to its previous state.

10. The Restart/Load Text Input command causes tsmod_run.ox to be edited to include the directive '#define TEXT_INPUT ' and the statement ' Text_Input(){ #include "[chosen file]" } ', which include the contents of the chosen text file as lines in the Text_Input function. After the restart, these lines are deleted to restore tsmod_run.ox to its previous state.

11 Normally, the text file will contain either a complete specification for the (non-default) program settings, or a complete (non-default) model specification. In these cases it is usually important to set the defaults before importing the new settings. In the first case, do this by setting the first line of the file to
Set_Defaults();
In the second case, set the first line of the file to
Set_ModDefaults();
Failure to include one of these lines may lead to unpredictable results.

12. If the new settings include parameter values, then the last line of the file must be set to
Load_TextValues();
This function converts the new values to the internal program format. If it omitted, the new values will be ignored. Note, it is not possible to read in a subset of new parameter values and retain others. Running this function resets all the parameters.

13. The Restart / Load Text Defaults option edits tsmod_run.ox with permanent effect. To restore the standard default settings, edit the file and comment out the '#define TEXT_DEFAULTS' line. The defaults file can also be modified in a text editor. See the programming manual for information on text commands.

14. If the Restart command is given after editing the user's code, resulting in a compilation error, the program will fail to restart. In this event, check the console window for error messages. TSM can always be restarted using the default shortcut if required. This shortcut uses the unchanged copy of tsmod_run.ox in [TSM-HOME].

15. Use 'Simple Restart' for option changes that require a restart, such as hiding/showing advanced estimation features.

## *2.1 File / Settings*

Settings files (extension ".tsm") save the current state of the program, including optional program settings, model specifications, and parameter values. Loading a settings file recreates the exact state of the program when it was written, including operational settings, model specifications and starting values. .

© James Davidson 2024

The available commands are

Open…              Loads settings from a selected file. Only files with extension
                   ".tsm" can be opened.

Import Model...     Adds a model specification stored in a ".tsd" file to the
                   current settings.

Save               Saves settings to the default file, "settings.tsm".

Save As...          Saves settings to a named settings file. The extension ".tsm" is
                   added to the file name if not supplied.

Export...           Exports settings, listings and data to a .tsm file.

Settings as Text       >>
Display/Save Settings...   Writes all the non-default settings as text, in TSM script
                       format, either to the results window or to a file.

Save Current Model...   Writes the non-default specifications for the currently
                       loaded model to a file

Save System Defaults...   Writes the current system settings as text.

Clear all Settings     Clear the memory, replacing all settings with defaults. (See
                       Options / General Information and Defaults for a list.) Delete
                       all stored models, results, graphics, listings and associated .
                       tsd files

Notes:
1.     Changes to settings are written continuously to "settings.tsm" as the program
       runs. Data, program listings and graphics are similarly saved to the file
       "settings.tsd". If the program is closed down, restarting it restores the working
       environment exactly as it was before. The current settings can be saved to a
       different named file using  "Save as...", for subsequent reloading by the
       command  "Open...".

2.     The default settings files are read at start-up if they exist. The current data
       set is loaded directly from "settings.tsd" even if the original spreadsheet file
       has been moved or deleted in the meantime, data being sourced from the
       spreadsheet file only if "settings.tsd" is absent. Optionally this feature can
       be disabled; see Options / General / Operational Settings / Save Data
       with Settings.

3.     TSM can be started by double-clicking on any settings file. This loads the file
       in question, as if loaded by the "Open..." command. System settings are
       always written to "settings.tsm", however; named settings files are changed
       only by the "Save as..." command.

IMPORTING MODELS
4.     The "Import Model..." command can be used to access model specifications
       stored in .tsd files (which can include data sets) created with a different

settings file, or a different installation. This option (which replaces a superseded import command that operated on .tsm files) only works for .tsd files created with TSM 4.27 and later releases.

5.      If a model is imported which has the same name as an existing model, its name is modified by adding a % character.

EXPORTING SETTINGS

6.      The "Export..." option allows model specifications to be distributed to other users. This option has the following features:

     *      All relevant information is saved in the .tsm file, including the current data set(s), model listings (usually saved in separate .tsd files), and supplied Ox code.

     *      All local path information, including MRU lists, is omitted from the file.

     *      When the file is loaded, the stored data and listings files are automatically saved in the local "Start-in" directory. In this way, the local installation can exactly reproduce the set-up existing on the machine where the 'export' file was created.

7.      To be exported using "Export...",  supplied Ox code must be contained in one of the following locations:
     1)      The standard file "usercode.ox" located in the Start-in directory;

     2)      The file named in a "#include" compiler directive in "usercode.ox". In this case,
         *     any executable code in "usercode.ox" itself is ignored.
         *     only the first "#include" directive in "usercode.ox" is recognized. Any additional inclusions are ignored.

8.      Any text in the results window which is highlighted when the File / Export... command is given is stored in the settings file, and written back to the results window at start-up. This trick allows instructions and notes about the contents of the exported file to be easily conveyed to recipients of the file.

9.      The Export function is distinct from the saving of program data in the file "settings.tsd". (See 2.3 File / Listings and 11 Directories and Files.) 'Export' bundles everything (including the contents of settings.tsd) in a single file.

SETTINGS AS TEXT

10.    The "Display/Save Settings" command opens a choice box, with the options of printing the text in the results window, and opening the file dialog for saving the text. In the latter case, a ".txt" file extension is recommended. See the programming manual for details of the scripting language.

11.    The "Save Current Model" command saves just the loaded model settings to

         © James Davidson 2024

a file. To view these settings in TSM, store the model in Model Manager with the "Automatic Descriptions" option checked, then click the "Model Description" button. The main function of this command is to allow the settings to be imported into another instance of TSM using "File / Restart / Load Text Input".

12. The "Save System Defaults" command saves a file for loading at start-up. This options allows the user to control the state of the program set by the File / Settings / Clear All command, or renewing the settings file. Use the File / Restart command to set up the system to read the designated file.

## 2.2 File / Data

Data can be read and written in one of eight formats. The format is specified by the file extension.

".xls", ".xlsx"   Excel 2003 and 2007 worksheets. The first row of the sheet must contain variable names.

".csv"   Comma delimited text file. Format as for .xls.

".in7"   GiveWin file. See OxMetrics documentation for details.

".dht"   GAUSS data file.

".dta"   STATA (Versions 4-6) data file.

".dat"   OxMetrics "data with load information" file. See the OxMetrics documentation for details.

".mat"   ASCII file containing a matrix, with variables in columns and observations in rows. The first line of the file must contain two integers, number of rows followed by number of columns. In this case the variable names assigned are "Var1", "Var2", etc. These can be changed in the Data Transformation and Editing dialog.

The available commands are

Open…   Opens the operating system's file dialog, for opening. Any data currently in memory are over-written.

Merge…   Merges new file with data already loaded. After selecting the file, a dialog is presented to enter the 'row offset'; that is, the relative position in the current file of the first observation of the new file. The offset is:
>   zero if start dates match,
>   positive if start date of new series is later, or
>   negative if start date of new series is earlier.
Close the dialog either by pressing "Return" or clicking the Close symbol on the menu bar.

© James Davidson 2024

Save          Saves data under the name specified in the last Open or Save As command.

Save As…        Opens the operating system's file dialog, for saving. The type of file saved is determined either by the extension or, if no extension is added, by the current selection in Options / Output and Retrieval Options.

Edit Spreadsheet...
         This option opens a "raw" data file for editing, effectively duplicating the role of Excel or similar software.

Notes:
1.      Data files may be loaded by the program automatically, as when a settings file or a stored model is loaded. If the file in question is not found on the stored path, other paths known to the program are searched. Messages are printed to show that this has occurred, or if the file is not found anywhere. Since the file found might not be the one intended, it is important to check the program messages.

2.      If data are saved in .MAT format, the variable names are listed in the results window.

3.      When loading a data set a prompt appears by default, giving the option to retain or clear the current model specifications. In most cases, these won't be appropriate to the new data. To make clearing automatic and so bypass this prompt on future occasions, without disabling the other prompts, choose the option "Always". To reinstate this prompt, deselect and then reselect the option "Enable Saving/Clearing Prompts" in Options / General.

DATA DESCRIPTIONS
4.      In addition to names for the series, data files can optionally store text containing a description of the variable. The description field can be used, among other things, to contain formulae specifying transformations of variables to be performed "on the fly" as part of a model specification. A user-specified, reserved character, @ by default, is used to separate a variable name from the description field. All characters following the @ form the description. Descriptions for the current data file can be viewed by the command Help / Data Descriptions, and created and edited in Setup / Data Transformations and Editing / Edit; see 'Edit Description'. If desired, the separator character can be changed during installation by editing the file tsmgui4.oxh and setting the function Get_NameCharacter().

5.      Ox allows a maximum of 64 characters to appear in the name field (the first row of the spreadsheet), but descriptions can nonetheless be of arbitrary length. If the name and description together exceed 64 characters, the additional characters are stored numerically at the bottom of the corresponding spreadsheet column (a Double value occupies 8 bytes, and so can represent up to 8 text characters.) The description zone is separated from the data proper by two rows containing, respectively the constants "DBL_MAX" and "DBL_MIN"  (the largest and smallest storable numbers). These additional rows can be inspected by loading the file into a program

      

such as Microsoft Excel.

6.    The input file may have missing observations (e.g. blank cells in an Excel file) and these are replaced in the data matrix by NaN (= Not a Number). If the variable in question is included in a model, the sample period is automatically truncated, as necessary, to exclude NaN observations.

PERIODIC OBSERVATIONS AND DATES

7.    If provided, date information (years, quarters, months etc.) is used to label graphs and printouts, and is presented in the Setup / Set Sample dialog for selecting the sample. Ox dating information is used by default. This is the first column of any spreadsheet file saved by TSM, and has the format "Year-Period". This format is suitable when the data have a regular periodicity, but if other dating schemes are used thus column is ignored, and can be deleted from the spreadsheet if desired.
For data sets other than panel data, there are three ways in which the user can create dates.

      \*    Year-Period: Enter information on the start date and the periodicity and

             let TSM generate the dates, which are saved in the default Ox format. This is the best scheme for annual, quarterly and monthly data, where the periodicity is fixed.

      \*    Calendar Dates: Create a literal date series in the data file using the "yyyy-mm-dd" date format. This is the best scheme for daily data, where the periodicity is irregular due to week-ends and holidays.

      \*    Date Labels: These are arbitrary numerical values, which need not be increasing, or regular, stored as a special series in the data file. These are suitable for irregularly spaced data and historic long period data series.

    Note: If a file contains more than one type of date information, calendar dates take precedence over date labels, and date labels take precedence over Ox-format year-period dates.

PERIODIC DATES

8.    These use the Ox dating format, and is the easiest form of dating to set up. The date information can be added in TSM using the "Set Dates" command under Setup / Data Transformation and Editing / Edit. A column is created with the reserved name !STARTDATE!, and having entries in the first three row positions, as follows:
1.    The year of the first observation
2.    The period (quarter/month/week/weekday/day) of the first observation
3.    The frequency: one of
        2 (half-years)
        4 (quarters)

                        

12 (months)
52 (weeks)
260 (weekdays)
365 (days).

The rest of the column can contain anything, or be blank.
This item is removed after being read, so it will not appear in the variable list.
If the data file is then saved in TSM, it will contain the date information in Ox format - the first column of the file will contain the dates in the format "year - period". This column will be used to create date information next time the file is loaded.

9.     If files are merged, the date information of the first-loaded file is used unless the merged file contains a !STARTDATE! column, in which case this information is used to date the combined observations. Ox date information in the merged file is ignored.

10.    The periods must have equal length (e.g. no missing holidays, or leap years, in daily data). Ignoring these irregularities may be permissible for labelling graphs, for example.

CALENDAR DATES

11.    To set individual daily dates, load the data file using the command File / Data / Edit Spreadsheet... Create a new variable with the name "Date" (for example, use the "Make Zeros" command) and move it to the first position in the file using the Edit command "Move Up List". Now enter the dates in the format yyyy-mm-dd (the "Copy" button is useful to propagate the template from cell to cell). Optionally, times of day can also be entered using the 24-hour format yyyy-mm-ddThh:mm.

12.    Dating information is used automatically if it is present. To over-ride this behaviour and set periodic dates in the default manner, use the "Set Dates" command in Setup / Date Transformation and Editing. In this case, the existing date information is deleted. If dated files are merged in TSM, the dating information in both files is ignored.

DATE LABELS

13.    Create a variable in the data file with the reserved name !DATELABELS!. the elements of this column are used to label the observations. They must be numeric, but are otherwise arbitrary. For example, in the case of irregularly spaced data, including just the year, repeated as required, may provide sufficient information to label a time plot. The labels could also be in descending order to denote (e.g.) "Years ago".

14.    The label series can be created in TSM using the command "Set Date Labels" in Setup / Data Transformation and Editing (Edit). The series is not visible on the variable list, but appears in the dates column when series are listed, and is saved to a spreadsheet file with the data

15.    If files with date labels are merged in TSM, the dating information in the files is merged. Labels from the merged file replace existing labels when both exist. Otherwise, missing labels are replaced by zeros.

© James Davidson 2024

16. To load a panel sample, the data file must be formatted according to the following specifications. The data for each variable must be input as a single column. The observations on 'individuals' should appear in row blocks, with the dated observations in date order, within each block.

17. Panels can be unbalanced, with different start and finish dates for each block. A date must take the form of either a year, or a period, or a year + period (quarter, month etc.) Dated daily data are not supported.

18. Panels can also be irregular, with missing periods as well as different start-finish dates. However, irregular panels must have a seasonal frequency of one (years, or periods). The year + period format is not supported.

19. The file must contain the following guide columns. They can appear anywhere in the file.
    * A column with the name !PANEL! (or equivalently !INDIVIDUAL!) in the first row. This contains the individual identifiers. The same number or text string (see para 27 below) should appears in all the rows of an individual block. Numbers need not be consecutive. (Block boundaries are detected by these entries changing.)
    * A dates column headed with the name !YEAR! or !PERIOD!. This will typically contain the year of each observation, hence the sequence is repeated within the individual blocks.

    Also, optionally:
    * If a !YEAR! column exists with repeated entries, a column headed !PERIOD! will contain the quarter/month counts within each year.
    * A column headed !GROUP! can be used to label groups of individuals, such that all individuals in a group get the same identifier. Any numbers can be used as identifiers, but the number 0 is always used to flag that the individuals in question are to be omitted from the analysis.

    The guide column names can be written in upper or lower case.

20. When both !YEAR! and !PERIOD! columns appear in the file, dates are counted in years and periods, so the !YEAR! column should contain repeated entries. For example, suppose a panel contained observations for the five quarters 2001Q2 to 2002Q2 inclusive. The first few rows of the file should then look like this:

| !PANEL! | !YEAR! | !PERIOD! | ...(data) |
|---|---|---|---|
| 1 | 2001 | 2 | ... |
| 1 | 2001 | 3 | |
| 1 | 2001 | 4 | |
| 1 | 2002 | 1 | |
| 1 | 2002 | 2 | |
| 2 | 2001 | 2 | |
| 2 | 2001 | 3 | |
| 2 | 2001 | 4 | |
| 2 | 2002 | 1 | |
| 2 | 2002 | 2 | |

|  |  |  |
|---|---|---|
| 3 | 2001 | 2 |
| ... | ... | ... |

21. NOTE: the dating system for panel data is distinct from the usual case, because Ox cannot maintain panel dates automatically. Saved panel data sets contain an Ox 'dates' column, but this is ignored by TSM when the file has the format specified above.

22. Individuals in a panel may be identified by names instead of numbers. See 27 below for details.

23. To load a panel data set conventionally (for example, so as to estimate time regressions for individuals one at a time) simply load the file as a spreadsheet (see below), delete the guide columns, and save the file under a different name.

SPREADSHEET EDITING

24. Use this option to edit a "raw" spreadsheet or data file. TSM functions in this mode as a simple alternative to Excel or equivalent spreadsheet program. This mode is chiefly useful for setting up series of calendar dates, date labels and indicators, and also for preparing files in panel data format.

25. When a data file is loaded by this command, the data editing dialog opens automatically, and all other program functions, except data graphics, are disabled. Guide series that are normally concealed from the user - having special names enclosed in !! pairs - are visible, and can be created and changed as required. When the dialog is closed, the file is processed in the usual way, and TSM then functions as normal.

26. In spreadsheet editing mode (but not otherwise) rows of the data matrix can be deleted or inserted at the position of the selected cell. Use the "Insert Row" and "Delete Row" buttons that appear in the left-hand panel. Inserted rows have their elements set to .NaN, and must be filled with data by hand.

27. In files formatted for panel data, a column headed either !PANEL! or !INDIVIDUAL! (upper or lower case) contains indicators of the individuals in the data set. Entries in this column can be either numbers (positive integers) or text (names not exceeding eight characters in length). These entries are stored as equivalent double-precision values in the spreadsheet and converted to text when the file is processed. Simply type the names into the cells of a column so headed in the List/Edit table, or use the "Edit Observation" field as usual. Then use the "Copy" button to fill the required number of rows with the same entry.
    Note:
    * Change in the contents of these rows signals the next individual in the sample, so the same text (or number) must appear in each individual block.
    * Non-numeric text can only be entered in a column headed as above.

37                                      © James Davidson 2024

\*      In programs such as Excel, the default numerical formatting may show the text fields as zeros! They can only be viewed correctly in TSM.

28.      In spreadsheet editing mode (but not otherwise), the Edit command "Set Panel Guides" allows the !PANEL!, !PERIOD! and (if required) !YEAR! guide columns to be inserted automatically in a panel data set. Enter the number of periods per individual in the text box which opens or, in the case of years + periods, enter the number of years followed by the number of periods in the box, separated by a comma. For this option to work, the panel must be complete, with equal numbers of time periods for each individual. In the case of incomplete or irregular panels, with missing time periods, "complete" the panel by adding in the required rows temporarily (see 25 above). Delete them again following the insertion of the guide columns.

29.      A column headed !GROUP! (upper or lower case) contains group identifiers. These are constructed in the same way as the !PANEL! identifiers (see 27), but they only need to appear in the first row of each individual block. The contents of the remaining rows of each block are optional.

30      If the !GROUP! column is present, any individual block containing 0 as its first element is omitted from the analysis. This allows individuals to be included in/excluded from the data set with minimum difficulty. To disable the column without deleting it, simply rename it. It then appears in the data set as a "variable" when the data are reloaded.

TROUBLESHOOTING

31.      Although Ox should read spreadsheet files such as the Excel .xls format and .xlsx files, it is occasionally found that such files contain incompatible features. The message "Data input failed" appears. In this event, use Excel to save the file in comma delimited (.csv) format. These files are always readable and contain all the information relevant to TSM functionality. If TSM is then used to save the data in .xls or .xlsx format, the file will be readable by both TSM and Excel.

## *2.3 File / Results*

Results and information appear in the program's text window.
The output can be saved to disk files in two ways.
1.      Select part or all of the window's contents, using the mouse, and append these lines to a designated file.
2.      Have all the output saved automatically to a file, in the background.
Both methods can be used at the same time, if desired.

The available commands are

\* Enable Background Saving         Switches on Background Save mode

\* Disable Background Saving        Switches off Background Save mode.

          © James Davidson 2024

* New Results File…                Opens the file selection dialog, showing
                                   the current results file. Choice can be an
                                   existing text file or a new one.

Notes:
1.      Results are always appended to the current results file. To start a new file,
        choose a new name, or delete the existing file first.

2.      The default results file is "results.txt", in the working ('Start-in') directory.
        This will be created automatically if it does not exist. There is no default name
        for selections. You are prompted for a file name if one is not currently
        specified. If background saving is enabled, the New Results File command
        designates a new file for background saving. If background saving is disabled,
        the chosen file will be used for selected text.

3.      If the folder specified for a new results file is different from the working
        directory, the results folder is changed accordingly. Listings files and graphics
        will be saved thereafter to the same location. However, the settings file must
        always reside in the 'Start-in' directory, so it can be found at launch.

4.      To guard against attempting to append text to an existing binary file, results
        files must have extension ".txt". This is appended to the chosen name if it is
        not present.

5.      Yet another way to save results is to use the Windows clipboard. Just highlight
        the desired text in the results window, Copy (right-click for the context menu),
        and Paste into your favourite text editor.

6.      The 'Locate Results Folder...' command opens a text field containing the
        current path to the results folder. This can be edited to change the folder.
        Press the "Browse" button to navigate the directory tree and select a folder,
        else type the path directly in the text field.
        Notes:
        *       Selecting a new results file automatically sets the results folder to the
                one containing the file.
        *       Graphics are always saved in the current results folder.
        *       The 'Start-in' directory is used as the results folder by default, but note
                that the current path is saved between sessions in "settings.tsm".

## 2.4 File / Listings

"Listings" refer to outputs generated by the program in tabular or graphical
form. The commands
*       Open Listings File...
*       Save Listings File...
allow listings to be saved and reloaded in a format that can be displayed and/or
further processed by the program. These files have extension ".tsd".

Listings can also be exported in spreadsheet format, for processing by other
software. The command "Export Spreadsheet" gives access to a submenu

© James Davidson 2024

containing the following items:

*      Series:                       actuals and fitted values, residuals, conditional variances, regime probabilities.
*      Correlograms:         also Q Statistics for the residuals and squared residuals corresponding to each lag.
*      MA coefficients:      equivalently, impulse and/or step impulse responses.
*      Forecasts:                including standard errors and/or confidence bounds where available.
*      Recursive Estimates:  coefficients and standard errors, summary statistics and test statistics.
*      Recursive forecasts:
*      Criterion plots:
*      Monte Carlo replications (or optionally, relative frequencies).

Notes:

LISTING FILES

1.      ".tsd" files are created automatically when a model is stored with Model / Save Current Model..., or with the Model Manager, and receive the name assigned to the model. Such files are likewise loaded by the command Model / Load a Model.... Files created in this way can also be read using the Open Listings File... command, but in this case there is no change to model specifications. Saved files can be named in any way desired. These commands are intended to be useful mainly for storing and retrieving the results of Monte Carlo experiments.

2.      The command File / Settings / Import / Import Model... also reads ".tsd" files, but in this case model specifications as well as listings are loaded, with the imported model receiving the name of the file containing it. Use the command Open Listings File... to load the listings, but without creating a new model. The "Save Listings File..." command does nonetheless include the current model specifications in the file. Reading the file with the "Import Model" command would allow these to be retrieved, if desired.

3.      A ".tsd" file contains all the listings currently held in the program at the moment it is saved. When the file is loaded, all such items get replaced in memory. Be sure to save anything you want to keep before loading a file.

4.      The file with default name "settings.tsd" is special, containing the data and results associated with the current model specification. This It is paired with "settings.tsm" and (if the feature is enabled, see 8.8 Options / General) preserves the current working environment between sessions. It is saved automatically, and read automatically at start-up, there is no reason to open it with this command.

5.      Files with the generic names "Batch_Run??_Inst??.tsd are created when parallel processing is used to run Monte Carlo experiments. These files have a special format, and can only be read using the command File / Data / Tabulations / MC Batch Results... An error message is printed if an attempt is made to open these files with "Open Listing File".

6.    Spreadsheet files are saved by default to the currently selected results folder, under a descriptive name appended with the run ID number, or optionally through the file dialog.

7.    Available file formats are the same as for data, as selected in Options / Output and Retrieval. When saving in ".MAT" format, the columns headings are listed in the results window.

8.    There is an option to save all listings automatically following a run, see Options \ Output and Retrieval.

9.    By default, the output to the 'Forecasts' spreadsheet contains a selection of quantiles of the forecast distributions, either from the bootstrap distribution in Monte Carlo forecasting case, or otherwise Gaussian. To simplify this output to contain only the point forecasts (means or medians in the Monte Carlo forecasting case) go to Options / General / Special Settings, and select "Export Point Forecasts Only" from the pull-down menu. Double-click the text field to toggle between settings "True" and "False".

10.   In the 'Recursive Forecasts' spreadsheet, the rows of the table match the rows of the 'Recursive Estimates' table. The entries in the first column correspond to the dates from which the forecasts are projected forward. The j-step-ahead forecasts, for j = 1,2,..., then appear in successive columns.
      This arrangement is in contrast to that of the 'Forecasts' spreadsheet in which the row count corresponds to the number of steps ahead. Here, the first column shows the dates of the observations being forecast.

## 2.5 File / Tabulations

These commands handle special spreadsheet files, formatted to be read as tables for plotting or critical value calculations.

Load EDF Table...    Loads a previously prepared spreadsheet file containing an empirical distribution function. This can be displayed in Setup / Look Up Probability and Setup / Look Up Critical Value, and used by the program to generate test p-values. Files with the correct format are generated by the Monte Carlo module, see Setup / Monte Carlo Experiment.

Merge EDF Table...   Combines a selected file with the resident EDF table. When combining tabulations for different sample sizes, the names and specifications for the two tests must be identical.

Save EDF Table...    Converts an EDF table held in memory back to a file, reversing the "Load..." command. Optionally a new name can be entered using the file dialog, the stored file name is used by default.

© James Davidson 2024

Clear EDF Table...     Removes table from memory. The table is saved with model
                       listings, ('.tsd' files) so use this command before storing a
                       model to avoid bulky files .

Load Density Plot...   Loads a previously saved spreadsheet file containing data for
                       generating density plots in the dialog Graphics / Stored MC
                       Distributions. The file contents are merged with any existing
                       plot data. If a loaded plot has an existing name, the name is
                       augmented with "(1)".

MC Batch Results...    This command duplicates the action of the "Results" button in
                       the Setup / Monte Carlo Experiment dialog.

Notes:
1.     The files read by these commands must have a spreadsheet format and
       extension, but also require special formatting to be read successfully. TSM
       optionally creates these files in the course of running simulation experiments.
       Alternatively, given suitable data the files might be constructed manually in
       a spreadsheet program. The format specifications are given in Appendix G.

2.     EDF tables are stored in the listings (".tsd") file created for a model, and
       are loaded along with the model if this file is available. If listings file is
       unavailable when a model specifying EDF critical values is loaded, TSM
       looks for the specified spreadsheet file. This might be the case if the
       estimation job is being run as an external process, in a different location,
       for example.

3.     An exported settings (".tsm") file bundles all material relevant to each
       stored model, including data sets and EDF tables. When such a file is
       loaded, both the data and EDF spreadsheet files are created
       automatically in the working directory.

## 2.6 File / Graphics

Save last Graphic >          Saves the most recently displayed graphic to a file.
                             Optionally, one can either open the file dialog to enter
                             a name for the graphics file, or accept the default file
                             name.

Import Graphics Settings...  Open the file dialog to select for import a previously
                             exported store of the settings for the Options / Graphics
                             dialog

Export Graphics Settings...  Open the file dialog to name a file to receive the current
                             settings in the Options / Graphics dialog. These can be
                             subsequently re-imported with the Import Graphics
                             Settings... command.

Notes:
1.     The default file type for saving is set in Options / Output and Retrieval. The

© James Davidson 2024

type can also be specified interactively in the file dialog, by the choice of file extension. The options are '.png', '.gif' (bitmap formats) and '.eps', '.fig', '.tex' (vector graphics formats).

2.    Font specifications, and bitmap dimensions for '.png' and '.gif' files, are selected in Options / Graphics.

3.    There is an option to save all graphics automatically as they are displayed; see Options / Graphics.

4.    Graphics settings files receive the extension '.tsd', and the blue 'TSM Document' icon, as for model data and batch output files. However, they can only be read with the Import command in this menu. Attempting to read them through the File / Import Model... command produces an error message.

## *2.7 File / Folders*

Use these commands to show and/or change the storage locations of different types of files used by TSM. A command opens a text field containing the current path to the folder.  To change the path, you can

o    edit the field directly,

o    press the "Browse" button to navigate the directory tree and select or create a folder,

o    press "Home Folder" to insert the path to the folder where TSM was started, also called the "Start-in" folder.

Press OK to confirm the change, or Cancel.

The file categories are:

*Results...    This folder holds the automatic results window listing enabled by File / Results / Enable Background Saving, and other text files, also graphics files of all types.

*Data          Default location of data files.

*Model Files   Temporary files, with .tsd suffix, holding data associated with models.

*EDF Files    Empirical distribution function files in spreadsheet format, created to provide test p-values. Typically an EDF file is associated with a model.

*User Code    Files with extension .ox, supplied by the user, to be compiled with the main program.

*Batch Files    Code (.ox) files and output (.tsd) files generated for external jobs, such as Monte Carlo experiments run in parallel, including Condor jobs. These files are temporary and are normally deleted after their contents has been consolidated.

© James Davidson 2024

Notes:

1.     The file locations are stored in the TSM settings file and remembered between sessions.

2.     The same folder can be used for any or all of the file categories, By default, this is the home folder. Using separate folders becomes the more convenient option when large numbers of files are generated. These locations can be subfolders of the home folder, and can also be located elsewhere in the file system.

3.     If the model, EDF, and user code file locations are changed, the files associated with the current program settings are moved physically to the new locations.

4.     Data files can be loaded from anywhere, but if edited they will be saved to the current data folder. This ensures that files read from other locations (such as data sticks) are preserved intact by default. (They can of course optionally be saved anywhere.)

5.     If the data file location is changed, the data file, or files, currently loaded into TSM are saved to the new location. If the old location is the home folder the file(s) will also be deleted from there. Otherwise it is left to the user to tidy up the old location manually, as desired. This is to guard against unpredictable behaviour in the event that (for example) two users are working with a common data store.

6.     Default model, EDF and batch file locations can be set permanently by editing the file tsmgui4. However, these default locations must be within the home folder. Files can be stored anywhere, but locations outside the home folder can only be set dynamically.

# 3. Setup

## *3.1 Setup / Set Sample*

Select the sample for estimation and other data analysis procedures using the scrollbars. Date information is given if provided with the data. The maximum sample is selected by default.

Notes:
1.     Seven different sample selections can be set concurrently, for different program functions. This dialog is accessible from within every dialog where a sample must be specified, and the sample to be selected depends on where it is opened from. The locations are

        Setup / Recursive/Rolling Estimation.    .        Sample 1
        Setup / Summary Statistics   .     .        Sample 2
        Setup / Nonparametric Regression   .     .        Sample 3
        Setup / Semiparametric Long Memory .     .        Sample 4
        Setup / Cointegration Analysis .     .        Sample 5
        Model / Linear Regression   .     .        Sample 1
        Model / Dynamic Equation   .     .        Sample 1
        Graphics / Show Data Graphic .     .        Sample 6
        Setup / Data Transformation and Editing   .        Sample 7
        Options / Simulation and Resampling   .     .        Sample 8

    When this dialog when opened from the menu bar, it's function depends on which dialog is currently open. By default, it sets Sample 1.

2.     Sample selection 1 is displayed on the status bar. Sample selections 2-7 are displayed at the top left of the relevant dialogs. Sample selections 1, 3, 4, 5 and 6 are stored as part of a model, and retrieved by reloading the model. Selections 2, 7, and 8 are saved with general program settings.

3.     By default, the chosen sample is automatically truncated to exclude missing observations (denoted NaN) of any variable in the model specified. Remember that missing values are created by taking lags, leads and differences. The actual sample used is the sequence of non-missing observations that terminates with the last non-missing observation of the selected sample. Note that this is not necessarily the longest available set of non-missing observations.

4.     Use the "All Available Observations" button to check for missing observations. The largest sample available for all the variables in the currently specified model is selected. The actual sample used for estimation is restricted to this set.

5.     The sample count is initialized at the first line of the data file, not the first available (i.e. non-missing) observation.

6.     Click the "Simulate" button if the sample is being set for a simulation

        © James Davidson 2024

run, not for estimation. In this case, the "Available Sample" to be checked corresponds to included regressors only, ignoring the 'dependent' variable to be simulated.

* The button is highlighted when the option is selected. Cancel it by clicking again, or by closing the dialog.
* The button is disabled unless the dialog is opened from either the menu bar or one of the Model dialogs.

7. Check the "Filter Out  Missing Observations" checkbox to create the sample by simply dropping any missing observations from the chosen interval, rather than looking for a consecutive non-missing sequence as described in 3. This setting is in most cases inappropriate for time series, and is disabled if any time-series model features are selected.

8. In panel data models, the setting in this dialog determines the time periods to be included in the estimation or simulation for each of the individuals in the sample. The number of individuals to be included can only be changed by editing the grouping indicator in the data file.

9. The sample can be specified arbitrarily by means of an indicator variable. The indicator variable must have the name "!selectobs!" and is applied as follows:
* Any non-zero value: observation included in sample,
* Zero: observation omitted from sample.
This variable can be created using the option Edit / Set Sample Indicators in the Setup / Data Transformation and Editing dialog. Alternatively, any existing data series can be designated as the indicator by simply renaming (a copy of) it as !selectobs!.
Note:
* When !selectobs! exists in the data set, a checkbox is displayed in this dialog, to select/deselect its use. Otherwise the checkbox isd hidden
* The indicator is not named explicitly in variable lists. To remove it, use the Set Sample Indicators option as described, and change all zeros to non-zero values.
* Only one indicator named !selectobs! can exist in a data set at one time. Maintain two or more by assigning them different names, and use the Edit / Rename function to activate and deactivate.

## 3.2 Setup / Data Sets

By default, the program holds one data set at a time. Loading a new data set causes the currently held data set to be removed.

Selecting the option "Load Multiple Data Sets" in Options / General (requires a program restart to activate) allows data sets to be retaned in memory and rapidly swapped. When a new file is loaded, the current data set is moved to a holding area ("in the background"). It is moved back to the "foreground", for

© James Davidson 2024

analysis either by selecting it using the menu (see below) or automatically, by loading a model which specifies it. The data set currently in foreground takes its place in the holding area. The stored data sets are remembered between sessions and reloaded at start-up, provided the disk files containing them have not been moved or deleted.

Switch To...    Opens a popup menu with the list of data sets held in background memory. Click a name to bring the selected data set to the foreground for analysis. The data set currently in foreground (if any) moves to the background.
Note: This menu is also opened by the "Open File" button on the tool bar. When "Load Multiple Data Sets" is disabled, this button opens the popup menu for disk file access.

Remove...    Opens a popup menu with the list of data sets in background. Click a name to remove the data set from memory.

Remove All    Removes all data sets from memory.

New Data    Clears the foreground data space, moving the current foreground data set to the background.

Notes:

1.    Data sets must have unique names corresponding to the name of the file containing them, excluding the extension. It is not possible to load files with the same name simultaneously, even if they have different extensions (file types) or reside in different directories. A warning is shown, and unless the action is cancelled the new data set will overwrite the existing one.
Tip: to avoid a clash, save the current data set with a new name before opening its namesake.

2.    Use the New Data command to allow a new data set to be created in the Setup / Data Transformation and Editing dialog. This command is available in the menu whether or not "Load Multiple Data Sets" is enabled.

3.    When the New Data and Remove All commands are given, all menu items and toolbar buttons that operate on data are disabled.

4.    The Switch To... and Remove... popup menus include the name of the foreground data set in disabled (grayed out) mode. Clicking this item has no effect. Remove the foreground data set using the New Data... command.

5.    Clearing a data set does not delete the disk file. The data can be reloaded at any time.

6.    When the File / Settings / Export... command is given, all the currently stored data sets are saved in the specified settings (.tsm) file. Opening this file will cause the data sets contained in it to be re-loaded. They are written to individual files in the working directory, using the currently selected data file format.

7. By default, a maximum if 20 data sets can be stored at one time. Change this default by editing the tsmgui4.h file (see Appendix B).

## *3.3 Setup / Data Transformations and Editing*

In this dialog, data series can be edited, renamed, deleted, and created as dummies or by transformations of existing variables. The data can be rearranged, sorted, reversed and written to new files. Observations can be pasted from external sources via the Windows clipboard.

* Commands selectable from the pull-down menu are arranged under the headings "Edit" and "Transform".
  o Press a button to bring up the desired list of choices.
  o Choose the desired operation or transformation.
  o Where appropriate, select one or more variables from the list.
  o Press "Go" to execute the command.

* Press "List/Edit Series to view and/or edit observations in spreadsheet format, after selecting one or more variables from the list. Alternatively, double-click the selection. This option is also available (without the variable selection option) through the menu item
  Setup / Data Spreadsheet... See 3.4 for further details.

* To mark a variable, highlight it and press "Mark". To remove the mark, press "Mark" with no variable highlighted. Marking enables operations involving a pair of variables.

* "Select All" highlights the whole list, "Clear" removes all highlights.

* The "Sample..." button brings up the sample selection dialog. Editing operations and transformations are performed for the selected subset of the data, where appropriate.(This option is not available for panels.)

* The "Formula..." button opens a text field where a formula for a new variable can be entered directly using the notation described in 1.5. This option is also available through the menu item Setup / Make Data Formula... See 3.5 for further details.

* If changes are made, the "Save Modified Data" button is enabled. Use this to open the choice box for saving, either to the original file or as a new file.

EDIT COMMANDS

Note: commands marked with * do not operate on individual variables. Simply press 'Go' to execute. Any highlights in the variable list are ignored in these cases.

"Rename"

After selecting a variable and pressing 'Go', a text field appears for typing the new name. The dialog is frozen until 'OK' or the [Enter] key is pressed to

© James Davidson 2024

complete the entry. If multiple variables are highlighted, each is presented for renaming in turn.

* Characters that are illegal in variable names appearing in formulae, (see Section 1.5) can be automatically removed or replaced in existing names by entering special strings (upper or lower case) in the New Name text field:

    o    If the New Name is entered as #DELSP, the existing name is modified by removing illegal characters.

    o    If the New Name is entered as #REPSP, the existing name has illegal characters replaced by underscores.

    o    If the New Name is entered as #DELSPFR, the existing name is truncated at the first space.

"Delete"

A confirmation box appears. Press the 'All Selected' button to delete all highlighted variables at once, otherwise each is presented in turn for confirmation.

CAUTION! This command cannot be undone.

"Edit Description"

Allows optional description(s) of the selected variable(s) to be created or edited. View these descriptions in Help / Data Descriptions.

NOTE: If the total length of the name + description string (including separator character) does not exceed 64 characters, the string is stored in the name field of the data spreadsheet file. If this limit is exceeded in one or more cases, only the truncated strings are saved with the data. The complete strings are saved to a text file with matching name and extension ".names.txt".

"Save Selected"

Allows a subset of the variables to be saved in a new named file. Highlight any number of variables in the list before giving the command. This opens the file dialog for saving, provide a new name and type for the file. The existing data set is unaffected.

"Make Seasonals" *

For frequency N, create N dummy variables with every Nth observation = 1, otherwise 0. Set N = 4 for quarterly data, N = 12 for monthly, etc.

"Move Up", "Move Down"

Use these commands to re-order the variables in the data set. Pressing "Go" repeatedly will move a single highlighted variable up or down as many positions as required. The commands are circular, "Move Down" takes the last variable in the list to the top, and "Move Up" takes the first variable to the end. Alternatively, these commands move a block of highlighted variables, but in this case the highlighting is cleared for all but the last.

"Resize Sample" *

Use this command to create additional space for data to be entered by hand, or created by simulation. Type the desired number of observations into the text box provided.

* If this number is less that the number of observations already existing,

© James Davidson 2024

the sample is trimmed to the specified length.
CAUTION: Excess observations are deleted, this operation is not reversible.
* If the number is greater than the current length, existing series are extended with NaN values. A new variable "Zeros" is created automatically with the full number of observations, with value 0.

Notes:
1. Use Zeros as the dependent variable when specifying a model for simulation, renamed as required.
2. If the data set is initially empty when this command is given (give the command File / Data / Clear to clear the current set) the new data set receives the associated file name 'newdata ' with the currently selected file type.

"Set Sample Start" *
Either trims the number of observations specified from the start of the sample, or creates new rows (set to .NaN) if the number specified is negative. In the latter case the option allows the data set to be extended manually with new initial observations.
CAUTION: this operation is not reversible.

"Delete Selected Rows"
This command deletes complete rows from the data matrix for which the the selected variable equals the constant entered in the box. The constant value can be .NaN, to delete all rows where the selected variable is missing. E.g., to extract the observations for a particular quarter from a quarterly data set, create seasonal dummies and delete the cases where the dummy for the chosen quarter is 0.
CAUTION: this operation is not reversible. Be careful to save the reduced data set under a new name!

"Make Zeros", "Make Intercept", "Make Trend" *
These commands create dummy variables. Intercept has values 1, Trend has values 1,2,3,...,T for a sample of T observations.
Notes:
1. To create an event dummy, first create Zeros, then edit it observation by observation in List/Edit/Compare.
2. Use the built-in intercept and trend dummies for estimation wherever possible.

"Make Normal(0,1)", "Make Student t(df)", "Make Uniform"  *
Uses the built-in random number generator to create series of independent observations with the specified distribution.

"Make PDL"
Creates polynomial distributed lag variates (see Models and Methods for details) for highlighted variable(s), with lag length specified by the scroll bar. Five new variables are created, corresponding to terms of order 0 up to 4 in the lag polynomial. Delete any unneeded cases.

"Make .NaN"*

© James Davidson 2024

Creates an empty variable into which data can be entered in the editing spreadsheet. Useful for importing a new variable, in conjunction with the clipboard pasting option.

"Reverse Data" *

Operates on the whole data set, listing the observations in reverse order.

"Sort Data By" *

Operates on the whole data set, sorting the observations according to the ascending order of the highlighted variable. Don't select more than one - If two or more are selected, only the last choice is effective.

"Undo Sort" *

Reverses the action of the Reverse Data and Sort Data By commands, restoring the data set to its original order.

"Shape"

Re-assembles, by rows, a data series of length T into c columns of length ceil(T/c), where c > 1 is selectable. In case T is not divisible by c, the columns are padded to equal length with zeros, and then with NaNs to fill the rest of the data matrix.

Uses for Shape include the following:

1    Separate the real and imaginary components of a discrete Fourier transform. Set c = 2.
     Note: the jth periodogram point, for j = 1,..., [T/2], is the sum of squares of the jth row elements of these columns.

2.   Separate seasonal data into its seasonal components. For example, with quarterly data, set c = 4. If the first observation is from Quarter 1, then the first column created will contain the Quarter 1 observations, the second column the Quarter 2 observations, and so on.

"Vectorize Columns"

Arranges the selected variables (two or more) into a single column by stacking the transposed rows. The new variable is saved to a file.
For example, use this command to create a single series from a table of monthly observations where the columns contain observations for each month and the rows correspond to years.
Note: Use "Shape" to recreate the original table.

"Set Sample Indicators"*

Opens the data editing dialog showing a variable "!selectobs!" which initially has all observations set to 1. Change a value to 0 to indicate that this observation should be omitted from the estimation sample. If an indicator set already exists, this command allows it to be edited.
Note:

1.   Once this option has been selected, an indicator variable named "!selectobs!" is saved with the data set, and is enabled on re-loading. It can be edited in a spreadsheet program if desired.

2.   The indicator does not appear in the variable list, but its values are displayed in the editing dialog following the

© James Davidson 2024

observation numbers/dates.
3.      To use the sample indicators, check the checkbox in
        Setup /Set Sample. (This checkbox is hidden unless
        indicators exist.)

"Set Dates" *
        Allows the dates of observations to be set or changed. This works exactly like
        the method for inserting date information into a data file prior to loading;
        see File / Data. for details. A new column of the data set is created with name
        !STARTDATE!, and opened for editing. Enter the initial year in the first row,
        the initial sub-period in the second row, and the sub-period frequency in the
        third row. Enter 2 for half-years, 4 for quarters, 12 for months, etc.
        Notes:
        1.      The new column is deleted immediately after the new information is
                processed.
        2.      The date information is saved with the data in GiveWin or other
                spreadsheet formats, and also in ".dat" files, although not in matrix
                (.mat) files.
        3.      CAUTION! This operation is not reversible. If the observations
                are already dated, the existing date information will be lost, including
                individual daily dates.
        4.      If the period is set to 0, the series is treated as undated. Observations
                are numbered from 1.

"Set Date Labels" *
        A variable with the reserved name "!DateLabels!", is created if it does not
        exist, and opened for editing. Enter the required date labels for each row of
        the table. Arbitrary numerical values are allowed. This variable is not
        displayed in the variable list, but appears in the Date column in the data table,
        taking precedence over Ox-format Year-Period dates. To reinstate the latter,
        give the "Set Dates" command, as above.
        CAUTION: Calendar date information will be deleted permanently.

"Print to Results Window"
        This command allows observations to be cut and pasted into other
        applications. Variables are listed in single columns, even if two or more are
        selected. Use the Sample dialog to select the required range, before giving
        this command.

TRANSFORM COMMANDS
*       Most commands in this menu can also be executed using the Edit / Make
        Formula command. Here, they can be performed with a single click on
        any number of highlighted variables.
*       Created series is/are added to the end of the list, with a self-explanatory
        prefix or suffix appended to the name(s). Use Edit / Rename to change
        the automatically created name, if desired. Use Edit / Delete to remove
        any unwanted series.
*       Xt, t = 1,...,T denotes a highlighted variable in the data set.
*       Mt, t = 1,...,T denotes the 'marked' variable, if any.
*       j represents the scrollbar setting for leads/lags/differences.
*       C, d, E, Lamda represent values entered in the text field.

© James Davidson 2024

| | | | | |
|---|---|---|---|---|
| "Lag" | . | . | . | Xt-j |
| "Lead" | | . | . | Xt+j |
| "Difference" | . | . | . | Xt - Xt-j |
| "Log-Difference" | . | . | log(Xt) - log(Xt-j) |
| "% Difference" | | . | . | 100.(Xt - Xt-j)/Xt-j |
| "Fractional Difference" | | . | Sum{i=0,...,t-1}Bi*Xt-i, where Bi = Gamma(i-d)/(Gamma(d).Gamma(i+1)) |
| "Cumulate" | . | . | . | Sum{j=0,...,t-1}Xt-j |
| "Logarithm" | . | . | . | log(Xt) |
| "Exponential". | | . | . | exp(Xt) |
| "Logistic" | . | . | . | exp(Xt)/(1 + exp(Xt)) |
| "Log-Odds" | . | . | . | log(Xt/(C - Xt)) (C = 1 or 100) |
| "Box-Cox" | . | . | . | (Xt^lamda - 1)/lamda (lamda > 0) |
| "Square" | . | . | . | Xt^2 |
| "Square Root" | | . | . | Xt^(1/2) |
| "Reciprocal" | . | . | . | 1/Xt |
| "Power" | . | . | . | X^E (E an integer), \|Xt\|^E (otherwise) |
| "Sine". | . | . | . | sin(Pi*C*Xt) |
| "Cosine" | . | . | . | cos(Pi*C*Xt) |
| "Absolute Value" | | . | . | \|Xt\| |
| "Max(X,C) (C#Marked)" | | . | max(Xt, C) (max(Xt, C*Mt)) |
| "Min(X,C) (C#Marked)" | | . | min(Xt, C) (min(Xt, C*Mt)) |
| "Sign" | . | . | . | +1 if Xt >= 0, -1 otherwise |
| "Add C (C#Marked)". | | . | Xt + C (Xt + C*Mt) |
| "Subtract C (C#Marked)" | | . | Xt - C (Xt - C*Mt) |
| "Multiply by  C (C#Marked)" | | Xt*C (Xt*C*Mt) |
| "Divide by C (C#Marked)" | | . | Xt/C (Xt/(C*Mt)) |
| "Centre (Mean-Devs)" | | . | Xt - Mean(X) |
| "Standardize". | | . | . | (Xt - Mean(X))/SD(X) |
| "Detrend" | . | . | . | Xt - a - b*t (by regression). |
| "Moving Average" | | . | . | j-point moving average. |
| "Epanechikov Kernel" | | . | Kernel smoother. |
| "Gaussian Kernel" | | . | . | Kernel smoother. |
| "Triangular Kernel" | | . | . | Kernel smoother. |
| "Uniform Kernel" | | . | . | Kernel smoother. |
| "Biweight Kernel" | | . | . | Kernel smoother. |
| "Fourier Transform". | | . | . | See 9. below. |
| "Hodrick-Prescott Filter" | | . | HP filter with penalty Lamda |
| "Cubic Spline" | | . | . | Cubic spline, see 10 below. |
| "Principal Components" | | . | PCs of the selected variables. |
| "Reverse Series" | | . | . | Time reversal. |
| "Tanaka Filter" | | . | . | See 11. below. |

Additional Notes:

1. By default, missing observations resulting from forming lags, leads or differences are replaced by NaN, and these rows are omitted from estimation automatically. Select the "Pad with Zeros" radio button to replace with zeros instead, so that the rows are retained for estimation.

2. Note that when leads or lags are created, the excess observations are

not discarded. The end (start) of the complete data matrix is extended with extra rows, the missing observations being filled with either NaNs or zeros, depending on option selected.

3. If a sub-sample is selected using the "Sample..." button, the excluded observations are set to NaN in the new series.

4. NaN is returned in the case of illegal values (e.g., negative arguments of log(.) and (.)^(1/2), x < 0 and x > C in log-odds.)

5. The 'Power' transformation returns the power of the absolute value for non-integer exponents. The square root transformation is an alternative way to take the power of 1/2, but behaves differently, assigning NaN to negative observations.

6. The symbols # and % are used in the new variable names to denote multiplication and division in variable names. The more logical choices '*' and '/' cannot be used in file names.

7. "Moving Average" creates the MA of series coordinates from lag 0 to lag j-1. To create a centred moving average, set j to an odd value and apply the (j-1)/2-point "Lead" transformation.

8. The five alternative kernel smoothers compute weighted moving averages of the series using the indicated weight functions and chosen bandwidth H. Except for the Gaussian case, the jth smoothed element is the weighted average of 2H+1 points for $H < j < T - H$, and $H + j$ points for $1 <= j <= H$ and $T - H <= j <= T$ where T is the series length. For the Gaussian kernel, H is the standard deviation.
For the kernel formulae see (e.g.) the Probability package in the Ox language documentation.

9. "Fourier Transform" computes the FFT of the selected series. The real and imaginary parts of the transform are vectorized to give a real series of the same length of the original, with cosine and sine terms alternating. The components can be separated again if required, using Edit / Shape.
Note: The series as generated are of length 2T, and are trimmed to fit the existing data set. To avoid losing these points, first extend the data set using Edit / Re-Size Sample.

10. Set the cubic spline bandwidth to zero for the default setting. Positive values represent the number of parameters used in the fit, so larger values yield a closer fit to the data - matching the sample size yields a perfect fit.
For irregularly spaced observations, use Mark to select the variable containing observation dates, before giving the command.

11. The option "Tanaka Filter" operates on the selected series X to create the series X* where $X^*_{1} = 0$ and
$X^*_{t} = \text{Sum}_{1}^{t-1}(X_{t-j}/j), t = 2,...,T$
This construction features in various popular tests of fractional

© James Davidson 2024

integration order as proposed, notably, by Tanaka (1999). See also Breitung and Hassler (2002), inter alia.

## *3.4 Setup / Data Spreadsheet*

The command opens a new dialog in spreadsheet format, allowing individual observations to be viewed and edited.

The List/Edit dialog provides a simple spreadsheet-style editing interface.
* To select a cell, click with the mouse or use the navigation keys (Enter, Arrows, Page-Up, and Page-Down).

* To edit the contents of a cell "in place", double-click the cell to open for editing. Click a second time to highlight the contents.
  o To discard the edit and retain the existing value, press the "Escape" key.
  o To change the cell contents, press any navigation key or select another cell with the mouse.

* Alternatively, press 'Edit Observation'. The value in the currently selected cell appears highlighted in the Edit box. Type the new value. Then press:
  o 'Next' (or the [Enter] key) to save the new value, move the highlight to the next observation, and load this for editing.
  o 'Copy' to save the new value, move the highlight to the next observation, but retain the new value in the Edit field.
  o 'Cancel' to discard the new value and close the Edit box.

* To search a series for a particular entry, highlight any cell in the column and press Search. Enter the target string in the text window which opens. The "Search >>>" button searches down from the current cell, and the "<<< Search" button searches upward. The first occurrence of the target string is highlighted.

* To paste one or more data values from the Windows clipboard into cells in a column, select the first cell in the range and press the "Paste Clipboard" button, or enter the keyboard command Ctrl-V. Press Close, or the [X] button, to close the editing dialog.

Notes:
1. The left-hand column shows row numbers and dates, and is read-only. These cells cannot be selected or edited.

2. Editing "in place" may round the data value to fit the cell - six significant digits is the maximum. If precision is an issue, the 'Edit Observation' text field allows exact editing of the value, with up to 12 digits retained after the decimal point.

3. Use the Copy button to repeatedly insert the same value to successive observations. This option allows easy creation of (e.g.) dummies for sub-periods.

4.  Observations can be copied via the Windows clipboard from sources such as text editors, word processors, spreadsheet programs and the TSM results window. The clipboard must contain a character string consisting of data values and separators (files cannot be pasted)

    *   A data value is a string containing digits 0-9, optionally a decimal point (.) optionally starting with a minus (-), and may include an exponent with the form "e+nnn" or "e-nnn" where "n" denotes a digit and "E", "e" are equivalent.
    *   A separator may consist of any number of arbitrary characters except 0-9 and ".-+eE". Spaces, tabs, punctuation marks (except ".") and carriage returns are all eligible separators.

    For example: a variable listed by observation in columns (to be read left-to-right and down, a common format) can be copied complete to the clipboard and then pasted into a column of the table, appearing in the correct order.

5.  If the number of pasted values exceeds the number of available rows in the spreadsheet, the excess values are discarded. Remember that the spreadsheet dimensions match the sample selected for editing, not necessarily the full data set.

6.  CAUTION: Editing cannot be undone! To undo edits, re-load the original data file.


## *3.5 Setup / Make Data Formula*

This command gives direct access to a text box where formulae can be entered to create new variables as functions of existing variables in the data set.

The formula should be typed in the box in the format
    [New Variable] = [Formula]
where [New Variable] represents a variable name and [Formula] represents text formulated as described in Section 1.5, 'Entering Formulae' (Case A).

 Notes:
1.  These actions can also be performed as an option in the dialog Setup / Data Transformation and Editing - press "Formula".

2.  If [New Variable] matches an existing variable in the data set, a choice box is displayed with the option to over-write it. Otherwise, the new variable is added to the data set.

3.  Variables can be referenced either by name or in the format X[j] where j denotes the column number in the data set. These numbers are displayed in the variable list in Setup / Data Transformation and Editingwhen this option is accessed there, as noted in 1. Be careful not to overlook that these numbers may change, relative to the corresponding names, if variables are deleted and/or

© James Davidson 2024

re-ordered. Reference by name is the safer option, especially when stored formulae are re-evaluated after data modifications..

4.     To evaluate a recursive formula, use E#{-j} to denote the function value lagged j periods.
By default, zero initial values are assigned to E#{-j}. To assign different start-up values do as follows, assuming the recursion contains M lags.

1)     Create a new variable with the name to be assigned to the formula.
2)     Edit this series, as necessary, so that the desired start-up value(s) appear in the M initial positions.
3)     Press "Sample..." and advance the initial date for the calculations by M steps.
4)     Create the formula, selecting "Yes" when prompted to overwrite the series.
5)     Reset the initial date to view the complete series in the spreadsheet, including the start-up values.

5.     If the formula as entered cannot be parsed, an error message appears and nothing is changed.

TEXT BOX CONTROLS

"Close"         Closes the text box.
"Delete"        Deletes the currently displayed formula, the clears field.
"Next"          Displays the next formula, if it exists.
"New"           Opens a new blank formula where there is no "Next".
"Previous"      Displays the previous formula, if it exists.
"Go"            Creates new variable from the displayed formula.
[Return]        Equivalent to pressing "OK".
[Escape]        Discards current editing, restores existing formula.
Closing the text box with the Windows Close button [x]  is equivalent to pressing "Cancel".
"Run All"       Evaluates all the currently stored formulae in sequence.

The button bar also displays a formula number (i) in the format [ i / n] where n denotes the number of formulae currently stored.

## 3.6 Setup / Automatic Model Selection

Two options for estimating a sequence of models automatically are controlled from this dialog. Note that they cannot be run simultaneously.

AUTOMATIC REGRESSOR SELECTION

When this option is selected in the Actions menu, or by pressing the "Run Regressor Selection" button, the program chooses the set of regressors to optimize the chosen model selection criterion, out of the specified regressors in a baseline model. The option is described in detail in 6.4 Actions / Automatic Regressor Selection. Use this dialog to choose the selection criterion and which regressor 'Types' are to be included in the search.

For example, select 'focus' variables (to be included in every specification) as Type 1 regressors, and 'nuisance' variables, whose inclusion is optional, as Type 2 regressors. To conduct the search over the latter set only, uncheck the box "Include Type 1 Regressors" and check the box "Include Type 2 Regressors".

Notes.
1.     The selection criterion, one of Akaike, Schwarz and Hannan-Quinn, is chosen with the radio buttons. This selection is independent of that in Options / Tests and Diagnostics, which controls the choice of lag lengths in ADF cointegration tests and Saikkonen-Stock-Watson efficient estimation.

2.     Tip: to make the lags of a focus (Type 1) variable optional, set Type 1 lags to 0, and then create the one-period lag in Setup / Data Transformation and Editing. Include this variable as a Type 2 with Lags set to the desired maximum.

3.     A confirmation box opens to guard against accidental launches. This reminds you what selection criterion is currently chosen.

MULTIPLE ARMA MODELS

When this option is selected in the Actions menu, or by pressing the "Run Multiple ARMA" button, the program automatically estimates all the ARMA(p,q) or ARFIMA(p,d,q) specifications, for all values of p and q in the indicated ranges.

*     Enter max p, max q and max p+q in the text fields provided. With the default settings (all entries = 2) the program estimates the following cases of (p,q), in the order shown: (0,0), (1,0), (2,0), (0,1), (1,1), (0,2).

*     A confirmation box appears to guard against accidental launches.

*     See Actions / Estimate Multiple ARMA Models for additional details.

## 3.7 Setup / Recursive/Rolling Estimation

This dialog sets up the options for recursive estimation. For details on the output produced by this option, see Contents / Actions Menu / Recursive/Rolling Estimation.

*     Choose rolling estimation (fixed sample size) or incremental estimation (extending sample size) with the radio buttons.

*     Set the initial sample in the usual way, as for the estimation sample. Access the Set Sample dialog either using the button provided, or in Setup / Set Sample or in a model dialog.

*     The 'Go' button duplicates the Actions Recursive/ Rolling estimation button. If this dialog is open, the "Run" button on the toolbar is a third way to launch the run. A confirmation box opens to guard against accidental

© James Davidson 2024

launches.

Notes:
1.      'Step size' is the number of extra observations added at each step. Set by dragging the scrollbar. If this setting is greater than 1 the sequences of estimates are plotted as step functions.

2.      The terminal observation required will in most cases be the maximum available. The scrollbar allows it to be set smaller. The lower limit is constrained by the initial sample and step size settings.

3.      The first check box allows the option of saving regression diagnostics and test statistics to be selected. This is on by default.

4.      The second check box allows saving to a file of all the forecasts for each estimation period.
        NOTE:  The forecasts, other than the N-step-ahead, cannot be displayed interactively.

5.      By default the number of forecast steps (N) is fixed, and the N-step forecast is reported for each recursion step. If the second check box is checked, the terminal forecast date is fixed, so that N contracts as the end of the sample advances. This option can be used to compare the forecasts of a particular date from different standpoints.

## 3.8 Setup / Compute Summary Statistics

This dialog presents the variable list. Select one or more variables and press 'Go' to print tables of summary statistics. Alternatively, just double-click a variable. The standard outputs include sample range, moments and median, and the Jarque-Bera statistic for test of normality.

Optionally, the following statistics relating to serial dependence can be computed:

HAC Variance:
        Kernel-weighted estimator of omega, the variance of the series sum divided by square root of the sample size.
                omega =  sigma + 2*lambda
        where sigma = variance and lambda = sum of autocovariances, finite under weak dependence (short memory) and positive unless the series is over-differenced.
        If autoregressive prewhitening is specified in the calculation, the kernel estimator is computed for AR(1) residuals. Omega is the resulting variance multiplied by $1/(1 - alpha^2)$, where alpha is the reported AR coefficient.


Tests of I(0):

1.      The Robinson-Lobato (1998) nonparametric test of I(0). The bandwidth m is calculated by the rule described in the paper.

© James Davidson 2024

2.      Lo's (1991) R/S test for I(0) against I(d), for  d > 0 or d < 0.  A bounding
        p-value (probability of the statistic exceeding the computed value under
        H0) is given in braces.

3.      The Kwiatkowski et al. (1992) (KPSS) test of I(0), against I(1). A
        bounding p-value (probability of the statistic exceeding the computed value
        under H0) is given in braces.

4.      The Harris-McCabe-Leybourne (2006) N(0,1) test of short memory (see
        8.2 Options / Tests and Diagnostics for additional details.

Tests of I(1):

1.      Phillips-Perron (1988) test of I(1) against I0). A bounding p-value (probability
        of the statistic lying below the computed value under H0) is given in braces.

2.      The augmented Dickey-Fuller test of I(1) against I(0). The number of lags is
        chosen to optimize the Schwarz information criterion over the range 0 to
        $[T^{1/3}]$.

3.      Elliott-Rothenberg-Stock (1996)  GLS-DF test of I(1)

4.      Elliott-Rothenberg-Stock (1996)  P (feasible likelihood ratio)  test of I(1).

Robinson's d.
        Semiparametric estimator of the long memory parameter d, see Robinson,
        Anns. Stat. 22 (1994) 515-539. The statistic corresponds to eq. (4.2) of
        the paper with lamda_m = 2 pi $n^{-0.35}$ and q = 0.5.

Notes:
1.      Check the 'Detrend?' box to compute the statistics for the series after
        detrending by OLS.

2.      Check the 'Differenced?' box to compute the statistics for the differenced
        series in addition to the undifferenced series.

3.      Choose the order of correlograms of the series and squared series, with
        associated Q statistics, using the scroll bar. If this is set to 0, no
        correlograms are computed. Choice of Box-Pierce or Ljung-Box Q test
        variants is selected in Options / Tests and Diagnostics.

4.      Check the "Data Correlations" checkbox with Correlogram Order set to 0,
        to see the contemporaneous correlation matrix of two or more series.
        In this case, no other statistics appear and checkboxes are "greyed out".
        If only one series is selected with this option set, no output is produced.

5.      Check the "Data Correlations" checkbox with Correlogram Order set to
        a positive value to compute the correlograms and cross-correlograms of a pair
        of series. Denoting the selected variables as 'X' and 'Y' respectively, the four
        columns of the table show, for lags j > 0:

© James Davidson 2024

* The autocorrelation sequences  Corr(X_t, X_t-j) and Corr(Y_t, Y_t-j)
* The cross-correlation sequences  Corr(X_t , Y_t-j) and Corr(Y_t , X_t-j)

Note that stationarity is assumed so that Corr(X_t, Y_t+j) is taken to be the same as Corr(Y_t, X_t-j).

6.  Check the "Partial Correlogram" checkbox to print the partial autocorrelation functions (PACF) instead of the simple autocorrelations.

7.  Optionally, the number of correlogram and partial correlogram points plotted in Graphics / Show Data Graphic can be controlled by the setting in this dialog.  Check "Make Graph Order" to enable this feature.

8.  Check the "Quantiles" checkbox to report the following quantiles of the data distribution: 0.01, 0.05, 0.1, 0.3, 0.5, 0.07, 0.09, 0.95, 0.99. For small samples the extreme cases are omitted.
    * If detrending  is specified, the mean of the distribution is zero by construction.

9.  The "Test and HAC Options..." button gives access to the Options / Tests and Diagnostics dialog, where lags for the ADF test, and kernel and bandwidth for computation of the KPSS, V/S, RS, PP and ERS-P tests, and the HAC variance, are selected. Choose these manually or select information criterion and/or plug-in automatic choices.
    NOTE: If the Summary Statistics dialog is open when these selections are made, they are saved until either changed similarly, or program settings are reset to defaults. They are independent of settings made in the same dialog when specifying a model.

10.  The "Sample..." button opens a dialog to set the sample period to be analysed (Sample 2). See Setup / Set Sample for details.

11.  The test p-values are computed from the published tables and are not available for every ordinate. The values given are upper bounds. Dickey-Fuller's table for the maximum sample size is quoted for the ADF and Phillips-Perron tests, and in this case the probabilities refer to the lower tail.

12.  Check the "Report LR Variance" checkbox to see the estimates of the long run variance of the series differences in an I(1) test. For the ADF and ERF-GLS test, this is computed from the fitted autoregressive model parameters in ADF regression. In the PP and ERF-P tests, the spectral estimate (HAC variance estimate) is reported.

BOOTSTRAP TESTS OF I(1)

Bootstrap versions of the I(1) tests are implemented by resampling the differenced and centred test series and integrating the resulting draws.

13.  With this dialog displayed, open Options / Simulation and

Resampling and put "Bootstrap Replications" to a positive value.
The option to compute bootstrap tests is then offered in a choice
box when "Report I(1) Tests" is checked.

14.     It is the user's responsibility to choose an appropriate resampling
        method. Sieve autoregressive prewhitening with automatic lag
        selection by Akaike's criterion is selected by default.

15.     Selections made in Options / Simulation and Resampling while
        this dialog is open are saved independently of settings for bootstrap
        inference for a model. Hence, these settings should only need to be
        made once. Disable the option by putting "Bootstrap Replications"
        to 0.

16.     An alternative way to perform these tests (useful for Monte Carlo
        experiments) is to run an estimation in "Raw Series Diagnostics"
        mode. See 6.1, Actions / Run Estimation.

## 3.9 Setup / Nonparametric Regression

Opens a dialog to specify a Nadaraya-Watson estimate of the conditional mean,
using the Gaussian kernel. Two variables must be selected from the list, the
dependent variable y and regressor x. Use the radio buttons to toggle between
these choices.

Press "Sample ..." to open the Set Sample dialog and select the sample to be
analysed (Sample 3). This sample is shared with the semiparametric long
memory estimators, and can be set in either dialog.

By default, the bandwidth used is $S.n^{-1/5}$ where n = sample size and S is
the sample standard deviation of the regressor. Use the scroll bar to increase
or decrease this setting by factors of 2.

The function is plotted with the y-x scatter optionally superimposed.

## 3.10 Setup / Semiparametric Long Memory

This dialog computes estimates of the long-range dependence parameter d, using
one of three semiparametric methods:

*       "Narrow-band" log periodogram regression (Geweke and Porter-Hudak,
        1983) The log-periodogram points for j=1,...,M are regressed on
        -2log(sin(pi.j/T)) where T is sample size, and M/T --> 0.

*       "Broad-band"  log periodogram regression (Moulines and Soulier, 1999)
        The log-periodogram points for j=1,...,[T/2] are regressed on
        -2log(sin(pi.j/T)) and P Fourier terms (cosine functions of pi.j/T) to
        approximate omitted short-range components, where P --> infinity but
        P/T --> 0.

*       Local Whittle Gaussian maximum likelihood (Kunsch 1987, Robinson 1995)
        Maximizes the  frequency domain Gaussian likelihood as function of d, using

© James Davidson 2024

the first M periodogram points where M/T --> 0.

The outputs from this option includes point estimate, asymptotic standard error and t-test of significance.

Notes:
1. To run the estimation, select a variable or variables from the list, and press Go. Multiple selections are permitted, and will be computed in succession. Alternatively, just double-click on a variable in the list.

2. The series may be either differenced, or detrended by regression, prior to computing the periodogram. Choose option with the radio buttons.

3. Press "Sample ..." to select the sample for estimation (Sample 3). This is stored independently of the sample selection for other program functions, but is shared with Nonparametric Regression, and can be set in either dialog. See Setup / Set Sample for details.

4. For the GPH bandwidth and trim options, and MS Fourier terms, the power of T being selected (to nearest integer below) is shown as a guideline. Note that the optimal bandwidth is typically a power of T times a constant factor, that may be small. The weight here is 1.

5. The GPH bandwidth setting is chosen with the scroll bar GPH recommended $M = O(T^{(1/2)})$. The minimum MSE setting is known to be $M = O(T^{(4/5)})$, but the bias can be large with this choice. Experimentation is recommended to see how sensitive the estimate of d is to the selection.

6. The GPH trimming option also omits the lowest frequencies. See the relevant literature for guidance on this setting. If in doubt, keep at 0.

7. GPH estimation optionally returns a Hausman-type test for the presence of bias (Davidson and Sibbertsen 2009), by comparing broad and narrow bandwidths. The broad bandwidth should be set to the maximum (T/2) to test the hypothesis of a pure fractional process. Use this setting if in doubt. Setting the bandwidth to $CT^{beta}$ for $4/5 < beta < 1$ yields a test of a more general hypothesis, allowing any short-run dependence that does not induce bias. The test is inconsistent with $beta <= 4/5$.

8. GPH can also return a skip-sampling test of the null hypothesis of long memory (Davidson and Rambaccussing 2015). This is most effective in combination with the Wald test (t-test) that d is zero. To enable this option, set the parameter delta (see cited paper for details) to a positive value; thus, 0.3 may give a reasonable trade-off between size distortion and power.

9. The choice of MS Fourier order is also a trade-off between bias and efficiency - see MS paper for recommendations, or experiment.

10. The smoothing option, available for both estimators, uses local non-overlapping averages of periodogram points in the regression. If in doubt,

keep this setting at 1.

11.    To do Monte Carlo experiments on the estimators specified in this dialog, have the dialog open when the model is stored in the Model Manager, or with Model / Save Current Model...

12.    The local Whittle estimator uses a line search algorithm. The convergence criterion and the maximum iterations are set in Options / General / Special Settings. These settings are shared by other program features, but should not generally need changing.

## 3.11 Setup / Cointegration Analysis

This dialog allows a Johansen-type multivariate cointegration analysis to be carried out on any subset of the data. The available procedures are:

*     Testing hypotheses of I(0) and I(1) on the selected variables.

*     Lag length selection for the Johansen CIVAR, by Akaike, Schwarz or Hannan-Quinn selection criteria.

*     Johansen maximum eigenvalue and trace tests for cointegrating rank.

*     Wald tests of structural exclusion restrictions on the cointegrating vectors, including MINIMAL analysis.

Select the variables to be included and CIVAR options, then select the desired procedure using the radio buttons. To execute the procedure press Go, or the Run or Evaluate buttons on the tool bar.

Notes:

1.    Open the Set Sample dialog using the "Sample ..." button to set the desired sample (Sample 4). This selection is independent of those set for other program functions. See Setup / Set Sample for details.

2.    The KPSS test of I(0) and P-P test of I(1) are also available in the Setup / Summary Statistics dialog, but here, for convenience, are computed for all the selected variables at once.

3.    The p-value inequalities reported are taken from the tables of the eigenvalue and trace tests in Osterwald-Lenum (1992)

4.    A check box allows the inclusion/exclusion of drift terms in the CIVAR. If the drift is suppressed, intercepts of the estimated cointegrating vectors can be computed. However, note that standard errors are not computed for these.

5.    If the 'Drift' option is selected, a chi-squared statistic for the significance of the drift is reported following the maximal eigenvalue and trace tests. The degrees of freedom are $p - r$ where $p$ is the dimension of the cointegrating VAR. (See e.g. Davidson (2000) Section 16.4.3 for details.)

Note: that this test is only valid if the cointegrating rank is actually equal to r, the row of the table in which the statistic appears.

6.    MINIMAL analysis is described in Davidson (1998a). A sequence of chi-squared tests is computed to estimate the set of 'irreducible' cointegrating relations (such that cointegration is not a property of any subset of the included variables.) Identified, structural cointegrating relations must belong to this set. If the cointegration hypothesis is not rejected, estimates of the vectors are reported, with asymptotically valid standard errors, and Phillips-Perron statistics for the corresponding 'cointegrating residuals'.

7.    There is the option to specify a single test of the hypothesis that the cointegrating space contains an element subject to specified zero restrictions. When the radio button for this option is selected, the variable list shows only the currently selected CIVAR variables. Select a subset of these to test the hypothesis that they form a "cointegrating subset".

8.    Scroll bars allow setting the number of additional lags in the CIVAR, and also the cointegrating rank to be assumed for computing the 'beta' matrix of cointegrating vectors. This setting is taken as given in the MINIMAL analysis.

9.    See Davidson (1998a) for details of the "rule of thumb" adjustment to the chi-squared critical values.

## 3.12 Setup / Monte Carlo Experiments...

This dialog provides the options to set up and run one or more simulation experiments. To make use of it, there should normally be at least one model stored using the Setup / Model Manager dialog. If two different models are stored, one can be used to generate the artificial data and the other to specify the estimation, allowing a flexible approach to misspecification analysis.

*    Select one or more Data Generation Models (DGMs) from the list

*    Select one or more Models for Estimation (EMs) from the list (can be the same as or different from the DGMs).

*    Use the scroll bar to select the number of replications, and set other options with the checkboxes.

*    Press "Run" to perform the experiment(s). To interrupt a run in progress, press the "Stop" button on the toolbar, with the options of aborting or continuing.

*    Press "Clear Models" to clear all highlights from the model list.

*    Press "Results" to
     o    Re-display the tables from the experiment last run with the currently loaded DGM.
     o    Load and display Monte Carlo results from file, including batch output.

Notes:

1.     The model list can be rearranged, and also sorted alphabetically, in the Model Manager (see Section 4.1).

        Note:   When the model list is sorted, block selection by dragging or right-clicking the mouse (see Section 1.4) is disabled.

2.     Multiple experiments can be set up by selecting two or more models in either category.

      o       If one EM is selected and two or more DGMs, then the experiment is performed on the EM using each DGM in turn.

      o       If two or more EMs are selected and one DGM, then the experiment is performed on each EM using the specified DGM.

      o       If multiple models are selected in both categories, then all the possible experiments are performed - each DGM with each EM in turn.

      o       The same model(s) can be selected in both categories.

3.     In a single experiment, if "Run" is pressed a second or further time without changing any run settings, the option is presented to add further replications (the number currently selected). If the last run was aborted, then the options of either resuming or extending the last run are presented similarly.

        This feature is not available when multiple experiments are specified.

4.     The sample period for estimation is always defined by the DGM. The sample settings for the EM are ignored. To have a dynamic simulation run for a number of start-up periods, select the "Presample Data Random in Simulations" option in Options / Simulation and Resampling. Then set the DGM sample period as desired for estimation.

5.     If no models are defined, the current specifications are stored as a model with the default name (run ID) and used for both DGM and EM.

6.     The DGM specification MUST include a choice of shock distribution (Gaussian/Stable, Model or Bootstrap) selected in the Options / Simulation and Resampling dialog. Set the shock variance in the same dialog for the Gaussian option, and appropriate parameter values in the Values dialogs for the Model option. Make these selections before saving the DGM.

7.     Between 1 and 1 million replications can be specified. The selectable values are in the ranges 1,2,..,5,10,...,50, 100,...,1000, 1500,...,10000, 15000,..., 100000, 150000,...1000000.

8.     If a test takes a variable parameter, such as degrees of freedom or number of tested restrictions, the tables display this value enclosed in square brackets appended to the test name. The suffix is omitted if the parameter has a default value, such as 1 in the case of t-ratios.

OPTIONS

9.     The parameter moments reported are the mean, standard deviation, skewness and kurtosis of the Monte Carlo distribution of parameter estimates.

                                                   © James Davidson 2024

10. Optionally, moments can be computed for the estimated standard errors of the estimates, using whatever formula is specified in the EM (standard, robust or HAC). The mean of this distribution may, for example, be compared with the SD of the Monte Carlo distribution of the estimate.

11. To calculate parameter biases and RMSEs, the assumed "true" parameter values must be set as part of the EM. These may match those of the DGM, but note that the two models need not even be comparable. Set these values for unconstrained parameters using the Values dialogs in the usual way.
Note: this option is not available for semiparametric long memory estimation, which does not make use of the Values dialog. The bias and RMSE can be hand-computed from the information reported in this case.

12. Any optional tests that are specified in the estimation model will be included in the tabulations. The routinely computed diagnostic tests (Jarque-Bera test and Q tests) can be optionally included in the tabulations. By default, they are omitted.

13. The upper-tail quantiles of the distributions of the test statistics (50%, 10%, 5%, 2.5%, 1%) are optionally reported in the case of parameter t values. By default, the absolute t values are tabulated, so that, for example, the 5% tail quantile should be approximately 1.96 in large samples when the null hypothesis is true.

14. Also optionally reported are the empirical distribution function (EDF) points for the nominal p-values. If the test criteria are correctly sized and their null distributions are continuous, these distributions should be Uniform [0,1] if the null hypothesis is true, and 100x% of replications should lie below x. Over-rejection yields a value larger than the nominal one. The cases tabulated are x = 0.01, 0.025, 0.05, 0.1, 0.5.

15. The option "2-sided" causes the p-value EDF to be computed for the upper and lower tails combined, as well as the upper tail. For example, in the case "5%", the reported value is the fraction of p-values in the replications falling outside the equal-tails range of [0.025, 0.975]. When combined with signed t-statistics (see next paragraph) this option gives a result comparable to tabulation of the absolute t-statistics, but differing in the case of asymmetry about zero. It also shows the coverage probabilities for any user-defined statistic.

16. Optionally, the signed t statistics can be tabulated instead of the absolute t statistics, to allow the evaluation of one-tailed tests. In this case, the 5% tail quantiles should be approximately 1.64 in large samples when the null hypothesis is true. Note that the upper tail is tabulated, so be sure the parameters are signed appropriately.

17. The t statistics can optionally be centred on the parameter values specified by the EM (see para. 11 above). Otherwise, they are centred on 0. When this option is selected with the EM and DGM values matching (as when

© James Davidson 2024

the same model takes both roles) the distributions of the test statistics can be studied under the (true) null hypotheses.
Note: this option is available only when the bias and RMSE are available.

18. When the distribution of test statistics follows the distribution specified for the null hypothesis, and this distribution is continuous, the p-values should be uniformly distributed, so that their c.d.f. is the 45-degree line of the unit square. The Kolmogorov-Smirnov statistic (the maximum absolute difference between the p-value EDF and the uniform distribution) can optionally be calculated. This provides an alternative to the quantiles to assess how closely the distributions match. Nominal critical values for the K-S distribution, assuming 80+ bins, are 1.22 (10%), 1.35 (5%), 1.63 (1%).
Note: this option is only available if the frequency table is enabled, see para 20 below.)

19. If GMM is simulated, there are two options available, the "one-step" estimator, and the efficient GMM estimator using the first-round estimates to compute the weight matrix. Check the "Iterate GMM" checkbox to select the latter option. Further iterations are not possible, in this release.

20. If "Print Model Descriptions" is checked, the complete non-default model settings are printed in the output. These are codes in the TSM scripting language, see the Programming Reference document for details.
Note:   Thee same lines appear in the model description field in Model Manager, if "Automatic Descriptions" is checked when the model is saved. However, they are generated independently and do not need to exist in the stored model.

21. By default, the individual replications are stored. These can be plotted as kernel densities or histograms (see Graphics menu) or written out to a spreadsheet file for further analysis using Files / Listings / Save Monte Carlo Replications. For 1000+ replications, the option Save Frequency Table is enabled, together with a field to enter the desired number of bins for histogram construction. This option results in the actual replications being discarded. The distribution of the first 1000 replications is used to construct the data range and bin widths for construction of the frequency table, given the number of bins selected. As well as putting no constraint on memory for large experiments, this option also allows higher quality graphics, with control over the smoothing bandwidth for kernel density estimation.

22. Increasing the number of bins used to generate the frequency distributions gives more accurate estimates of quantiles and test sizes, at the cost of greater computing time. The default setting of 100 gives sizes estimated to the nearest percentage point.

23. A function coded by the user in Ox can be used to examine the result of a replication and optionally discard it, for example, because parameter estimates take illegal values or otherwise represent a false maximum of the estimation criterion. Activate this option with "Censor MC Replications" in Options / General / Special Settings.

See Appendix C for coding details.

24. Check the "Save EDFs" checkbox to save tabulations of the empirical distribution of test statistics. If the t statistics are centred on the true parameter values, or otherwise if the null hypothesis of interest is set to be true in the simulation, these distributions can be used to provide critical values, and hence to estimate true test powers by simulating cases of the alternative. These tables can also be used to generate p-values for tests in observed data. EDF tabulations are smoothed using the Gaussian kernel with the currently selected bandwidth. See 7.3 Monte Carlo Graphics, for information on selecting the bandwidth and setting it individually for different series.

25. EDFs are saved as spreadsheet or matrix files of the type specified by Options / Output and Retrieval / "Export Listings to Files of Type". The assigned file names are of the form "EDF_Run" + run i.d. + extension. These files can be renamed, merged and/or edited either in TSM or a spreadsheet program (see below). See Appendix G for details of the EDF file format.
Note: The prefix "EDF_" is recommended for file names, since such files can be loaded easily by drag-and-drop.

OUTPUTS

26. The results of the experiment, including tables and plots data, are automatically stored with the DGM at the end of the run. Reloading this model allows the tables (using the "Results" command)  and graphics to be re-displayed at any time. When performing a sequence of experiments, save the DGM under a new name before modifying or re-using it, so as to preserve these results for future study.

27. The command Graphics / Monte Carlo Distributions opens a dialog to select and display the empirical densities of each parameter, test statistic and p-value. These are optionally displayed in both histogram and kernel density form. The normal curve with matching mean and variance can be optionally superimposed for comparison; see Options / Graphics. See 7.3 Graphics / Monte Carlo Graphics for further information about creating plots, including the ability to combine density plots from different experiments, for comparison.

28. Either the complete set of replications, or the frequency table if this option is selected, can be exported to a spreadsheet file for further analysis. Give the command File / Listings / Export Spreadsheet / Monte Carlo Replications. The spreadsheet format for frequency table (histogram) files is as follows:
Row 1:      Name of item
Row 2:      Bandwidth index (integer 0:32), used for drawing kernel density
Row 3:      Initial bin ordinate
Row 4:      Increment (bin width)

Rows 5...      Histogram values.

29.   Check the "Run as External Process" checkbox to launch the run as a batch job in a console window. This option allows the TSM GUI to remain under the user's control for other tasks while a long Monte Carlo job is performed. A temporary executable Ox file is created to be run in batch mode in a console window. The file has the generic name "Batch_Run[ID].ox" where [ID] denotes the run ID number. The file is deleted at closedown if "Delete Temporary Files on Exit" is checked in Options / General.

30.   When the run is completed, the results are sent to a text file with the generic name "Batch_Run[ID].txt". This can be loaded in the TSM window using the command File / Load Text File.
Note:
*      The listings for the run are stored with the Data Generation model. Reload this model to see the plots.
*      If "# Parallel Runs" is set to 2 or greater the action is different, see 35ff below.

31.   If the checkbox "Defer External Jobs" is checked in Options / General, the Ox file containing the job is created, but not started automatically. Use this option if the job is to be run at a different time or on a different machine, for example.

PARALLEL PROCESSING

32.   A batch job consisting of a single experiment can be divided up into several runs in parallel, according to the setting of "# Parallel Runs". This feature allows a multi-core processor to be exploited to speed up the run time of the experiment, for example. The job can also be run on different machines in parallel, a task which Condor implements automatically but can also be set up manually.

33.   The executable files in this case receive the generic names "Batch_Run[ID]_Inst??.ox", where [ID] stands for the run ID, and ?? denotes the instance number. To run instances on different machines, check "Defer External Jobs" in Options / General, and copy and run the Ox files by hand in console windows.

34.   When the runs are completed, the outputs from each instance are written to temporary files with names matching the Ox files and the ".tsd" extension. Press the "Results" button, choose "Load Results from File" and select any file called "Batch_Run[ID]_Inst??.tsd" in the file dialog. The complete results for the runs are automatically imported and aggregated into a single set of tables for display. The option is presented to save the aggregated results in a new ".tsd" file and delete the individual instance files (".ox" and ".tsd"), in addition to output and log files if these have been created.
Note:
*      There is no need to wait for all the instances to complete to

perform the aggregation step. Select the highest-numbered file to be sure that all those available are loaded.

* However, don't choose the temporary file deletion option until all instances are returned! Aggregation must be done as a single operation.

* An alternative way to load the batch outputs is by the command Files / Tabulations / MC Batch Results..... or simply by dragging any of the files to the results window with the mouse.

* An alternative way to save the aggregated results is by File / Listings / Save Listings File...

35. By default, the instances return the actual replications. Sorting and frequency calculations are done at the aggregation stage. This exactly replicates the procedure for a single run, and is recommended for most purposes. High-quality plots are available with this option provided the total number of replications, in all instances, is at least 1000. If the option "Save Frequency Table" is set (only permitted with at least 1000 replications per instance) the frequency tables returned by each instance are simply averaged. This method is appropriate where each instance involves a large number of replications, and the minor distortion induced by averaging can be tolerated. (Note: bin widths for sorting are chosen dynamically and may differ slightly across instances).

36. To ensure the parallel experiments are statistically independent of each other, each is supplied with a different seed for the random number generator. If the current seed is S, the seed sent to each instance is calculated as floor(M*U[0,1]), where U[0,1] is a uniform random number generated successively for each instance, using S, and M = 2147483647 is the largest positive integer available in Ox. In the case where the seed is user-supplied (see Options / Simulation and Resampling), this scheme ensures reproducibility of the complete parallelized run. By default, S is generated from the system clock as the number of seconds since midnight.

CONDOR

37. If running on a network with the Condor HTC system implemented, the procedure for launching a Condor job is effectively the same to external processing on the local machine, as described above. Simply check the "Run Condor Job" checkbox, when this is displayed. Like external jobs to be run locally, Condor jobs can be either started from within TSM, or deferred to be launched manually. See Appendix H for information on working with TSM in the Condor environment.

38. Condor can optionally copy the Ox files needed to run the jobs from a different Ox installation. The obvious application here is where the user is running the 64-bit versions of Ox and TSM under Windows, while the Condor cluster consists of 32-bit workstations. Then, the files need to be taken from a 32-bit Ox installation. Go to Options / General, "Special Settings" and set the item 'Condor Alternative Executable' to TRUE to enable this option, and then enter the path to

© James Davidson 2024

the Ox installation as a text string in 'Condor Executable Path'. For the case given, the entry required would typically be "c:\program files\oxmetrics8\ox\". (Type without the quotes.)

39.     Condor will optionally return files after a run containing diagnostic information. These include log files generated by Condor itself, and also the standard output that would appear in the DOS console window if the job were running locally. To switch on the writing of these files, go to Options / General, "Special Settings", and set either or both of 'Condor Log Files' and 'Condor Output Files' to TRUE (equivalently, enter 1 in the text field). To switch off the outputs set the relevant fields to FALSE (equivalently, enter 0).

WARP SPEED BOOTSTRAP

40.     When this option is turned on, bootstrap tests are computed by doing just one bootstrap replication in each Monte Carlo replication. The p-values are then computed by locating the sample statistics in the distribution of bootstrap statistics over the experiment.

41.     To activate this option, go to Options / General / Special Settings, scroll to the bottom of the choice list and double-click the text field to change the toggle setting. Set the bootstrap options as usual for the model for estimation, but remember that the "Bootstrap Replications" setting is ignored. Also note that the actual replications must be recorded to construct the bootstrap distribution when the experiment terminates, hence the "Save Frequency Tables" option is also ignored.

42.     Under this option, the bootstrap test statistics are available for plotting and/or exporting to a spreadsheet along with the other experimental outputs. Look for the items with a "WpBoot" prefix. This makes possible a direct comparison of the bootstrap distributions with the original sample distributions generated in the Monte Carlo run.

## 3.13  Setup / Tail Probabilities and Critical Values...

This dialog gives access to tabulations and plots of standard distributions: the standard normal and Student t, either signed or absolute, and the chi-squared and F distributions.

*       Select the radio button for the distribution required.

*       Set degrees of freedom as appropriate (except for "EDF from File", see Note 4 below).

*       Select the tail probability required. The upper tail is the default.

*       Depending on the dialog opened, enter either the critical value to get the tail probability, or the tail probability to get the critical value.

© James Davidson 2024

\* To show the probability/critical value (also printed in the results window) press the right-hand button. To plot the density with the specified tail area shaded, press "Plot PDF".

Notes:
1. The upper tail probability is the area under the density function to the right of the critical value. Use the critical value to perform a 1-tailed test with the corresponding significance level. The lower tail probability is the complementary area, equal to 1 - upper tail probability,

2. The distributions "|Normal|" and "|Student t|" allow calculation of p-values for 2-tailed tests. These correspond to the p-values reported with the estimation output.

3. Use the signed distributions "Normal" and "Student t" to calculate p-values for 1-tailed tests. Just cut and paste the test statistic into the "Look Up Tail Probability" dialog. Change its sign for tests of the lower tail.

4. Select "EDF from File", then press "Select", to choose a test statistic from the currently open EDF file - the file dialog opens automatically if no EDF file is currently open. To open a different EDF file, use the command File / Tabulations / Load EDF.

5. The density plot is the same in either dialog. The shaded region is the area corresponding to the tail probability, and equivalently, the area to the right of the critical value.

6. Tests of the unit root hypothesis (ADF, PP and ERS tests) have critical regions in the lower tail by convention. To maintain the usual rule of rejection in the upper tail, as assumed in these dialogs, EDF tabulations of these statistics are reversed. The changes of sign are indicated by a prefix "-" attached to statistic names.

## *3.14  Calculator...*

This utility substitutes for a hand calculator  The menu item opens a text box. Enter a numerical expression and press <<< Go >>>,  or the Return key. The expression and its value are printed in the Results window, as well as appearing below the text field. For example, to calculate the area of a circle with radius 5, type the formula Pi*5^2 and the output has the form

Calculator: Pi*5^2 = 78.539816339745

See 1.5 User Interface / Entering Formulae for details of the syntax and available functions. In this mode, only operators, functions, numeric constants, and "Pi" should appear in formulae.

The utility also draws plots of functions of one or two variables. This feature is implemented by including either one or two variable names in a formula. X and Y are obvious choices, but any valid names are allowed (see 1.5). The symbol(s) are replaced by the specified

range of numerical values when the function is evaluated.

* To plot two or more functions of one variable on the same graph, enter the formulae in the text box separated by ";" (semi-colon).

* Functions of two variables can be plotted as contours and/or as a three-dimensional surface in perspective. Make this choice in the Options / Graphics dialog.

Notes:
1. Optionally, this feature can be accessed by a dedicated toolbar button. Go to Options / General / Interface Settings and check/uncheck the item "Show Calculator Button". This button is hidden by default.

2. Text can be copied from either the results window or the Values fields and pasted into the text field. Computing functions of program outputs (estimates, standard errors, etc.) is therefore quick and straightforward. Confidence bounds are an obvious application.

3. Use only parentheses, "(" and ")", to enclose expressions and arguments. The symbols []{} are not allowed, and result in a parsing error.

4. To implement the plotting feature, the range of plot points for the variables must be specified. In this case pressing <<<Go>>> (or hitting Return) re-opens the text box. Enter the minimum value for the first variable, press <<<Go>>> again and enter the maximum value. In the case of two variables, repeat this cycle for the second one. The final press of <<<Go>>> yields the plot.

5. The plot bounds can be entered either as fixed values or as formulae - for example "-2*Pi" and "+2*Pi" are valid entries. However, including names in formulae at this point produces a syntax error.

6. The number of plot points is displayed on the button bar of the text box at this stage, and can be changed. Setting this value to N results in N-1 line segments connecting N function values. The default is 50.

7. The function formula and its numerical integral over the specified range, calculated by the composite trapezoidal rule, are printed in the results window. Increase N for greater accuracy, which is also dependent on the smoothness of the function.

8. To evaluate a power series, enter the formula for a term as a function of x, and set the plot bounds at 0 and N-1. In this case, the plot nodes are the integers. With this specification, the simple sum of

the terms is reported.

9.    The plot points can optionally be listed in the results window.
      Go to Options / Output and Retrieval / Print in Results Window,
      and check "Calculator Plot Data".


## 3.15  Matrix Calculator...

This dialog allows the direct calculation of matrix expressions, based
on matrices constructed as:
*      moment matrices of mean squares and mean products from
       variables in the data set.
*      imported outputs from the latest estimation run, or the currently
       loaded model.
*      typed or pasted values in a table with specified row and
       column dimensions.

Choose the "Select Matrix" radio button to see the list of stored
matrices.

"Save File" writes stored matrices to a disk file in spreadsheet
format.  "Load File" retrieves saved matrices, optionally merged
with the current store.

Select from the list and press "Name" to rename and/or edit the
description of a stored matrix

Select from the list and press "Delete" to delete a stored matrix.

Notes
1.    Construct a sample moment matrix as follows:
      1)     Choose the radio buttons "Row Variables" and "Column
             Variables" in turn, and make selections from the variable
             list.
      2)     Set the observations to be used with "Sample...".
      3)     Press "Evaluate".  A text field opens to type a name for the
             new matrix. An optional description can also be entered.

2.    For a square, symmetric moment matrix, simply choose the
      row variables and press "Evaluate". If no column variables
      are selected these are set to match the row variables
      automatically.

3.    To see the list of matrix descriptions, go to Help / Data
      Descriptions while the Matrix Calculator dialog is open.

4.    Matrices are saved in vectorized form as spreadsheet columns,
      with the first two elements containing the row and column
      dimensions. These files can be edited in the usual way in a
      spreadsheet package.

© James Davidson 2024

5.  Saved file names are given the prefix MAT_ by default. Files whose names have this prefix can be dragged and dropped into TSM, and will be opened as matrix files.

Choose the required item from the pull-down menu, and press <<< Import >>> to add the matrix or vector to the store. These objects were either created in the latest estimation run, or retrieved from an earlier run by loading the corresponding stored model. The automatically assigned name has the run ID number appended to it.

Notes:
6.  The OPG matrix is calculated only if the robust or HAC covariance matrix is specified in Options / Tests and Diagnostics. Otherwise, nothing is returned. The reported OPG matrix is normalized (divided by sample size) and hence is $O\_p(1)$.

7.  When estimation is by OLS or IV in the Linear Regression dialog, the returned Hessian matrix is positive definite. In the Dynamic Equation dialog, on the other hand, the estimation criterion is a maximand by default and the Hessian is negative definite. The sign can optionally be switched to report a minimand in the Options / Optimization and Run dialog. The Hessian matrix is a sum of terms and so is $O\_p(T)$.

8.  The criterion gradient is reported at the current parameter point. and is zero when evaluated at the unrestricted optimum! Interesting cases arise on evaluating without optimization, after imposing restrictions on the parameters. Do as follows:
    o   Estimate the model with restriction(s) imposed - see the Parameter Constraints dialog.
    o   Remove the restriction(s).
    o   Give the menu command "Evaluate at Current Values". The vector so obtained is the basic ingredient of the score (LM) test of the restriction(s).

EDITING MATRICES

To view a matrix in the editing box, double-click its name in the list, or select from the list and press "View/Edit Matrix". If no matrices are stored, the option "Create New Matrix" is presented. In the editing box the following operations are available:

*   Double-click a cell to edit its contents. Save the change by either pressing Enter, or clicking another cell. Discard the change by pressing Escape.

*   To paste cell entries via the clipboard, first type them in the Results window or a text file, separated by spaces, tabs or carriage returns. Highlight the list, right-click and choose Copy. Then highlight a cell

of the matrix and press "Paste". The clipboard contents are pasted
into cells from the highlighted cell onward, from left to right and row
by row, until either the list is exhausted or all cells are filled.

*       Press "Clear" to fill all cells with zeros.

*       Press "Print" to print the matrix contents as text in the Results window.

*       Press "Code" the print the matrix with Ox formatting, ready to be cut
        and pasted into an Ox program.

*       To insert a row or column of zeros at position 1, press "Ins." and, in the
        dialog that opens, choose "Row" or "Column" as required.
        To insert a row or column at other positions, first select a cell. The new
        row/column will be  inserted following the highlighted position.

*       To delete a row or column, highlight a cell in the row/column by clicking  it.
        Then press "Del." and in the dialog, choose "Row" or "Column" as required.

*       If the matrix is changed, pressing Close opens a choice box to keep
        or discard the changes.

*       Pressing "View/Edit" when no stored matrix is highlighted opens the
        editing box with a single row and column, containing 0. To create a
        new matrix from scratch, first insert the required number of rows
        and columns, then either type or paste entries as required. Press
        "Close" to store the new matrix with a name and optional description.

MATRIX CALCULATOR
Pressing this button opens a text field where formulae can be entered. This works
just like the "Calculator..." option, except that matrices are entered by name,
as well as (optionally) numerical scalar values. The formulae must conform to
the rules of matrix algebra.

*       The general syntax is of the form "name = formula". The result is
        saved in the matrix list under "name".
        If "name" is omitted, the text box opens to allow a name to be
        entered. If the name matches an existing name, the options to change
        it or over-write the existing matrix are offered.

*       Permitted matrix operators are
         + (sum) , - (difference), * (product), ' (transpose),
        ** (Kronecker product), @ (Hadamard product).

*       Matrix functions are
        inv()   .       .       inverse
        gnv()   .       .       generalized inverse
        cho()   .       .       Cholesky decomposition
        eig()   .       .       eigenvalues (returned as diagonal matrix)
        evc()   .       .       eigenvectors (normalized to unit length)

| | | | |
|---|---|---|---|
| eim() | . | . | moduli of eigenvalues (returned as column) |
| vec() | . | . | vectorization |
| vcr() | . | . | row vectorization |
| dia() | . | . | matrix diagonal (returned as column) |
| det() | . | . | determinant |
| tra() | . | . | trace |
| nrm() | . | . | Frobenius norm |
| len() | . | . | Euclidean lengths of columns (returned as row) |

Notes:

9. All the functions specified in the "Calculator..." option can also be used, and are applied elementwise. The operators / and ^ can be used with scalar divisor or exponent, and are applied elementwise.

10. For matrices A and B the expressions A'*B and A'B are equivalent.

11. To return the eigenvalues of a matrix x as a column, enter the formula dia(eig(x))

12. Complex-valued matrices are not handled. The algorithm generates an error message if eigenvalues/eigenvectors have imaginary parts. In this case, eim() returns the moduli.

13. If the argument of inv() is not square, the elementwise inversions (reciprocals of matrix elements) are returned.

14. The inversion and Cholesky decomposition functions result in errors if the arguments are not nonsingular or positive definite, respectively.

15. If A is a matrix and b a vector with matching row (column) dimension, the products A@b and b@A are defined and identical, having the dimensions of A. The rows (columns) are the scalar products of the rows (columns) of A with the corresponding elements of b.

16. Scalar outputs from det(), tra(), nrm() and len() are printed in the results window. The values are optionally saved as 1x1 matrices, but when the pop-up dialog requesting a name for the matrix appears, press Cancel to discard it.

# 4. Model

## *4.0 Model / General*

This page provides general information on setting up models using the set of dialogs grouped under the "Model" menu. The "Linear Regression" dialog controls models that are estimated by evaluation of a formula (analytic solution of normal equations) without the need for numerical optimization. This is also true of most of the panel data options presented in "Panel Data". By contrast, all the methods set up in the "Dynamic Equation" dialog are estimated numerically, although there are some cases (linear models) in which the resulting estimator is the same.

MODEL MANAGER

Model specifications, including fitted or imposed parameter values and other model outputs in graphical or listing form, can be stored under a supplied name, and subsequently retrieved from a list, using the following commands in the Model menu. More detailed control of model storage and retrieval is available through the Model Manager dialog.

Load a Model...        Opens a pop-up menu listing models in order of last retrieval. The default maximum number of list entries is 20 If the model required is not on the list, access it though Model Manager.

Save Current Model...Opens a text field where the name for the new model can be entered. The current model name appears in the field for editing or confirmation. If the name entered already exists in the model list, a confirmation box opens with options "Overwrite" and "Cancel".

Quick Model Save...  Saves the current model under the current model name, without prompting. Use this command to skip the naming steps under "Save Current Model...".

BUTTONS

1.    The "Values" button either opens the corresponding Values dialog, or refreshes it with new settings if it is already open. If any parameters are "fixed" in this dialog, the button is highlighted as a reminder.

2.    The "Clear" button sets all specifications in the dialog to the defaults.

3.    The "Go" button, where provided, is an alternative way to launch an estimation, equivalent to selecting the "Run" button on the toolbar, or the Actions / Run Estimation menu item.

4.    Where provided, the "Sample" button opens the Setup / Set Sample dialog.

5.    Where provided, the "Options" button opens the last-opened estimation options dialog from the Options menu. By default, this is the Tests and

© James Davidson 2024

Diagnostics dialog.

Notes:

1)     Pressing the Options button repeatedly cycles through all the accessible options dialogs.

2)     Only relevant options are accessible from this button, depending on whether it is located in the Linear Regression or Dynamic Equation dialog. All options dialogs are accessible in the usual way, through the menu or the >>> and <<< bars.

ACTIVATING SPECIAL FEATURES

The special model features specified in these dialogs (conditional variance, supplied function, regime switching, equilibrium relations, parameter constraints) can be turned "on" and "off" in two ways:

*     Checking/unchecking the appropriate checkbox in the main "Dynamic Equation..." dialog.

*     Checking/unchecking the "Active" checkbox in the relevant dialog.

These pairs of switches are connected, so always show the same status

Notes:

6.     The activation status of a model feature is shown by check marks against the menu item, as well as by the check boxes.

7.     If it is not possible to activate a feature (e.g. no options have been selected) the "Active" checkboxes are disabled.

8.     The current model feature settings are retained even when the feature is de-activated.

DIALOG CONTROL

All dialogs are closed when an estimation is run. Restore them to their original positions by the Actions / Restore Dialogs command, or by clicking the "Windows" button on the tool bar. Alternatively, select the option to have them restored automatically following a run; see Options / General.

SPECIFYING SYSTEMS OF EQUATIONS

9.     A system of equations is specified by checking this option in the main Model dialogs, and then selecting a set of two or more dependent variables. All the equations are given the same specifications (AR/MA components, regressors, etc.). The specification of individual equations is changed by selecting restrictions on individual parameters in the Values dialogs.

10.     To modify the specification of an equation, first open the corresponding Values dialog. Use the left and right arrow buttons at the case of the dialog box to switch between equations. To exclude a regressor (or lag) from an equation, select the "Fixed" checkbox for the parameter(s) in question, and then make sure that 0 is entered in the Values field. If desired, any other fixed value can be set.

11.     Variables selected as endogenous can also appear on the right-hand sides of the equations as Type 1 regressors, to allow a simultaneous system. Three Stage Least Squares can be selected as an option in the Model

                            © James Davidson 2024

/ Linear Regression dialog. In the Dynamic Equation dialog, the presence of such variables automatically causes the FIML estimator to be computed, provided suitable identifying restrictions have been imposed on the equations. See Section 4.4 for further details.

12.      Equilibrium relations can be embedded in systems of equations to create an error correction model or cointegrating VAR. The Model / Equilibrium Relations dialog allows the selection only of a set of variables to be included in all the relations. The Values dialog is used to impose restrictions, which must include a normalizing restriction (fixed at -1, typically) on one variable.

AUTOMATIC DATA TRANSFORMATION

13.      Variables in a model may be subject to transformations. Logarithms are the most popular case, but absolute values, squares, logistics and others may also have applications. While one way to do this is to create and store the new series in the database, a convenient alternative is to perform the specified transformation "on the fly", after setting up the equation(s) in terms of the original variables.

14.      To implement this option, go to Options / General / Special Settings and select "Automatic Data Transformations" from the pull-down menu. Double-click on the text field to toggle the option on and off.

15.      To indicate to the program the transformation required for each variable, edit the data description field (see Setup / Data Transformation and Editing / Edit / Data Descriptions).
   *      The data description field should contain the flag #FN (case sensitive) followed by the formula to be implemented, which is written using the syntax detailed in Section 1.5, "Entering Formulae".  (This is "case E".)
   *      The reserved variable names s# and S# may be used to represent the variable being transformed, as an alternative to its actual name which is also valid. (Note that the transformation may involve other variables in the data set as well as constants.)
   *      For example, the text
         #FN log(s# + 100)
      transforms the data by taking the logarithm after adding 100. Note how spaces can be freely used to improve readability.
   *      This text should follow any other material appearing in the data description.

16.      If the transformation is 1-1 with a unique inverse, fitted values and forecasts may be optionally reported for the original series, not the transformed ones. To implement this option, a formula for the inverse transformation must be provided.
   *      Following   formula, enter the text #INV followed by the formula for the inverse function.
   *      In this case, use either of the reserved names s# and S# to represent the transformed variable.
   *      For example, the text

     © James Davidson 2024

$$\text{\#FN log(s\# + 100) \#INV exp(s\#) - 10}$$
specifies the transformation and its inverse.
Only include this component if forecasts of the original variables
are required.

17.    To preserve backwards compatibility, the text #LOG (case
sensitive), appearing anywhere in the data description field,
specifies the simple logarithmic transformation and its inverse.
Thus, the strings "#LOG" and "#FN log(s#) #INV exp(s#)"
are equivalent.

CONDITIONAL VARIANCE SIMULATION

18.    If the string #CVAR appears in a variable description, followed
by the name or column number of another variable in the data set,
this is ignored unless the variable is simulated in a Monte Carlo
experiment and simulating the model in question produces a
conditional variance series, for example, a GARCH or Markov-
switching model. In this case, the generated conditional variance
series is written to the column indicated. This is then available
for analysis at the estimation stage of the experiment.


## 4.1 Model / Model Manager

A "model" is the complete set of specifications, values and results from estimation
or simulation of an equation or system of equations. At any time, the program holds
in memory the specifications that the user has selected, and either starting values
entered, or estimated values following an estimation run. This dialog allows
model specifications and results to be stored under identifying titles, and
subsequently recalled. Material stored includes parameter estimates with information
on restrictions or bounds, estimation and testing options, generated series, forecasts,
recursive estimates and criterion plots.  The latter series can be used to produce
additional plots and test statistics without re-running the estimation. Stored models
are also used to set up Monte Carlo experiments.

| | |
|---|---|
| "Load Model" | Loads the model whose title is highlighted in the list. |
| "Store Current Model" | Opens a text field for the entry of a model title. To complete the entry and store the model, either press Return or click the "OK" button, otherwise click "Cancel" . |
| "Model Description" | Opens a text box where details of the highlighted model can be entered by the user for later reference. Click the button again to close the box. |
| "Move Up"/"Move Down" | Use these buttons to arrange the order of the models in the list. |
| "Sort" ("Unsort") | Sorts the model list. There is the choice of sorting by name, and by the date of last modification (most |

recent first). The "Move Up" and "Move Down" buttons are disabled. If the list is currently sorted, the button restores the original ordering.
(NOTE: Sorting by date is only approximate for models created with TSM versions prior to 4.43.)

"Delete"                        Opens a confirmation box, and removes the model whose title is currently highlighted in the list if the action is confirmed.

"Clear All"                     Opens a confirmation box, and removes all models in the list if the action is confirmed.

STORING MODELS

1.    By default, the model title "Run ###" is displayed in the text box, where ### denotes the current run ID number. The entry can then be edited as desired. If an existing title is first highlighted, this name appears in the text box. If "OK" is pressed without changing it, the current specifications overwrite the ones stored under that title. This option allows a stored model to be modified or updated. Rename a model by loading it, saving it under the new title, and then deleting the original copy.

2.    The Store command  places the model specifications in a memory buffer, but does not save them to disk. Give the command File / Settings / Save or Save As... to save them permanently in a .tsm file. When a settings file is reloaded, the stored models it contains will be accessible as before.

3.    A stored model contains the complete current settings of all model dialogs, including Model / Linear Regression, Model / Dynamic Equation, Setup / Nonparametric Regression, Setup / Semiparametric Long Memory and Setup / Cointegration Analysis. For the purposes of  Monte Carlo experiments, the model is recorded as a linear regression, a dynamic equation or a semi-parametric long memory model, depending on which dialog is open when the Store command is given. Note that estimation is always done numerically when launched from the Dynamic Equation dialog. To have the desired estimator recorded as part of the stored model, have the appropriate dialog open when the Store command is given.

4.    There is an option to store the estimated model automatically following a successful estimation run - see Options / General. Use this option with care, since the settings file could potentially become very large.

5.    In addition to parameter values, results and listings associated with a model are stored in a file with the name of the model and the suffix ".tsd". These files are not intended to be read or manipulated outside TSM. To save results to spreadsheet files, use the commands under File / Listings. However, note that the latter data cannot be read back into the program for purposes such as graphing, or calculating tests.

6.    If  the Store Current Model command is given following a Monte Carlo experiment, the replications or frequency tables (whichever have been

© James Davidson 2024

specified) are stored in the ".tsd" file under the given model title.

7.  If the "Store Data with Model" box is checked, a copy of the data set is saved in the ".tsd" file. This will be loaded automatically when the model is re-loaded. If this box is not checked, you are prompted to save the current data file if it has been edited, and the file with this name is reloaded with the model - however, there is no protection against it being altered in the interim, in this case.

8.  If "Store Results with Model" is checked, the text of the last-run estimation output is stored as a string. When the model is re-loaded, the next tile Actions / Evaluate at Current Values is chosen (or "Calculator" button pressed) causes a choice box appears, presenting the option of re-displaying the results.

9.  If the "Save as Generic Model" checkbox is checked, the data file is not reloaded with the model, which is the default action. This setting is appropriate for models intended to be adapted to different data sets. For example, such a model may specify inputs for a user-coded estimate or test.
    Note:   If this option is selected, the "Store Data with Model" and "Store Results with Model" options are disabled.

10. Stored models include the current settings for the Semiparametric Long Memory and Cointegration Analysis dialogs. Different settings for these modules can therefore be stored and recalled in the same way as for the main equation specifications.

11. As models are created and modified, settings and estimates are not discarded when specifications change -- if a regressor, or equation, is deleted, its position in the model is recorded and it can then be reinstated at a later stage, with the associated parameter values. It may be desirable to delete these hidden components when the model is finalized, to conserve memory and file space and ensure speedy and robust processing. To finalize a model (permanently remove unused components) Check the "Finalize Model" checkbox before giving the command "Store Current Model". Note that the check is cancelled automatically.

LOADING MODELS

12. Loading a model overwrites the current settings. Store these first, if they will be needed subsequently.

13. See 8 above: press the "Evaluate" button to re-display the estimation output, if this has been saved with the model. This saves time if re-evaluating the results would be time-consuming, as when output of bootstrap inference is selected.

14. The data file specified in the model will be reloaded automatically, provided it has not been deleted or moved. It can have been edited, or had series added to it, but any variables that appear in the model must of course be present in the file, otherwise an error occurs. It is the user's responsibility to ensure this,

© James Davidson 2024

and the safest way is to check the "Store Data with Model" box so that a dedicated copy of the data is saved.

15. When program settings are cleared and restored to defaults, a model called "Defaults" is created automatically. It is recommended to set up your preferred options for routine operations, and use Defaults to store these settings. They can then be easily restored, by loading this model.

MODEL DESCRIPTIONS

16. When the descriptions box is open, other dialog functions are disabled. However, the description displayed can be changed by selecting a different model from the list. Use the PgUp and PgDn keys to scan rapidly through the list.

17. If the "Automatic Descriptions" checkbox is checked, non-default model settings are written to the descriptions box automatically when the model is stored. These are coded in the scripting language detailed in the programming manual. This provides a convenient way to check that the model specification is as intended.

    Note: If either the Semiparametric Long Memory or the Cointegration Analysis dialogs are open when the Store command is given, the settings written relate to those modules; otherwise, the usual model settings are written.

18. You can type your own comments in the model description field. However, these will be saved with the model only if "Automatic Descriptions" is unchecked. To save the model settings and personal comments together, save the model first with "Automatic Descriptions" checked, then add your comments and save again with the option unchecked

19. If the "Print Descriptions" checkbox is checked, the model description, if any, is written to the results windows when the model is loaded.

DELETING MODELS

20. When a model is deleted, the associated ".tsd" file is deleted from the disk. Also note that these files are overwritten when a new model is stored with the same name.

21. There is an option to remove all the ".tsd" files automatically at shutdown. Their contents, except for Monte Carlo results, can be regenerated by running the command Evaluate at Current Values with appropriate options set.

## 4.2 Model / Linear Regression...

This dialog is for specifying a linear equation, for estimation by Ordinary Least Squares (OLS) or Instrumental Variables (IV/2SLS). When a system of equations is specified, it also permits estimation by the "seemingly unrelated regressions" (SUR) and three stage least squares (3SLS) system estimators.

* Select the type of estimator required under Select Estimator. The options

offered depend on whether the System of Equations checkbox is checked. For IV estimation, select the instruments in the Model / Select Instruments dialog.

* Variables for inclusion in the equation are chosen by highlighting them in the list. First, use the radio buttons to choose which type variable to specify, either the dependent variable(s), or regressors of Types 1 or 2. Any number of regressors can be selected, of either type, but not as both types at once. Only one dependent variable can be selected, unless the "System of Equations" checkbox is checked.

* Set the 'Lags' scrollbar to automatically include lags, up to the set order, of all variables of the Type whose radio button is currently selected.

* Use the check-boxes to include an intercept and/or a trend dummy.

* To run a Wald test of parameter constraints, first enable this option using the check-box. Open the Model / Parameter Constraints dialog by clicking the square [...] button. Parameters for zero and linear restrictions must be specified in the Values / Equation dialog.

* To run a cointegrating regression, first enable the "Cointegrating Regression" options by checking the checkbox.
  Then:
  o check either or both of the Cointegration Test checkboxes to compute the residual Phillips-Perron and/or augmented Dickey-Fuller statistics.

  o Check "Fully Modified Least Squares" to compute the Phillips-Hansen (1990) FMLS efficient estimator.

  o Check "SSW Efficient Least Squares" to compute the least squares estimator augmented by leads/lags of regressor differences as proposed by Saikkonen (1991) and Stock and Watson (1993) (SSW-LS).

* If a panel data set is loaded, the dialog is automatically formatted in panel mode, with the "Panel Data..." button enabled. This button opens the Model / Panel Data dialog, which can also be accessed from the menus.

* For Instrumental Variables estimation (IV/2SLS/3SLS), use the "Instruments..." button to open the selection dialog.

* The "Diagnostic Tests..." button opens the diagnostic test selection dialog, or closes it if it is open. This dialog can also be accessed from Options / Tests and Diagnostics. The show/hide status is remembered when this dialog is closed and re-opened.

* The "Sample ..." button opens the Set Sample dialog, or closed it if it is open. Use this to select the observations to be used for estimation. The setting is the same as if opened from Setup / Set Sample. Pressing the "Simulate" button allows the "All Available Observations" button to show the available sample of exogenous variables for a simulation of the dependent variable. The

show/hide status is remembered  when this dialog is closed and re-opened.

* The "Options..." button opens the most recently opened estimation options dialog from the Options menu. By default, this is the Tests and Diagnostics dialog. Click the horizontal <<<< and >>>> bars in these dialogs to switch between them.

* The "Values..." button opens the equation Values dialog, or refreshes it if it is already open.

* To compute the estimates, press the <<< Go >>> button. This duplicates the 'Evaluate' button on the toolbar, and the Actions / Run Estimation menu item.

* The "Clear" button clears all settings in this dialog.

Notes:

1. For models specified in this dialog the estimates are computed by the usual analytic formulae. Equations specified in the Model / Dynamic Equation dialog are estimated by numerical optimization, whether or not their form is linear.

2. When this dialog is opened, the Dynamic Equation, Conditional Variance, Supplied Function and Regime Switching dialogs are all closed automatically. These features are all disabled ("greyed out"), until it is closed again. All estimators except Least Squares and IV are also unavailable.

3. See Model / Select Instruments for details of the instrument selection procedure.

VARIABLE AND LAG SPECIFICATION

4. When one or more variables are selected in any category, the corresponding radio button is highlighted (lighter grey). To see the actual variable(s) (highlighted in the variable list) and the lag order selected, click on the radio button in question.

5. Regressors of Type 1 can have a different lag order specified from those of Type 2. It is recommended to (e.g.) make variables for lagging of Type 2, and others, such as dummies, of Type 1. If lags are not specified, then regressors of Types 1 and 2 are equivalent.

6. In single equation mode, if the dependent variable is selected as a regressor of Type 2, and Lags > 0, only its lagged value(s) are included in the set. This provides an alternative way to estimate an autoregression, without numerical optimization. In all other cases (as a regressor of Type 1, or with Lags = 0) the dependent variable is ignored.

7. In Equations System mode, selecting the same set variables as the "Dependent Variables" and as "Regressors of Type 2", with k > 0 lags selected, results in the unrestricted VAR(k) system being estimated.

8.      The trend term, if selected with the checkbox, is a regressor of Type 1. With this option selected, lags of Type 1 regressors are disabled to avoid collinearity.

WALD TEST OF CONSTRAINTS

9.      Estimation subject to constraints is not available for linear regressions. LM tests of parameter constraints and fixing restrictions are likewise disabled, although parameters can be fixed in the usual way. However, note that Wald and LM tests are equivalent in this case.

10.     To perform an F test of "significance of the regression", check the Wald Test checkbox, and also the "Test Joint Significance of Regressors" box in the Model / Constraints dialog. Note that lagged dependent variables, current endogenous variables in simultaneous systems and the trend term are not included in this set by default, but can be added to it by checking the Wald Test checkboxes in the values dialog..

COINTEGRATING REGRESSION

11.     The cointegration options are all disabled if the Systems checkbox is checked, or the Estimator selection is IV/2SLS. The Phillips-Perron and ADF statistics are only available with OLS estimation. These options are disabled if FMLS or SSW-LS are selected.

12.     The Phillips-Perron test and FMLS estimator both make use of a Heteroscedasticity and Autocorrelation Consistent (HAC) covariance estimator. Set the desired kernel and bandwidth in Options / Test and Diagnostics Options.

13.     The SSW-LS estimator includes leads and lags of the differences of all variables included as Type 1 regressors. The coefficients of these ancillary regressors are not reported in the output. Lagging Type 1 regressors is disabled. Type 2 regressors can be included, with lags chosen by the user (not chosen automatically). For valid application of the method, the dependent variable and Type 1 regressors must all be I(1), while the Type 2 variables must all be I(0) (stationary).

14.     The ADF test and the SSW-LS estimator depend, respectively, on a choice of lag length, and of lead/lag length. The lengths L can be selected in two ways:
        *       Automatically, by optimizing a model selection criterion over the interval $0 <= L <= [n^{(1/3)}]$ . The selection criterion is is set in Options / Test and Diagnostics. It is the Schwarz IC by default.
        *       Manually, using the lag selection scrollbar. To enable this option, set the model selection criterion to "None" in Options / Test and Diagnostics.

15.     FMLS and SSW-LS cannot both be selected at once. Uncheck to latter to select the former.

## 4.3 Model / Panel Data...

This dialog is not accessible unless a panel data set is loaded. It sets the available options for panel data estimation, in conjunction with settings in the "Model / Linear Equation...", "Model / Select Instruments..." and "Setup / Set Sample..." dialogs.

Notes.
1.  Choose the first and last date for estimation or simulation of panel models in the Setup / Select Sample dialog. In the case of an unbalanced panel, the first date used for a particular individual is the later of this setting and the first available date, while the last date used is the earlier of this setting and the last available date.

2.  The options in this dialog cannot be combined arbitrarily. When an option is disabled or "greyed out", this means that it is not compatible with the other current selections.

DATA TRANSFORMATIONS

See for example Arellano (2003) for details of the transformations available here. Also note:

3.  If "Deviations from Time Means" is selected, the intercept is suppressed automatically, if selected.

4.  The "Time Means" transformation creates a data set in which each of the individual data points for each time period is replaced with the time average of that individual's observations. Running a regression with these data is identical to running a regression with the N individual means in the case of a balanced panel. In an unbalanced panel, however, the means effectively receive a weight Ti (number of observed periods). Be careful to note that the number of degrees of freedom in this regression is N-p, not O-p where O is the total number of observations (i.e. O = NT in a balanced panel).

5.  In "Time First Differences", the first-dated observation of each individual is set to 0.

DUMMY VARIABLES

When these boxes are checked, the corresponding dummies are created and added to the regression. They appear in the list as regressors of Type 3.

6.  If dummies are selected in conjunction with an intercept, the first individual/date/group is treated as the reference category, and the corresponding dummy omitted.

7.  Group dummies are set according to the grouping indicated by the !GROUP! guide column in the data file. If this option is selected, the "Individual Dummies" option is automatically cancelled.

8.　To make use of these options, first choose "OLS/GLS" in the Model \ Linear Equation dialog. If the "IV/2SLS" option is selected, the radio buttons in this group are disabled.

9.　GLS and ML estimators for random effects are only available for single equations, in this release. These options are disabled if the "System of Equations" checkbox is checked in Model / Linear Equation.

10.　The GLS and ML estimators employ the same transformation of the data, but the parameter Tau is computed in the first case by a the ratio of "within" and "between" regression variances, and in the second case by an iterative procedure, using a univariate line search to optimize the concentrated Gaussian likelihood.

11.　The text fields "SigV" and "SigEta" are provided to set up stochastic simulations with random effects. SigV represents the within-individuals variance, and SigEta the between-individuals variance. The contents of these fields, which can be edited by the user, are utilized to create the disturbances when the "Gaussian" option is selected under "Shock Distribution" in Options / Simulation and Resampling.
Note: These values not used to initialize estimation. However, the results from the latest GLS or ML estimation are automatically written to these fields, as well as to the printed output, allowing an estimated model to be easily simulated.

12.　The ML estimator uses a line search algorithm. The convergence criterion and the maximum iterations are set in Options / General / Special Settings. Note that these settings are shared by other program features, but should not generally need changing.

TESTS AND DIAGNOSTICS

13.　Standard linear and nonlinear Wald tests of restrictions on the parameters are available for panel regressions. Choose the covariance matrix formula, either standard or robust, in the Options / Tests and Diagnostics dialog. The other diagnostic tests controlled from this dialog are not available.

14.　Except in the "individual means" case, the modified Durbin-Watson statistic due to Bhargava et al. (1982) is reported as "BFN-DW". This provides a test for serial correlation in the within-individual disturbances. Partial tabulations can be found in the cited paper for the case of OLS estimation, assuming that fixed effects have been fitted by dummies or taking individual mean-deviations. For other cases, an EDF table might be created for statistic by Monte Carlo. Note: in the Monte Carlo module the statistic tabulated is 2 - DW, such that the upper tail quantiles are relevant to the positive autocorrelation alternative.

15.　The Breusch-Pagan LM test for the null hypothesis of no random effects is reported in OLS estimation without transformations.

This tests the hypothesis that random effects are absent. The null distribution is Chi-squared(1) with large N.

16.    The 'variable addition' implementation of Hausman's specification test is reported for random effects models. This tests the null hypothesis that the regressors in individual mean-deviation form are insignificant when added to the GLS-transformed equation. The distribution is Chi-squared with degrees of freedom equal to the number of regressors, with large N.

## *4.4 Model / Dynamic Equation...*

This dialog sets up the specifications for a dynamic model. As well as setting the basic ARIMA and ARFIMA specifications, as specified in equation (4.1) of Models and Methods, it also specifies which additional model features will be activated.

*    Use the radio buttons to select an estimator. The estimation criteria (see Models and Method for details) are:
  - o    Least Squares / Least Generalized Variance / FIML
  - o    Instrumental Variables / GMM
  - o    Conditional Maximum Likelihood with Gaussian (normal) errors
  - o    Conditional Maximum Likelihood with Student's t errors
  - o    Conditional Maximum Likelihood with skewed Student's t errors
  - o    Conditional Maximum Likelihood with General Error Distribution
  - o    Whittle (frequency domain) ML with Gaussian errors
  - o    Probit and ordered probit
  - o    Logit and ordered logit
  - o    Poisson (count data)
  - o    Negative Binomial I (count data)
  - o    Negative Binomial II (count data)

*    Use the scroll bars to choose the order of ARMA model.

*    Use the checkboxes to select the ARFIMA (long memory) specification, impose a unit root, select a nonlinear moving average (see 16 below) and include intercepts of Type 1 or 2, and a trend dummy. Only one type of intercept can be selected at a time.

*    Use the radio buttons to switch between selection of the four categories of variable: the dependent variable, and regressors of Types 1, 2 and 3. Then highlight the variables required in the list.

*    Set the 'Lags' scrollbar to a positive value to include lags, up to this order, of all the variables of the Type whose radio button is currently selected.

*    Use the lower scrollbar to select the number of lags of the error to include in the bilinear specification. If set to zero, this corresponds to the usual AR(F)IMA(p,d,q) model.

*    Use the other check-boxes to select or deselect additional model features

(conditional variance model, coded nonlinear function, regime switching, equilibrium relations and parameter constraints). Press the [...] button to the left of the checkbox to open the corresponding dialog, if currently closed.

* For Instrumental Variables estimation (IV/GMM), use the "Choose Instruments..." button to open the selection dialog.

* The "Diagnostic Tests..." button opens the diagnostic test selection dialog, or closes it if it is open. This dialog can also be accessed from Options / Tests and Diagnostics. The show/hide status is remembered when this dialog is closed and re-opened.

* The "Sample ..." button opens the Set Sample dialog, or closed it if it is open. Use this to select the observations to be used for estimation. The setting is the same as if opened from Setup / Set Sample. Pressing the "Simulate" button allows the "All Available Observations" button to show the available sample of exogenous variables for a simulation of the dependent variable. The show/hide status is remembered when this dialog is closed and re-opened.

* The "Options..." button opens the most recently opened estimation options dialog from the Options menu. By default, this is the Tests and Diagnostics dialog. Click the horizontal <<<< and >>>> bars in these dialogs to switch between options dialogs.

* The "Values..." button opens the equation Values dialog, or refreshes it if it is already open.

* To compute the estimates, press the <<< Go >>> button. This duplicates the 'Run' button on the toolbar, and the Actions / Run Estimation menu item.

* The "Clear" button clears all settings in this dialog.

Notes:
ESTIMATION CRITERIA
1. Note limitations on the application of some estimators.
    * GARCH and Markov switching options are unavailable with least squares and GMM.
    * Systems of equations are unavailable with skewed Student errors, and all discrete data options.
    * The Whittle method is only implemented for univariate ARFIMA models.

2. If no conditional heteroscedasticity or regime-switching features are specified, the difference between Least Squares/LGV and Gaussian ML is that in the latter case the error variance (or covariance matrix, for systems) is estimated jointly with the other parameters .

3. Multi-stage GMM, with efficient weights matrix based on first-stage residuals, is effected by running the estimation repeatedly. See Actions / Run Estimation. For this option to be active, the type of covariance matrix desired ('Robust', or 'HAC') must be selected in Options / Test and

© James Davidson 2024

Diagnostics Options. If 'Standard' is selected, only one-step GMM is available.

4.    In the probit/logit and ordered probit/logit models, the dependent variable must assume one of a fixed number of nonnegative integer values, including 0. In the standard binary case, the values must be either 0 or 1. "Count data" are also nonnegative integers, but in this case there is no fixed upper limit to the permitted values.

5.    To select the ordered probit/logit option, enter the number of additional cases in the Options / ML and Dynamics dialog. Setting this value to 0 (the default) gives the standard binary (0-1 data) cases. The data must be integer-valued with the reference (minimum) value equal to 0, and maximum value equal the number of cases less 1 (equivalently, 1 + the Options / ML and Dynamics dialog entry). See Section 4.7 of Models and Methods for the parameterization adopted. The shift parameters are cumulated to get the total intercept shift of each case, relative to the reference case. If an intermediate value is not represented in the data set, the model is under-identified, but a restricted version can be estimated by fixing the corresponding parameter to 0. For example, if the sample data take values 0, 1 and 3, fix "Ordered probit Gamma_2" at 0. This assigns zero probability to case 2.

MODEL SPECIFICATION

6.    Check the 'Impose Unit Root' option to estimate an ARIMA(p,1,q) model, and also an ARFIMA(p,1+d,q) where $-0.5 < d < 0.5$. This option is distinct from performing the estimation on the pre-differenced series, because the regression statistics, plots and forecasts are all generated for the original series of levels.
Note:
*    Any Type 1 regressors included are differenced along with the dependent variable(s).
*    If a trend is included, it becomes in effect an intercept for the differenced model. A Type 1 intercept is unidentified, and is not permitted. A Type 2 intercept can enter as an alternative to the trend, but will be suppressed if a trend is specified.
*    Type 2 and Type 3 regressors enter as usual, and are not differenced.
*    Lagged dependent variables, included as either simple regressors of Type 2 or as components of equilibrium relations, are not differenced.
*    In fractional cointegration models, the program modifies the reported value(s) of d3 in the appropriate way, such that 1 is added to the order of difference d1, as reported.

7.    Use the 'Values' button to open the relevant Values dialogs, and update them with new settings if already open. Use the 'Clear' button to remove all selections in this dialog. The 'Go' button launches an estimation run (duplicates the Run button and Actions / Run Estimation).

VARIABLE SELECTION

8.    The three regressor Types correspond to the vectors denoted $x\_1t$, $x\_2t$ and $x\_3t$ in equation (4.1) of Models and Methods. Therefore, the estimated coefficient of a particular regressor can be different depending its Type, when

© James Davidson 2024

AR and/or MA components are specified. Type 1 and 2 intercepts must be interpreted similarly.

9.  When one or more variables are selected in any category, the corresponding radio button is highlighted (lighter grey). To see the actual variable(s) selected (highlighted in the variable list) click on the radio button in question.

10.  Only one variable can be selected as dependent, unless the "Systems of Equations" box is checked. Multiple selection is allowed in the regressor categories.

11.  The 'Lags' setting allows different orders of lag for the regressors of each Type. For example, include variables for lagging in the Type 2 category, and non-lagged variables such as dummies in the Type 1 category. Individual lags can be suppressed by 'fixing' the coefficients at zero in the Values dialog. If neither lags nor ARMA components are selected, and there are no current endogenous variables, the regressor Types are all treated equivalently.

12.  If the dependent variable is selected as a regressor of Type 2, and Lags > 0, only the lagged values are included. This provides an alternative way (from setting AR Order > 0) of including the lagged dependent variable as a regressor. This feature is for compatibility with the Linear Regression option.

LINEAR REGRESSION

13.  A linear equation can also be specified in the Model / Linear Regression dialog, and in that case will be estimated by the usual analytic formulae. Models specified in this dialog are always estimated numerically.

14.  This dialog and the Linear Regression dialog cannot both be open at once. One is disabled whenever the other is open. The variable selections (Type 1 and 2 regressors) are common to both.

15.  An autoregression can be estimated either in this dialog, numerically, or as a linear regression by using the Lags feature (see 5 above).

BILINEAR MODEL

16.  Remember that if the bilinear order (r) is increased from 0 to 1, this adds p parameters to the model , where p is the AR order. However raising r further only adds one additional parameter at each step (not p) due to the imposed restrictions. This is different behaviour from the commonly described unrestricted bilinear model.

NONLINEAR MOVING AVERAGE (SPS)

17.  Selecting this feature replaces the error term Ut with the expression
        Ut - Ut-1(alpha + beta/(1 + exp(gamma*(Ut-1 - c1)*(Ut-1 - c2))))
    where alpha, beta, gamma, c1 and c2 are additional parameters.
    Combine this option with an imposed unit root to get models which can switch stochastically between I(0) and I(1).

18.  Variants:
    *       To implement a close approximation to the STOPBREAK

model, fix alpha = -2, beta = 2, c1 = c2 = 0 .

* To obtain the STIMA model, fix gamma =100 (or any sufficiently large value), impose the constraint c1 = -c2. Then, alpha is the MA coefficient for small deviations ($|Ut-1| < c1$) and alpha + beta for large deviations.

* Alternatively, all the parameters can (in principle) be freely estimated.

19. Nonlinear MA models can also be implemented by coding the required formula, see Models / Coded Equations. If both options are selected, this one is over-ridden.

OTHER MODEL FEATURES

20. The selectable model features are conditional variances (ARCH/GARCH), coded functions, regime switching, equilibrium relations, and constraints. Select these features to be included in the model with the checkboxes at the lower left side of the dialog.

21. If a feature checkbox is "greyed out" (disabled) this means that the feature is unavailable with current settings, usually choice of estimator. For example, conditional variance models need an ML estimator to be selected. Also, "Coded Function" is greyed out if coded equilibrium relations are specified in the Equilibrium Relations dialog. For this case the Coded Function dialog is dedicated to specifying equilibrium relations.

22. Model features generally require options to be set in the corresponding dialogs to be active. It is not possible to select a feature with the checkbox, unless some model options have been set.

23. Click the square buttons to the left of the checkboxes to open the corresponding dialogs. If options are set, these buttons are highlighted. The dialogs can also be opened from the Model menu.

24. NOTE: Values dialog(s) corresponding to model features (equilibrium relations, switching regimes) are not accessible unless the feature is checked.

25. To compute a Wald test as part of the run, as well as to estimate subject to constraints, the "Parameter Constraints" box must be checked. Wald tests can also be performed separately, see Actions / Compute Test Statistics / Wald Test.

SYSTEMS OF EQUATIONS

26. Select system estimation by checking the "System of Equations" checkbox.

27. When a system of equations is specified, the estimator computed depends on the estimator selection.

* If the option "LGV/FIML" is selected, the log-determinant of the error covariance matrix is minimized (Least Generalized Variance), or FIML in the simultaneous equations case (see 22 below). In these cases, the covariance parameters are concentrated out, and do not

© James Davidson 2024

appear in the vector of parameters.

   *   If the option "IV/2SLS/GMM/3SLS" is selected, the estimation
       method is system GMM with the instruments selected. If the
       estimation is launched with the Actions / Run Estimation commands
       or the "Run" button, the error covariance matrix set equal to the
       identity matrix. However, if the command Actions / Multi-Stage
       GMM is given twice in succession, the second run uses the efficient
       weights matrix estimated from the current residuals. This iteration
       can be repeated as often as desired. The form of the weights matrix
       used depends on the Covariance Matrix Formula setting (see
       Options / Tests and Diagnostics), either standard, robust (allowing for
       heteroscedasticity) or HAC.

   *   If the Gaussian ML, Student t ML or GED ML estimation options are
       selected, the error variances and contemporaneous correlations
       are estimated as additional parameters. The variances may be modelled
       by ARCH and other processes by selecting the "Conditional Variance
       Model" feature. These parameters (variances and correlations) can
       also be subject to regime switching.

28. If current endogenous variables are included as Type 1 regressors, the FIML
    variant of the LGV and ML estimators is automatically selected. That is, the
    log-determinant of the square matrix of coefficients is incorporated in the
    maximand. Suitable identifying restrictions must be imposed on the
    coefficients for this option to proceed. The "coefficient" of the normalized
    variable in each equation is automatically fixed at 0 in this case.
    Note: Endogenous variables specified as Type 2 regressors are ignored. If
    lags are specified, only the lagged values are included.

29. In a system, coded functions (see 4.6 Model / Coded Functions) may not be
    simultaneous - only one endogenous variable may validly appear in each
    function. However, coded functions may be embedded in linear simultaneous
    systems. Note that when coded functions are specified, the residuals of these
    functions replace the measured variables specified as dependent. The same
    substitution also occurs if these variables appear as contemporaneous
    regressors of Type I. Identifying restrictions must hold as normal.

30. In a VARFIMA model, it is possible to parameterize the fractional difference
    parameters for Equations 2,... as the differences from the parameter of
    Equation 1. This makes it easy to constrain and test equality of the
    parameters across equations. To select this option, set the third checkbox
    in the Model / Equilibrium Relations dialog. This setting applies even if error
    correction terms are not specified.

ESTIMATING TYPE I (V)ARFIMA PROCESSES

31. It is possible to correct for the omission of the presample data in a fractionally
    integrated model by including some generated regressors (depending on the d
    and variance parameters). This option is enabled by setting the "Type I
    Fractional" text box in Options / General to a positive integer value,
    corresponding to the number of generated regressors to be included. In the

© James Davidson 2024

output the estimated coefficients are denoted typeIFrac-Z1, typeIFrac-Z2,....
These are akin to principal components of the omitted stochastic terms.

32.     Recommended settings are for the number of components are 1 or 2. The
        columns of higher order are very small, so including too many may produce
        unstable results.

## 4.5 Model / Conditional Variance...

This dialog sets up the specifications for equations (6.1)/(6.2) or equations (7.1)/(7.2)
of Models and Methods. In the latter case, the specification set applies to the
conditional variances of each equation.

*       Use the scroll bars to choose order of GARCH specification.

*       Use the radio buttons and checkboxes to select other model options.

*       Use the radio buttons to switch between selection of the three possible types
        of GARCH regressor, then click on the variable list to make the selections.

*       Distributed lags of the GARCH regressors can be specified using the 'Lags'
        scrollbar, just as in the conditional mean model.

*       Check the "Abs." checkbox to take absolute values of the regressors
        automatically, or the "Sq." checkbox to take the squares of the regressors.
        Note that you cannot check both boxes at once.

Notes:
1.      This menu item is checked if a conditional variance model is specified. Its
        settings are ignored unless a ML estimator is selected and "Conditional
        Variance Model" is checked in the Model / Dynamic Equation... dialog.

2.      The 'regressor' radio buttons are highlighted (lighter grey) to show a selection
        has been made.
3.      The Options button opens a dialog also accessible from Options / ML and
        Dynamics. See 8.6a for information on these choices.

4.      Clear all specifications with the Clear button.

5.      If the 'GARCH' radio button is selected, the 'Asymmetry' checkbox selects
        the Threshold GARCH (GJR) model. In the other cases, the signed terms
        of the APARCH or EGARCH are included/suppressed, depending on the
        setting of this checkbox.

6.      The dependent variable(s) in the equation(s) can be chosen as GARCH
        regressors of Type 2, provided the lag length is set to 1 or greater. In this
        case the zero-order lag is suppressed, and the lagged dependent
        variable(s) act as GARCH regressors. This is similar to the option in the
        mean model. Dependent variables are NOT permitted as regressors of
        Types 1 or 3. These options are cancelled automatically.

© James Davidson 2024

7. If the 'Abs' checkbox is checked, the output shows the variable names enclosed in "|" . It is recommended to check this option when lagged dependent variables are selected, especially when running simulations.

8. The GARCH-M panel allows either the conditional variance, or conditional standard deviation, to be added to the selected regressors. Choose the regressor Type, and Mode, using the radio buttons. Type U denotes that the conditional variance or s.d. is to be included in a user-coded function. NOTE: To simulate GARCH-M, the regressor Type must be either 2 or U!

MULTIVARIATE GARCH

9. If the equation system option is selected, the default is to have fixed correlations between the variance-adjusted residuals. An alternative is the DCC model, in which the elements of the correlation matrix are generated by a common GARCH(1,1)-type model. The interpretation of the coefficients depends on the setting, standard vs. "ARMA in squares" form, see Options / ML and Dynamics. In the first case the fitted parameters are alpha and beta, in the second case they are delta = alpha + beta and beta. Note:
    1) The DCC parameters are displayed in the Values / Distribution dialog.
    2) The MA form setting in Options / ML and Dynamics does not affect the DCC specification.

10. If a multiple equation model is specified, the BEKK-Multivariate GARCH model can be selected. This is the case K = 1 as defined by Engle and Kroner (1995). Be careful to interpret the parameters correctly in this case. Selecting Equation i in the Values / Conditional Variance dialog, using the choice widget, displays the ith columns of the Aj and Bj lag polynomial matrices, and although note that these matrices really define the dynamic structure of the vectorized conditional covariance matrix, Vec H_t. See Models and Methods for details of the model. The estimates are reported in the results window in the same style.
    Note: Neither the MA form nor the GARCH form options (Options / ML and Dynamics) affect the BEKK parameterization.

11. Evaluation of the DCC and BEKK likelihood functions both require T NxN-matrix inversions, where T is the sample size and N is the number of equations. Function evaluation could therefore be very slow for large systems.

12. If the BEKK option is selected, the GARCH_M variables consist of the conditional variances and covariances of all equations. The SD mode is not available in this case.

13. In a system of equations with multivariate FIGARCH or HYGARCH disturbances, it is possible to parameterize the fractional difference parameters for Equations 2,... as the differences from the parameter of Equation 1. This makes it easy to constrain and test equality of the parameters across equations. To select this option (which will also apply

© James Davidson 2024

to VARFIMA parameters, if specified) set the checkbox in the Model / Equilibrium Relations dialog. This works even if error correction terms are not specified.

## *4.6 Model / Coded Function...*

This dialog is used to set up nonlinear functions coded by the user. Coded functions can form complete models for estimation, or components of models. Test statistics can also be coded. See Models and Methods, Section 4.6, for the details. The dialog is also used to specify coded functions for simulated data generation.

\*        Specifications created in this dialog are ignored unless the "Coded Function" checkbox in Model / Dynamic Equation is checked. The checkbox cannot be checked unless a specification exists.

\*        The "Clear" button clears all text fields, both the parameter names and the contents of the formula fields. Use with caution!

\*        The "Clear Built-In Components" button is provided as a convenience, for the case where the coded function represents the complete model. Then, all other model features need to be disabled. In particular, make sure that items such as the built-in intercept are deselected, to avoid duplication and an unidentified model.

METHODS

There are three ways to code a model for estimation:
I.       Type coded formulae for equations or equation components into the text box, interactively.

II.      Supply the code for the model equations as an Ox function, and compile this with the program.

III.     Write Ox code as in II, but for the complete likelihood function.

Method I is easiest to implement, but limits the function to a single line of code. Methods II and III allow any degree of generality, but require a minimum of programming expertise. Since the function is compiled with the program, TSM must be stopped and restarted to make changes to the specification.

INTERACTIVE CODING

To set up the model, do as follows:
1)      Open the Model / Coded Equations dialog, and select one of the radio buttons "Equation", "Residual", "Nonlinear MA" and "Nonlinear ECM".
2)      Select the dependent variable(s), in the usual way, in the Model / Dynamic Equation dialog.
3)      Type the equation formulae in the text fields. Use "Previous" and "Next" to navigate the fields, "Clear" to initialize the formula and "Cancel" to discard changes and revert to the stored text.
4)      Press "Test" to check the model parses OK, and list the parameter names in the name fields below.

Notes:

1. See Section 1.5 (General / Entering Formulae) for detailed instructions on entering codes into the text field.
2. Parameters removed from the equation are not deleted automatically, but are 'suspended'. Clear the name field manually to delete a parameter permanently from the model.
3. Increase the 'Maximum Parameters' value and refresh the dialog to display additional fields.

TYPES OF FORMULAE

"Equation" The equation is entered explicitly, in the format " = [Formula]", where the formula corresponds to the explained part of the model, and implicitly the dependent variable named in the heading appears on the left of the "=". The residuals (corresponding to the function f1(.) in equation (4.29) of Models and Methods) are computed as the difference between the left- and right-hand sides of the equation. Use this style whenever the model can be represented in the form required - which is most cases.

"Residual" Enter an expression which itself equates implicitly to the residuals of the model (f1(.) in equation (4.29)). In this case, no "=" sign should appear in the formula. The purpose of this option is to allow models in which the dependent variable(s) enter(s) nonlinearly. Since the model cannot be solved automatically, no Actual/Fitted plots, forecasts or simulations are available in this case. This option is also used for the formulation of equilibrium relations for ECM models.

"Nonlinear Component"
This is a nonlinear function of regressors, f2, augmenting or replacing the "Type 2" regressor selection. This component plays a different role in the model from the f1 specification when there are autoregressive or fractionally integrated components. Otherwise, f1 (equation) and f2 are effectively equivalent. However, since the dependent variable(s) are not specified here do not include the "=" in this case.

"Nonlinear ECM"
In error correction models, equations can contain nonlinear functions f3(Z) of the lagged equilibrium relation Z. Three popular forms are pre-programmed. For other cases, enter the function(s) f4 in implicit form, using the reserved variable name Z# to denote the equilibrium relation or, in case of two or more relations, the names Z#1, Z#2, ...
Lags can be included as Z#{-j} for j>0. Note that the lag j is relative to the minimum lag specified in the Model / Equilibrium Relations dialog. Presample lags are replaced by 0.

"Nonlinear MA"
With this option, the model disturbance term Ut is replaced by the moving average formulation

© James Davidson 2024

$$U_t = E_t - f4(E_{t-j}, j > 0).$$

Enter the formula for the function f4(.) using the reserved variable names E#{-j} to represent the disturbances lagged j periods.
(E here corresponds to v in equation (4.29) of Models and Methods.)
As in "Residual" the formula should be in implicit form, do not include "=".
In multi-equation models, use the variable names E#1{-j}, E#2{-j}, ..., to represent the disturbances for each equation.
Presample lags are replaced by 0.
Note:

1. With this option the function has to be evaluated recursively. Iterative estimation of these models could be time-consuming!

2. A pre-programmed nonlinear MA model can be accessed from the Model / Dynamic Equation dialog, and should be faster to estimate than the coded version. Use this option except to implement other variants. If both options are selected, the coded version takes priority.

Notes:

1. For the Equation option, an equation code must pair with a selected dependent variable - see the field heading. It is not possible to add an equation without first selecting a variable. Equations can be validly entered in the explicit format "Dependent Variable = [Formula]". The name of the dependent variable is added implicitly, if omitted.

2. In the Residuals option, the dependent variables must still be selected even though there is no explicit normalization. The number of equations formulated must be equal the number of selected variables. Checking that the formulated equations have a solution for the chosen variables is the responsibility of the user.

3. For the Residuals option, the typed equations define the model to be estimated. Dependent variable selections in Model / Dynamic Equation are ignored! Be sure to delete any unwanted equations using "Clear".

4. When "System" is checked in Models / Dynamic Equation, and two or more endogenous variables are selected, coded specifications can be created for each equation. Switch between these with the "Next" and "Previous" buttons. If no function is specified for a variable, enter "= 0" in the corresponding field.

5. While linear simultaneous equations systems can feature coded components (set the matrix B in equation (5.1) of Models and Methods)  simultaneity in coded equations is not allowed for estimation. That is:
   * "Equation" specifications should contain only exogenous or lagged endogenous variables on the right-hand side.
   * "Residual" specifications may be estimated, subject to the caveat of containing only one endogenous variable each.

It is the user's responsibility to respect these restrictions. The program cannot monitor violations, and incorrect estimates and simulations will result!!
Note, nonlinear simultaneous equations can be simulated using the W#1, W#2,... reserved names to denote shocks, see User Interface / Entering Formulae.

6. When the "Coded Equilibrium Relations" option is selected in the Model / Equilibrium Relations dialog, only the Residuals option is selectable in this dialog.

TEXT BOX CONTROLS

"Cancel"             Discard changes and restore existing formula.
"Clear"              Clears displayed formula. Enters left-hand side for "Equation" entry.
"Next"               Displays the next formula, where present, disabled otherwise.
"Previous"           Displays previous formula where present, disabled otherwise.
"Test"               Tests parsing of the current model equation(s).
Closing the dialog with the Windows Close button [x] is equivalent to pressing "Cancel".

PARSING THE MODEL

Parsing fails in cases including the following:
*       Variable name does not match any variable in the data set.
*       Non-matching or misplaced brackets/parentheses/braces.
*       Missing/misplaced operators.
*       Wrong left-hand side variable ("Equation").
An error message is displayed in a pop-up window, and should assist the debugging process
Note:   A name not appearing in either the variable list or current parameter list is assumed to be a new parameter, and added to the list. Check the list carefully to see if it's what you intend.

PARAMETER NAMES
*       "P[i]" where "i" is the parameter number shown beside the text field, is always a legitimate choice of name. It must still be typed explicitly in the field, so that the program knows how many parameters are specified.

*       Nine parameter name fields are displayed by default. To change the number, type the value in the "Maximum Parameters" field. To redraw the dialog box click any button, e.g. "Cancel".

*       Be careful that parameter names do not duplicate variable names. The convention of naming a coefficient by the variable it multiplies must be avoided here!

*       In multi-equation models, the same names appear in each equation, but refer to different parameters. They are distinguished in the output by prefixes "Eq1:", "Eq2:" etc.

*       If a parameter is named but does not appear in an equation, it is fixed at

© James Davidson 2024

0 in that equation. It appears in the Values dialog, but is suppressed in the output. This allows equation-specific parameters to be named distinctively

VARIABLE NAMES

* These can be spelled out in full or replaced by the short form "x[i]" where "i" is the position of the variable in the list, starting from 0. The line numbers are displayed in the variable list in Models / Dynamic Equation whenever the text field is open.

* Under the "Equation" coding option, select the dependent variable(s) in the Model / Dynamic Equation dialog first. Unless the formula fields already contain formulae (press Clear to delete these), opening a field will automatically insert the variable name and " = ", ready for the rest of the formula to be typed.

SIMULATING CODED EQUATIONS

Any coded functions of "Equation" type can be simulated. Normally, the generated disturbance is added to the coded function to simulate the dependent variable. However, a special feature available for simulation models is to include the disturbance as a coded component, which can therefore be transformed or enter the model nonlinearly.

* After selecting "Equation" in this dialog, represent the disturbance by the reserved name W#, or W#{-j} for lagged values. See 1.5 General / Entering Formulae for further details. If the reserved name W# appears in one or more equations of a model, shocks are not added to any coded function(s). Therefore, in a multi-equation model, every equation must have its shock included explicitly in the formula if any one equation does.

* Models with these features cannot be estimated. If the W# reserved name appears in an equation for estimation, an error message is generated. However, it will often be possible to estimate the same model in a different formulation. Code the function separately for estimation, by selecting "Residual" in this dialog and creating the counterpart nonlinear equation for W#. The reserved names E#{-j} can be used to represent the lagged residuals, in this context.

* In equation systems, the generated shocks are written W#1, W#2, etc. It is possible to simulate models having a recursive (triangular) structure. The equations are solved sequentially, and equations numbered 2 or above can contain solutions from the preceding equations. These are coded using the reserved names E#1, E#2, etc.
NOTE: Such equations cannot be estimated, in general! TSM can handle only linear simultaneity in estimation.

SIMULATING WITH RANDOM EXPLANATORY VARIABLES

Monte Carlo experiments may need to be conducted under "random regressor" assumptions, so that the estimators simulated are marginalized with respect to the distributions of exogenous variables. One way to achieve this is to set up a recursive (triangular) multi-equation model, in which one or more equations generate the explanatory variables. Another way is to insert normally

distributed random numbers directly into the relevant columns of the data matrix prior to the simulation. This method is generally more numerically efficient, and is particularly useful in models such as the probit, logit, Poisson etc. where only single-equation estimation is currently supported,

* To enable this feature, simply reference a variable in the coded equation as X#[i] where i represents the column number of the data matrix. (Referring to the variable by name is not supported.)

* Only standard normal data can be generated in this way, but the equation coding might apply transformations to yield modified distributions.

* Don't forget that the previous contents of the data set are over-written each time a simulation is run. It's recommended to create "placeholder" variables to hold the generated data. Do this using the "Make Zeros" and "Rename" commands in the Setup / Data Transformation and Editing dialog.

OX CODING

See Appendix C for information on compiling supplied code into the program. Supplied Ox code can be used to return five types of object, according to the selection made with the radio buttons -

Residuals: Return a vector/matrix of equation residuals, as a function of data and parameters from UserEquation(). Return solved values and forecasts from UserSolve().

Log-likelihood: Return a vector of log-likelihood contributions as a function of data and parameters from UserLikelihood().

Statistic: Return a statistic or vector of statistics as a function of data and optional constants (passed as parameters) from UserStatistic().

Data Generation: Return a generated data set as a function of existing data and constants passed as parameters from UserGenerate().

Test Only: Choose this setting if the only coded function to be returned is a test statistic from UserTest(), based on the regular estimation outputs, selected with the "Test" checkbox. If this option is set, the "Coded Function" checkbox is automatically deselected.

Notes.
1. If the coded functions have been set up for selection by name, as described in Appendix C, they can be selected in the dialog using the Previous and Next buttons to navigate the list. Preset names cannot be edited. If no naming function has been defined, a function can still be selected by typing its name into the field.

1. The test statistic option, from UserTest(), accessed by checking the "Test" checkbox, is compatible with any estimated model. This includes linear equations, and the Ox Residuals and Log-likelihood options, but not Statistic

© James Davidson 2024

or Data Generation.

2.      If a coded test is the only programmed option, select "Test Only" and do NOT check the  "Coded Function" checkbox in Model / Dynamic Equation.

3.      If "Coded Equilibrium Relations" is selected in Model / Equilibrium Relations, and "Equilibrium Relations" is checked in Model / Dynamic Equation,  the "Coded Function" checkbox in in Model / Dynamic Equation is disabled. Note that in this case the parameter values for the coded relations are displayed in Values / Equilibrium Relations

4.      For all options except the coded test and coded equilibrium relations, the "Coded Functions" checkbox must be checked in Model / Dynamic Equation to activate the option.

5.      Enter names of parameters in the text fields provided in the dialog. Optionally, an identifying name can be entered in the 'Function Name' field. This is used to identify the output and can also select the function from a code library maintained in the run file. Alternatively, a name can be passed from the function itself. To activate this option, leave the Function Name field blank.

6.      When residuals or the log-likelihood are returned, setting dependent variable(s) in the Model / Dynamic Equation dialog is optional, though recommended. If set, it/they should correspond to the normalized variable(s) in the supplied functions. These settings are not used to compute the estimates, but allow the correct plots, series and forecasts to be generated.

7.      It is the user's responsibility to set the sample period (see Setup / Set Sample) to ensure that all observations are available for the variables in the model. This is done automatically for models specified through the program options, but the program cannot know whether variables included in the user's function have missing observations.

8.      For statistics and data generation, there is the option to enter names for constants in the parameter name fields, then set values for these in the Values / Equation dialog. These values are not modified by TSM, but are passed through directly to the user's function. If this feature is utilised, it provides a means for quick changes of specification, without editing the code.

9.      To generate data, "placeholder" variables must exist in the data set. The contents of these columns is unimportant, being replaced by the generated series when the simulation runs. For example, they can be created using the Extend Sample, Make Zero and Rename commands in Setup / Data Transformation and Editing. Then select these variables as the Dependent Variable(s) in Model / Dynamic Equation.

## 4.7 Model / Regime Switching...

This dialog allows the choice of
* number of regimes,
* type of switching - simple Markov switching, explained switching, Hamilton's model, or smooth transition (ST).
* which model components to switch,
* specification of the explained switching probabilities or transition function, where these options are selected.

Notes:
1. Regime switching is not activated until at least one model component is selected for switching!

2. Except for the smooth transition model, regime switching requires a time-domain maximum likelihood estimator. Options are "greyed out" until a valid selection is made.

3. The option "Estimate Regime Differences" selects the parameterization that allows individual significance tests of differences between the regimes. However, note that in a test for *no* differences between regimes (= test for number of regimes) the switching probabilities or ST parameters are unidentified under the null hypothesis. This is a well-known testing pitfall, and the standard tests don't apply.

MARKOV SWITCHING
4. If "Estimate Regime Differences" is selected with Markov or Hamilton switching, then rows 2,...,M of the Markov transition matrix are likewise expressed as differences from row 1. This allows a test of the hypothesis that transition probabilities are independent of the current regime.

5. In the Markov and Hamilton models, both the matrix of estimated transition probabilities $p\_{ji}$, and the unrestricted parameters $t\_{ji}$ with standard errors are reported in the output. The latter are labelled "Logistic of P(j|i)". See Models and Methods for details.

6. In Hamilton's model, the "Intercept" selection in Models / Dynamic Equation is ignored. The intercept is specified as shown in Models and Methods, Section 8.3.

7. Analytic forecasts are available for the Markov switching models. However, the number of calculations needed to compute the multi-step forecast 2-standard error bands increases exponentially with the number of steps, and can become excessive. Standard error calculations are terminated after 100 seconds has elapsed, and only the point forecasts are reported for further steps. A warning message is printed in this case. The SE bands can also be omitted, see Options / Graphics.

EXPLAINED SWITCHING
8. In the explained switching case, the radio buttons allow the choice of variables to explain the probability of switching into the given regime at

© James Davidson 2024

date t. An intercept can be optionally added to the equation.

9. Distributed lags are not implemented for explained switching and smooth transition models, but the 'Lag' scrollbar allows a fixed lag to be applied to the variable(s). If and only if the lag is positive, the dependent variable(s) can be included in the set.

10. There are two versions of the explained switching model. In the first version, the equation for the probability of switching into regime j at date t may depend on the current regime (i) only through an intercept shift, by inclusion of a dummy variable for the intercept in question, with Regime 1 as the reference case. In the second version, the switching equations may be fully regime-dependent, and switch just like other model components. Check the box "Regime-Dependent Coeffs." to enable this case.

11. The second explained-switching model contains the first one as a special case, but the latter is set up in a slightly different way, with regime dummies, and is retained for backwards compatibility and simplicity. If dummies (or switching intercepts) but no explanatory variables are specified in explained switching, the model represents an alternative parameterization of the regular Markov-switching model.

SMOOTH TRANSITION

12. In the smooth transition (ST) case, select the transition indicator ($z\_t$) from the variable list after clicking the first radio button. Only one selection is possible. If no transition indicator is selected, the smooth transition option is cancelled and the type setting reverts to Markov switching.

13. The inclusion of an ST intercept is selected by a checkbox. If both this and explanatory variables are omitted, the dominant regime depends on the sign of $z\_t$. The second set of parameters is unavailable unless 'Double Transition' is checked.

14. The number of regimes is fixed at 2 for the ST model.

15. NOTE: The explanatory variable selections for the explained switching and smooth transition options share a common storage location. If the type of switching selected is changed, the existing variable selections for these options are cancelled.

16. Analytic forecasts are not available for ST models in this release. Use the Monte Carlo forecast option.

17. The ST model has two identical maxima of the likelihood, since interchanging both the regimes and the sign of the smoothness parameter (gamma) yields the same likelihood, by construction. Be careful to interpret the estimates correctly.

## 4.8 Model / Equilibrium Relations...

An Equilibrium Relation (Eq.R.) is defined as a linear combination of lagged

© James Davidson 2024

variables that is inserted as a "regressor" in one or more equations of a system, with the combining coefficients being either restricted or estimated. This is sometimes called an Error Correction Model (ECM). Most often, an Eq.R. has the interpretation of a cointegrating relation in a system of I(1) variables, although it could also, in principle, appear in a model of stationary processes. Eq. R.s can also appear in single equations, in which case they implement a re-parameterization of the dynamic model in "ECM" form.

*       Choose one of the three options for specifying the equilibrium relation(s).
      A:      Select "Linear Equilibrium Relations" to specify the relation(s) by choosing variables from the list.
      B:      Select "Nonlinear AR / Closed VECM", to include the set of dependent variables as lagged regressors automatically.
      C:      Select "Coded Equilibrium Relations" to specify nonlinear equilibrium relations using the Coded Function dialog.

*       Use the first scroll bar to select the number of Eq.R.s to be included in the model.

*       Use the second scroll bar to chose the lag to apply to the Eq. R.s (must be >0).

*       Restrictions on each relation are set up in the Values \ Equilibrium Relations dialog.

Notes:
Options A and B
1.      Every Eq.R. must contain at least one fixed coefficient, otherwise the relation will be unidentified. Set the fix in the Values / Equilibrium Relations dialog. If no restriction is set manually, the first coefficient in the list will automatically have its coefficient set to 1 when the estimation is run.
Note: The Values button is highlighted if two or more coefficients are restricted in any relation.

2.      With two or more Eq.Rs, note that this dialog specifies the variables to appear in all of them. To create distinct relations, exclusion restrictions (fixing coefficients at 0) must be set up in Values / Equilibrium Relations.
Note: For consistent estimation, the relations should satisfy the standard rank conditions for identification of a linear system.

3.      When "Closed VECM" is checked, the lagged variables included in the VECM are the dependent variables LESS the fitted combination of Type 1 intercept and/or Type 1 regressors. See Models and Methods, Section for 5.3, details of the specification.

Option C
4.      The equilibrium relations are formulated as the function f5 in the Model / Coded Function dialog. Note that "Residual" format must be used (no "=" in the field). This is the only permitted option when option C is selected in this dialog.

          © James Davidson 2024

5.	When "Coded Equilibrium Relations" is checked, and "Equilibrium Relations" is also checked in the Model / Dynamic Equation dialog, the "Coded Function" checkbox is disabled - coded functions can appear either in the main equation(s) or in equilibrium relation(s), but not both at once.

6.	The number of relations in the model is determined by the number specified in the Model / Coded Function dialog. This overrides the setting in this dialog.

7.	Formulating the relations with enough restrictions to ensure identification is entirely the responsibility of the user. Use this option only if you know what you are doing!

8.	The specification of  the relations can be different when there are two or more of them. They nominally contain the same parameters, but these can be automatically fixed at 0 if not included in a particular relation. See Model / Coded Functions for details. They can also be restricted in the The Values / Equilibrium Relations dialog, if desired.

All Options
9.	Eq.R.s are only included in the model if this feature is selected in Model / Dynamic Equation.

10.	The radio buttons in the "Types of ECM" area allow selection of nonlinear response mechanisms as well as the usual Linear case. See Models and Methods for details. Three popular variants are pre-coded. Use the Model / Coded Function dialog to define other cases.

11.	Check the first checkbox in the "Fractional Cointegration" area to implement a fractional cointegration (F-VECM) model. Note that this option is ignored unless the VARFIMA / Fractional VECM box is checked in the Dynamic Equation System dialog. In this case, setting the "Impose Unit Root" option causes the estimated value(s) of d1 to be smaller by 1 than otherwise. This is taken into account in the calculation of the cointegration order(s) d3, such that the equilibrium relations appear differenced to order 1+d1-d3 in the system.

12.	Check the second checkbox to select the generalized cointegration model. In this case, the cointegrating variables may be fractional differences of the measured variables. See Models and Methods for details.

13.	The checkbox labelled "Diffs Spec'n for FI & ECM Pars" controls the option of estimating certain parameters in equations 2,..., N as the differences from the corresponding parameters in Equation 1. The parameters in question are the fractional integration parameters d1 and d3, and the parameters of the nonlinear ECM adjustment mechanisms. The option allows the restriction of equal parameters to be both tested and imposed, in the latter case by putting the parameters for equations 2,...,N to zero and setting the "Fixed" checkboxes, in the values dialog.

NOTE: It is not necessary to specify a fractional VECM to set this option. It applies also to a regularly-cointegrating or non-cointegrating VARFIMA model, and also to the fractional lag parameters of a system FIGARCH or HYGARCH specification.

14. If "Impose Unit Root" is specified in Model / Dynamic Equation, the VECM is estimated with unit roots imposed in the short-run dynamics, while the Eq. R.s enter as lagged levels.

## *4.9 Model / Parameter Constraints...*

This dialog sets options for imposing restrictions on the coefficients.

Three types of restrictions can be set up:
a)      zero restrictions,
b)      multiple linear restrictions, and
c)      general restrictions specified as coded formulae.

In case (b), only the constants (intercepts) of the constraint functions are set in this dialog. The linear functions themselves, like the parameters for zeroing, are set in the Values dialogs.

Restrictions can be implemented in either of two ways:
*       Tested on the unconstrained estimates, using the Wald principle,
*       Imposed in estimation.

In the latter case, the LM test for the restrictions is computed, if this option is selected in the Options / Tests and Diagnostics dialog. In this case, any fixing restrictions on the parameters will be treated as part of the maintained model and not tested. To test fixing restrictions, make sure the Parameter Constraints option is deselected in Model / Dynamic Equation... or Model / Linear Regression

Notes:
1. Up to 10 linear restrictions can be specified with nonzero constant terms. However, up to 20 additional restrictions could be set with zero constant terms.

2. Estimation subject to constraints is not available with OLS and IV estimation. This option is disabled ("grayed out") when the Model / Linear Regression dialog is open.

3. By default, the statistic reported is asymptotically chi squared. Set the check box to report the F statistic, with the reported p-value taken from the F tabulation. Note that this setting is shared with the user-specified LM and moment tests (see Actions / Compute Test Statistics). It can be selected/deselected in any of the relevant dialogs.

REGRESSOR SIGNIFICANCE TEST
4. The option "Test Joint Significance of Regressors" provides a valid time series implementation to the "F test of the regression" computed by many

© James Davidson 2024

packages.

5.	In a single equation it sets up a Wald test of the joint significance of all regressors in the mean model, all Types and components of equilibrium relations (if any) but excluding the intercept, trend and seasonal dummies, and lagged endogenous variables. The null hypothesis can therefore be set up as a univariate time series representation of the dependent variable. Note that this test can be performed in the usual way by selecting the parameters individually in the Values dialogs, but this option provides a handy shortcut.

6.	The program "knows" a regressor is a seasonal dummy if the string "seasonal dummy" or "seasonal dummies" is part of the data description. This is set up automatically if the dummies are generated in Setup / Data Transformations and Editing.

7.	In equation systems, all exogenous regressors of all types in all equations and in equilibrium relations are included in the test, but intercepts, trends and current and/or lagged endogenous variables are excluded. "Endogenous" means any variable explained by the system.

8.	Additional parameters can be added to the test set in the usual way, by checking the relevant "Wald Test" box in the Values dialogs.

9.	Variables having parameters "fixed" are excluded from the test.

CODED_RESTRICTIONS

10.	A coded restriction on the parameters is set up by typing the equation into a text field. After selecting the restriction type with the radio button, press the Code button to open the text box.
TEXT BOX CONTROLS
"Cancel"	Close box, discarding changes.
"Delete"	Delete the currently displayed formula.
"Next"	Displays the next formula, where present.
"New"	Opens a new blank formula where there is no "Next".
"Previous"	Displays previous formula, where present.
"OK"	Closes the text box and stores current formulae.
[Return]	Equivalent to pressing "OK".
Closing the dialog with the Windows Close button [x] is equivalent to "Cancel".
For information on entering formulae, see Section 1.5

11.	Parameters can be written in the notation "a[i]" where "i" is the number appearing in the left-hand column in the Values dialogs. "p[i]", "A[i]" and "P[i]" are all permitted variants, but the square brackets are mandatory, note.
CAUTION: Parameters can also be identified by name. However, some parameters are named differently in the Values dialogs and in the output. If in doubt use the number notation, which is unambiguous.

12.	Suppose there are P restrictions. Using the numbering notation

© James Davidson 2024

described in Section 1.5 (in practice, parameter names can also be used) the restrictions must be entered in the solved form

a[j1] = f1(a[k1],...,a[kM])

...

a[jP] = fP(a[k1],...,a[kM])

for k1,...,kM not equal to j1,...jP, where f1(),..,fP() are differentiable algebraic expressions. The program imposes the restrictions by replacing a[j1],...,a[jP] with the given functions of a[k1],...,a[kM].

13. Restrictions cannot be entered in implicit form!
    Example:
        a[0] = a[1]*a[3];
        a[2] = exp(a[1] - a[4]);
    is OK.
        a[0] = a[1]*a[2];
        a[2] = a[1] + a[3];
    is NOT valid! (Not solved form.) A valid form of the first equation would be
        a[0] = a[1]*(a[1]+a[3]);

## 4.10 Model / Select Instruments...

This dialog is enabled only if IV/GMM is the selected estimator.

* Highlight by clicking with the mouse those variables to be included as instruments.

* Use the checkboxes to indicate whether the intercept and trend dummies are to be included as instruments.

* Use the "Additional Instruments" scrollbar to set the number of lags of additional instruments to be used.

* Set the "Initial Lag" scrollbar to a positive value to include only lags of this at least this order.

* Use the "Lagged Endogenous" scrollbar to set the number of lags of endogenous variables to be used as instruments.

* The Clear button unselects all variables in the list.

Notes:
1. Any non-lagged variable appearing in the equation but not appearing in the instrument list is treated as endogenous.

2. Any variable appearing in the instrument list but not in the equation is an additional instrument. Set the first scroll bar to select the number of lags of these variables to include in the instrument set.

3. To use lagged endogenous variables as instruments, set the second scrollbar

to the number of lags to be included:

- o      If this scrollbar setting is zero, no lagged endogenous variables are used as instruments.
- o      If this scrollbar setting is positive, the number of lags of a variable included as instruments is the greater of this setting and the number of lags specified to appear in the equation, according to the lag setting of the variable's Type.

4.      The "Initial Lag" setting applies both to additional instruments and, if positive, to lagged endogenous instruments. The number of lags is not affected by this setting, so the longest lag is increased commensurately.

5.      The normalized (left-hand side) variable (or variables, in a system) cannot be selected in this dialog, since they are necessarily endogenous. Its (their) lags are included automatically if the second scrollbar is set to a positive value.

7.      In panel data models the intercept is suppressed under certain data transformations such as individual mean deviations and first differences. When these options are selected the instruments are subjected to the same transformation. The intercept is not available as an instruments in these cases, and the checkbox is disabled.

# 5. Values

## 5.0 Values: General

The Values dialogs are used to operate on model parameters. They allow
*       setting starting values for the search algorithm.
*       fixing parameters at chosen values in the estimation.
*       setting bounds on parameters, for estimation subject to inequality constraints.
*       setting a range for a grid evaluation of the estimation criterion, in one or two dimensions.
*       selecting parameters for Wald significance tests.
*       setting up linear functions of parameters for Wald restrictions tests.


Unlike the other dialogs having fixed layout, Values dialogs can be resized.
Grab the border and drag with the mouse.


BUTTONS

1.      The "Refresh" button rewrites the values field to reflect changes to the model specification. Pressing the "Values" button in the relevant Model dialog has the same effect.

2.      Press the "Clear" button once to set all free values to defaults. Fixed values are left unchanged. Press the button a second time to clear all settings, including parameter fixes. Pressing the Refresh button resets the function of the Clear button.

3.      In the Equation and Equilibrium Relation dialogs, when more than one equation/relation is specified, [>] and [<] buttons are displayed on the button bar. Use these buttons to cycle through the equations/relations, respectively forwards and backwards.

4.      All dialogs except Markov Probabilities and Smooth Transition have "Regime:" [>] and [<] buttons that appear when switching regimes are specified. Note that if the parameters in question have not been selected to switch, then Regime 1 shows the common values for all regimes.
        The dialogs for Regime 2 and above are then blank.

5.      An alternative navigation method is to use the choice widget(s) on the tool bar. First choose the equation/relation/regime values to be displayed.
        Change the display by either double-clicking the widget, or pressing the "Refresh" button in the Values dialog.

6.      All dialogs have a checkbox for selecting the parameter bounds and grid plotting layout. This control is located at the bottom of the panel, it may be necessary to scroll the display view it. It shows/hides two columns of text fields for entering bound values - see paras 11-15 below. Checking/ unchecking the box in any dialog effects the setting in all the Values dialogs.


VALUES FIELDS

7.      The values attained in an estimation run are displayed in the Values fields to double-precision machine accuracy. These provide the starting values for a

© James Davidson 2024

new run, and can be edited.

8.    Check 'Fixed' to fix the parameter at the value set, during estimation. Applications of this feature include:
*     Suppressing intermediate lags in ARMA/GARCH models (fix the values at zero).
*     Temporarily restricting a model, without changing the specifications.
*     Computing LM tests of the corresponding restriction(s) - see Options / Test and Diagnostics Options.
*     Evaluating the optimized likelihood over a grid of points - see paras. 13-14 below and also Actions / Plot Criterion Grid for more details.

9.    If the model specification is changed, values are reset to defaults. To remove a variable or lag from the model temporarily, "fix" its parameter at zero. This preserves other stored values.

10.    As a reminder that one or more parameters are fixed, the Values buttons in the corresponding dialogs are highlighted, and check-marks appear against the Values menu items.

PARAMETER BOUNDS

11.    To constrain a parameter to lie in an interval, enter the upper and lower bounds in Values but do not check the 'Fixed' box. The reported parameter is then a logistic transformation of an underlying, unconstrained value. The bounds are ignored unless the upper strictly exceeds the lower. Cancel the setting by putting both bounds to 0.

12.    If the estimate is in the interior of the interval, the standard error is approximated using the delta method. However, if the constraint binds in the sample, the covariance matrix may be near singular and standard errors unavailable or unreliable. It may be best to re-estimate with the parameter fixed at its boundary value, in this case.

13    It is not recommended to use this technique routinely to impose, e.g., positivity and stability constraints on lag polynomials. The search algorithm should usually work OK without this. The method is implemented more as a last resort for difficult cases.

GRID PLOTS

14.    If bounds are set and the 'Fixed' box is also checked, the bounds will be used as the limits of a Grid Plot. Either one or two parameters can enter the grid. If the 'Fixed' box is checked and both bounds are 0, then the parameter in question is simply held fixed when the grid is drawn. If three or more parameter are both fixed and have bounds set, the first two encountered are the plotted parameters. The others are just treated as fixed in the usual way.

15.    The number of grid points must be set in the Options / Optimization and Run dialog. The number is counted from zero. Thus, if the bounds are 0 and 1 and it is desired to evaluate at intervals of width 0.1, enter 10 in this text field. The likelihood will then be evaluated at 11 points, 0, 0.1,0.2,…,1.

16.     To compute a Wald test of joint significance, enable "Wald Tests" in the Model / Linear Regression dialog, or "Parameter Constraints" in the Model / Dynamic Equation dialog. Then check the checkboxes of the variables to be included in the test.

17.     If the option to test joint significance of the regressors is selected, these are included automatically, and do not need to be checked her. Any additional parameters checked here will be added to the test set.

18.     To compute a Wald test of linear restrictions Rb = c, or impose these restrictions in estimation, the number of restrictions and the elements of c must first be entered in the Model \ Constraints dialog. Then, refresh the Values dialog and enter the columns of R' (i.e., the rows of R, transposed) in the fields displayed.

19.     A restriction can involve any subset of model parameters, and hence can be spread over two or more number of Values dialogs. It may be helpful to open the relevant dialogs together and arrange them on the screen in a column. Restrictions may also be applied across equations and across regimes. Since only one equation/regime Values dialog can be displayed at a time, use the choice widgets or "Next..." buttons to switch quickly between them.

20.     For each linear restriction, one parameter must be designated to be "solved" as a function of the remainder. Make this choice by checking the relevant "Constrained"/"Wald test" checkbox. There must be same number of checks in this column as restrictions, otherwise an error message is generated. It is also the user's responsibility to be sure that the restrictions can be solved uniquely for the checked parameters - in other words, the square sub-matrix of R' defined by the checked rows must be nonsingular. If this property does not hold, an error message is generated.

## 5.1 Values / Equation...

This dialog sets the parameters specified for an equation, including specifications set in either the Linear Equation dialog or the Dynamic Equation / Conditional Variance / Supplied Function dialogs.

Notes:
1.      The Parameter Bounds / Grid Plotting checkbox affects all the Values dialogs. When checked, a text field is displayed to set the number of grid points for criterion plotting.

2.      In multi-equation models, use the "Equation:" [>] and [<] buttons to cycle through the equations, either forwards or backwards. Alternatively, select an equation using the "Equation" choice widget on the tool bar and and then press the "Refresh" button. In regime-switching models, to see the different regimes use the "Regime:" [<] and [>] buttons, or the "Regime" choice widget, similarly.

© James Davidson 2024

3.    If a BEKK multivariate GARCH model is specified, the ith columns
      of the matrices A_j and B_j are displayed in the dialog for equation i,
      as Alphaj(i,k) and Betaj(i,k), for k = 1,...,[number of equations].
      Be careful to note how these matrices enter the model - see the main
      document for details. Although entered and reported in the same
      style, these matrices are not comparable with the multivariate GARCH
      specification. Be careful to interpret them correctly.

4.    DCC dynamic parameters are displayed in the Values / Distribution
      dialog.

## 5.2 Values / Equilibrium Relations...

This dialog shows parameters of the equilibrium relation(s).

Note:
1.     If more than one relation is specified, use either the [>] and [<] buttons
       in the dialog, or the "Equation" choice widget on the tool bar, to switch
       between them.

2.     At least one of the parameters in this dialog must be fixed at a non-zero
       value (the normalized variable). If none is specified manually, the first entry
       is automatically fixed at 1 when the equation is run.

## 5.3 Values / Distribution...

This dialog shows parameters defining the distribution that are not associated
with any one equation in a multi-equation model. These are:
*      the Student t parameters (degrees of freedom and skewness)
*       parameter of the GED distribution
*      parameter of the negative binomial distributions
*      degrees of freedom for the chi-squared probit model.
*      The intercept of the excess zero probability function in a "zero-inflated"
       Poisson or ordered probit/logit model. (Regressors are of Type 3.)
*      error correlations.
*      Parameters of the DCC multivariate GARCH model

Notes:
1.     In regime-switching models, use the [>] and [<] buttons to cycle through
       the regimes, or the "Regimes" choice widget on the tool bar, with
       "Refresh", to select a regime directly.

2.     Covariance parameters of the BEKK multivariate GARCH model,
       while not associated with any one equation, are nonetheless displayed in
       the Values \ Conditional Variance dialog. See Help page 5.2 for
       correct interpretation of these parameters.

## 5.4 Values / Switching Regimes...

This dialog displays parameters specified in the Model / Regime Switching dialog.

© James Davidson 2024

\*    In the Markov switching case (including Hamilton's model) these are the transition probabilities $p_{ji} = P(j|i)$, forming a $M'(M-1)$ matrix. These are the probabilities of switching from regime $i$ to regime $j$.

\*    In the Explained Switching and Smooth Transition models, the dialog shows the parameters of these switching models.

Notes:

1.    In the Markov case, the switching probabilities are constrained to be positive with their sum in [0,1], such that the implicitly defined Mth probabilities are nonnegative. These are transformed to unconstrained values $t_{ji}$ for estimation using a logistic map. See Models and Methods for the mapping formulae. These values are displayed in the output, with standard errors.

2.    To show the $t_{ji}$ values in the dialog, with upper and lower bounds transformed similarly when shown, check the "Display Logistic Transformations" checkbox.

3.    By selecting the "Estimate Regime Differences" option, the Markov switching probabilities can (for example) be constrained not to depend on the current state. Under this hypothesis, the sample is simply a random mixture of regimes prevailing with the unconditional switch probabilities. Note that the null values of the "incremental probabilities" of regimes 2,...,M are 1/M, corresponding to $t_{j2} = ... = t_{jM} = 0$.

4.    If regime-dependent coefficients are specified in Explained Switching, use the "Next Regime" button to cycle through the regimes, or the "Regime" choice widget on the toolbar to select a regime directly.

5.    Note that specifying only the intercept and regime dummies in the explained switching model provides another way to estimate the simple Markov switching model, in an alternative parameterization.

# 6. Actions

## *6.1 Actions / Run Estimation*

Closes all dialogs, saves the current configuration in "settings.tsm", and starts estimation of the currently specified model.

To have the dialog box configuration restored automatically after the run is finished, select this option in General Options. Alternatively, re-open by hand with Actions / Restore Dialogs.

Notes:
1. This command can also be started with the "Run" button on the toolbar, and with the "Go" buttons in the Model / Linear Regression and  Model / Dynamic Equation dialogs.

2. Sequential and iterative calculations can be interrupted while in progress, by clicking the "Stop" button on the toolbar. A choice box appears, possibly after a short delay, with the option of aborting the run, or continuing.

BATCH JOBS
3. If the checkbox "Run Next Estimation as External Process" is checked in Options / Optimization and Run before the command is given, the run is performed externally by starting a new instance of Ox.

4. If the checkbox "Defer External Jobs" is checked in Options / General, the Ox file containing the job is created, but not run. Use this option if the job is to be run on a different machine, for example.

RAW SERIES DIAGNOSTICS
5. If the "Run Estimation" command is given when a dependent variable is selected but NO model features are selected (i.e. no parameters specified), any selected series diagnostics, such as tests of autocorrelation and weak dependence, are computed for the raw data series. Specify desired tests in the Diagnostic Tests dialog under the headings "Q Statistic" and "Tests of Residuals".

6. Tests of I(1) including ERS tests, using either asymptotic or bootstrap criteria, can be computed for raw series in the Setup / Compute Summary Statistics dialog. The same tests can be run as raw series diagnostics, allowing them to be evaluated in Monte Carlo experiments.
   Specify the settings for these tests as follows.
   1) Enable "Raw Data I(1) Tests" in Options / General / Special Settings.
   2) Make sure that NO model features are selected.
   3) Open the Setup / Compute Summary Statistics dialog. Select the variable to be tested, and differencing and/or detrending options if required. Do not close the dialog!
   4) Set test options in Options / Tests and Diagnostics. For bootstrap tests, also set options in  Options / Simulation and Resampling.

5)      Go to Model / Save Current Model.
Keeping the Summary Statistics dialog open causes the settings to be stored separately in the named model, without changing the existing test and bootstrap specifications.

## *6.2 Actions / Evaluate at Current Values*

This command is the same as Actions / Run Estimation except that the optimization algorithm is not called. The post-estimation output is simply created at the parameter values currently stored in the Values fields. Note that these are the estimates, following an estimation run. This option can be used, for example, to obtain test statistics, forecasts, listings etc. for a model that has just been estimated, without repeating the optimization.

Notes:
1.      This command is also launched by the "Evaluate" button on the toolbar.

2.      If either the 'Semiparametric Long Memory', 'Cointegration Analysis' or 'Monte Carlo Experiment' dialog is open, the command has the same effect as the 'Go' buttons in those dialogs.

3.      Most evaluations cannot be interrupted, once begun. The exceptional cases are bootstrap and subsampling runs, and Monte Carlo forecasts.

## *6.3 Actions / Estimate Multiple ARMA Models*

Estimates a sequence of models with increasing ARMA orders, starting with ARMA(0,0). All other features of the model are as currently specified, although regime switching options are disabled. The limits of the sequence are set in the Setup/Multiple ARMA Models dialog.

Notes:
1.      Starting values are set to either the estimates from the preceding specification, or zero, as appropriate. The ARFIMA(0,d,0) is initialized with Robinson's (1994) nonparametric estimator of d.

2.      Other selected model features (regressors, conditional heteroscedasticity, supplied function) are estimated as usual, with the supplied starting values, for each case.

3.      Post-estimation options (forecasts, correlograms, graphics, impulse responses, tests, series output and retrieval) are suspended. Re-estimate the selected model to get these outputs.

4.      This option cannot be used in conjunction with regime switching, nor with multiple-equation models.

5.      The model optimizing each of the selection criteria (Schwarz, Hannan-Quinn and Akaike) is reported following the run.

## 6.4 Actions / Automatic Regressor Selection

Given a baseline model containing N regressors of one or more Types (including current and lagged values of some variables, typically), this option estimates all the 2^N possible models formed by taking subsets of the full regressor set. It then reports the case which optimizes the chosen selection criterion, which can be either the Akaike (AIC), the Schwarz (SIC) or the Hannan-Quinn (HQC).

Notes:
1.  Optional settings are selected in Setup / Automatic Model Selection (regressor Types), which also features an alternate launch button

2.  The program excludes regressors from the baseline specification by turning their "Fixed" flags on, and setting the coefficient values to 0. In other words, the baseline specification is not actually changed during the run. To use the selected model for further work, it must be specified manually in the usual way. If the baseline specification includes regression parameters that are already "Fixed", this setting is retained and the variable in question is excluded from the search.

3.  The procedure optimizes exhaustively over the model space, looking at every case, and hence it is path independent and does not depend on test outcomes. Thus, it is not a "general to specific" search, nor "specific to general", but in fact both. The limitation is that it cannot feasibly examine such large model spaces some other automatic specification search packages that are currently available, which are selective in the cases they examine.

4.  The procedure can be used with any model except regime switching, including equation systems, although it only searches over regressor choices. Any other model features will correspond to the baseline specification, and the search is conditional on these. Note that the number of models needing to be estimated may be very large. For example, with 4 variables and 4 lags (+ current values) N = 20 and the choice is over 2^20 = 1048576 cases. This is feasible for OLS, but probably not for iterative estimation.

5.  There is no option to print all the estimation results. The output includes the baseline model, and the optimal case.


## 6.5 Actions / Recursive/Rolling Estimation

This option estimates the same model for a succession of samples, either of increasing size with fixed starting point (incremental estimation) or fixed size with moving starting point (rolling estimation). We call these procedures "recursive estimation", and while they are not recursive in the strict sense of using an updating formula to modify the estimator with the extra observation at each step, (it is computed afresh) the current estimates provide starting values for each new step.

© James Davidson 2024

See the Setup / Recursive/Rolling Estimation... dialog for details of setting up the run.

The sequence of results for each sample period can be exported to a file, according to the setting in Options / Output and Retrieval Options, or viewed as graphics. The items available are the parameter estimates with 2-standard error bands, plus (optionally) regression statistics including criterion values, residual moments, autocorrelation tests, stability (ex-post forecast) tests, and also any tests of parameter restrictions that have been set up.

Notes:
1.      The sequential model selection criteria are not recorded. These would differ from the estimation criterion by the same fixed amount in each run, so the latter contains all the relevant information.

2       If N-step-ahead ex ante forecasting is specified (see Options / Forecasting and Simulation) the Nth forecast step and standard error (or percentiles for Monte Carlo forecasting) are reported in the output. By default, N is fixed so that the actual date forecast moves forward with the sample. There is an option to keep the terminal forecast date fixed, so that N contracts as the end of the sample advances. See Setup / Recursive/Rolling Estimation for details.

3.      This command is launched by the "Run" button on the toolbar if the Setup / Recursive/Rolling Estimation dialog is open.

4.      On checking the "Report Convergence Status" box, a sequence of symbols is printed in the results window to show the result of each numerical optimization.
      +      denotes strong convergence of the search algorithm
      ?      denotes weak convergence of the search algorithm
      !      denotes failure to converge
(See the Ox documentation for the definitions of these cases.)
This option has no effect in the case of non-iterative estimation.

## 6.6 Actions / Plot Criterion Grid

Use this option to create a contour plot of the concentrated criterion function. "Parameter Bounds / Grid Plotting" must be enabled in the Values / Equation dialog.

The choice of parameters for plotting and the bounds of the grid are set in the appropriate Values dialog(s). The operation is carried out on the first one, or two, parameters satisfying the following conditions:
1.      the "fixed" flag is set;
2.      the upper bound exceeds the lower bound.
The criterion function is maximized with respect to all other parameters, holding the selected one(s) at the grid values. Normal output is suppressed. A progress bar shows the progress of the calculations.

The number of grid points can be set in a text box in either in Options / Optimization and Run, or in Values / Equation.

The grid of criterion values can be
* Printed in the results window:     Check "Criterion Grid" in Options /
                                     Output and Retrieval / Print in Results
                                     Window, before running the option.


* Exported to a spreadsheet:         Select File / Listings / Export
                                     Spreadsheet / Criterion Grid after
                                     running the option.


* Plotted on a graph:                Select Graphics / Criterion Plot after
                                     running the option.


Notes:
1.    Other parameters can be either fixed or free for this operation. If a pair
      of grid bounds for a fixed parameter are both set to zero, the parameter
      is treated as fixed, in the usual way.

2.    A choice of contour plot and surface plot is available for the two-
      dimensional case. Select the option  required in Options / Graphics.

3.    Default output to the results window includes the grid location of the
      criterion maximum, and locations of the largest values of the test
      statistics including absolute t ratios, diagnostic tests and any specified
      tests.

4.    Remember that the criterion is optimized (# Grid Points)^2 times
      in the case of a two dimensional grid. Choice of the number of
      points may need to depend on considerations of computing time.

5.    This option can be used to implement tests where parameters are
      unidentified under the null hypothesis. The usual procedure is to base
      a test on the supremum of the statistic over eligible values of the
      unidentified parameter(s)

## *6.7 Actions / Multi-Stage GMM*

This command allows the computation of efficient GMM estimates, where the
weights matrix is estimated to allow for heteroscedastic or autocorrelated errors.
This is a two- or multi-stage procedure, performed by issuing this command
repeatedly. Proceed as follows:

*     Select the covariance formula required ('Robust' or 'HAC' in Options \
      Tests and Diagnostics.)

*     Select GMM as the estimator in Setup / Estimators and Sample.

*     Formulate the equation(s) in Model / Dynamic Equation.

*     Select the instruments in Model / Select Instruments

© James Davidson 2024

\*       To compute the first-stage estimates (equivalent to regular IV/GMM) give the Actions / Multi-Stage GMM command.

\*       To compute the second-stage estimates, using the first-stage estimates to construct the required residuals, give the Actions / Multi-Stage GMM command a second time, before any other command.

\*       To iterate the procedure, repeat the last command as often as required.

Notes
1.      Giving any other estimation command (for example, Actions / Run Estimation, or the "Run" button) resets the program, so that the next Actions / Multi-Stage GMM command yields Stage 1 again.

2.      It may be possible to iterate the procedure to "convergence", so that no further change in the estimates occurs. However, this procedure does not optimize the GMM criterion function.

3.      This option is not available for IV estimation from the Model / Linear Regression dialog.


## 6.8 Actions / Compute Test Statistic

Opens a sub-menu listing test options
\*       Specify Score (LM) Test…
\*       Specify Moment (M or CM) Test…
\*       LM Tests of Parameter Restrictions
\*       Wald Test of Set Restrictions
\*       Specified Diagnostic Tests
\*       Bootstrap Test of I(0)

Notes
1.      By default, the statistics printed by selecting these options are asymptotically Chi-squared with degrees of freedom indicated in parentheses. The nominal p-values of the tests are shown, enclosed in braces, following the statistic values.

2.      The option of printing the "F form" of the tests, with associated p-values from the F tabulation, can be selected with the checkboxes in the relevant dialogs.
        Note:   In most cases the F form is to provide an alternative approximation to the asymptotic Chi-squared distribution. These tests are exact only in special cases.

SCORE (LM) TEST
3.      This option opens a dialog to allow specification of a test of the significance of additional regressors selected from the data set. Radio buttons allow selection of location of the test variables, either the Conditional Mean model (Equation (1), included as variables of Types 1, 2 or 3) or the Conditional Variance model (Equations (2)/(3), included as variables of Types 1, 2 or 3).

© James Davidson 2024

4.    If the selected Type is assigned to include one or more lags (see Model /
Dynamic Equation or  Model / Linear Regression, and Model / Conditional
Variance) then the same number of lags of the test variable(s) are included in
the test set. It is not necessary for the null (estimated) model to contain lags.
Just set the number of lags required (for the relevant Type) with the Lags
scrollbar in Model / Conditional Mean or Model / Conditional Variance.
The degrees of freedom of the test are 'number of test variables x (1+ number
of lags)'.  Select the "F form" of the test with the checkbox.

5.    There are two ways to run the test. Pressing "Go" will execute it immediately
using the currently stored parameter values. Otherwise, the test will be
performed following the next estimation run. To cancel the test, press "Clear".

MOMENT (M or CM) TEST

6.    This options opens a dialog to allow specification of a moment test of the
correlation of variables selected from the data set with either
*        the model residuals, or
*        the squared model residuals.

7.    If the Conditional Moment Test option is selected, the joint covariance matrix
of scores and moments is computed using the Robust formula (see Options /
Tests and Diagnostics). If the Moment Test option is selected, the selected
HAC formula is used to compute this covariance matrix. This selection
implies that serial correlation is permitted under the null hypothesis.
Select the "F form" of the test with the checkbox.

6.    Perform the test immediately with the current values by pressing "Go".
Otherwise it will be performed following the next estimation run. Cancel it by
pressing "Clear".

LM TESTS OF PARAMETER RESTRICTIONS

7.    By default, this option computes the LM test of the "fixed parameter"
restrictions in the current model, if any are set.
CAUTION: this test requires that the stored parameter values have been
estimated subject to the set restrictions. Otherwise it is invalid! It is
safer to run the test by checking the "LM Tests of Parameter Restrictions"
checkbox  in Options / Tests and Diagnostics, and then run the estimation
of the model. The test is then reported automatically.

8.    If the model has been estimated subject to parameter constraints set up
through the Model / Parameter Constraints and Values dialogs, (select
the option "Restricted Estimation") then the LM test is performed on these
constraints, ignoring any fixes. This allows the option to be used to test
restrictions while other parameter fixes are in force.
Note: Parameters cannot be simultaneously fixed and subject to
constraints. Uncheck any fixes on "constrained" parameters, before
running the test

WALD TEST OF SET RESTRICTIONS

9.    This option does not open a dialog, but simply computes the test of the

© James Davidson 2024

constraints currently specified in Model / Constraints and the Values dialogs. The option allows new constraints to be tested on the current estimated model without re-estimation. If no constraints are specified, it does nothing.

SPECIFIED DIAGNOSTIC TESTS

10.    This command evaluates the optional tests that have been specified in the Options / Tests and Diagnostics dialog. These are computed automatically following an estimation or evaluation run, but this method allows additional tests to be computed without generating the complete estimation outputs. This command also runs the raw series diagnostics option; see 6.1 for details.

TEST FOR I(0) DEPENDENT VARIABLE

11.    This option simulates the model currently stored and computes Breitung's (2002) statistic for the partial sums. The reported Kolmogorov-Smirnov statistic compares the bootstrap distribution of this statistic with the same statistic computed for the corresponding i.i.d. shocks. Its object is to check whether the model dynamics are compatible with sample averages adequately approximating their asymptotic distributions, in the selected sample size.

12.    The I(0) test is for single equation models only, and is properly intended for use with univariate time series models. If the equation contains exogenous variables, these are held conditionally fixed in the simulations. If they are nonstationary, this will in general yield a case of the alternative hypothesis.

13.    Choose a number of replications large enough to validate the asymptotic tables from Smirnov (1948) - say 500.  Set the number of replications in Options / Simulation and Resampling, after checking the "Inference by Resampling" checkbox to activate the replications scrollbar.
NOTE: This checkbox does not activate the test itself. It can be unchecked again after the setting is made.

## 6.9 Actions / Compute Forecasts & IRs

These commands generate and display either ex-ante forecasts or impulse responses after a model has been estimated. Specify the outputs required in Options / Forecasts.

Notes:
1.    If outputs are specified prior to giving the 'Run Estimation' or  'Evaluate at Current Values' commands, they are computed automatically as part of the run and printed along with the other outputs. These options allow results to be generated, according to the current settings, without re-running the estimation.

2.    "Ex ante forecasts" are generated using only exogenous variables and observations pre-dating the forecast period. Multi-step forecasts are computed recursively. Closed models, including models depending on deterministic trends, can be forecast beyond the range of the available data set.

3.  Tabular outputs are printed in the results window if this option is selected in Options / Output and Retrieval..., "Print in Results Window". Caution: This option is selected by default. If it is deselected, these commands may appear to have no effect!

4.  Plots are displayed automatically as the commands are run, if this option is selected in Options / Forecasting... They can also be generated subsequently with the commands Graphics / Ex Ante Forecasts > and Graphics / Impulse Responses >. These commands are disabled (greyed out) if no outputs are available for display.

## 6.10 Actions / Retrieve Series

This command adds items such as residuals, fitted values, conditional variances, switch probabilities and other generated series to the data set. To use it, first open the Options / Output and Retrieval dialog and check the items to be retrieved.

Notes:
1.  A fitted model must be present in memory when the command is given, or else it has no effect. Either an estimation run has previously been performed, or a previously stored model has been loaded using the Model Manager.

2.  When the Retrieve command is given, the checks in the Options / Output and Retrieval dialog are removed so that the same items are not saved twice.

3.  Items can be retrieved automatically following a run by checking them in the Options dialog before launching the run. Choose from the "Once" and "Always" options in this case. The former removes the checks after saving.

## 6.11 Actions / Simulate Current Model

This selection generates and displays a stochastic simulation of the current model. The user is prompted to choose to discard the series after viewing or write to the data set and, in the latter case, to either add the series as new variable(s), or over-write the existing dependent variable(s).

Notes:
1.  To select the mechanism generating the artificial shocks, see Options / Simulation and Resampling Options. The choices are Gaussian, model (according to likelihood function adopted) and a bootstrap of the current model residuals.

2.  Create  model for simulation in one of three ways: estimate a model, retrieve a stored model, or type parameter values direct into the Values dialogs.

3.  A dependent variable (or variables) for the simulated model must exist in the data set, and provides fixed pre-sample values when these are specified in a dynamic model. This series can be optionally over-written by the simulation, or the latter can be appended to the data set under a new name. In the latter case, the conditional variance series is also appended if this is generated in (e.g.) a GARCH or Markov-switching model.

© James Davidson 2024

4. To bypass the discard/save prompts, assign this command to one of the two user-selectable toolbar buttons (see Options / General). When the button is pressed most recent choices are applied automatically. Use the menu command to change them.

5. To simulate without a data set loaded, first use the 'Extend Sample' option in Setup / Data Transformations and Editing to specify the sample size required Then create a placeholder variable using 'Make Zeros'.

6. If estimation is by least squares or GMM, the residual variance is not in the parameter set, and is estimated extraneously from the residuals. If simulation follows estimation, the current estimate is used. To specify a model for simulation without prior estimation, select one of the conditional maximum likelihood (time domain) estimators. This will allow all the model parameters, including the error variance, to be entered in the Values dialogs.

7. It is possible to simulate equations in which the generated random shock enters nonlinearly, using the Coded Equation option. See Model / Coded Equations and General / Entering Formulae for details. Such equations cannot be estimated in this form, but must be coded separately.

8. A related option implemented through coded equations is to replace one or more data series by columns of standard normal random drawings before a simulation is run. This facility allows simulation experiments under "random regressor" assumptions without formulating a multi-equation model. See Model / Coded Equations and General / Entering Formulae for details.

PANEL DATA

9. To simulate panel data, a formatted panel data set must be loaded so that the program is in "panel data mode". Shocks can be generated either as bootstrapped residuals or as Gaussian pseudo-random variates. In the latter case, the "within" and "between" variances are set in the Model / Panel Data dialog. Random effects are produced by setting the "between" variance to a positive value.
Note: the "within" variance" setting overrides the variance setting in Options / Simulation and Resampling.

10. In the residual bootstrap option, individuals are drawn randomly with equal probability, and then the dated observations are resampled from the residuals for the selected individual, using the resampling scheme selected in Options / Simulation and Resampling.

11. The Data Transformation settings in Model / Panel Data are ignored in simulations. The actual data are simulated, using the actual exogenous variables.

12. Systems of panel equations can be simulated. However, if the Gaussian shocks option is adopted, in this release all equations have the same variances, as specified in Model / Panel Data, and zero covariances.

© James Davidson 2024

## 6.12 Other Commands

SET DEFAULT VALUES

This command sets all free parameter values to defaults. Fixed values, and all fixes, bounds and constraint settings, are retained.

CLEAR VALUES

This command has the same effect as pressing the Clear button in all the Values dialogs. Parameter values are set to defaults, and all fixes (except defaults), bounds, and constraints are removed.

CLEAR RESULTS AND SPECIFICATIONS

Clears the current model specifications, values, results, listings and graphics from memory. Stored models and associated outputs are retained. To delete these as well, use the command File / Settings /Clear all Settings

CLEAR RESULTS WINDOW

If the results buffer becomes full and will accept no more text, use this command to clear it. Note that its contents can first be first saved to a file, see File / Results / Save Selected Text.

SEARCH RESULTS WINDOW

A text window opens. Enter the target string and choose the search direction, from the current cursor position, with "Search >>>" or "<<< Search" as appropriate.

CLOSE ALL DIALOGS

This command is given automatically before estimation operations are carried out. Current window positions are stored.

RESTORE DIALOGS

Reverses the action of Close All Dialogs. Use it after an estimation run, and also after starting the program, to restore the previous window configuration. This action can optionally be made automatic, see Options / General Options.

RESET DIALOG POSITIONS

Resets the stored horizontal and vertical screen positions of all dialogs to (0,0), relative to the main window. This command is useful in case dialogs get "lost" from the visible screen area, which can happen if program settings are transferred between machines with different screen resolutions. This command cancels the setting "Remember Dialog Positions" in Options / General.

Notes:
1.  The actions of Close All Dialogs and Restore Dialogs are duplicated by the toolbar "Windows" button, which toggles between them.

2.  Dialog positions for open dialogs are saved between runs. Restore Dialogs (or the "Windows" button) restores your dialog configuration at shut-down.

3.  Speed up multiple text searches by assigning a toolbar button to the "Search Results Window" menu item -  see Options / General / Assign Toolbar Buttons. Searches are case-insensitive, but targets may not

contain line breaks.

# 7. Graphics

## 7.0 Graphics: General Information

Graphics are displayed in a window whose size is linked to screen resolution. As many graphs as desired can be open at once. Close them by clicking the "Close" button in the top-right hand corner, or from the Windows taskbar.

* The Gnuplot window offers a number of options. Options available as toolbar icons in Gnuplot 5 include copying the image to the Windows clipboard, sending to a printer, and saving in Windows EMF format. Choose "Options" to open a submenu of further useful commands. These include changing the text font, line style and background colour. In Gnuplot 4.2, click on the top-left corner icon to open a context menu with similar options.

* Saving to the clipboard allows the image to be pasted directly into (e.g.) Microsoft Word, or to any graphics manipulation package. For bitmaps, IrfanView is a highly recommended freeware package, which can convert an imported image to a variety of bitmap formats such as .gif, .png and .jpg. The Gnuplot window can be resized by dragging the corner with the mouse, and a saved bitmap image will have the same dimensions. Alternatively, pasting the clipboard into a graphics program such as CorelDraw results in a vector graphics image that can be edited further, and saved in various formats.

* An alternative procedure to save a graph is with the command File / Graphics / Save Last Graphic. Select the type of file to be saved in Options / Output and Retrieval. PNG bitmaps are saved in 640x480 pixel resolution. Vector graphics formats (EPS or EMF) can be imported into various graphics programs for further editing. The PNG and EPS formats have a smaller range of colours than the Windows terminal, and the colours may not match those selected. Desired line and text characteristics can be selected in Options / Graphics, but whether these selections are effective can depend on the format selected.

* Given knowledge of the Gnuplot command language, the text files with .PLT extension created by TSM can be edited by the user for more precise control over the image. These files are deleted by default but keeping them is an option; see Options / Graphics.

## 7.1 Graphics / Show Data Graphic

Opens a dialog for selection of graphical representations (time plot, correlogram, partial correlogram, spectrum, histogram and normal QQ-plot) of any variable in the data set.

SIMPLE PLOTTING
* Select the type(s) of plot required, using the check boxes.

\* Select one or more names in the variable list.

\* Optionally select for the series to be either detrended by OLS, or differenced, before plotting.

\* Press "Sample ..." to open the Set Sample dialog and choose the sample required.

\* Press "<<<  Plot  >>>". Alternatively, just double-click the name of any series (convenient  for a single plot).

\* Press "List" to display the the currently selected sample in spreadsheet format.

Notes:

1. Unselect a selected variable by clicking it second time. Press 'Clear' to unselect all selected variables.

2. Double-clicking a variable on the list just displays the plot. Selecting with a single click and then pressing 'Go' plots the variable(s) and leaves them selected.

3. Three modes of display are available for time plots, correlograms and partial correlograms, spectra and QQ-plots. The mode is selectable by the radio buttons in the Options / Graphics dialog:
   \* "Individual": each variable plotted in its own frame.
   \* "One Frame": multiple individual graphs in one frame.
   \* "One Graph": multiple series plots overlaid in one graph, with common axes.

4. In "Individual" mode, multiple representations of a series (e.g. time plot, correlogram, etc.) are still displayed in one frame. Draw these separately to get them in their own frames.

5. In  "One Frame" mode, up to 36 graphs can be displayed in one frame. However, with a large numbers of plots the quality of the display on a standard monitor is not too good.

6. In "One Graph" mode, line colours and styles are assigned according to the line style selections in Options / Graphics, variables being ordered by their position in the list. For more flexible assignment of styles,  see "Multi-series Plotting"  below. This mode is not available for Histogram/ Density plots, if these are selected the mode defaults to "One Frame".

7. If either the Scatter Plot or Bivariate Density options are selected, the first two highlighted variables in the list are plotted together - the first on the list is assigned the 'vertical' axis, (Y) and the second the 'horizontal' axis (X). Reverse the ordering, if desired, in Setup / Data Transformation and Editing / Edit / Move Up-Move Down. If only one variable is highlighted with these options, nothing is plotted. Other options can be selected at the same time, and will be displayed for all highlighted variables.

8.    By default, T/2 correlogram and partial correlogram points are plotted, where T is the length of the selected plotting sample. Optionally, the number of points to be plotted can be set in the Setup / Compute Summary Statistics dialog. The effective number of plotted points is the minimum of T/2 and this selection, provided if it is greater than zero and the option is selected.
        NOTE: Only the correlogram order is selected by this option. To plot the actual points listed in  Setup / Compute Summary Statistics, the samples selected in the two dialogs must match!

9.    In a scatter plot, the two regression lines (Y on X and X on Y) are superimposed by default. Note that these both pass through and jointly indicate the point of sample means. The more nearly parallel these are, the higher the correlation between the variables. Unselect this feature in Options / Graphics.

10.    Plotting options for the bivariate density can be selected in Options / Graphics. The options that can be selected are
        *    Contour/Surface/Both in "3D Plot Type",
        *    Density, Histogram and Normal in "Distribution Plot Options".

11.    The sample of observations selected for plotting (Sample 5) is independent of the samples set concurrently for other program functions. See Setup / Set Sample for details. If data are missing for part of the specified period, nothing is plotted for these observations.

MULTI_SERIES PLOTTING

This option allows up to 8 series to be plotted on one graph with line styles selected by the user. All the series representations except Histogram/Density and the bivariate plots can be drawn using this option. It gives the user more control over plotting styles, and allows multiple series plotting without changing the basic setting in Options / Graphics.

*    Check the "Multi-Series Plot" Checkbox.

*    Choose a line style (1-8) using the radio buttons, and then a variable from the list - at most one variable can be selected for each style.

*     Repeat for up to 8 series.

*    Press 'Go'.

Notes:
13.    The 'Centre' and 'Standardize' checkboxes allow the variables to be relocated and rescaled for comparability. Centre subtracts the mean, and Standardize also divides by the standard deviation.

14.    The 'RH Scale' checkbox is enabled when two or more series are selected. If checked, the series with the largest line number is plotted against the right-hand axis. If more than two series are selected, all the

© James Davidson 2024

others are plotted against the left-hand axis as usual.

15. The 'C-Bands' checkbox enables the option of drawing 2SE confidence bands around a series, based on a companion series of standard errors. Either one or two series can have confidence bands. These must be assigned to positions 1 and 3, and the companion series must be assigned to positions 2 and 4 respectively. Positions 5 to 8 can be assigned to additional series in the usual way.
Note: If the designated 'standard error' series contain negative values, an error message is shown.

16. If the line type for Line 8 is chosen as 'Band Fill', the selected variable is used to divide the sample into 'regimes'. If the selected variable is positive the regime is "On", and if the selected variable is zero or negative the regime is "Off". The observations of the "On" regime are indicated by light gray shading in the plot. Typically, the indicator variable would be a zero-one dummy.
NOTE: Only Line 8 can be given the Band Fill style.

LISTING

17. The 'List' button invokes the data spreadsheet that is also accessible from Setup / Data Transformation and Editing, although here editing functions are suspended. It simply allows a view of the values of the plotted points. The sample listed matches the plotted sample.

## *7.2 Graphics / Equation Graphics*

The following items can be plotted following an estimation run:
* 'Actual and fitted' series (submenu, time and scatter plots)
* Residuals (sub-menu)
* Variance-adjusted residuals (sub-menu)
* Conditional variances (GARCH and regime switching models only)
* Regime probabilities (sub-menu, regime switching models only)
* All the time plots for the model in one frame.
* Equilibrium relations (sub-menu, error correction models only)
* Ex ante forecasts (sub-menu)
* Solved moving average coefficients, impulse response and step response (submenu)
* A 2D or 3D plot of the criterion function if computed, see Actions / Plot Criterion Grid.
* Plots of the recursive or rolling estimations, if computed (dialog)
* Kernel density plots of the bootstrap distributions of parameter estimates and test statistics.

Notes:
1. In the Residuals, Adjusted Residuals and Equilibrium Relations submenus, the available options are time plot, correlogram, spectrum, correlogram and spectrum of absolute values, histogram and kernel density, and normal QQ-plot. There is an option to display all the representations in the same frame.

2. The Actual vs. Fitted scatter plot includes the "Fitted vs. Fitted" line, which has slope 1 by construction and passes through the point of sample means. The residuals correspond to the vertical deviations around this line, but ordered by magnitude of the fitted values instead of the sample time ordering.

3. Regime probabilities are summed over lags in Hamilton's switching model. Smoothed regime probabilities are computed by Kim's algorithm, see Kim and Nelson (1999), pp68-70. When switching probabilities are explained by exogenous variables, the variable transition probabilities can be plotted; more precisely, the plotted series are the simple averages over i of the probability of jumping from regime i to regime j, at time t. For smooth transition models, the transition function can be plotted.

4. Forecasts and solved MA weights can be computed for the equation and the conditional variance, where this is estimated.

5. The interpretation of ex-ante forecast time plots depends on the forecasting method. With analytic forecasting, the plots show the estimated mean forecast, with optional 2-standard error bands. Standard errors are not available for bilinear models and user-supplied dynamic models, nor for conditional variance forecasts in GARCH models.

6. Monte Carlo forecasting generates the distribution of stochastic simulations, and the plots show the median forecast and the 2.5% and 97.5% percentiles. These therefore define a true 95% confidence band, regardless of the error distribution, and are available for all models, and for conditional variances. Kernel density plots of the empirical forecast distributions are also available for any point in the forecast period (selected in Options / Forecast and Simulation Options).

7. To change the type or number of forecasts, make the desired selection in the Options / Forecast and Simulation dialog, then recompute the model, using either Actions / Run Estimation, or Actions / Evaluate at Current Values to avoid re-estimating.

8. To view a plot of the recursive/rolling estimations, either double-click a parameter/statistic in the list, or select it and press "Go". The rolling parameter estimates are displayed with 2-standard error bands (the usual regression standard errors, with covariance matrix formula selected in Options / Tests and Diagnostics.)

9. The 'Fill' symbol option is used for probability and variance plots. That is, the region between the plotted line and the time axis is shaded light gray. If for any reason a different plot style is required for these series, use the "Retrieve Series" command to add them to the data set, then plot using the options in Graphics / Show Data Graphic.

10. Select the type of plot desired for confidence and standard error bands in forecasts and recursive/rolling estimation (none, bands or bars) in Options / Graphics Options.

## 7.3 Graphics / Monte Carlo Graphics

The frequency distributions generated by Monte Carlo experiments can be displayed as kernel densities and histograms. Cumulative frequencies can also be displayed as an option.

* Select plots for display by clicking on the list, or pressing 'Select All'. Unselect an item by clicking again. Remove all selections with 'Clear'. Plot selected items by pressing 'Plot'.

* Double-clicking a list item displays the plot, but does not leave the item highlighted.

1. The graphic display dialog opens automatically at the completion of a Monte Carlo run. Alternatively, open it from the menu with Graphics / Monte Carlo Distributions...

2. The bandwidth used to compute the kernel density is user-selectable. Initially it is set to the default value, displayed as '1' on the scrollbar. Move the scrollbar to change the ratio of the bandwidth to the default, then redisplay the plot to see the effect. The chosen value is stored with the distribution data and used the next time the distribution is plotted, as shown by the scrollbar.

3. If the scrollbar setting is changed by the user, the new value is used for the next plot, and replaces the stored value(s) for the distribution(s) plotted.
Tip: to avoid changing a stored bandwidth after moving the scrollbar, close and re-open the dialog before giving the Plot command.

4. Monte Carlo graphics are stored in the .TSD file created with the Data Generation Model. If the .TSD file is present, reloading this model makes the results and plots of the last experiment performed with this model available to view, just as when the model was run. To save these outputs separately from the model, give the command File / Listings / Save Listings File, and enter a file name in the file dialog.

5. The distributions from a simulation run can be stored in an EDF file to supply critical values for tests. The currently stored bandwidths for each distribution are used to compute the EDF kernel densities.

   Note: EDF file creation can be done automatically by setting the option in Setup / Monte Carlo Experiment. This latter method uses the last-selected bandwidth for all distributions. Use the present option to fine-tune the bandwidths used for each distribution, if required.

6. Monte Carlo graphics can be stored and displayed independently of the DGM that produced them. This allows the results of different runs to be

plotted together for comparison.
* Highlight the plots to be stored.
* Click on "Add to Store".
* In the text box that opens, the current name is displayed. Edit or replace this name to identify the plot in a comparison of DGMs.
* Click "OK", or press the Return key, to close the text box.
* Repeat the naming procedure for each plot selected.

Note: The stored plots are saved in the settings file, not in the .TSD file associated with a model.

7. To display the stored plots, select Graphics / Stored MC Distributions ... from the menu.
* Select and display plots as usual
* Highlight a name and click "Rename" to re-open the text box and type a new name.
  Note: Only one plot at a time can be renamed.
* Highlight any number of names and click "Remove" to remove the graphics from the store.
* Highlight one name and click "Move Up" or "Move Dn" to change the order in which plots are displayed

Note: The bandwidths for stored plots can be changed in the same way as described in 3.

8. To show two or more density plots in the same graph for comparison select the Options / Graphics from the menu and then "One Graph" in Multiple Series Display. The MoveUp/Dn buttons allow control over the colours or line types used to display each plot, also the order in which legends appear.

Note: Histograms are not displayed for two or more densities in the same graph, even if this option is selected.

9. The saved plots can be exported to a spreadsheet file, for subsequent reloading and/or merging with those stored in a different settings file. Pressing the "Save Plot Data" button opens the file dialog for saving.

Notes:
* Reload plots with the command File / Data / Load Tabulations / Density Plot File.
* File names with the prefix "DNS_" are recommended. Such files can be loaded quickly using drag-and-drop.
* The spreadsheet format is as follows:
  Row 1:       Plot names.
  Row 2:       Kernel bandwidth index (integer 0-32)
  Row 3:       Initial bin ordinate.
  Row 4.       Increment (bin width).
  Rows 5,....  Histogram values.

# 8. Options

## 8.0 Options: General Information

By default, only one Options dialog is open at a time. Opening a second dialog causes the first to close. This behaviour is optional, see General Options.

The <<< and >>> buttons at the top of each window allow easy navigation between the options dialogs. Pressing them opens the next/previous options dialog, so that current program settings can be reviewed rapidly.

DEFAULTS

The default settings are as follows. These are selected if the file "settings.tsm" is not found in the working directory at start-up, or on the command File / Settings / New. "*" in the left-hand column indicates that this setting is stored with a model, and hence is liable be changed when a model is loaded.

Output and Retrieval Options:

| | | |
|---|---|---|
| Save data to files of type | . . . | XLS |
| Save listings automatically. | . . . | No |
| Export graphs to files of type | . . . | PNG |
| Export graphics automatically | . . . | No |
| Retrieve generated series | . . . | Once |

Test and Diagnostics

| | | | |
|---|---|---|---|
| * | Correlogram points | . . . . | 0 |
| * | Covariance matrix formula | . . . | Robust |
| * | Kernel for HAC covariances | . . . | Parzen |
| * | Bandwidth for HAC covariances | . . | N-W plug-in |
| * | Show residual moments | . . . | Yes |
| * | Show log-likelihood & criteria | . . | Yes |
| * | Tests in R^2 form | . . . . | No |
| * | Report F statistics | . . . . | No |
| * | Use EDFs | . . . . . | No |
| * | Use HAC covariance matrix in tests | . . | Yes |
| * | Model selection criterion | . . . | Akaike |

Diagnostic Tests

| | | | |
|---|---|---|---|
| * | Q Test order | . . . . . | 12 |
| * | Q-Test type | . . . . . | Box-Pierce |
| * | LM tests | . . . . . | No |
| * | Moment tests | . . . . . | No |

Forecasting Options

| | | | |
|---|---|---|---|
| * | Forecast type | . . . . | Ex Ante Multi-step |
| * | Forecast method | . . . . | Analytic |
| * | # Actuals in forecast plot | . . . | 50 |
| * | # Monte Carlo replications | . . . | 1000 |
| * | Report median. | . . . . | Yes |

© James Davidson 2024

| | | | | | | |
|---|---|---|---|---|---|---|
| * | Compute forecast SEs . | . | . | . | Yes |
| * | Compute confidence bands | . | . | . | Yes |
| * | Impulse responses | . | . | . | Impulse |

Simulation and Resampling Options
| | | | | | | |
|---|---|---|---|---|---|---|
| * | Shock distribution | . | . | . | . | Gaussian |
| * | Random number seed . | . | . | . | 0 (system clock) |
| * | Shock variance (Gaussian) | . | . | . | 1 |
| * | Presample data | . | . | . | Fixed |
| * | type I fractional | . | . | . | No |
| * | Use likelihood model . | . | . | . | Yes |
| * | Enable bootstrap inference | . | . | . | No |
| | Once/always . | . | . | . | Once |
| * | Confidence intervals . | . | . | . | Equal Tails |
| * | Bootstrap replications | . | . | . | 99 |
| * | Fast double bootstrap | . | . | . | No |
| * | Bias correction | . | . | . | No |
| * | Newton-Raphson algorithm | . | . | . | No |
| * | Maximum iterations . | . | . | . | 50 |

Graphics Options:
| | | | | | |
|---|---|---|---|---|---|
| Font . | . | . | . | . | Arial |
| Text size | . | . | . | . | 10 |
| Dates format . | . | . | . | YY |
| Keep PLT files | . | . | . | No |
| Edit title | . | . | . | No |
| No title | . | . | . | No |
| Plot zero axes . | . | . | . | Yes |
| Gray background | . | . | . | No |
| Confidence interval style | . | . | . | Bars |
| Multiple series display . | . | . | One Frame |
| Legend placement | . | . | . | TL |
| Extended actual-fitted | . | . | . | No |
| Regression lines in scatter plots | . | . | Yes |
| 'Fill' probs and var plots | . | . | . | Yes |
| Criterion 3D Plot pype | . | . | . | Both |
| Distribution plot options: | | | | | |
| Density | . | . | . | . | Yes |
| Histogram | . | . | . | . | Yes |
| Normal Curve . | . | . | . | No |
| CDF . | . | . | . | . | No |

Line and Symbol Styles
| | | | | | |
|---|---|---|---|---|---|
| Plot type | . | . | . | . | Lines |
| Colour/monochrome . | . | . | . | Colour |
| Line width | . | . | . | . | 1 |
| Symbol size | . | . | . | . | 1 |

ML and Dynamic Options:
| | | | | | | |
|---|---|---|---|---|---|---|
| * | Root of ML variance . | . | . | . | 2 |
| * | Root of Student t d.f. . | . | . | . | 2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| * | Log of skewness parameter | . | . | . | Yes |
| * | Ordered logit/probit cases | . | . | . | 0 |
| * | AR latent process in discrete data | . | . | No |
| * | Zero-inflated logit/probit | . | . | . | No |
| * | Chi-squared probit | . | . | . | No |
| * | Exponential Poisson/negbin mean | . | . | No |
| * | Maximum ARMA/GARCH order | . | . | 10 |
| * | MA polynomial with '+'s | . | . | . | No |
| * | Restrict pre-sample lags | . | . | . | No |
| * | Compute roots | . | . | . | Yes |
| * | Type 1 fractional | . | . | . | 0 |
| * | Difference spec for FI and ECM | . | . | No |

GARCH Settings
| | | | | | | |
|---|---|---|---|---|---|---|
| * | Root of GARCH intercept | . | . | . | 2 |
| * | GARCH coeffs. in ARMA-in-squares form | . | Yes |
| * | Type 1 GARCH intercept | . | . | . | Yes |
| * | No. Gauss-Seidel its. in GARCH-M | . | . | 5 |
| * | GARCH-M trimming factor | . | . | . | 10 |
| * | Compute EGARCH Likelihood by G-S | . | No |
| * | No. G-S Its. to compute EGARCH | . | . | 50 |
| * | Omit EGARCH mean of abs. shocks | . | . | No |

Optimization and Run Options:
| | | | | | | |
|---|---|---|---|---|---|---|
| * | Report minimand | . | . | . | . | No |
| | No. grid points for criterion plot | . | . | 10 |
| * | Absolute bound on dynamic params. | . | . | 3 |
| * | Max no. of BFGS iterations | . | . | -1 = Ox Default |
| | Iteration print frequency | . | . | . | 0 |
| * | Strong convergence crit. | . | . | . | -1 = Ox Default |
| * | Weak convergence crit. | . | . | . | -1 = Ox Default |
| * | Parameter rescaling factor | . | . | . | 0 (= off). |
| * | Use numerical derivatives | . | . | . | No |
| * | Simulated annealing | . | . | . | Yes |
| * | # SA iterations | . | . | . | 500 |
| * | Initial temperature | . | . | . | 5 |
| * | Temperature reduction factor | . | . | 0.85 |
| * | # Iterations before temp. reduction | . | . | 10 |
| * | Restart on non-convergence | . | . | . | Yes |
| * | Skip estimation outputs | . | . | . | No |

General Options:
| | | | | | |
|---|---|---|---|---|---|
| Restore dialogs automatically | . | . | No |
| Remember dialog positions | . | . | . | No |
| Enable savings prompts | . | . | . | Yes |
| Automatically close options dialogs | . | Yes |
| Echo results to console | . | . | . | No |
| Show tool tips | . | . | . | . | Yes |
| Results window text size | . | . | . | 12pt |
| Font and icon size | . | . | . | 100% |
| Store model after each run | . | . | . | No |

| | |
|---|---|
| Save data with settings          .          .          . | Yes |
| Keep results window contents          .          .          . | No |
| Delete temporary files on exit .          .          . | No |
| Omit data descriptions on saving          .          . | No |
| Include user code in exported settings          .          . | No |
| Enable error recovery          .          .          . | Yes |
| Enable optimization estimators          .          . | Yes |
| Defer external jobs          .          .          .          . | No |
| Load multiple data sets          .          .          . | No |

## *8.1 Options / Output and Retrievals ...*

SAVING DATA FILES

1.    The type of data file saved (.XLS, .XLSX., CSV, .WKS, .IN7, .DAT or .MAT) is determined by the extension added to the file name. If no extension is added by the user, the one selected in this dialog is included automatically. However, note that the user's choice over-rides the selection in this dialog. For example, if the name is selected from the list of existing files in the file dialog, with extension .XLS, the selection here is ignored, if different.

EXPORTING AND PRINTING LISTINGS

2.    Select a radio button for the type of spreadsheet file or ASCII file to be saved by the command File / Listings / Export Spreadsheet. By default the files are identified by a descriptive name and the current run ID.

3.    Check the checkbox "Save Listings Automatically" to have all the series associated with an estimation run saved to a file of the selected type. If this box is not checked, specific outputs from the latest run can be saved individually as required, by selecting the menu items under File / Listings /.

4.    To distinguish sample and ex-post forecast periods in the output file, an extra column is printed with an indicator variable, 0 = sample, 1 = forecast.

5.    Check the boxes in "Print to Results Window" to have the items appear on the screen. Large screen listings are best avoided, since the results window buffer may fill up rapidly.

EXPORTING GRAPHICS

6.    Graphics can be saved to file in both bitmap and vector graphics formats. PNG is a standard bitmap format providing an exact image of the displayed window.
EPS (Encapsulated PostScript) and EMF (Windows Enhanced MetaFile) are vector graphics formats. Vector graphics can  be edited in drawing packages such as Corel Draw.
PNG,  EPS and EMF files can all be loaded directly into Word and TeX documents. They can also be viewed and converted into other formats, such as JPEG, using freeware packages such as the recommended IrfanView.

© James Davidson 2024

7.	If the checkbox "Export Graphs Automatically" is checked, every graph is saved to disk in the chosen format, at the same time it is displayed. Otherwise, the menu item File / Graphics / Save Last Graphic saves the graph most recently displayed.

8.	.GWG (GiveWin format) files are created if the compiler directive #DEFINE OXDRAW appears in the run file. This option loads the OxDraw graphics package, and disables GnuPlot, so display within the program is not available. The format settings in this dialog are ignored. See Appendix B for more details.

OUTPUTTING TABLES

9.	Items that can optionally be printed to the output include the forecasts and impulse responses, residual correlograms, the parameter covariance matrix, the gradient vector and its covariance matrix, and the Hessian matrix of the estimation criterion.

10.	Ex ante forecasts are tabulated together with either standard errors of forecast or 95% confidence bounds, depending on the option selected in Options / Forecasting. If ex ante forecasts are computed for periods for which some or all of the actual data are available in the data set, so that forecast errors are known, the root mean square of the forecast errors is shown beneath the table.

11.	The gradient and Hessian may have diagnostic uses.
	*	The gradient is displayed only for iterative estimators, in the cases where either the optimization algorithm fails to return the state "Strong Convergence", or the output is evaluated at the current point, without optimizing.
	*	The covariance matrix of the gradient is not computed when the standard covariance matrix formula is selected in Options / Tests and Diagnostics.
	*	In ordinary least squares the Hessian is evaluated as the matrix of sums of squares and products of the regressors divided by the residual variance, hence is positive definite. Criteria accessed in the Dynamic Equation dialog are maximands by default, and in that case hence the Hessian is negative definite.

12.	The grid of criterion values following a call to Actions / Plot Criterion Grid can be optionally printed to the results window, provided this checkbox is checked first. The results can also be displayed graphically and exported to a spreadsheet, and these latter options can be effected after running the estimation.

OUTPUTTING SERIES

13.	The following series can be either printed in the results window, or exported to a file.
	*	The actual data series, $Y\_t$.
	*	The fitted series $Yhat\_t$, followed by the ex post 1-step forecasts, if specified.
	*	The residuals $Y\_t - Yhat\_t$ (solving equation (1) for $u\_t$ at the

© James Davidson 2024

estimated parameters) followed by ex post 1-step forecast errors, if specified.

If conditional heteroscedasticity options are selected, also

*    The variance-adjusted residuals , followed by ex post 1-step adjusted forecast errors, if specified.

*    The estimated conditional variances, followed by ex post 1-step variance forecasts, if specified.

If a smooth transition model specified, also the transition weights.

If Markov-switching regimes are specified, also

*    The filter probabilities

*    The smoothed probabilities

*    For explained switching models, the transition probabilities for Regimes 1 to M-1.

In regime-switching models the residuals are computed as weighted averages of the regimes, using the filter probabilities as weights.

Note: In the case of the Probit model, generalized residuals replace variance-adjusted residuals.

RETRIEVING GENERATED SERIES

14.    "Retrieving" series means adding them to the current data set for further analysis in the program. The radio buttons allow you to choose whether to retrieve the next run only, and then cancel the option, or keep it selected. The series are allocated standard names, with numbers appended to distinguish successive retrievals. They can be renamed more descriptively in the Data Transformations dialog, if desired.

15.    Retrieving the series of log-likelihood contributions allows non-nested models to be compared using the Vuong (1989) test. Since the test requires estimates of two different models, it must be computed by hand. Use the Data Transformation and Editing dialog to compute the differences of the series for two comparable models, use Setup / Compute Summary Statistics to compute the mean and standard deviation of the differences. Finally use Setup / Calculator to compute the Vuong statistic from these components.

16.    Following a recursive or rolling estimation the sequences of estimates can be also retrieved to the data set. In this case the initial sample points, over which no estimates are available, are filled with NaNs.

## 8.2 Options / Tests and Diagnostics ...

CORRELOGRAM POINTS

The scroll bar controls the number of correlogram points to be computed when this option is selected in Options / Output and Retrieval. The upper limit is set at (number of observations)/3.

The first two checkboxes, which are checked by default, control the display of various standard statistics following an estimation run. Unchecking these boxes simplifies the estimation output, which may be appropriate for elementary

© James Davidson 2024

tutorial use.

LM TESTS OF SET RESTRICTIONS AND "FIXED" VALUES

Check the "LM Tests of Parameter Restrictions" checkbox to compute the score (LM) statistic for the validity of restrictions imposed on the parameters in the "Values" dialogs.

There are two modes of operation of this option.

* If parameter constraints have been activated in the Parameter Constraints dialog with the option "Restricted Estimation" selected, these restrictions are tested.
* If not parameter constraints are set, but parameters have been "fixed" in the Values dialogs, a test of the fixing restrictions is generated.

Notes:

1. In first case, parameter fixes are not tested. This allows testing one set of restrictions while simultaneously including other restrictions in the maintained hypothesis.

2. These tests can also be computed for a fitted model with the option Actions / Compute Test Statistics / LM Tests of Parameter Restrictions

3. This option is not available for linear regressions. Use Wald tests to test restrictions in this context.

HAC COVARIANCE MATRIX IN TESTS

A number of test statistics, especially the tests of weak dependence, depend on HAC estimates of the covariance matrix of scores or residuals. HAC is also an optional choice for the calculation of standard errors and test statistics via the residual covariance matrix. This option, set checked by default, makes the choice of HAC for the additional tests independent of the selection in 'Covariance Matrix Formulae' (see below). Uncheck this option to make all selections subject to the dialog setting.

FORM OF THE LM STATISTICS

The 'T.R^2 form' option gives a choice of small sample approximations for diagnostic LM tests on the mean model. The default is to report the statistic $W/(1 - W/T)$, where $W$ is the usual LM statistic, and $T$ is the sample size. If $W$ takes the usual form of $TR^2$ in the artificial regression of the residuals on the gradients under H0, the default is therefore $TR^2/(1 - R^2)$. In a linear model, this is equivalent to normalizing the statistic with the unrestricted residual variance under the alternative. The asymptotic distributions are the same under H0. Check the option to report $W$ itself. When the 'F-form' of the statistic is selected, this option is unavailable.

LAGS FOR THE ADF AND ERS TESTS OF I(1)

These tests for the integration order of series are accessible in the Setup / Compute Summary Statistics dialog. The number of lags to compute the statistic can be chosen either manually using the scrollbar, or automatically using a model selection criterion selected in the right-hand panel of this

dialog.

4.      The 'ADF Lags' scrollbar is disabled unless the "No Criterion"
        option is selected.

5.      The ADF lags can also be set in the Model / Linear Regression
        dialog, when the "Cointegrating Regression" option is selected.
        In this context, the setting is applied to the ADF-based test of
        cointegration as well as the test of I(1). Note that the setting in
        either dialog is mirrored in the other - the same internal counter
        is set in either location.

DIAGNOSTIC TESTS

These options are presented in a separate dialog, accessible from this dialog
and also the Linear Regression and Dynamic Equation dialogs. The details
are given in Section 8.2a of this manual.

COVARIANCE MATRIX FORMULAE

Four methods of calculation are available for computing covariance matrices:
*       'Standard' corresponds to the information matrix in ML estimators, and is
        generally appropriate for correctly specified models with i.i.d. errors.
*       'Robust' allows for mis-specified likelihoods and heteroscedastic but
        independent errors.
*       'HAC' is for models with residual autocorrelation as well as possible
        heteroscedasticity.
*       'KVB' selects the inconsistent estimator proposed by Kiefer, Vogelsang
        and Bunzel (2000).

Notes:
6.       The method selected here is used for equation standard errors and also
        Wald and LM tests. M statistics use HAC, and nonparametric tests of I(0)
        (see Actions / Compute Summary Statistics) use HAC. Whittle estimates
        use HAC if KVB is selected.

7.      The KVB estimates correspond to HAC using the Bartlett kernel and
        bandwidth equal to sample size. Note that they are inconsistent, but yield
        asymptotically pivotal tests based on a nonstandard distribution. P-value
        inequalities are reported from the published tables, and are identified in the
        output with the @ symbol .
        NOTE: KVB can be implemented for nonlinear models and GMM at the
        user's discretion, but take care to note that asymptotic distribution results
        for these cases are not available to date. See the citations in Models and
        Methods for further details.

HAC KERNEL FORMULAE AND BANDWIDTH

The kernel formula selected here is used when the HAC covariance matrix is
selected. It also determines the choice for Phillips-Perron, KPSS and RS tests, and
Fully Modified Least Squares estimation. It does not effect the kernels used in
density
plotting and nonparametric regression, which are Ox defaults and cannot be changed
by the user.

© James Davidson 2024

There are three options for selecting the HAC bandwidth.

*   Manual selection using the scroll bar
*   The automatic 'plug-in' method due to Andrews (1991)
*   The automatic 'plug-in' method of Newey and West (1994).

The plug-in bandwidths are set as the minimum of the value returned by the formula and (sample size)/2. Larger bandwidths can be set manually. Check the 'Prewhitening' checkbox to prewhiten the data by a VAR(1) regression, prior to computing the HAC covariance matrix, V. V then is 'recoloured' as DVD' where $D = (I - A)^{-1}$ and A is the VAR coefficient matrix.

NOTE: Use prewhitening with caution! It may inappropriate, e.g., in construction of test statistics for I(0), by overcompensating under the alternative.

MODEL SELECTION CRITERIA

A model selection criterion is needed for three automated model selection features.

*   Automatic regressor selection (see the Actions menu and Setup / Automatic Model Selection).
*   Lag selection in the augmented Dickey Fuller test.
*   Lag selection in the Stock-Watson/Saikkonen augmented least squares estimator for cointegrating relations.

The radio buttons allow a choice of the Akaike, Schwarz and Hannan-Quinn criteria for these purposes.

SUPPLYING AN EDF FILE FOR P-VALUES

Check the box "Use EDF from File for p-Values" if tabulations of test distributions has been created by Monte Carlo simulation of the null hypothesis. The currently loaded EDF file is used to supply p-values for each test for which a tabulation exists in the file. In the case of significance tests (t-tests) the parameter's name must match a name in the file.

Notes:

8.   P-values taken from the EDF are marked with a '$' in the output.

9.   See Appendix G for details of the EDF file format.

10.   EDF files can be read and plotted in Setup / Look Up Tail Probability and Setup / Look Up Critical Value. To read off a p-value (for a 1-tail test) type the test statistic into the Critical Value field in the 'Tail Probability' dialog.

11.   EDF files are spreadsheet files. They can be loaded as data files, inspected and merged in TSM.

12.   If tabulations for several sample sizes are created, these can be merged into a single EDF file. Use the command File / Data / Tabulations / Merge EDF File to merge a file on disk with the one resident in memory. These tables will be used to calculate p-values by interpolation.
For example, suppose tabulations for T = 50, 100, 200 and 500 observations have been created and merged. If the actual T is 150, a weighted average of the tabulated p-values is returned, with weight
$$w = ((200 - 150)/(200 - 100))^2 = 0.25$$

     © James Davidson 2024

on the case T = 100, and 1 - w = 0.75 on the case T = 200
(i.e., greater weight on the larger sample size than with a linear interpolation).
 If T exceeds the largest tabulated sample size or is less than the smallest,
then the nearest values are used with weight 1.

13. Tests of the unit root hypothesis (ADF, PP and ERS tests) have critical
regions in the lower tail by convention. However, the rule of rejection in
the upper tail is adopted for EDF critical values.  Note, to accommodate
this requirement the EDF tabulations of these statistics are reversed.
Calculations in Setup / Look Up Tail Probability and Setup / Look Up
Critical Value are adjusted likewise.

14. Bootstrap tests yield a p-value which, when the test is correctly sized,
is uniformly distributed on [0,1] under the null hypothesis. A Monte Carlo
experiment can tabulate the p-values for a bootstrap test under the null
("warp speed" Monte Carlo is a feasible method). To save the p-value
tabulations in an EDF file, set "Tabulate p-values in EDFs" to TRUE
in Options / General / Special Settings. The name of the resulting file
contains the string "EDFpv_". An EDF file so named is assumed to
tabulate p-values, and can be used to generate corrected p-values in a
bootstrap test. This method might provide a feasible implementation of
the double bootstrap.

## 8.2a Options / Tests and Diagnostics / Diagnostic Tests...

Tests selected in this dialog are calculated automatically following an estimation run.
They can also be computed independently using current estimation outputs from the
Actions / Compute Test Statistics menu.

The main preset diagnostic tests are available in score test and conditional moment
test versions. Unless otherwise indicated the statistics, by default, are asymptotically
Chi-squared with degrees of freedom shown in parentheses. The nominal p-values of
the tests are shown, enclosed in braces, following the statistic values. F-versions of
the tests can optionally be substituted, although note that these are not exact except
in special cases.

Q TESTS FOR SERIAL DEPENDENCE
Q tests for residual autocorrelation in levels and squares are printed after an
estimation run by default. The Box-Pierce (1970) formula is used by default.
Check the "Ljung-Box" radio button to use the Ljung-Box  (1978) formula, or
"None" to disable this option.

Notes:
1. The upper limit to the Q test order is set at (number of observations)/3.

2. In models where variance-adjusted residuals are defined (GARCH, regime
switching), the Q statistics are computed for these series.

3. When ARMA components are fitted, the degrees of freedom of the 'levels'
Q-test, as used for computing asymptotic p-values, are reduced by p+q
accordingly. No such adjustment is made in respect of fractional integration

© James Davidson 2024

models, nonlinear dynamics, GARCH models etc, where the required corrections are generally unknown. The "nominal" p-value must be interpreted with caution in these cases.

SCORE (LM) TESTS

The following pre-set score tests are available. In each case, the score of the extended model at the null parameter values is tested. The extensions are:
*       Autocorrelation: lagged residuals in mean equation(s).
*       Neglected ARCH: lagged squared residuals in conditional variance equation(s),
      or lagged absolute normalized residuals, if EGARCH is specified.
*       Nonlinear Functional Form (RESET): Integer powers of the fitted values of the null model in the mean equation(s).
*       Heteroscedasticity: square of the fitted values in conditional variance equation(s).
*       White's heteroscedasticity test: squares and cross-products of the explanatory variables in conditional variance equation(s).
*       AR common factors: p lags of Type 1 regressors in a model with AR(p) errors specified.
Select the number of lags to be included and the order of polynomial in the fitted values with the scroll bars. These statistics are indicated in the output by "(LM)" after the description.

Notes:
4.       If no conditional variance is specified in the null model, the LM tests for neglected ARCH and heteroscedasticity is performed by regressing the squared residuals on the test variable(s). Otherwise, the test variable(s) are added to the conditional variance model as regressors.

5.       The test variables are designated regressors of Type 3 if a MA (GMA) component is selected, and of Type 2 if a AR or FI (GAR or FIGARCH) component is selected, and otherwise of Type 1.
      NOTE: Test variables cannot be assigned to a Type for which lags are already selected. If lags are specified for Type 1 variables, the test may be unavailable. A warning message is printed in this case.

6.       In the implementation of White's heteroscedasticity test, the possibility of collinearity between the squares and products is allowed for automatically. In this case, the principle components with positive eigenvalues are extracted.

7.       The AR common factor test is enabled in the following case: a model with Type 1 regressors, current values only; no Type 2 regressors; and an AR(p) error process is specified, but not a fractional process. The null hypothesis is that p lags of the regressors are not independently significant, in other words that there are common AR factors shared by the dependent variable and regressors in the model, so that AR errors are valid representation of the dynamics.

8.       In multiple-equation models, the full set of test variables is added to each equation of the model.

                         © James Davidson 2024

9.    LM tests are computed by creating an extended model with test variables added to it. In some situations the resulting model is unidentified, with a singular covariance matrix. In such cases the statistic cannot be computed and is assigned the value ".NaN", standing for Not a Number.

CONDITIONAL MOMENT (CM) TESTS

Eight conditional moment tests are pre-set. The hypotheses under test are the following.
*    Autocorrelation: correlation between current and lagged normalized residuals.
*    Neglected ARCH: correlation between current and lagged squared residuals.
*    Nonlinear Functional Form: correlation between the normalized residuals and integer powers of the fitted values.
*    Heteroscedasticity: correlation between the squared normalized residuals and the squared fitted values.
*    White's heteroscedasticity test: correlation between squared normalized residuals and the squares and cross-products of explanatory variables.
*    Test for AR common factors: correlation between the normalized residuals and lags of the Type 1 explanatory variables. (see para. 3 above).
*    Information Matrix Test: expected value of the difference between the information matrix and outer product of the score.
*    Bierens' Specification Test: correlation between the normalized residuals and a bounded nonlinear function of the exogenous variables; this test is supposedly consistent against all nonlinear functional form alternatives.

Notes:
10.    CM and LM tests are not available for use with the Whittle estimator.

11.    In GARCH-type models, CM tests use the normalized (variance-adjusted) residuals as the test covariates. In tests for neglected ARCH, at least fourth-order moments must exist, and failure of this assumption can invalidate test results - caution is advised when interpreting the findings.

12.    CM diagnostic tests are always computed with the joint covariance matrix of scores and moments computed by the robust method (see 21. below). Moment tests can also be specified using one of the HAC estimators, see Actions / Compute Test Statistics / Moment Tests.

14.    It may prove impossible to evaluate the White heteroscedasticity tests because the columns of the indicator matrix are linearly dependent. In this case principal components of the test indicators are used, with a corresponding reduction in the degrees of freedom of the test.

RESIDUAL-BASED TESTS

The following tests are applied to a single equation's residuals. Hence, in multi-equation models,  a test statistic is computed for each equation.

*    The Durbin Watson test is included only for heritage reasons. It is asymptotically equivalent to the Box-Pierce Q statistic for 1 lag. The LM or CM tests for residual autocorrelation are recommended in a regression context.

\* The CUSQ (or cusum-of-squares) test is for unconditional constancy of the variance of the residuals in a regression model, and hence acts as a test for parameter stability. It is computed from the absolute maximum of the partial sums of the deviations of squared residuals from their mean, and the location of this maximum in the sample is reported along with the statistic, as a diagnostic aid. The asymptotically valid p-values for the test are computed using the known analytic formula for the supremum of a Brownian bridge, and hence are reported exactly.

\* The KPSS, RS and Harris-McCabe-Leybourne (HML) tests are alternative tests of the hypothesis that the residuals are I(0). These are also available for data series from the Setup / Summary Statistics dialog, and may be useful in conjunction with the ADF and PP tests in the context of cointegrating regression. Two parameters are selectable to define the HML test, truncation parameter c and bandwidth parameter L, These are selectable in the 'Special Settings' area of the Options / General dialog. Use the default values c = 1 and L = 2/3, or see Harris et. al (2006) for guidance.

OTHER DIAGNOSTIC TESTS

\* The Information Matrix test is an M-test that compares the inverted information matrix with the robust covariance matrix estimator. The test has $k(k+1)/2$ degrees of freedom when k parameters are estimated. This can be used as a guide to correct model specification, and can also indicate best choice of covariance matrix formula. When least squares estimation is specified, the matrices computed are for the concentrated Gaussian log-likelihood function. In other words, in Gaussian ML the test is performed with the variance parameters included in the parameter set. In Least Squares (including OLS and LGV) the variance parameters are concentrated out, so that the test has fewer degrees of freedom. The test is not available for 2SLS and GMM estimation. NOTE: if the columns of the indicator matrix are linearly dependent the statistic cannot be computed. In this case the indicators are replaced by their independent principal components, reducing the test degrees of freedom.

\* The Nyblom-Hansen model stability test is available as a whole-model test (degrees of freedom equal to the number of fitted parameters, including residual variance) and, optionally, as one-degree-of-freedom tests for individual the parameters. Check the "Individual N-H Tests." checkbox for the latter option. These are not printed by default since the output is potentially bulky. The tests use the table given in Hansen (1990), and the p-values are available in inequality form.

\* The Andrews (1993) LM test of structural change with unknown change point is available, like the Nyblom-Hansen test, in whole-model and single-parameter variants. The tests use Andrews' table with 10%, 5% 1% critical values, hence the p-values are available in inequality form. These tests require an interval of the sample period to be specified (in the form of upper and lower fractions of sample size) and the default values are pi1 = 0.15 and pi2 = 0.85. To change these settings go to Options

/ General / Special Settings, and enter the values desired in the fields "Andrews Test Lower Bound (pi1)" and "Andrews Test Upper Bound (pi2)".

* Use the 'Coded Test' checkbox to activate/deactivate a test statistic that has been supplied by the user as Ox code (see Appendix C). This checkbox has the same function as the 'Test' checkbox in the Model / Coded Function dialog. It is disabled unless code is loaded and the 'Test Only' option is selected in Model / Coded Function. In the case where other Ox code is being compiled, this test option should be controlled from the same dialog.

* The bootstrap test of I(0) compares the generated values from a time series model with i.i.d. data, to verify the validity of "short memory" assumptions underlying asymptotic tests in finite samples. It is only available for univariate time series models.
Note that bootstrap inference for the equation itself is disabled if this option is selected.

Notes:
14. The Andrews LM test uses a simple line search algorithm to locate the maximum test statistic, as a function of the break date. The convergence criterion and the maximum iterations are set in Options / General / Special Settings. Note that these settings are shared by other program features. They should not generally need changing.

15. For details on the implementation of the I(0) test see 6.8, Actions / Compute Test Statistics / Bootstrap Test of I(0).

CONSISTENT SPECIFICATION TESTS

Note: to simplify the interface, dialog controls for the advanced consistent tests of functional form and dynamic specification (including Bierens tests) are not displayed by default. To switch these features on, go to Options / General / Special Settings, and set "Show Consistent Test Settings" to TRUE (double-click toggle).

The consistent tests of specification, which we generically refer to as CS tests, are based on the covariance between a target series derived from the fitted model and an exponential test function of explanatory variables. See Models and Methods for full details.

* Select the variables to appear in the test function (and a fixed number of lags) in the dialog accessed with the Test Variables button.
To include polynomial distributed lags of the test variables (permits long lags with fewer nuisance parameters) set the polynomial order (P) in this dialog.
    o If P = 0 the lags in the test function are set in the Test Variables dialog.
    o If P > 0 this setting is ignored.
* The possible target series are
    1. model residuals (Bierens tests))        (1 d.f. test)
    2. score contributions, full gradient vector    (p d.f. test)
    3. individual score-contribution elements    (1 d.f. test).

* The test statistics depend on nuisance parameters, and are computable in one of three modes, depending on how the nuisance parameters are treated:
  1. Sup test (S)
  2. Integrated test (ICM-A) (IA)
  3. Exponential integrated test (ICM-B) (IB)

  (The right-hand column indicates how the cases are identified in the output.)
* p-values for the tests can be computed by:
  1. "2-statistic" procedure(using standard chi-squared table)
  2. Hansen (1996) bootstrap (could be computationally burdensome).

Notes:

16. The number of Hansen bootstrap replications are set in this dialog. The Options / Simulation and Resampling replications setting is independent of this one. The usual bootstrap method could also be used to generate p-values. In this case the Hansen bootstrap setting is ignored.

17. Various test settings are found under Options / General / Special Settings; see options with prefix "Consistent". In general, these will not need to be changed from the defaults. They include
    * parameters for 2-statistic bound (gamma0, rho1, rho2, rho3)
    * upper and lower hypercube bounds for the weight set Ksi (Defaults are -1, +1).
    * Precision and maximum number of evaluations for the computation of integrated and sup statistics.
    * Maximum lag order for the polynomial lag option.
    * Scale factor for the test function, determines the range of variation of $\log(w_t)$ over the given sample. (Default = 3, such that $0.22 < w_t < 4.48$).

18. If gamma0 is set >100 only the base statistic $S_0$ will be reported, with weights ksi set to the unit vector.

## 8.3 Options / Forecasting ...

FORECASTS

1. The levels forecasts include 2-standard error bands. These are asymptotic and do not allow for parameter uncertainty. For conditional heteroscedasticity models, they are computed dynamically using the volatility forecasts.

2. The scroll bar label shows the number of forecast steps selected. The date or observation number of the final step, which depends on the sample selection (see Setup / Set Sample) is shown above the scroll bar.

3. There is no limit to the number of ex-ante forecasts. To increase the maximum, select it, then close and re-open the dialog. However, in models containing exogenous variables (other than the trend and GARCH_M terms) the limit is set by the available post-sample observations.

4.  "Ex ante" multi-step forecasts use the lagged forecast values to projected two or more steps forward. "Ex Post" 1-step forecasts are simply the fitted values of the model (using actual values of all lagged and explanatory variables) for the selected forecast period. This cannot exceed the available observations, so the estimation sample must be set appropriately.

5.  By default, the forecast period for ex-post forecasts is set to the maximum available observations following the 'Last Observation' set in the Setup / Set Sample dialog. If this setting is changed, the number of forecasts is adjusted to the default. To include fewer than the maximum number of observations in the forecast period, change this setting after setting the estimation sample.

6.  Ex post forecasts and forecast errors are displayed in the Actual/Fitted and Residual plots. The main purpose of this option is to check model stability. Two test statistics are computed, appropriate to short and long forecast periods respectively. The chi-squared (short-period) test (Forecast Test 1) is the generalization of Chow's (1960) second test. and depends on the additional assumption of Gaussian disturbances under the null hypothesis.

7.  The second test reported depends on the model. In single equation linear regression models, Chow's stability test statistic is reported (choose between F and Chi-squared forms by checking the box in Options / Tests and Diagnostics.) In other cases, where the need to re-estimating the model for sub-periods precludes generalizing the Chow procedure, a difference-of-means test applied to the squared residuals and forecast errors. This is asymptotically N(0,1) as both periods tend to infinity. This is called Forecast Test 2. If it is desired to compute this statistic for the linear regression case, simply open the Dynamic Equation dialog and select Actions / Evaluate at Current Values.

MULTI_STEP FORECAST METHOD

8.  Two methods of computing forecasts and associated standard error bands are provided. The analytic method computes formulae for the mean and standard error of the forecast, and the plots show two-standard error bands corresponding to 95% confidence bands when the errors are Gaussian.

9.  Two methods of computing analytic forecasts are available. New code in version 4.38 is used by default. This method solves the model simulation algorithm numerically, with zero shocks, and hence is available even for nonlinear structures such as ECMs. In the case of nonlinear-in-variables dynamics, the result is a linear approximation to the true dynamics which may be contrasted with the stochastic simulation (Monte Carlo) method. The new code works with most models specified through the linear regression or dynamic specification dialogs, with the exception of Markov-switching models.
    The old code, which works only for VARMA/GARCH models, is still available by selecting this option in Options / General / Special Settings. It is selected automatically for Markov switching models, alternatively

© James Davidson 2024

use the Monte Carlo method.

10. The Monte Carlo method uses the fitted model to compute stochastic simulations of the process, and reports quantiles of the empirical distribution for each step. In this case all models, including those containing lagged dependent variables as regressors, are computed by dynamic simulation. Hence, true multi-step forecasts are reported in this case.

11. Two types of Monte Carlo plot are available. Time plots show the median and 2.5% and 97.5% percentiles, so that the plotted bands in this case show a true 95% confidence interval (neglecting estimation error) even when the disturbances are non-Gaussian. Alternatively, kernel densities can be displayed, for the complete empirical distribution of any forecast step.

12. Analytic standard error bands for regime switching models use a recursive algorithm, involving $m^K$ function calls for M regimes and K steps ahead. In Hamilton's model, the number of effective regimes is $M^{\wedge}$(number of lags). Unless they converge rapidly, these calculations can become very time consuming. They are therefore terminated automatically after 100 seconds has elapsed. Unless the bands have converged sufficiently to be extrapolated thereafter, only the point forecasts are reported for the remaining steps. (NB: diminishing returns set in rapidly! Even raising this factor substantially could yield at most a couple of extra steps.)

13. The display mode for standard error/confidence bands, including omitting them, can be selected in Graphics Options.

14. In dated (daily/weekly) data, it is not possible to plot forecasts with date information beyond the end of the observed sample, since the information is missing. An error message is displayed. In this case, either change the date type to "No Dates", or extend the data set and supply the additional dates manually (This must be done outside TSM).

MONTE CARLO SETTINGS
15. The default number of replications is 1000.

16. Select either the mean or the median of the Monte Carlo distribution for reporting. In the former case, 2-standard error bands are also shown in the plots, and in the latter case, the 2.5% and 97.5% quantiles are shown.

17. To select the simulation mode and change the random number seed see Options / Simulation and Resampling.

18. The scroll bar allows selection of one of the forecast periods for a density plot. Note that the maximum selection varies with the setting for number of forecasts.

ACTUALS IN FORECAST PLOT
19. It is usually helpful in a post-sample plot to show some, but not always all, of the series for the estimation period. The default is to include the last 50

© James Davidson 2024

observations. This can be changed, and to include all available observations, simply enter any value large enough. Note that when the forecast period includes or overlaps the data period (as for ex-post forecasts) the actual values are always shown.

20.  Forecast standard errors and confidence bands are enabled by default.
     It may be desirable to switch them off in certain circumstances.
     For example:
     *      If a user-coded function contains lagged adjustment, these
            dynamics cannot be taken account of in analytic variance formulae,
            which will therefore be incorrect.
     *      In Markov-switching models, calculation of multiple-step standard
            errors can be computationally very burdensome.
     *      With analytic forecasts the reported confidence bands are based
            on Gaussian critical values, and should be suppressed if the
            Gaussianity assumption is unwarranted.
     Note that standard errors are enabled automatically with confidence bands.

21.  By default, forecasts exported to a spreadsheet file include either quantiles
     or standard errors. For some purposes, particularly if further analysis is
     required, it may be more convenient to export only the table of point
     forecasts (e.g. median quantiles). Check the "Export Point Forecasts
     Only" checkbox to enable this restricted mode of exporting.

FORECAST ERROR VARIANCE DECOMPOSITION

22.  This option is available with analytic forecasts of multi-equation models.
     The values returned show, for each variable, the proportion of forecast
     error variance attributable to each equation in the system. The formula
     is taken from Lutkepohl (2007), page 63ff. The moving average coefficients
     are computed as for analytic forecasts by solving the model numerically, and
     hence are well defined even for nonlinear in parameters structures such as
     ECMs. In the case of nonlinear models, the sequences of weights generated
     provide a linear approximation to the model dynamics. The matrices of
     proportions are displayed for each horizon following the standard forecast
     outputs for each equation. The data are exported to a spreadsheet using
     the command Files / Listings / Export Spreadsheet / Forecasts, as for
     forecast outputs.

PLOT DISPLAY

23.  Ex ante forecasts can be viewed by listing values in the results window (see
     Options / Output and Retrieval) or graphically (Graphics / Ex Ante
     Forecasts). Check the "Display Forecasts Automatically" checkbox to
     have the main graphs displayed immediately a forecast is run. Use the
     Graphics menu to display plots individually, and additional items such
     as Monte Carlo distribution plots.

IMPULSE RESPONSES

24.  The  impulse responses (coefficients of the solved moving average process)
      for levels and also conditional (ARCH/GARCH) variances, will be plotted

according to the radio button setting, in either impulse form, or step (cumulated) form, or with both plots on the same graph.

25. In a system of N equations, the responses shown are those on a dependent variable of an impulse to its own shock process only (N plots). If the "System" checkbox is checked, then the responses on each dependent variable of impulses to each shock process are calculated (N^2 plots). This checkbox has no effect in a single equation model.

26. There is no limit to the number of steps possible. To increase the maximum, select it and close-open the dialog.

27. Only the own-shock responses can be listed in the results window. For a listing of the full set, export the data to a listing file (see File / Listings / Save MA Coefficients).
Note:   Only the impulse responses are exported. Compute the step responses manually by cumulation, if required.

## 8.4 Options / Simulation and Resampling ...

SIMULATION OPTIONS

1. There are nine ways to generate artificial shocks to drive the one-off simulations, Monte Carlo experiments and bootstrap tests. See Models and Methods for details and formulae.

    *   'Model' uses the distribution specified by the likelihood function, if a maximum likelihood estimator is specified. If the estimator specified is not ML, then Gaussian shocks are used with variance equal to the residual variance from the last estimation.

    *   'Gaussian' uses normal random numbers with zero mean and the standard deviation specified in the text field provided.

    *   'Stable' uses random numbers from a centred, infinite-variance stable distribution, with the three parameters specified in the text fields. Alpha (<= 2) is the stability parameter where 2 is the Gaussian case, and corresponds to the largest existing absolute moment. Beta controls skewness and 'Scale' has a role analogous to the SD.

    *   'Formula' generates shocks with a user-defined formula, depending on one or more sets of uniform and/or normal random numbers, and optionally on variables from the data set and parameter values defined in the model. In a multi-equation model the same formula is used with different shocks, and parameters (if specified) specific to the equation. Enter the formula in the dialog box provided, see section 1.5 Entering Formulae for details. (Note that this is "case D").

    *   'Simple Bootstrap/Block Bootstrap' randomly resamples the actual residuals, with replacement. The resampled series is centred and multiplied by a variance correction factor, sqrt(n/n-k).

© James Davidson 2024

The block bootstrap is implemented by choosing a block length > 1.

* 'Stationary Block Bootstrap' is like the block bootstrap except that the blocks have random lengths drawn from the geometric distribution.

* 'Wild Bootstrap' preserves the heteroscedasticity of the original sample. Instead of randomly resampling, the data are randomly transformed by a drawing from a two-point discrete distribution.

* The Fourier bootstrap resamples from the discrete Fourier transform of the residuals using the Rademacher wild bootstrap, and returns the inverse transform of the resampled series. An alternative to the sieve-AR, suitable for Gaussian homoscedastic residuals showing generalized autocorrelation.

* 'Data Resampling' is only appropriate for serially independent data (cross-section data, in general). Observations (both dependent and explanatory variables) are resampled randomly with replacement from the specified range of the data set. Click "Set Range..." to open the Sample Setup dialog and select the sampling range.

2. Optionally (see the checkbox) the 'Likelihood Model' method can be used for one-off simulations and Monte Carlo experiments, regardless of the selected resampling method. This setting allows the specification of a Monte Carlo experiment where the bootstrap is used for tests, without the need to set up different models for data generation and estimation.

3. If the Gaussian standard deviation (or stable scale parameter) is set to 0, the sample residual s.d. from the most recent estimation is substituted, if such exists. Otherwise, it is reset to 1.

4. Tip: if the Gaussian, Stable or Formula options are selected, scale and (in multiple equation models) correlation parameters can be set in the Values / Equation and Values / Distribution dialogs. The trick is to set the 'estimator' to Gaussian ML in the Model / Dynamic Equation dialog. Note that when the model is used for simulation the estimation criterion is not invoked, so this setting is compatible with non-Gaussian shocks. In this case the s.d./scale setting in this dialog is ignored.

5. Non-standard distributions can be simulated using the Formula option by setting u# as the argument of the inverse c.d.f. if this is can be easily formulated. Formulae can also contain model parameters and depend on variables in the data set, referenced in the usual manner.

6. The block bootstrap option allow valid bootstrap inference in mis-specified models with autocorrelated errors, provided the autocorrelation is not too persistent. The chosen block length value entered in the text box is also shown as a power of the current sample size. In the case of the stationary block bootstrap, the mean block length is selected.

© James Davidson 2024

7. All bootstrap methods can be combined with the sieve autoregression, in which the resampled series are passed through an autoregressive filter fitted to the original series. By default the sieve-AR method uses an increasing function of sample size as maximum lag order. For this case, enter "-1" in the Max Sieve-AR Lags text field. Setting a zero or positive value over-rides the default. Setting "0" is equivalent to the regular bootstrap method.

8. If 'Bootstrap/Block Bootstrap' is selected with block length 1, and the random number seed is set to -1, the shock-generation procedure returns the actual model residuals, in original order. The generated series should then be identical to the original data series, apart from possible rounding error. This feature provides a check that the simulation module is inverting the fitted model correctly. (Note that to generate Markov-switching models also requires a latent switching process, so this check will be valid only if the Markov transition matrix is fixed at the identity matrix.)

9. The wild bootstrap variants are not available for Monte Carlo forecasts. If these options are selected, the regular bootstrap is substituted.

10. The option "Wild Bootstrap Skewness" sets the parameter 'a' such that the t-th drawing is $e.u\_t$ where $u\_t$ is the sample value (fixed), and $e = a$ and $-1/a$ with probabilities $1/(1+a^2)$ and $a^2/(1+a^2)$ respectively. Note that this setup preserves the first two marginal moments of the distribution. The default is $a = 1$, which always yields a symmetric distribution, but preserves kurtosis. Setting $a = 1.618$ preserves the original skewness (if any) but doubles the kurtosis. The value of the Kolmogorov-Smirnov statistic, comparing the fit of the original data points to the bootstrap distribution, is reported to aid the best choice of this value.

SIMULATION PRESAMPLE MODE

11. If a dynamic simulation run starts at observation 2 or later, the initial observations can be generated in two ways, either fixed or random. If fixed, the actual values of the selected dependent variable are used. If random, the sequence is randomly set in each replication, using generated or resampled shocks and the specified model. In other words, the simulation actually initializes at date 1, whatever the nominal start date.

   * If a partial sum process (unit root) is specified, the cumulation begins at the start date, not at date 0.

   * The random presample option can be selected for Monte Carlo experiments. Fixed (actual) values are always used for bootstrap tests and Monte Carlo forecasts.

   * The setting makes no difference if the simulation start date is 1.

RANDOM NUMBER GENERATION

12. The random number generator is used in several program operations, including bootstrap estimation, Monte Carlo forecasting and Monte Carlo

© James Davidson 2024

experiments, as well as free-standing simulation runs initiated by the user. It is also used in optimization algorithms based on a random search, such as simulated annealing.

13.   The generator is initialized before running the following tasks:

  *       an estimation run, which may involve random optimizations, and bootstrap-based tests.

  *       a free-standing model simulation.

  *       a Monte Carlo experiment, which involves both simulation and estimation stages.

14.   The Random Number Seed initializes the random number generator, such that it produces an identical sequence from any given seed. If a positive integer (an arbitrary set of decimal digits) is entered in this field, the same sequence of random numbers will be used in successive runs, which are accordingly exactly reproducible. If the value is set to 0 (the default) the seed is taken from the system clock (number of seconds since midnight). Accordingly, every run will be different and not reproducible.

INFERENCE BY RESAMPLING

15.   Two methods are implemented for generating confidence intervals and p-values by resampling the data, rather than by asymptotic criteria.

  *       The bootstrap simulates the specified model using the TSM simulation engine with randomly drawn artificial disturbances. These can be either drawn from a known distribution, such as the normal or Student's t, or resampled from the empirical distribution of the estimated residuals. Monte Carlo simulation of the estimation procedure yields the bootstrap distribution.

  *       Subsampling is the method of fitting the model to successive contiguous subsamples of the data set, and constructing the empirical distribution of the subsamples. It does not involve drawing random numbers.

  Results are reported in the form of confidence intervals, p-values and plots of the empirical distributions generated.

16.   By default, resampling is deselected automatically after a single run ("Once"). There is an option to select it permanently as the inference method ("Always"). Selection is indicated by a flag on the status line, '(B)' for bootstrap and '(S)' for subsampling.

17.   The number of replications determines the accuracy with which the p-value is measured. However, because of the small-sample error in the p-value (due to replacing actual with estimated parameters) the accuracy attained from a large number of replications is typically spurious. The default of 99

© James Davidson 2024

(implying accuracy is bounded by 1%) is probably adequate in many cases. 399 would be sufficient in nearly every case.

18.    The 'Static' option generates the bootstrap draws by adding the resampled shocks to the fitted values of the estimated equation. This procedure is equivalent to the usual bootstrap method when the model contains only exogenous variables (fixed in repeated draws) although it is typically much faster. On the other hand, in models containing lagged dependent variables the resampled series are generated conditional on the lagged actual data, not the lagged resampled data.

19.    The 'm/T' option uses a bootstrap sample of length smaller than (a fractional power of) the observed sample. This option may yield a consistent bootstrap in cases of data with infinite variance. First select the number of bootstrap replications with the scrollbar. Then, select the option and choose the bootstrap sample size with the (re-assigned) scrollbar.
       Note:
       1.    The option is cancelled if the resampling method is changed to Subsampling.
       2.    The option is not available for models with exogenous variables,

20.    The bias correction option subtracts an estimate of the bias from each parameter point estimate. This is computed as the difference between the mean of the bootstrap distribution and the point estimate (estimating, respectively, the mean of the estimator distribution and the true value).

21.    The fast double bootstrap is a device aimed at reducing the error in rejection probability of bootstrap tests. It is not guaranteed to improve performance in all cases, but showing that test outcomes are invariant to this selection offers additional robustness.

RESAMPLING CONFIDENCE INTERVALS

22.    Selecting the "SEs" radio button results in the estimation output appearing in the usual format. Otherwise the "Std. Err." and "t-Ratio" columns are replaced in the results window output by lower and upper confidence bounds.

23.    Three schemes for constructing confidence intervals are implemented.
       Use the text field "Level" to select the percentage of bootstrap replications lying within the reported intervals. Set this to a value X between 0 and 100.

       Equal tails:     Bounds are computed such that (100 - X)/2% of the bootstrap point estimates fall in each tail.

       Equal tails percentile t:
                        Bounds computed such that (100 - X)/2% of the bootstrap t statistics fall in each tail, scaled by the standard errors of the resampling distributions.

       Symmetric percentile t:
                        Bounds based on the X% quantile of the absolute t statistics,

© James Davidson 2024

scaled by the standard errors of the resampling.

In the last case the intervals are centred on the sample estimates, by construction. In the equal-tailed cases their locations will depend on the skewness of the resampling distributions.

24. Reported p-values estimate the tail probabilities of the smallest confidence intervals containing the sample statistics. When equal-tailed confidence intervals are specified, these are the signed t-ratios, by default. Otherwise the p-values relate to the absolute t-ratios.
Optionally, as a legacy setting, p-values for the absolute t's can be reported in every case. See Options / General / Special Settings / Bootstrap Absolute t Statistics.

25. The model is re-estimated with simulated data in each replication. For models estimated through the Dynamic Equation dialog, this means numerical optimization of estimation criteria. A simpler search procedure using only BFGS is implemented in this case. Numerically precise convergence is not required to reproduce the sampling distribution adequately, and replication times can be reduced by limiting the maximum number of BFGS iterations. This can be set here independently of the main optimization settings.

TYPE I FRACTIONAL PROCESSES

26. A fractional (ARFIMA) model has long memory, and can depend significantly on remote pre-sample influences. If these processes are simulated by setting presample shocks to zero (the 'type II' model), this changes the distribution of estimators and test statistics even in large samples. The 'type I' model can in principle be simulated by selecting the random presample option and setting a long presample period. However, the number of lags required for a good approximation may be excessive. The option 'Generate Fractional Series as Type I' (set by default) simulates the effect of presample lags by adding a single vector (or matrix) of Gaussian random variables with the correct covariance structure. Selecting this option ensures that (e.g.) bootstrap distributions match the type I case.

Notes:
1) With this option the fractional process must be stationary, i.e., $d < 0.5$. To simulate a nonstationary process, replace d by $d - 1$, and check the option "Impose Unit Root" in Model / Dynamic Equation.

2) The option is not available in conjunction with regime switching. A type I process is always generated in this case.

3) If the Presample data are held fixed in the simulations (required in bootstrap and Monte Carlo forecasting) estimates or imposed values of the presample shock components must be available. The number of these to be included is set in the "type I Fractional" text box in Options / General; if the setting is zero, the type II

process is generated.

## *8.5 Options / Graphics ...*

Press the button "Line and Symbol Styles" to access a separate dialog
for controlling the appearance of series graphs. These options are
described in Section 8.5a of this manual.

TEXT ATTRIBUTES

1.      Select the font and size in points for titles, axis labels and legends in
        exported files. Note that the actual appearance will depend on the
        file type selected for export (see Options / Output and Retrieval.)

2.      Gnuplot graphs always use the Arial font when initially displayed. The
        text style for these graphs can be changed using the context menu accessed
        by clicking  the top-left corner of the window.

DATES IN PLOTS

3.      The default date display mode is selected automatically when a new data set
        is loaded. Alternative settings must then be re-selected in the dialog.
        The defaults are:
        *        YYYY for annual or quarterly data.
        *        MM/YYYY for monthly and higher frequencies.
        *        DD/MM/YY for dated (daily or weekly) data.
        *        'Labels' for undated data.

4.      By default, 'Labels' are set to the row numbers of the data matrix. An option
        also exists to create arbitrary observation labels as a column of the data file.
        For details see 3.3, Setup/Data Transformation and Editing / Edit / Set Date
        Labels.

5.      Dates are disabled in the cases of panel samples and samples selected by
        indicators. In these cases, consecutive observation numbers do not in general
        correspond to consecutive dates.

6.      With daily/weekday data, unless calendar dates have been loaded with the
        data (see 2.2 File / Data) the date indication is generally approximate, since
        there is no allowance for missing days such as holidays. In most daily data
        sets with large numbers of observations, the axis labels should provide an
        adequate approximation.

OTHER PLOTTING OPTIONS

7.      Files with extension .PLT are created by GnuDraw for processing by
        GnuPlot. These are text files that can be edited for further processing
        if you know the GnuPlot command language. If the box "Keep PLT
        Files" is checked, these files are retained in the results folder after
        creating a graphics file. Otherwise they are deleted automatically.

8.      By default, plots are displayed with an embedded title. This can optionally
        be either changed by the user or simply suppressed. Checking the
        "Edit Title" checkbox causes a text entry box  to appear when a graphic

is generated. This contains the default title, for either modification or replacement. Note that this option cannot be selected if the "No Title" checkbox is checked.

9.  When the origin falls in the range of the plot, the zero axis is drawn as a light grey broken line by default. This option can be turned off using the "Plot Zero Axis" checkbox. When the option "RH Scale" is selected in Graphics / Show Data Graphic, two zero axes may be drawn, for the left- and right-hand scales respectively.

10. Plots have a white background by default. A light grey background is a selectable option when plotting with Gnuplot 5 and later. NOTE: This option is not available in the bundled Gnuplot 4.2!

11. The "Double-Chart" button on the toolbar displays, by default, time plots of the actual and fitted values and the residuals. The button is disabled until an estimation has been performed. If the option "Extended Actual-Fitted-Residual Plots" is checked, two additional plots are shown; the actual-fitted scatter plot, and the histogram and kernel density of the residuals, together with the normal density curve with matching mean and variance, for comparison.

12. By default, scatter plots of data include the two regression lines (X on Y and Y on X). In the case of the actual-fitted scatter, the single line with slope 1 is drawn (a 45 degree line when the axes match). These lines can optionally be suppressed.

13. A "Fill" style is used by default to highlight the display of variances and probabilities, whose signs are always positive. This option can be switched off by unchecking the checkbox..

CONFIDENCE BANDS

14. By default, multi-step forecast plots show 2-standard error confidence bands, or 95% estimated confidence bands in the case of Monte Carlo forecasts.
    The following options available for confidence band plotting.
    None:   also effected by turning off calculations in Options / Forecasts.
    Bands: additional lines are plotted to show 95% region. The band colour is line style 4, light green by default.
    Bars:   the default - 95% confidence region marked by vertical bars. The bar colour is line style 4, light green by default.
    Fan:    bands of increasing colour intensity  mark six different confidence regions around the point forecast. The confidence levels indicated are 98%, 95%, 90%, 80%, 60% and 40%.
            Notes:
            *       Fan colour is set under "Line and Symbol Styles". It is worth experimenting with this setting since some colours work better than others.
            *       Fans are not available for monochrome plots, or recursive forecasts. The default option (bands) is used in

© James Davidson 2024

these cases.

15.    There are three ways to plot several series at once:
       *       Individual: each plot in its own frame.
       *       Separate plots in the same frame (default).
       *       All series in the same plot.
       The last option can also be selected in the Graphics / Show Data Graphic
       dialog, in which case the style of each line is selectable. Note the option
       to plot the last selected series against the right-hand axis in this case.

16.    If a system is estimated, the above three options are available for plotting
       the series associated with each equation: actuals, fitteds, residuals etc. If
       the "Individual" option is selected, only the series for the equation
       selected in the "Equation" choice widget on the tool bar is plotted.

17.    Legends are used to label series when two or more appear in the same
       frame; otherwise, the series name appears as the plot title. There are five
       options for placing the legend in the frame:
       X       No legend.
       TL      Top left corner (default)
       TR      Top right corner
       BL      Bottom left corner
       BR      Bottom right corner
       Note that these are general settings. It is not possible to place the legend
       differently in different plots in the same frame.

18.    Check the "Box" checkbox to enclose the legend in a box. By default,
       no box is shown.

19.    These plots are available to for univariate and bivariate data distributions,
       bootstrap and subsampling distributions and Monte Carlo distributions,
       and also for Monte Carlo ex ante forecast distributions at specific dates.
       Optionally, either or both of the kernel density estimate and the histogram
       can be selected, with both selected by default. Also optionally available are
       the normal curve with matching mean and variance and, for univariate
       distributions only, the cumulative distribution function (CDF) plotted in a
       companion frame.

20.    There are three options for the three-dimensional plots created by Actions
       / Plot Criterion Grid and Graphics / Show Data Graphic / Bivariate Density.
       If "both" is selected, the resulting plot is the same as the surface plot with
       contours drawn on the 'floor' ($p_1$, $p_2$ plane) of the three-dimensional
       graph.

## 8.5a  Options / Graphics / Lines and Symbols

Series can be plotted by lines connecting the data points (default), by symbols

© James Davidson 2024

at the points, or by both lines and symbols.

Plots can be selected as colour or monochrome where in the latter case, full lines and patterns of dots and dashed lines can be selected.

These are general settings. The choice of individual colours and styles for each line is selected in the line and symbol styles area.

Notes:

1.      If the "Lines" radio button is selected, the Symbol choice widget and size selector are disabled.

2.      If "Both" is selected, the inclusion of symbols can be overridden for individual lines by choosing "None" in the Symbol choice widget.

3.      In Microsoft Windows, line widths and patterns can be further edited by clicking by clicking the top left corner of the Gnuplot display window and choosing Options / Line Styles... from the menu.

LINE AND SYMBOL STYLES

Eight line styles for series plotting can be set by the user.

The selectable option are line colour (or pattern, for monochrome plots), line width, symbol type, and symbol size.

First select the line to be styled with the radio buttons. Then select options with the line and symbol choice widgets, and line width and symbol size fields.

*      Available line widths are 1 (thinnest, default) to 6 (thickest).

*      Symbol size can have any positive value, 1 is default.

*      Monochrome lines can take any width, if solid.

*      Patterned lines (dots, dashes, etc.) are always of thickness 1.

Press the 'Restore Defaults' button to set all line settings to defaults.

Notes:

4.      The Colour/Pattern widget shows different options, depending on whether the colour or monochrome plots option is selected. Note that both settings (colour and pattern) are remembered. Switching between colour and monochrome plots does not delete the options.

5.      The eight line styles are used for data series plotting in the multi-series plot option in Graphics / Show Data Graphic.

6.      If a multi-series graph is plotted by selecting 'One Graph' under Multiple Series Display in the Options / Graphics dialog, the line styles are used in order, starting with 1.

7.      Equation graphics are drawn using line styles 1-4.
        *      Style 1 is used for the main series in each plot, including the 'Actual' series in Actual/Fitted plots.
        *      Style 2 is used for 'Fitted' series.
        *      Style 3 is used for forecasts.
        *      Style 4 is used for confidence bands.

8.      If multiple equation series are plotted in a single graph, the line styles are incremented by two, or three, so as to cycle through the available styles.

9. The symbol option 'Fill' does not draw a symbol, but shades the region between the line and the zero axis using either light gray in colour plots, or a hatching pattern in monochrome plots.
Optionally, this style is used by default in the display of probability and conditional variance series. This option is enabled by default, uncheck the checkbox in the Options / Graphics dialog to use Line Style 1 for these series.

10. The "Band Fill" line style is a special option to indicate regimes or time periods, depending on the value of the selected variable. It displays light grey bands in those periods when the variable is positive.
NOTE: Band Fill can only be selected for Line 8. Attempting to select it for Lines 1-7 produces a warning message.

11. The "Fan" style is to select the colour to be used for fan charts. Note that the series itself is drawn in a contrasting colour, and the results are not always predictable from the description. Experimentation is recommended!

SCATTER_PLOT OPTIONS

12. Scatter plots styles allow the form, colour and size of symbols to be chosen. Scatter 1 is used for all data plots, and for actual-fitted plots. Scatter 2 is used only for the post-sample actual-forecast scatter when the one-step forecasting option is selected.

13. The "RGB Scatter" option colour-codes the points in a scatter plot according to their position in the sample. The initial points are coloured red, those in the middle green, and the end of the sample blue, with a smooth interpolation from one colour to the next. Try plotting the scatter of a variable against the trend dummy to see the effect. Select a 'filled' symbol (square, circle or triangle) for best results. If monochrome plots are selected, the coding is a greyscale, running from black to white.
Note: This option is not available for actual-fitted plots which include post-sample actual-forecast points.

## 8.6 Options / ML and Dynamics ...

Press the button "GARCH Settings" to open a separate dialog for selecting various options relating to models of the conditional variance. The dialog is also accessible from Model / Conditional Variance / Options...

ML PARAMETERIZATION

1. The residual variance and GARCH intercept are scale-dependent parameters, bounded below by zero. In some models, such as highly persistent FIGARCH models, the GARCH intercept can be very small and estimation close to the parameter boundary poses numerical problems. These are eased by making the square root (or other fractional power) the parameter to be estimated. The default value, which works well in most

cases, is 2. Try raising this to 4 if there is evidence of a problem. Set to 1 to estimate the variance/intercept directly. Be careful to set the starting value and interpret the estimate appropriately.

2.    The Student's t degrees of freedom (d.f.) parameter is bounded below by 2, and is infinite in the Gaussian case. A square root transformation (the default) better numerical performance in the vicinity of these can yield extreme values. Setting the order of root to a negative value allows the estimation of the inverse d.f., which takes the value zero in the Gaussian case.

3.    Note: the variance/GARCH intercept and Student t d.f. parameters are constrained to be nonnegative during optimization by taking the absolute value of the parameter element to evaluate the likelihood function.

4.    By default, the Student skewness parameter is non-negative and equal to1 in the symmetric case. Estimating the logarithm makes 0 the symmetry value, and may yield better numerical properties. (In the default case, non-negativity is enforced as in 3.)

DISCRETE DATA OPTIONS

5.    To activate the ordered Probit and Logit options enter the number of additional cases in the text field. A zero setting gives the standard Probit/ logit.

6.    In a time series model, an autoregressive component can be included in the argument of a discrete distribution by checking the "Autoregressive Latent Process ...". checkbox. The structure of the resulting model is somewhat analogous to a GARCH equation with the argument (e.g. Poisson mean) playing the role of the conditional variance. Positivity is not, of course, required in the Probit/logit latent process.
With this option, the argument must depend on regressors of  Type 2, typically including the lagged dependent variable. The autoregressive parameter(s) appear in the output with a [2] prefix, and the order of the autoregression matches the lag set for Type 2 regressors. Fine tune the number of included lags by fixing parameters at zero in the Values/ Equation dialog.

7.    Zero-inflated discrete models allow for an excess of zero observations relative to the distribution specified. The probability of a 'zeros' regime generating the observation is controlled by an auxiliary distribution (Gaussian for the Probit, logistic for the Poisson or negative binomial)
If this probability is fixed, the distribution quantile is an additional parameter appearing in the Values / Distribution dialog with a prefix [3]. If the quantile is varying, depending on regressors, these are specified as regressors of Type 3 in the Dynamic Equation dialog, and are reported with a [3] prefix in the output.
The auxiliary distribution can be made dynamic by enabling both this option and the Autoregressive Latent Process option, as in 6 above.
If only one dynamic component is desired, disable the other by fixing the relevant parameters at zero.

8. The chi-squared distribution can be used to define the probit latent variable, with the degrees of freedom estimated as an additional parameter. The value appears in the Values/Distribution dialog.

9. In the Poisson and negative binomial models, the distribution mean can optionally be specified as the exponential of the latent regression function. Otherwise, the regressors must be non-negative, and the argument is constrained to its positive part in estimation.


DYNAMIC MODEL SETTINGS

10. By default, up to 10 lags can be set in the ARMA and GARCH models. This maximum can be reset to any desired value in the text box.

11. There is no accepted convention regarding the signs of the coefficients of a moving average lag polynomial. The program default is to report the MA coefficients theta_j in
    theta(L) = 1 - theta_1 L - ... - theta_q L^q,
and GARCH coefficients as beta_j  in
    beta(L) = 1 - beta_1 L - ... - beta_r L^r .
This is consistent with the convention in equation (4.8) of Models and Methods, and also more natural. For example, in the ARMA(1,1) model equal roots cancel each other out, and in this case the estimates will be equal. Similarly, in the ARMA-in-squares representation of the GARCH(1,1) model, beta_1 = delta_1 corresponds to alpha_1 = 0, and the estimates are again equal in this case. However, the MA coefficients can be optionally reported as theta_j in
    theta(L) = 1 + theta_1 L + ... + theta_q L^q,
and GARCH coefficients as beta_j in
    beta(L) = 1 + beta_1 L + ... + beta_r L^r.

12. If the first observation for estimation is greater than 1, lag distributions will normally be computed using all available pre-sample lags. Checking 'Restrict Pre-Sample Lags' allows the number of presample lags to be set in the textbox provided. Thus, setting 'Lag Truncation' to 0 means that all pre-sample values are replaced by 0. This provides comparability with the situation when the first observation used for estimation is 1.

13. The roots of dynamic systems (autoregressive and moving average lag polynomials, also the GARCH model counterparts of these) are computed by default. In very large models, in particular large BEKK models, the algorithm to compute the roots can become unduly time consuming. These calculations can be optionally disabled by unchecking 'Compute Roots'.

14. For details of the option "Type I Fractional" see 4.4 Model / Dynamic Equation, and Davidson and Hashimzade (2009).

15. In systems of equations featuring fractional differencing  and/or a

© James Davidson 2024

nonlinear ECM, checking the checkbox "Difference Spec'n for System FI & ECM Parameters" causes the equations to be parameterized so that equation 1 is the reference case, and the parameters in the other equations are measured as differences from this case. This option makes it easy to test and/or impose the restriction that the parameters are equal across equations.
Note: This control is duplicated in the dialog Model / Equilibrium Relations.

## *8.6a  Options / ML and Dynamics / GARCH options*

GARCH PARAMETERIZATION

1.	As the default, GARCH coefficients are reported in 'ARMA-in-squares' form, in which delta(L) and beta(L) in equations (6.1) and (6.2) are the AR and MA components. In this case the coefficients are named "GARCH AR1", "GARCH MA1" etc., etc. Optionally they can be reported in 'standard' form, corresponding to the representation of Bollerslev (1986). In this case, the coefficients named "GARCH Alpha1" etc. are the coefficients of beta(L) - delta(L), and "GARCH Beta1" etc. denote the coefficients of beta(L).
	Note:
	*	Roots of the GARCH AR polynomial are not reported when this option is checked.
	*	The EGARCH model does not strictly have an "ARMA-in-squares" representation. The default corresponds the stated form of equation (6.2). The 'Bollerslev' form is the one more commonly reported.
	*	It is the user's responsibility to match starting values to the convention selected.

2.	The GARCH intercept can also be computed in two different ways. By default, it is 'omega' in equations (6.1) and (6.2). By unchecking the box "Specify 'Type 1' GARCH Intercept", kappa = omega*beta(1) will be computed instead, see equations (6.4)ff of Models and Methods.

OPTIMIZATION SETTINGS
The following options should not normally need changing, but the defaults may not suit all cases. Experimentation is recommended if in doubt.

3.	The GARCH-M likelihood is computed by Gauss-Seidel iteration of equations (4.1) and (6.1), (or (4.1) and (6.2) for EGARCH). A fixed number of iterations is used, the default is 5. Increasing the number will increase computation time, but may not improve the estimates significantly.

4.	To stabilize the GARCH-M calculations, $h\_t$ (or $h\_t^{1/2}$) is trimmed before inclusion in equation (4.1). The upper bound is set to a multiple of the sample variance (or standard deviation) of the data, called the trimming factor. Try reducing this setting in case of failure of the algorithm.

5.	The EGARCH likelihood can be computed in one of two ways; direct

© James Davidson 2024

nonlinear recursion of equation (6.2) (default) and Gauss-Seidel iteration; i.e., repeatedly solving the linear difference equation obtained by conditioning on $h_t^{1/2}$ in equation (6.2). There may be small differences between the two estimates due to treatment of initial conditions. The main consideration in this choice is one of speed. The nonlinear equation cannot be solved using Ox's vector manipulation capability, and has to be programmed as a loop. This could result in long solution times in large samples, and the iterative method will often be quicker. A convergence criterion is used to control the G-S iterations. The default maximum number (50) should not need changing.

6. The EGARCH specification is driven by the weighted sum of the absolute and signed value of the normalized shock. This is conventionally centred by subtracting the mean of the absolute process, whether specified as Gaussian, GED or Student t. Optionally, this mean can be omitted, allowing the centring to be achieved by the equation intercept. This is a backwards compatibility option, suspending the change made to the EGARCH model specification in release 4.32.08-06-14.

## *8.7 Options / Optimization and Run ...*

OPTIMIZATION SETTINGS: BROYDEN-FLETCHER-GOLDFARB-SHANNO (BFGS) ALGORITHM

1. The options Maximum Number of Iterations, Print Frequency, Strong Convergence Criterion and Weak Convergence Criterion are set to their Ox defaults with the value -1. They will not normally need to be changed. Print Frequency shows the optimand gradient at the current points, as BFGS optimization proceeds.

2. Due to rounding error, the efficiency of the search algorithm can be sensitive to large differences in scale between parameters. This can be due to inappropriate units of measurement, for example. Setting the Parameter Rescaling Factor to 1 adjusts parameters to have a similar order of magnitude, based on the starting values. Hence, this setting alone will not have any effect with default starting values. However, setting to a value greater than 1 scales all parameters up, to limit rounding error. The value of 10 appears to improve the performance of the algorithm in at least some cases. However, no verified guidelines can be offered. Experiment! (and report your findings, please.) To disable rescaling (the default) set the factor to 0.

3. Check the "Use Numerical Derivatives" checkbox to force the use of numerical derivatives. TSM uses analytic derivatives by default to compute the BFGS step, when these are available. (Analytic derivatives are not implemented for some models.) In the event of a convergence failure, TSM compares analytic and numerical derivatives and switches to the latter automatically if a discrepancy is found. This option is saved as part of the model's specification.

SIMULATED ANNEALING

4. Optionally, a simulated annealing algorithm can be run for a fixed number of iterations to provide starting values for BFGS. This option is selected

© James Davidson 2024

by default. By choosing a large number of iterations, SA can be effectively made to act as the search algorithm proper. However, given good starting values BFGS should be much faster, so a compromise is probably best.

5.  The other three settable SA parameters are the initial temperature, the temperature reduction factor, and the number of iterations before temperature reduction. None of the defaults are necessarily best. If the number of iterations before temperature reduction is set <= 1 it is replaced by the default of max(100, 5*(#parameters)). This setting appears as .NaN in the text field.

OPTIMAND CONVENTION

6.  Two of the estimators available (least squares and GMM) are conventionally regarded as minimizing a criterion, while the ML options are maximands. This can lead to confusion, especially when interpreting the model selection criteria. The "optimand" for least squares and LGV is reported for comparability as the Gaussian log-likelihood (a monotone-decreasing transformation of the sum of squares) but the GMM optimand has no such counterpart.  By default, all the estimation criteria are treated as maximands, (so, -GMM criterion). The 'Report Minimand' option reverses the signs (so, -Log Likelihood). The important issue is that the selection criteria are always signed in the same way as the estimation criterion, and so are to be minimized (maximized) if this option is checked (unchecked).

NUMBER OF GRID POINTS

7.  Enter an integer value here to choose the density of points computed in a grid plot. If N is set, then N+1 points are evaluated in a 1-dimensional grid, or (N+1)^2. points in a 2-dimensional grid. Note that grid plotting must be enabled in Values / Equation, and the run launched with Actions / Plot Criterion Grid. Set the grid bounds(s) in the relevant Values dialog(s).

SUPPRESSING THE OUTPUT

8.  Optionally, computation of estimation outputs including residuals, covariance matrix, test statistics and forecasts can be skipped. This option may be useful in big models where these calculations are time-consuming, and the user is performing a sequence of restricted optimizations. (This can be a good strategy for optimizing poorly conditioned criterion functions.) The best parameter values attained are stored and can be inspected in Values. The full outputs can then be computed using the "Evaluate at Current Values" command, if required. Likewise, if the search algorithm fails to converge or is interrupted by the user, the output at the attained point can be evaluated in the same manner.

RUNNING AN EXTERNAL PROCESS

9.  The option of running estimations as independent processes from the command line allows the TSM GUI to remain under the user's control for other tasks while a long numerical optimization is performed.

© James Davidson 2024

It also allows two or more jobs to run be concurrently. Since Ox Console is not multi-threaded, this may be the best way to make full use of a multi-core processor.

10.  If the option "Run Next Estimation as External Process" is checked before launching an estimation run, the calculations are run in a new instance of Ox, running in a console window. A temporary executable Ox file is created to be run in batch mode. The usual estimation results are written to a text file identified by the run number. Import this text to the results window using the command File / Load Text File.

11.  The option is cancelled once the run is executed. Reset it in this dialog as required.

12.  If the specifications are stored as a named model in Model Manager before running the new process, the associated values and listings can also be imported. A ".tsd" data file is created with the usual outputs. Import the data by RE-loading the model in Model Manager. The estimates can also be loaded for further processing by the command Restart / Load Text Input. The name and path of the file to load is given in the results window.

13.  If the option "Delete Temporary Files on Exit" is checked in Options / General, the temporary code and listing files created by this option are deleted at closedown. Results files are not removed, and must be deleted by hand, as required.

RUNNING A CONDOR JOB

14.  On a network equipped with the Condor HTC system, estimations can also be run on otherwise idle machines attached to the network, and the results returned to the TSM work-station, "as if" run as local external processes. See Appendix H for more details on the use of Condor.

## 8.8 Options / General Options ...

INTERFACE SETTINGS

1.  Dialogs are closed when an estimation run is started. Checking the option "Restore Dialogs Automatically" causes them to re-open in the same positions at the end of the run.  Dialog positions are stored in "settings .tsm", so the last configuration in the previous session will also be restored at start-up. If this option is not selected, dialogs can be restored manually using the Actions / Restore Dialogs command.

2.  The positions of dialogs on the screen are remembered during a session, but by default are lost when the program is closed. They are positioned centrally when first opened. Checking "Remember Dialog Positions" has the program remember the positions between sessions. This allows the user to maintain a preferred screen layout, but note that if the monitor resolution is changed, they could become invisible. In this event, use the command Actions / Reset Dialog Positions to re-centre them.

© James Davidson 2024

3. By default, the user is prompted to save data/results before over-writing them or closing the program. Deselecting "Enable Saving Prompts" suppresses these reminders.
Note: The prompt to clear model specifications when loading a data set can be optionally skipped (select "Always"). The prompt is re-activated by deselecting and then reselecting this option.

4. "Automatically Close Options Dialogs" is selected by default, and has the effect of closing any options dialog currently open when another is opened. This saves a proliferation of open dialogs on the screen, but can be disabled if it is desired to work on two or more options dialogs at once.

5. TSM output can be echoed to the console as well as appearing in the results window. To enable this behaviour check the "Echo Results to Console" box. Output from Ox not under the control of TSM, such as error messages, will also to appear in the console window.

6. Tool tips appear by default when the mouse pointer is passed over the tool buttons. These can be optionally switched off.

7. The default font size of 12 points should suit most monitors, but can be changed. This setting also effects the help pages, whose widths are increased to accommodate the larger text. The results window is resizable by dragging the corner with the mouse.

8. The size of the fonts and icons is set at installation, but can be changed (without administrator privileges being required) by entering a percentage in the text box. A program restart is required to effect the change. Zero is the default and any entry smaller than 100% results in the installation setting being used. 16x16 toolbar icons are used with sizes < 130%, 32x32 icons with sizes < 200%, and 48x48 icons with still larger sizes. Sizes exceeding 150% are unlikely to be useful except on high-resolution displays such as the Microsoft Surface Pro tablet.
Notes:
* The option is implemented by placing a file "framesize.txt" in the working directory to be read at start-up.
* The results window font (see 7 above) is relative to this setting and does not need adjusting separately.

9. The results window background can be either white or tinted, to make the TSM window easier to locate on a crowded desktop, or simply easier on the eye. The button cycles through five preset choices (white, red, green, blue, gray) and a user-selectable color. The latter (yellow by default) is set by editing the file tsmgui4.h and changing the RGB values in the function "Get_TextAreaBG()".

OPERATIONAL SETTINGS
10. Saving model settings in Model Manager can always be done manually

but the option "Save Model After a Run" saves model settings and results automatically after each run. Enabling this option allows an earlier estimation run to be re-created instantly, by re-loading the model.
NOTE: A maximum of 100 models can be stored at once. Use this setting sparingly to avoid a cluttered results folder!

11. By default, the program's state is saved to the disk dynamically as work proceeds. Data, listings, charts results are saved to a file called (by default) "settings.tsd". The contents of this file from the previous session are reloaded at start-up. In particular, the current data set is reloaded intact even if the spreadsheet file from which it was copied has been moved, or deleted. Temporary changes or additions to the data set (such as saved generated series) only need be saved to the original file by choice. Optionally this feature can be disabled, but this should only be desirable if the data set is very large.
Note: The "settings.tsm" and "settings.tsd" files are paired. The latter file can be opened only when the former file holds a record of the date/time it was saved.

12. The contents of the results window are saved to disk during a session, but by default are only reinstated following a forced or abnormal shutdown. Optionally, these can be restored routinely. Note that this will result in the results window buffer filling up. Once the buffer is full, the earliest results are discarded to make space The buffer size can be changed by editing the file tsmgui4.oxh. The default setting is 0.5Mb. Enlarging it will slow the operation of the program and produce large settings files. Note that the window contents also can be saved continuously to a separate text file, see File / Results / Enable Background Saving. This may be the preferred method of retaining a record of work.

13. Enabling the option "Delete Temporary Files on Exit" will cause .TSD (model results and data) files to be deleted automatically at the end of a session. Uncheck this if you wish to keep your model results, or check it to avoid excessive clutter in your results folder.

14. Data descriptions are saved as extensions to the series names, following a delimiter character (@ by default). These extended names may not be compatible with other programs. Set the option "Omit Data Descriptions" when exporting data for use in other packages.

15. The Files / Settings / Export command saves a complete image of the current session, including data, for export to another installation. If user-supplied Ox code is loaded, this can optionally be bundled in the export file. Then, loading the settings re-creates the files needed to run the supplied code at the target installation.

16. If the Enable Error Recovery option is checked (the default) the program restarts automatically in the event of a crash (Ox error). Note that the Ox error message appears in the DOS console in this case.

### LINEAR REGRESSION MODE

17.     To simplify the user interface for teaching purposes, the options for nonlinear estimation by numerical methods can be disabled, and the toolbar buttons hidden. For this option, uncheck "Enable Optimization Estimators". A restart is required to implement this selection.

### DEFER EXTERNAL JOBS

18.     Check this option to create batch jobs (estimations or Monte Carlo jobs) as executable Ox files, for starting manually. Use this option to run the job on a different Ox installation, or to defer running it until a convenient time. When this option is unchecked (the default) the job is started immediately the Run command is given.

### LOAD MULTIPLE DATA SETS

19.     By default, multiple data sets can be loaded simultaneously. Data are reloaded at start-up and are not removed from memory except by executing one of the Clear commands under Setup / Datasets. This feature can optionally be disabled, so as to limit the amount of memory used by an instance of TSM, or just to simplify the operation of the program.
A restart is required to implement this selection.

### ASSIGNABLE TOOLBAR BUTTONS

20.     Up to 6 toolbar buttons, showing the "Dialog" icon and a number, can be assigned to run any command and open any dialog that does not itself have a dedicated button. To assign a button, first run the menu item you wish to assign. Use the pull-down menu to select a button for assignment, and then click "Assign to last Dialog Opened". The new button appears on the toolbar. To remove the assignment, click "Clear".
Note:
   o     To assign an Options dialog to a toolbar button, first deselect "Automatically Close Options Dialogs" under Interface Settings. This allows both Options / General and the other dialog to be open together.
   o     With Tool Tips activated, view the current  assignment by hovering the mouse pointer over the button.
   o     Buttons 2-6 are removed from the toolbar if unassigned. Button 1 is set to the default of re-opening the last dialog opened.

### SPECIAL SETTINGS

21.     A pull-down menu gives convenient access to a range of parameters that control the performance of various tests and other procedures selectable in the program. These are gathered here to reduce the clutter in other options dialogs, since they are changed only infrequently. This facility is designed to be expandable as more tests and diagnostic procedures are added to the program in the future.

22.     If a parameter take a fixed number of discrete values, such as "TRUE" and "FALSE", simply double-click the text box to change the setting, or cycle through available values when there are more than 2. Otherwise, type

© James Davidson 2024

numerical values into the text box. Passing the mouse pointer over the box registers the new value. Be careful not to move the mouse while typing the entry.

23.    In "Initialize Run ID", enter a nonnegative numerical value to reset the run counter to that value. The change comes into effect the next time "settings.tsm" is saved (e.g., the next time an estimation is run). The entry displayed then reverts to the default value, "OFF". If the next action following this one is to give the command File / Settings / Export..., the run counter is set to that value in the exported settings, but the current counter is left unchanged. (By default, the counter is set to 0 in an exported file.)

24.    Other special settings may be explained in their context. At any time, some settings may relate to experimental program features and hence undocumented. The following are settings not documented elsewhere:

"Show CS Test Settings"
If set to TRUE, settings for the suite of consistent specification tests (including the Bierens test) are shown, here and also in the Options / Tests and Diagnostics / Diagnostic Tests dialog. By default these items are hidden to simplify the interface.

"Automatic Data Transformation"
Enable the option to let data series be transformed for inclusion in a model. (See 4.0 for details.)

"Annual Difference Forecasts"
Forecasts are reported as the percentage change from the observed data in the previous year (4 quarters or 12 months back). In this case, a maximum of 4-quarters or 12-month forecasts can be computed.

"Plot with GnuPlot 4.2 (restart)"
If the full Gnuplot package (version 4.6.3 or later) is installed, it will be used to create graphics in place of the compact bundled version, Gnuplot 4.2.6. Setting this option to TRUE forces TSM to use the bundled version even if later versions are present. This option must be set before showing a graphic, or else a restart is needed for the change to take effect.
NOTE: If this option is set to FALSE, the installed Gnuplot version MUST be v4.6.3 or later, or else plotting will not work!

"Censor MC Replications"
In Monte Carlo experiments, a function coded by the user in Ox can be used to examine the result of a replication and optionally discard it, for example, because parameter estimates take illegal values or otherwise represent a false maximum of the estimation criterion. See Appendix C for details.

"Tabulate p-Values in EDFs"
In bootstrap inference the computed p-value is for some purposes a natural test statistic, having a uniform distribution asymptotically under the null hypothesis. In Monte Carlo experiments on bootstrap tests,

this option tabulates the p-values instead of the test statistics.

"Show Stubs"
When this option is set to TRUE, the function prstub(), included in a
user's Ox function, writes its arguments to standard output, with
spaces automatically inserted between the printed items. This is a
code development tool. Use it to check on intermediate outputs
and diagnose crashes.
Note: this setting is not stored. It must be set each time TSM is started.

# 9. Help

*        "User's Manual" gives access to these pages.

*        "View Text Files" opens additional documentation in ASCII files, for reading only.

*        "Data Descriptions" opens a list of names of the currently loaded variables, followed by any description that has been created for the variable. The description is preceded by the symbol @.

*        "Register this Copy..." opens a dialog to enter the user's name and the unlocking key provided to registered users of the software.

*        "About" opens a box showing copyright information and the registered username.

USER'S MANUAL

1.        Navigate the User's Manual by the menu (sections) and sub-menus (subsections), and also by clicking the "Previous" and "Next buttons to move backwards and forwards through the pages. These pages duplicate the contents of the PDF file "tsmod4ghp.pdf", which is accessible through the Start Menu (Windows).

VIEW TEXT FILES

2.        "File Dialog" opens the file dialog, to select any ASCII file for reading.

3.        "Most Recently Opened" repeats the last selection. By default (if no file has yet been opened) it opens "notes.txt" if this file exists.

4.        'Imported Ox Code' refers to a source code file, with .ox extension, which has been loaded from a distributed settings (.tsm) file. A distributor of coded functions  has this opportunity to document their  code, by including text comments in the file before the code statements. The code itself can also be inspected.

DESCRIPTIONS

5.        Data descriptions can be created in two ways. In Setup / Data Editing and Transformations, the choice widget item "Description" opens a text field where the description of the selected variable(s) can be entered. Alternatively, when the data are prepared for loading in a spreadsheet or text editor, add the description text to the variable name, normally preceded by a @ symbol. The delimiter symbol can be changed by editing the tsmgui4.h file.

6.        Some entry formats, such as .DAT, limit the length of a variable name. Overflowing text will be lost if the data are saved in this format.

7.        "Show Settings Info." displays the header from most recenty opened

       © James Davidson 2024

settings file (with .tsm extension). This contains the TSM version used to create the file, the data and time of creation, and whether the settings were Exported (created with the File / Settings / Export... command and hence portable between installations.

8.    The username must be entered in exactly the format as given with the unlocking key. Changes in (e.g.) spelling, capitalization, embedded spaces, etc., will cause the registration to fail.

9.    The program saves the licence file "registration.txt" in one of the two valid locations for this file. It first tries the TSM Home directory, i.e., [ox home] /packages/tsmod4/. If either the user does not have permission to write in this folder, or if a licence file already exists there, it writes to the current Results folder. This allows several different licences to coexist on the same workstation, if each user maintains a different Results folder.

10.   By default, the Results folder is the "start-in" folder specified in the short-cut used to start the program. However this location can be changed (e.g. by the command "File / Results / Locate Results Folder"), and new location is saved in "settings.tsm". If the settings file is subsequently deleted, it will be necessary to either restore the Results location, or move the licence file manually to the new one

11.   TSM reads the licence file at start-up, and looks for it in the current Results folder first. If it is not found there, the program looks in the TSM Home directory. If a licence file exists in the current Results folder it will 'take precedence', and any licence file stored in the TSM Home directory will be ignored. To "see" the latter registration, change the Results folder.

# 10. Interpreting the Output

1.      The symbols AR1, AR2, ..., MA1, MA2,... are used to label the ARMA coefficients in the output.  The prefixes 'GARCH', 'EGARCH' or 'APARCH', as appropriate, are appended for the conditional variance coefficients in the default case of  the 'ARMA-in squares' parameterization. For the standard (Bollerslev) parameterization, 'AR' and 'MA' are replaced by 'Alpha' and 'Beta', respectively. These notations are also used for the BEKK multivariate GARCH model. Be careful to interpret the latter parameters correctly!

2.      For models specified in the Model / Dynamic Equation dialog, the 'Type' of a regressor appears as a prefix in square brackets, for example, "[1]Var1" denotes that Var1 is a regressor of Type 1. This prefix is omitted if only Type 1 regressors are specified, and also in equations specified in the Model / Linear Regression dialog.
Note: in the latter case the Type only controls the number of lags generated, not the interpretation of the coefficients.

3.      A GARCH regressor is indicated similarly, with an additional "G" prefix. For example, "[G1]Var1" denotes that Var1 is a GARCH regressor of Type 1. The GARCH (or EGARCH or APARCH) intercept of Type 2 ('kappa') is indicated with the prefix '[2]'. However, no indicator is used for the GARCH (EGARCH, APARCH)  intercept of Type 1 ('omega').

4.      To indicate the regime in an explained regime-switching model, the regressor has an additional "Rg" prefix. For example, "[Rg1]Var1" denotes that Var1 is a regressor explaining the probability of switching to Regime 1.

5.      In smooth transition models, the parameters of the single and double transition factors are indicated with the prefix "Sm". For example, "[Sm1]Intercept" denotes the location parameter of the first transition factor.

6.      The bilinear model is described in the output as 'Bilinear-AR(F)IMA(p,r,d,q) where r refers to the selected bilinear order, and p, d and q are as for the standard AR(F)IMA model.

7.      By default, the GARCH  intercept term is estimated in square root form. Correspondingly, in ML estimation of models without a conditional variance component, the error standard deviation is estimated by default. This is for reasons of numerical stability. Particularly in IGARCH models where the parameter can be close to zero, numerical derivatives can be highly inaccurate without a rescaling transformation. The transformation can be selected by the user in the Options / ML and Dynamics dialog. In certain cases, raising the order of root from 2 to 4 may be found to aid

convergence, while in others 1 will work satisfactorily

8.    Similarly, the parameter of the Student t distribution can optionally be estimated as a root of the degrees of freedom. The square root is the default. See the Options / ML and Dynamics dialog to set alternative values. This is again for reasons of numerical stability. The parameter is both bounded below, and infinite in the Gaussian case, and extreme values are potentially difficult for numerical differentiation. By setting the root to a negative value, the inverse transformation can be applied, such that the parameter is zero in the Gaussian case.

STANDARD ERRORS, TEST STATISTICS AND P-VALUES

9.    Every estimated parameter is shown in the output with either its estimated standard error , or the words "Fixed" or "Solved" when the parameters have been so designated in the Values dialogs. Following the standard error, most parameters show the t ratio and nominal p-value of the significance test. The degrees of freedom are calculated as "sample size - (total model parameters)/(number of equations)".

10.   Parameters for which there is no natural "zero hypothesis" to be tested, such as the variance, GARCH intercept, Student's t degrees of freedom, etc., are shown with standard error in parentheses but no p-value.

11.   Nominal p-values are also given for quoted test statistics, in many cases asymptotic chi-squared statistics. For tests with non-standard distributions (Dickey-Fuller, KPSS, R/S) these are computed from the available tables. Since the coverage is incomplete, these results are shown as inequalities. Thus, {<0.05} would denote that the p-value is less than 0.05 (leading to rejection at the 5% level) according to the table. The result {<1} simply indicates that the statistic falls within the bounds of all the tabulated points. The tables themselves are given as Appendix I.

.
12.   If standard errors and/or test statistics print as NaN, (stands for "not a number") this means the Hessian matrix of the optimization criterion could not be inverted at the current point. Typically, this means the local optimum is not unique, which will also cause convergence failure. Lack of identification is a leading cause; for example, a Markov switching model with a zero-probability regime. When Print Covariance Matrix is selected in Options / Output and Retrieval Options, the Hessian matrix itself is printed in this case, as a diagnostic aid. Checking for zero rows and columns could indicate the source of the problem.

SELECTION CRITERIA AND SUMMARY STATISTICS

13.   By default the model selection criteria (Schwarz, Hannan-Quinn and Akaike) are defined as "maximand less penalty". Therefore, larger values are preferred. This convention can be optionally reversed, see "Report Minimand" in Options / Optimization and Run.

14.   "R-Squared" is the squared correlation coefficient between the dependent variable Y, and the model predictions ( = actual Y - residuals). Therefore, it is well-defined for any model, and must lie in [0,1]. However, whether it

is a useful measure of model performance depends on the model - it is of no use in assessing GARCH models, for example.

15. "R-Bar-Squared" is calculated as 1 - (1 - R-squared)*(T-1)/(T-p) where p is the number of parameters fitted in the model.

16. When models are estimated by Instrumental Variables from the Linear Regression menu, the Sargan and DWH (Durbin-Wu-Hausman) statistics are given. These are asymptotic chi-squared statistics under H0, with degrees of freedom indicated, and nominal p-values shown in braces. The Sargan statistic is N times the ratio of the GMM minimand to the residual sum of squares, where N is sample size. This tests validity of over-identifying restrictions. The DWH statistic tests for exogeneity of the regressors and is N*(RSS0 - RSS1)/RSS1 where RSS0 is the sum of squares from the OLS regression, and RSS1 the sum of squares from the regression including the reduced form residuals for the included endogenous variables.

CONVERGENCE ISSUES

17. All estimations specified in the Model / Dynamic Equation dialog are done numerically. The optimization routine is the Ox implementation of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton procedure. This method reports strong convergence in a high proportion of cases. However, don't assume that the message "Strong convergence" means that the estimation criterion has been globally optimized. It simply means that a local optimum has been found. Restarting the search algorithm from different points is the only way to confirm a global optimum.

18. The messages "Weak convergence" or "No convergence" may also be returned by the search algorithm. A run may also terminate with the message "Estimation failed! Algorithm returns NaN or +/- infinity". This typically indicates overflows in the floating-point calculations, e.g. dividing by zero or attempting to take the logarithm of a negative number, probably due to inappropriate starting values. There is an option to have optimization restarted automatically at the default parameter values, if this occurs. This will often clear a problem with bad starting values.

19. Analytic derivatives are implemented for most models since these are faster than numerical approximation, but the algorithms can fail in certain situations. In the case of convergence failure from this cause, the iterations are restarted with numerical derivatives enabled, and a note to this effect appears in the output. This setting is saved with the model in question. Numerical derivatives can also be selected manually in Options / Optimization and Run.

GRAPHICS FOR REGIME SWITCHING MODELS

20. For simple Markov and Hamilton switching models, the M-1 series of independent filter probabilities can be listed and/or plotted. These are the probabilities of the regime being occupied conditional on information to date t, as used to weight the likelihood terms in estimation. In the case of Hamilton's model, these are summed over lags (p) to reduce $M^{p+1}$ series down to M.

© James Davidson 2024

21.    For the explained switching model, the M-1 series of variable switching
       probabilities are also plotted. When regime dummies have been specified
       there are up to M-1 different series for each regime, although differing
       only by the same offset at each t. These are averaged to create the series
       plotted.

22.    For the smooth transition model, the regime weights are plotted (logistic
       functions of the transition indicator).

BOOTSTRAP AND SUBSAMPLING P-VALUES

23.    Resampling computing  p-values for tests by tabulating the statistics for
       a "pseudo-true" model, where the statistics are centred using the relevant
       components from the sample counterparts. This method can be used for
       t-ratios and Wald tests, and also M and CM tests. It is not so easily
       implemented for LM tests, although diagnostic tests for i.i.d. residuals
       are valid by construction in the bootstrap context. Other LM tests,
       and other tests are given their conventional asymptotic p-values. Note
       that resampling p-values are always indicated with a  '*'.

24.    Resampling p-values are computed for the Jarque-Bera test only for
       the bootstrap using Gaussian shocks. Otherwise the hypothesis is true by
       construction, so the bootstrap p-values are uninformative.

25.    The cointegration tests (ADF and Phillips-Perron) cannot be resampled
       since there is no way to set up a pseudo-true version of the null hypothesis.
       Note that in single-equation regression models, the regressors (except
       lagged dependent variables) are held fixed in repeated samples. In
       cointegration models this is an unrealistic assumption even in large samples.
       It is better to consider a multi-equation model in which the unit root
       processes are generated endogenously. Monte Carlo experiments can
       be used to evaluate inference procedures in this case.

© James Davidson 2024

# 11 Directories and Files

DIRECTORIES (Windows folders):

Start-in        Must be specified by the operating system when TSM4 starts.
                This is usually set by the short-cut used to start the program,
                Options are presented by the installation program, or it can
                be set manually by editing the short cut. Otherwise, by default
                it is the directory where the Ox file used to start  the program
                resides.
                Notes:
                1.      On networks or shared systems, the user must have write
                        permissions in the Start-in directory. Write permissions in the
                        TSM Home directory ([ox-home]/packages/tsmod4/) are not
                        required. It is NOT recommended to designate TSM-Home as
                        the Start-in directory.
                2.      In the Windows installation of TSM, double-clicking on a .tsm
                        file in Windows Explorer starts the program with the settings in
                        question. The Start-in directory is one where the .tsm file
                        resides. If the file is selected from a Start Menu the Start-in is
                        the user's registered home directory.

Results         Selected by the user as an alternate location for program outputs.
                By default, it is the same as the Start-in directory. Change the results
                folder in one of two ways:
                *       By the command File / Results / New Results File ... The
                        folder containing the chosen file becomes the new results
                        folder.
                *       By the command File / Folders / Results... Browse to the
                        required path, or type in the text field.

FILES

These files may be created by TSM4 in either the Start-in or Results directory.
(S) indicates that a file can only reside in the Start-in directory. Otherwise, it will
be created in the Results directory if that is different. (C) indicates that these file
names can optionally be changed for a particular installation, by editing the file
tsmgui4.h.

settings.tsm (S,C)      Contains all the user's settings and specifications,
                        including the most recent estimates, and stored model
                        specifications. Updated automatically throughout a TSM
                        session.

settings.tsd (S,C)      Optionally, this file stores the current data set results, listings
                        and graphics, and is updated automatically when these are
                        changed. Enabling this option (see 8.8 Options / General /
                        Operational Settings) allows the complete working
                        environment to be saved from one session to the next.

                Note:   'settings.tsm' and 'settings.tsd' are paired by a time stamp.
                        The latter file cannot be opened except in conjunction with its '.tsm'

counterpart. The data and estimation outputs stored in the latter file are always matched with the model specifications stored in the former.

| | |
|---|---|
| registration.txt | Contains the user's licence code. By default, this is written to the TSM Home directory, subject to the user having write permissions, and there being no pre-existing licence file there. Otherwise, it is written to Start-in directory or, if this also contains a licence file, to the current results folder. This allows multiple licenced users to share the same installation.<br><br>NOTE. The program looks for registration.txt in reverse order, first looking in current Results directory, then in Start-in if this is different, and finally in TSM Home. The 'registered user' for the current run is therefore the user who starts the program. |
| tsmod_run.ox (S,C) | User-editable Ox file, used to start the program incorporating the user's own code. The program can be started by running this file in OxEdit.<br>(See Appendix A on changing module settings.) |
| usercode.ox (S,C) | This is the standard container for user's Ox code, although it is often convenient to have it contain just #include "..." compiler directives referring to files of code supplied by the user. |
| errchk (S) | Written by TSM4 at startup and deleted at closedown. Used to signal an error termination. |

FILE TYPES

| | |
|---|---|
| *.tsm | Named settings file with .tsm extension, same format as settings.tsm, saved by the user. Can be loaded manually to overwrite existing settings. When created using the File / Export... option, data and listings are bundled with the settings. The corresponding files are recreated automatically when the file is loaded, in this case. |
| *.tsd | Contain specifications, values, graphs, listings and (optionally) data associated with models stored using the Model Manager. The file name is the same name as the model name. The specifications and values are also stored in the settings (.tsm) file, and since the other content can typically be recreated by re-running estimations, they may be treated as temporary. They will be deleted when the model is deleted, and also at closedown, if the general option "Delete Temporary files on Exit" is checked. However, their contents can also be imported directly, so that models can be passed from one installation or project to another. |
| *.txt | The program creates various files in standard text format, including results files (images of the results window text) |

|                     |                                                                                                                                                                                                                                                                                                                                                                    |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                     | and settings files created with the Settings / Save Text... command. These files can be view and edited in Windows Notepad or other word-processing programs.                                                                                                                                                                                                        |
| *.xls, *.xlsx, *.csv | Microsoft Excel spreadsheet format is the default format for saving data sets, forecasts, and other listings. Other options available in the Options / Output and Retrieval dialog. *.csv files are text files, have the same capabilities as the Excel 2 spreadsheets created by Ox, and are widely portable.                                                         |
| *.ox                | Text files containing Ox source code. They can be executable, such as tsmod_run.ox, but can also be included in other .ox files using the compiler directive ' #include "..." '. In normal use, a user's code files should be "#include"d in the standard code file usercode.ox. which is in turn #included in the executable tsmod_run.ox.                            |
| *.plt               | Graphics file produded by Gnudraw, formatted for input to Gnuplot. These files are deleted by default after Gnuplot has processed them, but are optionally saved in the Results directory for further processing by the user. They are ordinary text files and can be viewed and edited in Notepad, OxEdit and similar programs.                                       |
| *.png               | The default bitmap graphics format. This format is read by a wide range of commercial graphing and word processing software, and converted to formats such as bitmap, Jpeg, Tiff, etc.                                                                                                                                                                               |
| *.eps               | Default vector graphics format. Can be read by various public domain software, and converted to .wmf, .emf  and other formats by most graphics packages. Can also be inserted in Microsoft Word documents.                                                                                                                                                           |

TSM saves data and graphics in a variety of other optional formats. See 2.2 File / Data..., 8.1 Options / Output and Retrieval and 8.5 Options / Graphics for details.

# 12. Hints and Tips

Here are some non-obvious pointers for getting the best results from the program. Suggestions from users for further tips are always welcome.

NUMERICAL OPTIMIZATION

1.     TSM's optimization algorithm uses analytic derivatives, by default, wherever these have been coded. A few of the less frequently used models including the GED likelihood, EGARCH, DCC and BEKK GARCH, explained regime switching and Hamilton's model, still depend on numerical derivatives. Numerical derivatives can be selected optionally as a backstop, in case an unforeseen problem is encountered - see Options / Optimization and Run... The option to use numerical derivatives is saved with a model. The option is enabled automatically if TSM itself detects any problem with the coded formulae.

2.     Whatever optimization method is used, success can greatly depend on the good choice of starting values. Some packages implement elaborate schemes to choose starting values automatically for specific models, such as ARIMAs, but this type of strategy is difficult to combine with a wide range of models. TSM implements a number of basic strategies, but these cannot entirely replace the user's initiative in exploring the parameter space to achieve successful optimization.

3.     A limited number of steps using the simulated annealing (SA) algorithm are optionally taken with a view to getting good starting values for the BFGS gradient-based algorithm. This option is selected by default, but switching it off may speed things up if the estimation criteria are well-conditioned. SA uses randomization to decide on successive steps, and with a multi-modal criterion function, it is possible for successive runs of the search algorithm to converge to different points, from the same initial values. Running SA for a limited number of steps before starting BFGS has the effect of randomizing the BFGS initial values. Repeating an optimization run from different starting points or with different algorithm settings is, in any case, always a sound strategy.

4.     The program makes a maximum of two attempts to perform an optimization, where the second attempt uses the default starting values. If both attempts fail, there are several program settings that may assist.

       *     Enable the parameter rescaling option in Options / Optimization and Run... Set the value to unity to limit the ratio of largest to smallest parameters (effective only if starting values are a guide to the relative magnitudes). Setting to a higher value (e.g.10) inflates all parameter values, to reduce the risk of rounding error in the evaluation of numerical derivatives.

       *     Tinker with the simulated annealing settings in Options / Optimization and Run... There are not many useful guidelines for good choices here; in most cases raising the number of iterations cannot hurt, but also try

          © James Davidson 2024

switching it off (see Note 3).

* The scaling of the variance/intercept parameter in GARCH models can be critical, especially in the IGARCH case. See Options / ML and GARCH... for details.

* Experiment with the settings for EGARCH and GARCH-M models in Options / ML and Dynamics...

* For dynamic parameters (AR, MA etc.) experiment with the bounding option in Options / Optimization and Run...

None of these fixes is guaranteed to improve matters, but setting starting values to good guesses of the estimates is always helpful. For example, if it is believed that a unit root exists in ARMA (or GARCH) models, start the AR parameter at 1, rather than 0. Using the "Impose Unit Root" option in Model / Dynamic Equation may help similarly.

5. Another good strategy with a complicated model is to simplify it initially, by deleting or fixing some components. Estimate the simple model, then add back the deleted components one at a time. The estimates of the simpler model often provide the best starting values for those parameters in the more general one. This 'stepwise' approach to the optimum is often successful when an attempt to optimize the full model from the default initial values ends in failure.

6. As a last resort, try keeping a parameter away from illegal values by imposing inequality constraints. Set them in Values / ... after checking the checkbox for Parameter Bounds. This won't work if, for example, the optimum is actually on the boundary of the feasible region. In this case, simply fix the parameters in question at the boundary values. Conventional inference won't be possible on these parameters.

7. Switching regime models (Markov switching and smooth transition models) generally exhibit multiple maxima of the likelihood. Good starting values, and careful experimentation with different starting values, is strongly recommended. As a switching probability tends to 0 or 1, the likelihood becomes flat due to the logistic mapping employed, and the algorithm can get stuck at these points. Restarting with "plausible" probability values (such as 0.9, 0.1 in the 2-regime case) often succeeds in this situation.

CHOICE OF SYSTEM ESTIMATORS

8. Don't overlook the distinction between the Linear Regression and Dynamic Equation options. The latter always uses numerical optimization to compute the estimator, even when a closed-form solution exists. Large unrestricted VARs may be computed in the Linear Regression dialog using SUR. Exclusion restrictions can be imposed by "fixing" parameters, although they can be tested only by Wald tests on the unrestricted model.

9. There are two procedures to estimate a Gaussian linear equation system

© James Davidson 2024

by maximum likelihood. Least generalized variance (LGV) optimizes the concentrated log-likelihood function (the log-determinant of the covariance matrix). The "Gaussian ML" option estimates the log-likelihood as a function of all parameters, including the variance matrix, and hence is much more computationally intensive than LGV. This method is only required for models with nonlinear features such as ARCH/GARCH or Markov switching.

USING THE RESULTS WINDOW AS A TEXT EDITOR

10.    Text can be typed or pasted freely into the window. To paste from the clipboard, right-click in the text area and choose Paste. To save text in a file, first name the file with File / Results / New Results File. Then select the desired text with the mouse and give the command File / Results / Save Selected Text. Note that subsequent saves will be appended to the same file. To avoid this behaviour, give the New Results File command again to create or reinstate a results file.

ENTERING DATA BY HAND

11.    Use one of the following methods to type data directly into the program,. The second method is recommended for entering several variables at once. Don't forget to save the data set afterwards, before doing anything else.

12.    Use the Setup / Data Transformation and Editing dialog to create a new variable. If no data are currently loaded, give the command Edit / Re-Size Sample and enter the desired sample size, in which case a new variable Zeros is created automatically. Otherwise, give the command Make Zeros. Use Rename to change Zeros to the name of the new variable. Then choose List /Edit, select the first observation, and press Edit Observation. To enter successive observations, type the values into the edit field and press "Enter", repeatedly.

13.    A sequence of observations on a variable can also be pasted into TSM using the Windows clipboard. The values can be columns in a spreadsheet, or typed into a text editor or word processor. In the latter case, spaces, tabs, commas and carriage returns can all be used as separators.
To copy the text in the source application, the usual routine is to highlight it, then right-click and choose 'Copy' from the context menu, or issue the keyboard command Ctrl-C. Then open the List/Edit dialog as above, select the first cell in the target range, and either press "Paste Clipboard" or give the keyboard command Ctrl-V. If the data have been typed in a text file by variable with observations in columns (spreadsheet format), read the file into OxEdit and select a single column using the right mouse button.

14.    The last paragraph describes a feature implemented only in the Windows version of TSM. An alternative procedure is to load a text file directly. In this case, type the data as a matrix with observations in rows and variables in columns, spreadsheet style, either into a text editor or into the results window as described in 11 above. The first line should contain two integers: the number of observations (= lines to follow the first) and the number of variables (= values on each line). Separate entries on a

© James Davidson 2024

line with a space, and terminate each line with [Enter]. Save the file with extension .mat or .txt. Next, select either File / Data / Open or File / Data / Merge, as appropriate. The new variable(s) are named Var# by default, where # denotes the column number(s). Give new names using 'Rename', and optionally enter descriptions using 'Description' in the Setup / Data Transformation and Editing dialog. Before doing anything else, save the data in a spreadsheet format.

MAINTAINING AND EXPORTING PROGRAM SETTINGS

15. To save a model specification, including data file name, estimator, sample size, parameter values, test options, etc., open the Model Manager ('double-M' button) select 'Store Current Model' and supply a name (the current run number is the default). This facility allows you to work on different projects in parallel during a session. You may have a complex multi-equation model set up, and need to do some exploratory regressions or swap data sets. Saving the model allows you to reload the settings later, using the 'Load Model' command in the Model Manager. Your stored models are saved with the other settings in the settings.tsm file.

16. To keep a permanent copy of your favourite operational settings, save them in a backup file with a name such as "mydefaults.tsm". If you then need to change the settings for a particular run, they are easily restored by the command File / Settings / Open, and selecting the backup file. (Note that restoring settings does not reset the run ID counter.)

17. Named settings files are also useful in a teaching context. Students can be supplied with a ".tsm" file containing a desired set of defaults and also stored models for estimation or simulation exercises. The students can easily undo any changes they have made by reloading the defaults file.

18. By default, settings files are saved with local path information to locate data and listings files, which may make them unsuitable for transfer from one installation to another. Use the "Export" option in File / Settings / Save As... to omit this local information. The search path for all files will then be set to the "Start-in" directory when the settings file is loaded. In addition, the currently loaded data sets and model outputs (.tsd files) are bundled in the exported file, and unbundled to the "Start-in" directory when the file is subsequently loaded. For example, a class exercise with preset model specifications and data can be distributed to students in a single exported ".tsm" file.

19. To over-ride the default behaviour of setting the run ID number to zero in a new settings file, go to Options / General / Special Settings and enter the desired number under "Initialize Run ID". This number will be instated at the next estimation run. Alternatively, add the line " RUN_ID = ***; " to a text file of settings before loading them as described in Appendix E.

DOING SIMULATIONS

20. A star feature of the program is the ability to stochastically simulate any model that it can estimate, using either randomly resampled residuals or computer-generated variates as shocks. The simulation module is the

© James Davidson 2024

basis for nonlinear forecasting, bootstrap testing, and Monte Carlo experiments. To gain insight into the properties of an estimator, it is always a helpful exercise to apply it to artificial data from a known model.

21. To create an artificial sample, do the following
    1) Clear the data (File / Data / Clear).
    2) Set a sample size. In Setup / Data Transformation and Editing, choose Edit, scroll down to 'Re-Size Sample', enter a value, and press 'Go'. A variable called Zeros with the specified number of observations is automatically created. (It can be renamed.)
    3) Open Options / Simulation and Resampling, and choose a shock distribution, one of Likelihood Model, Gaussian, Stable or Formula. See Section 8.4 for details on setting these options.
    4) Set the model specifications in the usual way, selecting Zeros as the dependent variable. Pre-sample values are used to start a dynamic simulation and, if desired, these can be inserted in Zeros in Setup / Data Transformation and Editing.
    5) Enter parameter values in the relevant Values dialogs. (Don't forget variance parameters in the case of 'Model' shocks.)
    6) Open the Model Manager and store your specifications under a suitable name.
    7) Select Actions / Simulate Current Model, and add the artificial data to the data set if desired.
    8) Alternatively, use your model to set up a Monte Carlo experiment. Note how a different model can be used to estimate from the simulated data, to permit misspecification analysis.

22. Whilst the simulation code has been carefully checked, there are so many combinations of model features available that it is not possible to guarantee that they will all work together as intended. To ensure the integrity of simulation results, it is a good idea to run a check on the selected model. Users may also wish to check out their own supplied code.
    To verify that a model is being correctly simulated, follow these steps:
    1) Evaluate the model with a data set and chosen or estimated parameters, to generate a set of residuals.
    2) Open Options / Simulation and Resampling, and select 'Bootstrap' for the Shock Distribution. In the box Random Number Seed, enter the value -1. This setting causes the simulation to be done with the actual residuals, without random resampling.
    3) Run Actions / Simulate Current Model and save the simulation. This should match the original series exactly apart, conceivably, from small rounding errors. Compare the series in Setup / Data Transformation and Editing.

    4) In a Markov switching model, the simulation involves random regime switches. To disable these, put the switching probabilities to 0 or 1 in the simulation. In the case of explained switching models, choose parameter values large enough to ensure that the switching probabilities are either 0 or 1 for nearly all values of the explaining variables.

© James Davidson 2024

Don't forget to reset the random number seed afterwards! Should a flaw in the simulation code be suspected, please notify the author.

23.  Double-clicking on a TSM settings file, with extension ".tsm" and red TSM icon has the same effect as starting the program from the Start menu and then loading the settings with the command "File / Settings / Open...". The "Start-in" directory is in this case the directory where the target TSM file resides.

24.  Any data file having a supported format can be selected in Explorer and loaded into TSM. In this case, right-click the file with the mouse and select the option "Open with ..." from the context menu. Select the "Choose Program..." option, and choose "start_tsm" (with the Windows batch file icon) from the list.
     If this option does not appear in the list, choose "Browse...", navigate to the TSM_HOME directory and select the file "start_tsm.bat" from the file list. It should only be necessary to do this once.
     Notes:
     1)   If the option "Always use the selected program to open this kind of file" is checked, simply double-clicking on the file icon will in future load the file into TSM. Only choose this option if you don't edit this type of file in your spreadsheet program by default..

     2)   The directory containing the data file is used as the Start-in directory. If it contains a "settings.tsm" file, the settings are loaded from there. Otherwise, the defaults are set and a new "settings.tsm" is created.

RUNNING SEVERAL INSTANCES OF TSM

25.  To run lengthy optimizations or Monte Carlo runs, without tying up the program interface, try the option for exporting batch jobs to be run from the command line. For details, see BATCH JOBS under 6.1 Actions / Run Estimation and RUNNING AN EXTERNAL PROCESS under 3.10 Setup / Monte Carlo Experiments. When these jobs terminate, their outputs are written to text files with the run number. Give the command File / Load Text File to write the contents to the results window in the usual way.

26.  By default, a batch job is launched automatically to run concurrently with the GUI. This is the efficient way to make use of a multi-core processor. Alternatively the run (or runs) can be postponed to a time when the machine is otherwise idle. In this case they must be launched manually, as regular Ox programs.

27.  It is also possible to run two or more instances of the GUI concurrently. The only major problem to be avoided in this case is the different sessions interfering with each other, by attempting to access and write to the same files. The trick is to create a different working directory for each instance. In Windows, do as follows:
     *    Use the ' File / Settings / Export ' command to bundle your data

and models into a portable '.tsm' file.
*              Create a new directory for each instance, and copy the .tsm file to each.
*              Start each instance by double-clicking on the .tsm file (see 23 above).

The directory containing the .tsm file becomes the Start-up directory for each instance, with data, graphics, '.tsd' (model) and results files all written there by default.

28.      For a more permanent multi-tasking setup, create multiple Windows shortcuts on the Start Menu, each edited to point to a different 'Start-In' directory. Give these shortcuts suitable names to distinguish them.

29.      Another use for multiple instances is where TSM is running the user's Ox code. The supplied Windows shortcut "TSM4 with User Code" provides an example. See Appendix C for further details.

TROUBLESHOOTING

30.      If the program appears to behave unpredictably, this may be because a setting has been changed inadvertently. For example:
*              The values of certain estimated parameters (e.g. MA parameters, GARCH parameters can depend on program settings.
*              If a parameter is "fixed" at zero in the Values dialog, no value is reported for it in the output. It might appear that the estimated specification is different from the chosen one. (Note that when variable(s) are fixed, the relevant Values button is highlighted and menu item(s) checked, as a reminder.
*              Corruption of the settings file (which is reloaded at start-up unless renewed) can occasionally occur.

If you cannot figure out what the problem is, reset the program to defaults with File / Settings / New. Alternatively, close the program, delete the "settings.tsm" file, and restart.

31.      Another option is to write the current settings to a text file (File / Settings / Save Text...) and inspect this output. Note that only non-default settings are listed. The usage of variable names can be checked in the programming manual.

32.      If the program attempts to execute an illegal Ox instruction and crashes, a message is displayed, and then it automatically restarts. The Ox error message appears in the console window (second TSM icon on the taskbar) and can be copied to the clipboard and saved (follow the onscreen instructions). This message can guide users in debugging supplied code. If the error is due to a TSM bug, please send in the error message as requested!

33.      The error recovery screen displays options for quitting, continuing, and exporting the current settings to a file. This file contains valuable diagnostic information, preserving the last recorded state of the program before the crash. If possible, please send in this file along with the error message.

        

34.    Although Ox should read spreadsheet files such as the Excel .xls and .xlsx format it is occasionally found that such files contain incompatible features. The message "Data input failed" appears. In this event, use Excel to save the file in comma delimited (.csv) format. These files are always readable and contain all the information relevant to TSM functionality. If TSM is then used to save the data in .xls or .xlsx format, the file will be readable by both TSM and Excel.