TIME SERIES DATA ANALYSIS USING EVIEWS

I Gusti Ngurah Agung

Graduate School Of Management
Faculty Of Economics University Of Indonesia

Ph.D. in Biostatistics and
MSc. in Mathematical Statistics from
University of North Carolina at Chapel Hill

John Wiley & Sons (Asia) Pte Ltd

TIME SERIES DATA ANALYSIS USING EVIEWS

TIME SERIES DATA ANALYSIS USING EVIEWS

I Gusti Ngurah Agung

Graduate School Of Management
Faculty Of Economics University Of Indonesia

Ph.D. in Biostatistics and
MSc. in Mathematical Statistics from
University of North Carolina at Chapel Hill

John Wiley & Sons (Asia) Pte Ltd

STATISTICS IN PRACTICE

Statistics in Practice is an important international series of texts which provide detailed coverage of statistical concepts, methods and worked case studies in specific fields of investigation and study.

With sound motivation and many worked practical examples, the books show in down-to-earth terms how to select and use an appropriate range of statistical techniques in a particular practical field within each title's special topic area.

The books provide statistical support for professionals and research workers across a range of employment fields and research environments. Subject areas covered include medicine and pharmaceutics; industry, finance and commerce; public services; the earth and environmental sciences, and so on.

The books also provide support to students studying statistical courses applied to the above areas. The demand for graduates to be equipped for the work environment has led to such courses becoming increasingly prevalent at universities and colleges.

It is our aim to present judiciously chosen and well-written workbooks to meet everyday practical needs. Feedback of views from readers will be most valuable to monitor the success of this aim.

A complete list of titles in this series appears at the end of the volume.

Visit our Home Page on www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, ONT, L5R 4J3, Canada

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Agung, Ign.
Time series data analysis using EViews / Ign Agung.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-82367-5 (cloth)
1. Time-series analysis. 2. Econometric models. I. Title.
QA280.A265 2009
519.5'5–dc22 2008035077

ISBN 978-0-470-82367-5 (HB)

Typeset in 10/12pt Times by Thomson Digital, Noida, India.
Printed and bound in Singapore by Markono Print Media Pte Ltd, Singapore.
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Dedicated to my wife
Anak Agung Alit Mas,
children
Ningsih A. Chandra, Ratna E. Lefort, and Dharma Putra,
sons in law
Aditiawan Chandra, and Eric Lefort,
daughter in law
Refiana Andries, and
all grand children
Indra, Rama, Luana, Leonard, and Natasya

Contents

Preface

Time series data, growth, or change over time can be observed and recorded in all their biological and nonbiological aspects. Therefore, the method of time series data analysis should be applicable not only for financial economics but also for solving all biological and nonbiological growth problems. Today, the availability of statistical package programs has made it easier for each researcher to easily apply any statistical model, based on all types of data sets, such as cross-section, time series, cross-section over time and panel data. This book introduces and discusses time series data analysis, and represents the first book of a series dealing with data analysis using EViews.

After more than 25 years of teaching applied statistical methods and advising graduate students on their theses and dissertations, I have found that many students still have difficulties in doing data analysis, specifically in defining and evaluating alternative acceptable models, in theoretical or substantial and statistical senses. Using time series data, this book presents many types of linear models from a large or perhaps an infinite number of possible models (see Agung, 1999a, 2007). This book also offers notes on how to modify and extend each model. Hence, all illustrative models and examples presented in this book will provide a useful additional guide and basic knowledge to the users, specifically to students, in doing data analysis for their scientific research papers.

It has been recognized that EViews is an excellent interactive program, which provides an excellent tool for us to use to do the best detailed data analyses, particularly in developing and evaluating models, in doing residual analysis and in testing various hypothesis, either univariate or multivariate hypotheses. However, it has also been recognized that for selected statistical data analyses, other statistical package programs should be used, such as SPSS, SAS, STATA, AMOS, LISREL and DEA.

Even though it is easy to obtain the statistical output from a data set, we should always be aware that we never know exactly the true value of any parameter of the corresponding population or even the true population model. A population model is defined as the model that is assumed or defined by a researcher to be valid for the corresponding population. It should be remembered that it is not possible to represent what really happens in the population, even though a large number of variables are used. Furthermore, it is suggested that a person's best knowledge and experience should be used in defining several alternative models, not only one model, because we can never obtain the best model out of all possible models, in a statistical sense. To obtain the truth about a model or the best population model, read the following statements:

> *Often in statistics one is using parametric models .... Classical (parametric) statistics derives results under the assumption that these models are strictly true. However, apart from simple discrete models perhaps, such models are never exactly true (Hample, 1973, quoted by Gifi, 1990, p. 27).*

Corresponding to this statement, Agung (2004, 2006) has presented the application of linear models, either univariate or multivariate, starting from the simplest linear model, i.e. the cell-means models, based on either a single factor or multifactors. Even though this cell-means model could easily be justified to represent the true population model, the corresponding estimated regression function or the sample means greatly depends on the sampled data.

> *In data analysis we must look on a very heavy emphasis on judgment (Tukey, 1962, quoted by Gifi, 1990, p. 23).*

Corresponding to this statement, there should be a good or strong theoretical and substantial base for any proposed model specification. In addition, the conclusion of a testing hypothesis cannot be taken absolutely or for granted in order to omit or delete an exogenous variable from a model. Furthermore, the exogenous variables of a growth or time series model could include the basic or original independent variables, the time *t*-variable, the lagged of dependent or independent variables and their interaction factors, with or without taking into account the autocorrelation or serial correlation and heterogeneity of the error terms. Hence, there is a very large number of choices in developing models. It has also been known that based on a time series data set, many alternative models could be applied, starting with the simplest growth models, such as the geometric and exponential growth models up to the VAR (*Vector Autoregression*), VEC (*Vector Error Correction*), System Equation in general and GARTH (*Generalized Conditional Heteroskedasticity*) models.

The main objective of this book is to present many types of time series models, which could be defined or developed based on only a set of three or five variables. The book also presents several examples and notes on unestimable models, especially the nonlinear models, because of the *overflow* of the iteration estimation methods. To help the readers to understand the advantages and disadvantages of each of the models better, notes, conclusions and comments are also provided. These illustrative models could be used as good basic guides in defining and evaluating more advanced time series models, either univariate or multivariate models, with a larger number of variables.

This book contains eleven chapters as follows.

Chapter 1 presents the very basic method in EViews on how to construct an EViews workfile, and also a descriptive statistical analysis, in the form of summary tables and graphs. This chapter also offers some remarks and recommendations on how to use scatter plots for preliminary analysis in studying relationships between numerical variables.

Chapter 2 discusses continuous growth models with the numerical time *t* as an independent variable, starting with the two simplest growth models, such as the geometric and exponential growth models and the more advanced growth models, such as a group of the general univariate and multivariate models, and the S-shape *vector autoregressive* (VAR) growth models, together with their residual analyses. This chapter also presents growth models, which could be considered as an extension or modification of the Cobb–Douglas and the CES (Constant Elasticity of Substitution) production functions, models with interaction factors and trigonometric growth models. For alternative estimation methods, this chapter offers examples using the White and the Newey–West HAC estimation methods.

Chapter 3 presents examples and discussions on discontinuous growth models with the numerical time *t* and its defined or certain dummy variable(s) as independent variables of the models. This chapter provides alternative growth models having an interaction factor(s) between their exogenous variable(s) with the time *t* as an independent variable(s). Corresponding to the discontinued growth models, this chapter also presents examples on how to identify breakpoints, by using Chow's Breakpoint Test.

Chapter 4 discusses the time series models without the numerical time *t* as an independent variable, which are considered as *seemingly causal models* (*SCM*) for time series. For illustrative purposes, alternative representation of a model using dummy time variables and three-piece autoregressive SCMs are discussed based on a hypothetical data set, with their residual plots. This chapter also provides examples of the discontinued growth models, as well as models having an interaction factor(s).

Chapter 5 covers special cases of regression models based on selected data sets, such as the POOL1 and BASIC workfiles of the EViews/Examples Files, and the US Domestic Price of Copper, 1951–1980, which is presented as one of the exercises in Gujarati (2003, Table 12.7, p. 499). The BASIC workfile is discussed specifically to present good illustrative examples of nonparametric growth models.

Chapter 6 describes illustrative examples of multivariate linear models, including the VAR and SUR models, and the structural equation model (SEM), by using the symbol *Y* for the set of endogenous variables and the symbol *X* for the set of exogenous variables. The main idea for using these symbols is to provide illustrative general models that could be applied on any time series in all biological and nonbiological aspects or growth. As examples to illustrate, three *X* and two *Y* variables are selected or derived from the US Domestic Price of Copper data, which were used for linear model presentation in the previous chapters. All models presented there as examples could be used for any time series data. Analysts or researchers could replace the *X* and *Y* variables by the variables that are relevant to their field of studies in order to develop similar models.

Chapter 7 covers basic illustrative instrumental variables models, which could be easily extended using all types of models presented in the previous chapters, either with or without the time *t*-variable as an independent variable.

Chapter 8 presents the autoregressive conditional heteroskedasticity (ARCH) models, generalized ARCH (GARCH), threshold ARCH (TARCH) and exponential ARCH (E_GARCH) models, either additive or interaction factor models.

In addition to the Wald tests, which have been applied in the previous chapters for various testing hypotheses, Chapter 9 explores some additional testing hypotheses, such as the unit root test, the omitted and redundant variables tests, the nonnested test and Ramsy's RESET tests, with special comments on the conclusion of a testing hypothesis.

Chapter 10 introduced a general form of nonlinear time series model, which could also represent all time series models presented in the previous chapters. For illustrative examples, this chapter discusses models that should be considered, such as the *Generalized Cobb-Douglas* (G_CD) model and the *Generalized Constant Elasticity of Substitution* (G_CES) model.

Finally, Chapter 11 presents nonparametric estimation methods, which cover the classical or basic moving average estimation method and the *k-Nearest Forecast* (*k*-NF), which can easily be calculated manually or by using Microsoft Excel, and the smoothing techniques (Hardle, 1999), such as the Nearest Neighbor and Kernel Fit Models, which should be done using EViews.

In addition to these chapters, the theoretical aspects of the basic estimation methods based on the time series data are presented in four appendices. In writing these appendices I am indebted to Haidy A. Pasay, Ph.D, lecturer in Microeconomics and Econometrics at the Graduate Program of Economics, the Faculty of Economics, University of Indonesia, who are the coauthors of my book on Applied Microeconomics (Agung, Pasay and Sugiharso, 1994). They spent precious time reading and making detailed corrections on mathematical formulas and econometric comprehension.

I express my gratitude to the Graduate School of Management, Faculty of Economics, University of Indonesia, for providing a rich intellectual environment and facilities indispensable for the writing of this text, as well as other published books in Indonesian.

In the process of writing this applied statistical book in English, I am indebted to Dr Anh Dung Do, the President of PT Kusuma Raya (Management, Financing and Investment Advisory Services) and Lecturer in Strategic Management at the Master Program of the Faculty of Economics, University of Indonesia. Dr Do motivated and supported me in the completion of this book. He spent a lot of his precious time in reading and making various corrections to my drafts.

I am also deeply indebted to my daughter, Martingsih Agung Chandra, BSPh, MSi, The Founder and Director of NAC Consultant Public Relations, and my son, Dharma Putra, MBA, Director of the PURE Technology, PT. Teknologi Multimedia Indonesia, for all their help in reading and making corrections to my drafts.

Puri AGUNG
Jimbaran, Bali

1

EViews workfile and descriptive data analysis

1.1 What is the EViews workfile?

The EViews workfile is defined as a file in EViews, which provides many convenient visual ways, such as (i) to enter and save data sets, (ii) to create new series or variables from existing ones, (iii) to display and print series and (iv) to carry out and save results of statistical analysis, as well as each equation of the models applied in the analysis. By using EViews, each statistical model that applied previously could be recalled and modified easily and quickly to obtain the best fit model, based on personal judgment using an interactive process. Corresponding to this process, the researcher could use a specific name for each EViews workfile, so that it can be identified easily for future utilization.

This chapter will describe how to create a workfile in a very simple way by going through Microsoft Excel, as well as other package programs, if EViews 5 or 6 are used. Furthermore, this chapter will present some illustrative statistical data analysis, mainly the descriptive analysis, which could also be considered as an exploration or an evaluation data analysis.

1.2 Basic options in EViews

It is recognized that many students have been using EViews 4 and 5. For this reason, in this section the way to create a workfile using EViews 4 is also presented, as well as those using EViews 5 and 6. However, all statistical results presented as illustrative examples use EViews 6.

Figure 1.1 presents the toolbar of the EViews main menus. The first line is the Title Bar, the second line is the Main Menus and the last space is the Command Window and the Work Area.

Then all possible selections can be observed under each of the main menus. Two of the basic options are as follows:

(1) To create a workfile, click *File/New*, which will give the options in Figure 1.2.

Time Series Data Analysis Using EViews IGN Agung

Figure 1.1 The toolbar of the main menus

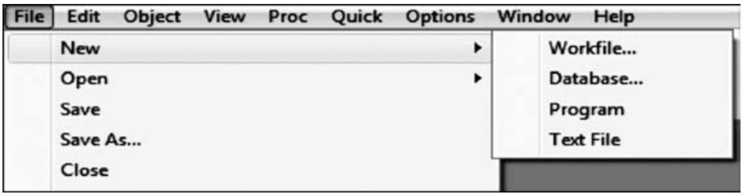


Figure 1.2 The complete options of the new file in EViews 4, 5 and 6

(2) To open a workfile, click *File/Open*, which will give the options in Figure 1.3 using EViews 4. Using EViews 5 or 6 gives the options in Figure 1.4.

Note that by using EViews 5 or 6, 'Foreign Data as Workfile. . .' can be opened. By selecting the option '*Foreign Data as Workfile. . .*' and clicking the '*All files* (*.*)' option, all files presented in Figure 1.5 can be seen, and can be opened as workfiles. Then a workfile can be saved as an EViews workfile.

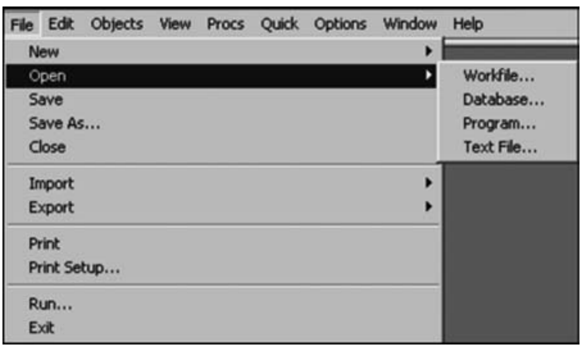


Figure 1.3 The complete options of the open file in EViews 4

Figure 1.4 The complete options of the open file in EViews 5 and 6

Figure 1.5 All files that can be opened as a workfile using EViews 5 and 6

1.3 Creating a workfile

1.3.1 Creating a workfile using EViews 5 or 6

Since many '*Foreign Data as Workfile…*' can be opened using EViews 5, as well as EViews 6, as presented in Figures 1.3 and 1.4, there are many alternative ways that can be used to create an EViews workfile. This makes it easy for a researcher to create or derive new variables, indicators, composite indexes as well as latent variables (unmeasurable or unobservable factors) by using any one of the package programs presented in Figure 1.4, which is very convenient for the researcher. Then he/she can open the whole data set as a workfile.

1.3.2 Creating a workfile using EViews 4

By assuming that creating an Excel datafile is not a problem for a researcher, only the steps required to copy Data.xls to an EViews workfile will be presented here. As an illustration and for the application of statistical data analysis, the data in Demo.xls will be used, which are already available in EViews 4.

To create the desired workfile, the steps are as follows:

(1) If EViews 4 is correctly installed, by clicking *My Documents…*, the directory 'EViews Example Files' will be seen in My Documents, as presented in Figure 1.6.
(2) Double click on the EViews Example Files, then double click on the data and the window in Figure 1.7 will appear. Then the file Demo.xls can be seen, in addition to several workfiles and programs. From now on, Demo.xls will be used.
(3) Double click on Demo.xls; a time series data set having four variables will be seen: *GDP*, *PR*, *NPM* and *RS* in an Excel spreadsheet, as shown in Figure 1.8. For further demonstrations of data analysis, three new variables are created in the spreadsheet: (i) *t* as the time variable having values from 1 up to 180, (ii) *Year* having values from 1952 up to 1996 and (iii) *Q* as a quarterly variable having values 1, 2, 3 and 4 for each year (see the spreadsheet below).

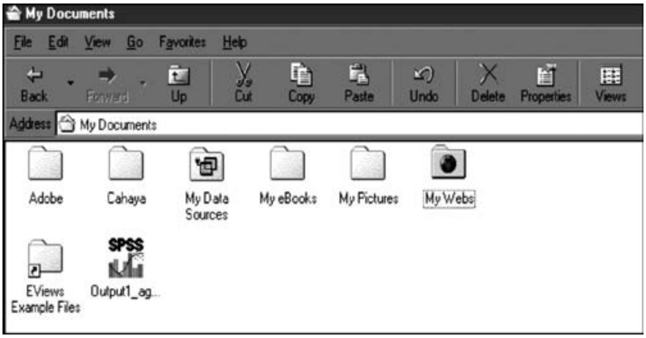
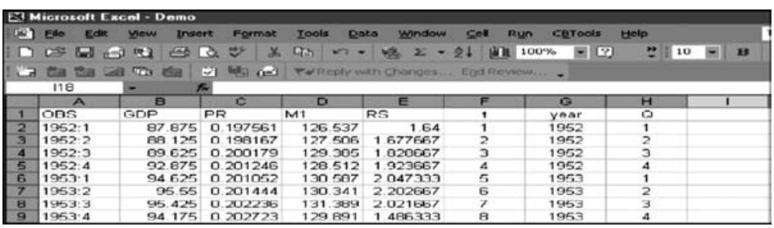


Figure 1.6 The EViews example files in My Documents

Figure 1.7 List of data that are available in EViews 4

(4) Block Demo.xls and then click *Edit/Copy*. . . .
(5) Open Eviews and then click *File/New/Workfile* . . . . This gives the window in Figure 1.9, showing the quarterly data set with starting and ending dates in Demo.xls. The rules for describing the dates are as follows:
- *Annual:* specify the year. Years from 1930 to 2029 may be identified using either 2- or 4-digit identifiers (e.g. ‘32’ or ‘1932’). All other years must be identified with full year identifiers.
- *Quarterly:* the year followed by a colon or the letter ‘Q,’ and then the quarter number. Examples: ‘1932 : 3,’ ‘32 : 3’ and ‘2003Q4.’



	A	B	C	D	E	F	G	H
1	OBS	GDP	PR	M1	RS	t	year	Q
2	1952:1	87.875	0.197561	126.537	1.64	1	1952	1
3	1952:2	88.125	0.198167	127.506	1.677667	2	1952	2
4	1952:3	89.625	0.200179	129.385	1.820667	3	1952	3
5	1952:4	92.875	0.201246	128.512	1.923667	4	1952	4
6	1953:1	94.625	0.201052	130.587	2.047333	5	1953	1
7	1953:2	95.55	0.201444	130.341	2.202667	6	1953	2
8	1953:3	95.425	0.202236	131.389	2.021667	7	1953	3
9	1953:4	94.175	0.202723	129.891	1.486333	8	1953	4

Figure 1.8 A part of data in Demo.xls

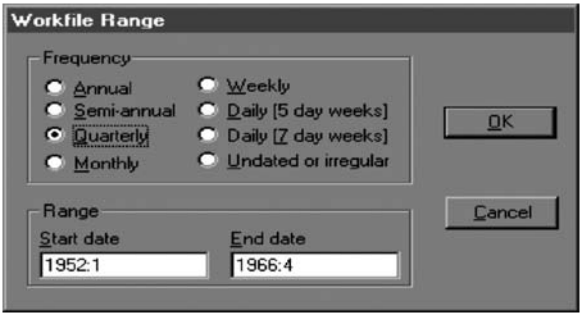


Figure 1.9 The workfile frequency and range

- *Monthly:* the year followed by a colon or the letter 'M,' and then the month number. Examples: '1932M9' and '1939 : 11.'
- *Semiannual:* the year followed a colon of the letter 'S,' and then either '1' or '2' to denote the period. Examples: '1932 : 2' and '1932S2.'
- *Weekly and daily:* by *default*, these dates should be specified as month number, followed by a colon, then followed by the day number, then followed by a colon, followed by the year. For example, entering '4 : 13 : 60' indicates that the workfile begins on April 13, 1960.
- Alternatively, for quarterly, monthly, weekly and daily data, just the year can be entered and EViews will automatically specify the first and the last observation.
- For other types of data, '*Undated or irregular*' is selected.

(6) Click *OK* produces the space or window, as presented in Figure 1.10. For every new data set or workfile at this stage, the window always shows a parameter vector 'C' and a space 'RESID,' which will be used to save the parameter and the residuals of the models used in an analysis.

(7) Click *Quick/Empty Group...* brings up the spreadsheet in Figure 1.11 on the screen. Put the cursor in the second column of the OBS indicator and then click so that the second column will block or darken.

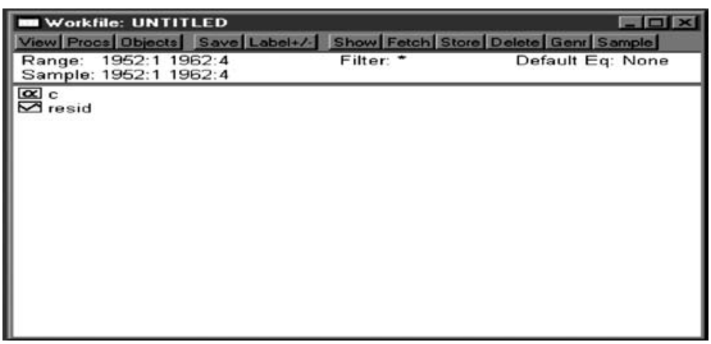


Figure 1.10 A workfile space of quarterly data in Demo.xls

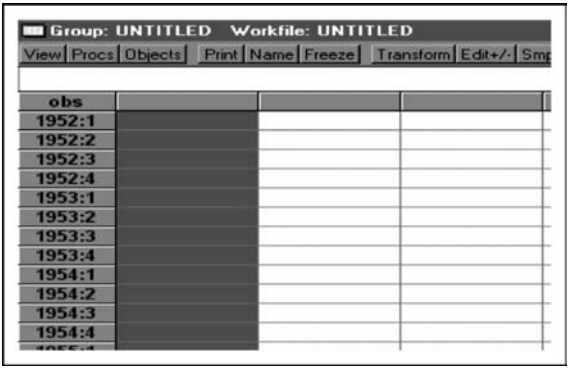
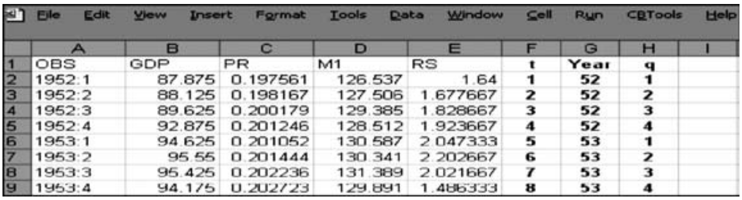


Figure 1.11 The group space to insert Demo.xls



	A	B	C	D	E	F	G	H	I
1	OBS	GDP	PR	M1	RS	t	Year	q	
2	1952:1	87.875	0.197561	126.537	1.64	1	52	1	
3	1952:2	88.125	0.198167	127.506	1.677667	2	52	2	
4	1952:3	89.625	0.200179	129.385	1.828667	3	52	3	
5	1952:4	92.875	0.201246	128.512	1.923667	4	52	4	
6	1953:1	94.625	0.201052	130.587	2.047333	5	53	1	
7	1953:2	95.55	0.201444	130.341	2.202667	6	53	2	
8	1953:3	95.425	0.202236	131.389	2.021667	7	53	3	
9	1953:4	94.175	0.202723	129.891	1.486333	8	53	4	

Figure 1.12 Demo.xls with additional data of the variables *t*, *Year* and *Q*

(8) Put the cursor again in column 2 and click the right button of the mouse; then click *Paste*. The spreadsheet in Figure 1.12 will be seen. In fact, additional variables, such as the variables *t*, *Year* and *Q* (quarter), can be created, entered or defined in the Excel spreadsheet, before the data set needs to be copied.

(9) Click *File/Saved As*... and then identify a name for the workfile. In this case, Demo_Modified is used, as shown in the following window (Figure 1.13).

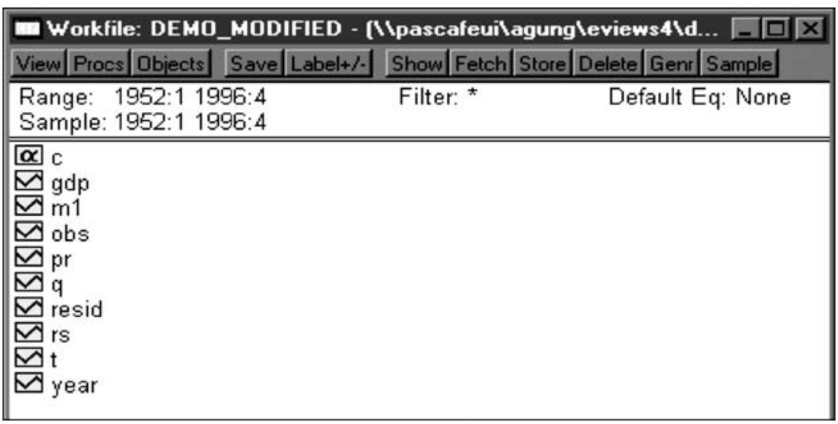


Figure 1.13 List of variables in the Demo_Modified workfile

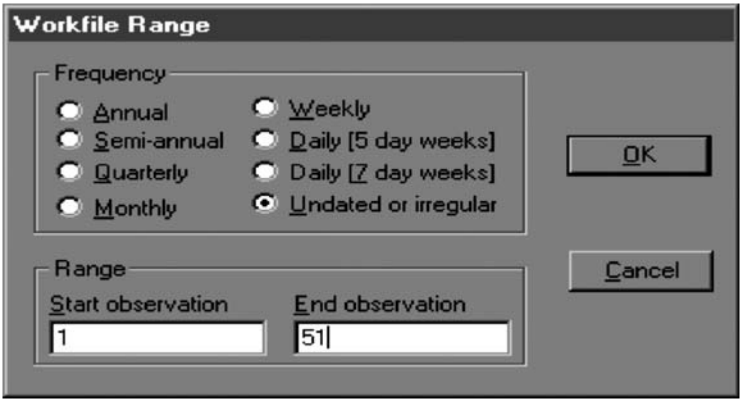


Figure 1.14 The option for creating a workfile based on an undated or irregular data set with 51 observations

1.3.2.1 Creating a workfile based on an undated data set

Figure 1.14 shows an example that can be used to create a workfile based on an undated data set. Using the same process as in the previous subsection, the workfile is created from an Excel datafile having 51 lines. The first line shows the names of the variables and the next 50 lines are the observation units.

1.4 Illustrative data analysis

The examples of the descriptive data analysis, as well as the inferential data analysis presented in this book, will be done using EViews 6. With reference to descriptive data analysis, it has been known that the statistical results are in the form of summary statistical tables and graphs. However, they have a very important role in data evaluation and policy analysis or decision making. Agung (2004) pointed out that summary descriptive statistics are one of the best supporting data for policy analysis. He also presented illustrative examples in selecting specific indicators, factors or variables, to show causal models in the form of summary tables.

However, in this chapter only a few methods are demonstrated in doing a statistical analysis, mainly a descriptive analysis using EViews 6 based on Demo_Modified.wf1.

1.4.1 Basic descriptive statistical summary

The summary statistics of the four numerical variables *GDP*, *M*1, *PR* and *RS* in Demo_Modified can be presented using the following steps:

(1) After opening the workfile, click the variable *GDP*; then by pressing the '*CTRL*' button click the variable *M*1. Make similar executions for the variables *PR* and *RS*; the result is that the four variables are blocked, as shown in Figure 1.15.
(2) Click *OK*...; the four variables will be seen on the screen, as presented in Figure 1.16. Then by clicking *OK*..., the data of the four variables will be seen on

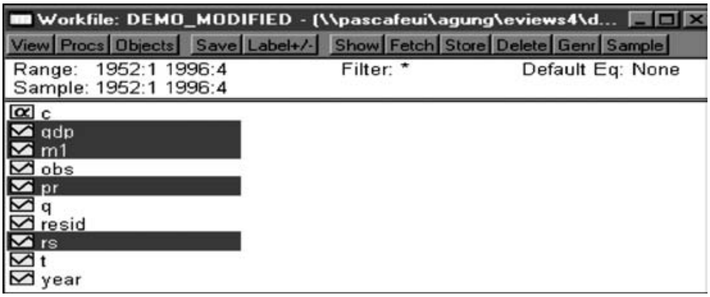


Figure 1.15 Blocked or selected variables that will be analyzed

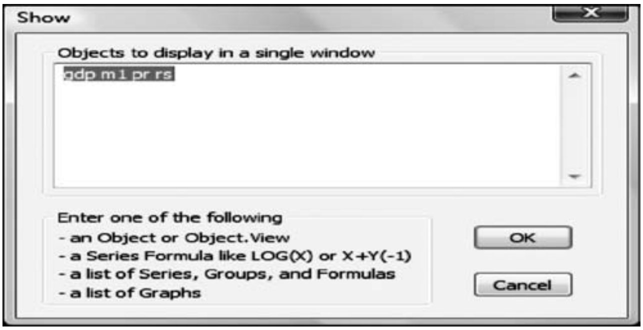
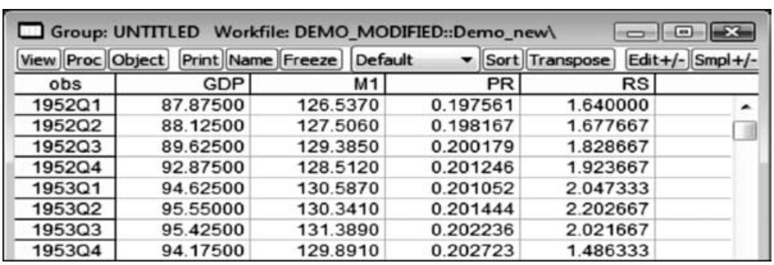


Figure 1.16 The variables whose data will be presented on the screen

the screen, as presented in Figure 1.17. This window should be used as a preliminary data evaluation, particularly for identifying new created variables and/or to edit selected values/scores, if it is needed.

(3) By clicking *View*..., the options in Figure 1.18 can be seen, which shows (14 + 2) alternative options, including two options for Descriptive Stats.

(4) Click *View/Descriptive Stats/Individual Samples*...; the summary descriptive statistics in Figure 1.19 are obtained. Selected computation formulas based on a


Group: UNTITLED Workfile: DEMO_MODIFIED::Demo_new\

View Proc Object Print Name Freeze Default Sort Transpose Edit+/- Smpl+/-

obs	GDP	M1	PR	RS
1952Q1	87.87500	126.5370	0.197561	1.640000
1952Q2	88.12500	127.5060	0.198167	1.677667
1952Q3	89.62500	129.3850	0.200179	1.828667
1952Q4	92.87500	128.5120	0.201246	1.923667
1953Q1	94.62500	130.5870	0.201052	2.047333
1953Q2	95.55000	130.3410	0.201444	2.202667
1953Q3	95.42500	131.3890	0.202236	2.021667
1953Q4	94.17500	129.8910	0.202723	1.486333

Figure 1.17 The screen shot of the data of selected variables in Figure 1.5