

# Macroeconomic Modeling and Forecasting Manual

Copyright 2012

Macroeconomic and Financial Management Institute of Eastern and Southern Africa

9 Earls Rd

Alexandra Park

Harare

Zimbabwe

Tel: +263 4 745988/9/91-94

Email: [capacity@mefmi.org](mailto:capacity@mefmi.org)

Web: [www.mefmi.org](http://www.mefmi.org)

All rights reserved

This manual is a product of MEFMI. The findings, interpretations and conclusions expressed in this manual do not necessarily reflect the views of the Board of Governors of MEFMI.

MEFMI does not guarantee the accuracy of the data included in this work. The material in this manual is copyrighted. Copying and / or transmitting portions or all work in this manual without permission is a violation of applicable law. MEFMI encourages dissemination of its work and will grant permission to reproduce portions of the work. For permission to photocopy or reprint any part of this manual, please send a written request to the Publications & Networking Officer.

ISBN #: 978-0-7974-5246-6

## CONSULTANCY TEAM

**Jesimen T. Chipika (Dr/Mrs), Team Leader**, Independent Senior Economic Development Consultant, Harare, Zimbabwe. Tel: +263-4-308032; Cell: +263 772214579, E-mail: [sjchip@africaonline.co.zw](mailto:sjchip@africaonline.co.zw)

**Jacob O. Oduor (Dr), Co-Leader**, Macroeconomic Policy Analyst, Kenya Institute for Public Policy Research and Policy Analysis (KIPPRA) and Lecturer in Econometrics, Macroeconomics and Mathematics, Kenyatta University, Nairobi, Kenya. Tel: +254 -20- 202241380 or +254-20-316243; Cell: +254722978186, E-mail: [joduur@kippra.or.ke](mailto:joduur@kippra.or.ke) or [oduorjacob@hotmail.com](mailto:oduorjacob@hotmail.com)

**Joseph Mverecha (Mr), MEFMI Fellow** and Chief Economist/ Head, Group Economics and Strategy, ZB Financial Holdings, Harare, Zimbabwe; Cell: +263773036193; E-mail: [jmverecha@rbz.co.zw](mailto:jmverecha@rbz.co.zw)

**Johnson J. Nyella (Mr), MEFMI Fellow and Principal Economist** and Head on Policy and Publication Division, Monetary and Financial Affairs Department, Bank of Tanzania, Dar es Salaam, Tanzania. Office Tel: +255-22-2110946-52 Ext. 2162, Home Tel: +255-22- 2627457 and Cell: +255744370542; Fax number: Office: +255- 22- 2111009; E-mail: [jjnyella@hq.bot-tz.org](mailto:jjnyella@hq.bot-tz.org); or: [jjnyella@hotmail.com](mailto:jjnyella@hotmail.com)

## COLLABORATING PARTNERS

**Margaret Chitiga Mabugu (Prof.)**, Executive Director for the Economic Performance and Development Unit, The Human Sciences Research Council, South Africa. Cell: +277733044985; E-mail: [Margaret.chitiga@gmail.com](mailto:Margaret.chitiga@gmail.com)

**Ramos Mabugu, (Dr )**, Program Manager and Head of Research and Policy, Macroeconomics and Public Finance Unit, Financial and Fiscal Commission of South Africa, Midrand, South Africa. Tel: +277 11 2072332; Cell: +2778 8073965; E-mail: [ramosm@ffc.co.za](mailto:ramosm@ffc.co.za)

## FOREWORD

In pursuit of its vision to build sustainable capacity in the region, the Macroeconomic and Financial Management Institute of Eastern and Southern Africa (MEFMI) initiated a program to develop and disseminate materials, tools and manuals for use by member states. The overall objective is to provide member countries with state-of-the-art, tools which assist them in designing, formulating and implementing economic policies in their respective economies.

Consistent with this vision, I am pleased to introduce the first edition of the *Macroeconomic Modeling and Forecasting Manual*, consisting of regional frameworks and models that can be applied by member countries. The *Manual* is a reference document, which provides a set of standard, sectoral and disaggregated modeling frameworks that can be applied in developing small macroeconomic and other small sectoral models. The *Manual* therefore, provides practitioners in the MEFMI region with a practical tool that enhances the understanding of the practical interface between theory and real economic situations. The reference document will be particularly useful in the modeling and forecasting functions in Central Banks, Ministries of Finance/Planning and Bureaus of Statistics.

The use of the *Manual* will capacitate MEFMI member countries to: (i) design, formulate and implement economic policy measures based on empirical evidence derived from the application of modeling and forecasting techniques; (ii) design and develop their macroeconomic modeling frameworks; (iii) develop specific models of inflation, exchange rate, consumption, investment, monetary transmission mechanisms, fiscal deficit management and financing, revenue and tax modeling and forecasting, interest rates and other small macroeconomic models; and (iv) build fully fledged robust macro models. The *Manual* will considerably strengthen MEFMI's current efforts to impart knowledge in Modeling and Forecasting in the region as an integral component of the Macroeconomic Management and Analysis Programme.

The *Manual*, which is user-driven, is fully owned by stakeholder institutions who contributed immensely to its development. Invaluable input was also received from specialised expert groups who carried out the whole development process, editing and proof reading of the *Manual*.

The consultative process involved in the development of this *Manual* demonstrates the spirit of mutual cooperation and sustained support MEFMI continues to receive from its member states. I would like to commend all the stakeholders that were involved in the development process for their invaluable support.

I would like to recommend the adoption and implementation of the *Manual* as a reference tool for modeling and forecasting in the MEFMI region.



Ellias E. Ngalande (PhD)  
MEFMI Executive Director

---

---

## PREFACE

The “MEFMI *Macroeconomic Modeling and Forecasting Manual (MFM)*” was developed in fulfillment of MEFMI’s mandate to provide tools for macroeconomic management to enhance sustainable human and institutional capacity in Central Banks, Ministries of Finance and Planning Commissions or equivalent and foster best practices in these institutions.

The region remains plagued with a plethora of economic and social challenges with relatively higher incidences of extreme poverty in the world. About 70 percent of the region’s population lives below the poverty line of \$2 per day and close to 60 percent are financially excluded.

Other key developmental challenges for the region are its vulnerability to global shocks, climatic change and limited participation of its people in decision making processes which have resulted in weak governance and policy institutions. This has been exacerbated by other internal constraints notably energy shortages, infrastructure bottlenecks, limited fiscal space and underdeveloped financial markets.

These challenges continue to hamstring the fragile economies and weigh down on growth and development in the region. Human and institutional capacity is severely undermined. The strategic focus of MEFMI is to design appropriate tools to efficiently and effectively address capacity challenges in the region. MEFMI endeavours to equip countries with the economic knowledge, skills, competencies, tools, experiences and resources to enhance the formulation and implementation of prudent macroeconomic policies. This is critical to accelerate progress towards poverty eradication in the region.

An assessment of capacity building needs indicates that there is growing demand for macroeconomic modeling as a tool for macroeconomic policy analysis and management. It is evident that countries have made very little progress in applying fully fledged models. Inadequate capacity in econometric modeling thus, remains one of the **critical missing links** in Macroeconomic Management in the region.

A survey of the status of modeling and forecasting in the MEFMI region revealed that:

- Most countries are using the Financial Programming (FP) framework of the IMF;
- Some countries have developed simple models in the areas of inflation, exchange rate and monetary transmission mechanisms;
- Some countries show slightly more advancement and have proceeded to establish inflation forecasting frameworks; and
- Most countries are yet to develop fully-fledged modeling and forecasting models.

To advance capacity building efforts in the area of Macroeconomic Modeling and Forecasting, MEFMI initiated a project to develop a *Manual* that will serve as a reference tool to guide Econometricians and Modelers in the region. This is critical for evidence based policy design and execution. This home grown Modeling and Forecasting Product will reduce exposure to

external development policies, tools and practices that do not necessarily take into consideration the unique situation, cultural heritage, knowledge systems and institutions in the region.

MEFMI embarked on this arduous journey in February 2010 with a clear vision to produce a world class modeling tool by February 2012. The completion of the *Manual* is a realisation of this vision and a fruition of the Institute's efforts to capacitate policy institutions in the region to deliver on their mandates.

MEFMI is highly indebted to the Team of Consultants which produced this world class product. The team was under the leadership of Dr/Mrs Jesimen T. Chipika, Lead Consultant and Dr Jacob Oduor Co-Lead Consultant. Other team members included Messrs Joseph Mverecha and Johnson Nyella. The *Manual* also benefited from the invaluable input of Professor Margaret Chitiga-Mabugu and Dr Ramos Mabugu.

The *Manual* produced by the consultants was critically reviewed and solidified in a regional workshop in August, 2011 under the supervision of Dr Christopher Malikane. This process further refined and improved the *Manual*, taking into account comments from stakeholders. In March 2012, the *Manual* was then subjected to a peer review mechanism which translated it into a world class product benchmarked to international best practices. The three peer reviewers were Professor Alemayehu Geda, University of Addis Ababa, Ethiopia; Dr Angelina Njuguna, United Nations Economic Commission for Africa (UNECA), Ethiopia; and Mr Konstantin Makrelov, National Treasury of South Africa.

The project was carried out under the overall guidance of Mr. Sydney Mabika, Programme Director, Macroeconomic Management Programme and Mr. Jean Baptiste Havugimana, Programme Officer.

## TABLE OF CONTENTS

FOREWORD.....	ii
PREFACE.....	iii
TABLE OF CONTENTS.....	v
LIST OF TABLES.....	viii
LIST OF FIGURES.....	viii
LIST OF BOXES.....	viii
LIST OF APPENDICES.....	viii
ACRONYMS.....	ix
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>1</b>
1.1 Purpose of the Macroeconomic Modeling and Forecasting Manual (MFM).....	1
1.2 Scope of the Manual.....	2
1.3 Macro Models and Economic Policy Formulation.....	3
1.4 Regional Developments in Economic Modeling.....	6
1.5 Macroeconomic Modeling Challenges and Constraints Experienced by MEFMI Member States.....	14
1.6 Evolution of Macro-Models.....	15
1.7 Criticisms of Macroeconometric Modeling and their Responses.....	18
1.8 Lessons Learnt.....	19
1.9 Structure of the Manual.....	20
<b>PART I: MACRO-ECONOMETRIC MODELING USING TIME SERIES DATA.....</b>	<b>21</b>
<b>CHAPTER 2: THEORETICAL/LOGICAL FRAMEWORK.....</b>	<b>22</b>
2.1 Need for Theory in Macroeconomic Modeling.....	22
2.2 Choosing the Right Model to Build.....	23
2.3 Key Interactions between Macro Blocks.....	23
2.4 Description of Blocks.....	24
2.4.1 Real Sector Block.....	24
2.4.2 Fiscal Block.....	30
2.4.3 Monetary Block.....	31
2.4.4 External Sector Block.....	35
2.4.5 Closure Rules.....	36
2.4.6 Consistency Framework in the Blocks.....	36
2.4.7 Possible Challenges when Modeling the Blocks.....	37
<b>CHAPTER 3: ESTIMATION OF THE EQUATIONS IN THE BLOCKS.....</b>	<b>39</b>
3.1 Pre-Estimation Issues.....	39
3.1.1 Identifying the Variables to be included in the Model.....	39
3.2 Data.....	40
3.2.1 Types of Macroeconomic Data.....	40
3.2.2 Getting Data into E-views.....	40
3.2.3 Data Cleaning and Transformations.....	44
3.2.4 Data Frequency Conversions.....	45
3.2.5 Data Gaps, Inconsistencies and Remedies.....	51
3.3 Diagnostics of Time Series Data - Testing for Stationarity.....	52
3.3.1 Meaning of Unit Roots and Stationarity.....	53
3.3.2 Requirements for Weak Stationarity.....	54
3.3.3 Why are Time Series mostly non-stationarity?.....	54
3.3.4 Types of Time Series Processes.....	54
3.3.5 Pure Random Walk.....	55
3.4 Testing for Unit Roots.....	56
3.5 Testing for Unit Roots in E-views.....	58
3.6 Making Non - Stationary Variables to be Stationary.....	62
3.7 Detrending a Trend Stationary Variable.....	62
3.8 Correcting for non-Stationarity by Differencing.....	65
3.9 Order of Integration.....	66

3.10 Estimation of Time Series Data.....	67
3.10.1 Estimation with Stationary Variables - Single Equation Approach.....	67
3.10.2 Estimation with Non-Stationary, but Cointegrated variables with No Feedback Effects - Error Correction Models.....	81
3.10.3 Estimation with non-Stationary, but Cointegrated Variables with Feedback Effects - Vector Error Correction Models (VECM).....	87
<b>CHAPTER 4: BUILDING A PROTOTYPE MODEL.....</b>	<b>98</b>
4.1 Introduction.....	98
4.2 A Simple Model of a Hypothetical Economy in E-views Platform.....	99
4.3 Bringing the Estimated Coefficients into the Model.....	100
4.4 Model Building.....	105
4.4.1 Creating the Model Object.....	105
4.4.2 In-Sample Forecasting with a Single Equation.....	110
4.4.3 Adding the Add-Factors.....	113
4.4.4 Impulse Responses.....	117
4.4.5 Making Assumptions.....	120
4.4.6 Out-of-Sample Forecasting with a Single Equation.....	125
4.4.7 Putting all the Equations Together.....	127
4.4.8 Adding Identities.....	129
4.4.9 Forecasting In and Out of Sample with All the Equations.....	130
4.4.10 Policy Simulations with All the Equations.....	133
4.4.11 Assessing Forecasting Accuracy.....	134
4.4.12 Improving Forecasting Accuracy.....	136
<b>CHAPTER 5: OTHER FORECASTING METHODS.....</b>	<b>138</b>
5.1 Introduction.....	138
5.2 Other Forecasting Methods.....	138
5.2.1 Unconditional Mean.....	138
5.2.2 Exponential Smoothing.....	139
5.2.3 Exponential Smoothing in E-views.....	139
5.2.4 Autoregressive Moving Average (ARIMA).....	143
5.2.5 Forecasting from an ARIMA Model.....	152
5.2.6 Forecasting from an Estimated Equation.....	153
5.2.7 Other regression-based forecasting approach.....	158
<b>PART II: ECONOMY-WIDE MODELING.....</b>	<b>159</b>
<b>CHAPTER 6: INTRODUCTION TO ECONOMY-WIDE MODELING.....</b>	<b>160</b>
6.1 Background.....	160
6.1.1 Introduction.....	160
6.1.2 Different Varieties of Economy-Wide Models.....	162
6.2 Input-Output (IO) Tables.....	163
6.2.1 Background.....	163
6.2.2 Input-Output Modeling.....	165
6.2.3 Solving IO Models.....	168
6.2.4 Impact Analysis.....	169
6.2.5 Multiplier Analysis.....	169
6.2.6 Strengths and Weaknesses of IO Models.....	171
6.3 Introduction to Social Accounting Matrix.....	172
6.3.1 What is a SAM?.....	172
6.3.2 SAM as a Modeling Devise.....	176
6.3.3 SAM Multiplier Analysis.....	178
6.3.4 Interpretation of SAM Multipliers.....	180
6.3.5 SAM Closure Rules.....	180
6.3.6 Numerical Example of a SAM Model.....	180
6.3.7 Comparison of IO and SAM Modeling.....	182
6.3.8 Advantages of SAM.....	183
6.3.9 Disadvantages of SAM.....	183

---

---

6.4 Solving IO Models .....	184
6.4.1 What are CGE Models?.....	184
6.4.2 Why and When to Use CGE Models.....	186
6.4.3 Theoretical General Equilibrium Model .....	186
6.4.4 CGE Models.....	187
6.4.5 Short Run CGE Model .....	189
6.4.6 Extensions of CGE Models .....	190
6.4.7 Advantages and Disadvantages of CGE Models .....	190
6.4.8 Solving the Models Using Software .....	191
6.4.8.1 Introduction to Basic Excel Operations .....	192
6.4.8.2 Solving a stylised SAM Model in Excel.....	203
6.4.9 Getting Familiar with GAMS Architecture.....	206
6.4.9.1 Solving General Equilibrium Model in GAMS .....	208
<b>SELECTED BIBLIOGRAPHY .....</b>	<b>215</b>
<b>APPENDIX A: Tables of Critical Values .....</b>	<b>226</b>
<b>APPENDIX B: Participants during the Manual Review Workshop in Harare-August 2011 .....</b>	<b>227</b>
<b>APPENDIX C: Developing a VECM.....</b>	<b>228</b>

**LIST OF TABLES**

Table 1.1 Status of Modeling and Forecasting in MEFMI Member Countries.....	6
Table 1.2 Key Features of the Different Categories of Macroeconomic Models (MEMs).....	16
Table 6.1 Input-Output Tables in Symbols.....	164
Table 6.2 Hypothetical Manufacturing - Agriculture Example.....	168
Table 6.3 Direct Coefficient Matrix of Manufacturing - Agriculture Example.....	169
Table 6.4 Stylised SAM Example.....	173
Table 6.5 One Sector SAM.....	175
Table 6.6 Hypothetical SAM with Numbers, Example 1.....	176
Table 6.7 Hypothetical SAM with Numbers, Example 2.....	177
Table 6.8 Comparison of IO and SAM Multipliers.....	182
Table 6.9 Hypothetical SAM with Numbers.....	203

**LIST OF FIGURES**

Figure 1.1 Basic Steps in Econometric Model Building.....	5
Figure 2.1 Key Interactions in a Macro Model.....	24
Figure 3.1 A Stationary Process.....	54
Figure 3.2 A Pure Random Walk Process.....	55
Figure 3.3 Random Walk with a Drift.....	55
Figure 3.4 Random Walk with Drift and Time Trend.....	56
Figure 4.1 Critical Steps in Model Building.....	99

**LIST OF BOXES**

Box 1.1 A Prototypical Example for a Macro-Econometric Model, the National Income Model by Granger.....	8
Box 1.2 Macroeconomic Modeling and Forecasting in the Banco De Mozambique, 2008 to 2012...9	9
Box 1.3 KIPPRRA-Treasury Macro Model (KTMM) - Kenya.....	11
Box 1.4 The IMF Model-the Polak Model.....	12
Box 1.5 World Bank Revised Minimum Standard Model (RMSM).....	13

**LIST OF APPENDICES**

Appendix A, Tables of Critical Values.....	226
Table A1, Critical Values of Cointegration Test.....	226
Appendix B, Participants during the Manual Review Workshop in Harare-August 2011.....	227
Appendix C, Developing a VECM.....	228



**ACRONYMS**

2SLS	Two-Stage Least Squares
3SLS	Three-Stage Least Squares
ACF	Autocorrelation function
ADF	Augmented Dickey-Fuller
ADL	Autoregressive Distributed Lag
AERC	African Economic Research Consortium
AfDB	African Development Bank
AGOA	African Growth and Opportunities Act
AIC	Akaike Information Criterion
AIDS	Acquired Immuno-Deficiency Syndrome
AIG	All Inclusive Government
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroscedasticity Models
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average Models
ARVs	Antiretroviral Drugs
ATMs	Automated Teller Machines
AU	African Union
BEST	Business and Enterprise Start-up Tool
BMC	Botswana Meat Commission
BoB	Bank of Botswana
BON	Bank of Namibia
BoP	Balance of Payment
BOU	Bank of Uganda
BOZ	Bank of Zambia
BTVET	Business, Technical and Vocational Education and Training
BURS	Botswana Unified Revenue Service
BVAR	Bayesian VAR Models
CBK	Central Bank of Kenya
CBL	Central Bank of Lesotho
CES	Constant Elasticity of Substitution
CFE	Cumulative Sum of Forecast Errors
CGE	Computable General Equilibrium models
CMA	Common Monetary Area
CPI	Consumer Price Index
CSO	Central Statistical Office
DDVAR	Double-Differenced VARs

---

---

DFGLS	Dickey Fuller GLS Test
DGP	Data Generation Process
DSGE	Dynamic Stochastic General Equilibrium Models
DSW	Department of Social Welfare
EAC	East Africa Community
ECAFE	Economic Commission for Asia and the Far East
ECM	Error-Correction Mechanism
FU	Fuel and Utilities
EG	Engle-Granger Two-Step Method,
ECM	Error Correction Models
ERS	Economic Recovery Strategy
ERS	Elliot, Rothenberg, and Stock Point-Optimal Test
EYM	Engle-Yoo Three-Step Modeling Method
FAR	Fiscal Adjustment Roadmap
FCAs	Foreign Currency Accounts
FDI	Foreign Direct Investment
FIML	Full Information Maximum Likelihood
FPE	Final Prediction Error
FPP	Financial Programming and Policy Framework
FY	Fiscal Year
GAMS	General Algebraic Modeling System
GDP	Gross Domestic Product
GEMPACK	General Equilibrium Package
GFEC	Global Financial and Economic Crisis
GFS	Government Financial Statistics
GMM	Generalized Method of Moments
GPA	Global Political Agreement
GRT	Granger Representation Theorem
HIV	Human Immuno-deficiency Virus
HQ	Hannan-Quinn Information Criterion
IIP	International Investment Position
ILO	International Labour Organization
ILS	Indirect Least Squares
IMF	International Monetary Fund
IO	Input-Output
IT	Inflation Targeting Frameworks
IV	Instrumental Variables
KIPPRA	Kenyan Institute for Public Policy Research and Analysis
KISS	Keep It Sophisticatedly Simple

---

---

KK	Keynes–Klein Model
KPSS	Kwaitkowski-Phillips-Schmidt-Shin
KTMM	Kenya’s ‘KIPPR- Treasury Macroeconomic Model
LCH	Life Cycle Hypothesis
LHWP	Lesotho Highlands Water Project
LIML	Limited Information Maximum Likelihood
LM	Lagrange Multiplier
LRA	Lesotho Revenue Authority
LSE	London School of Economics
MAD	Mean Absolute Deviation
MDGs	Millennium Development Goals
MEFMI	Macroeconomic and Financial Management Institute of Eastern and Southern Africa
MELT3	Medium-to-Long-Term Model
MEMBOT	Macro-Econometric Model for Botswana
MEMWOG	Macroeconomic Modeling Working Group
MEPD	Ministry of Economic Planning and Development
MEPM	Macroeconomic Policy Model
MERP	Millennium Economic Recovery Programme
MFDP	Ministry of Finance and Development Planning
MFM	Macroeconomic Modeling and Forecasting Manual
MOF	Ministry of Finance
MOFNP	Ministry of Finance and National Planning
MS	Muth–Sargent model
MSE	Mean Squared Error
NAMAF	Namibian Macroeconomic Framework
NARC	National Alliance of the Rainbow Coalition
NBER	National Bureau of Economic Research
NGOs	Non-Governmental Organizations
NHIES	Namibian Household Income and Expenditure Survey
NIESR	National Institute of Economic and Social Research
NIR	Net International Reserves
NP	Ng and Perron Tests
OLS	Ordinary Least Squares
PA	Proportional adjustment
PAYE	Pay As You Earn
PB	Phillips–Bergstrom model,
PPP	Purchasing Power Parity
QFAs	Quasi-Fiscal Activities

---

---

REER	Real Effective Exchange Rate
RESET	Ramsey Regression Equation Specification Error Test
RMP	Reserve Money Program
RMSE	Root-Mean-Square Error
RMSM	World Bank Revised Minimum Standard Model
RMSM-X	Revised Minimum Standard Model-Extended
ROW	Rest of the World
RSA	Republic of South Africa
RTGS	Real Time Gross Settlement
SAM	Social Accounting Matrix
SAMs	Social Accounting Matrices
SC	Schwarz Information Criterion
SEM	Simultaneous Equation Model
SNA	System of National Accounting
SPSS	Statistical Package for Social Sciences
STAR	Smooth-Transition Autoregression
SU	Supply and Use
SUR	Seemingly Unrelated Regressions models
UC	Unconditional Mean
UK	United Kingdom
UN	United Nations
UNDP	United Nations Development Programme
UNECA	United Nations Economic Commission for Africa
UNHS	Uganda National Household Survey
VARMA	Vector Autoregression Moving Average
VARs	Vector Autoregressions
VAT	Value-Added Tax
VECM	Vector Error Correction Models
VECMs	Vector-Equilibrium Correction Models
WASCOM	Wages and Salaries Commission
WJ	Walras-Johansen Model
WL	Walras-Leontief Model

---

---

# CHAPTER 1

## INTRODUCTION

- \* Aim of Manual is to present a set of techniques that can be used for model development, policy analysis and forecasting across the MEFMI member countries taking into account their data and capacity constraints.
- \* Manual is based on the 'Keep It Sophisticatedly Simple (KISS)' Principle.
- \* A hands-on, step-by-step process, 'learning by doing'.
- \* Manual does not cover complex methodologies such as dynamic stochastic general equilibrium (DSGE) models.
- \* Macroeconomic modeling efforts in the MEFMI region are encountering many challenges, including: shortage of modeling staff; computer software challenges; data challenges; overall modeling skills challenges; and institutional and coordination challenges.
- \* These challenges, if not addressed by MEFMI member states, can constitute serious risks to the success of the Manual.
- \* If the Manual is to be useful in the region, MEFMI member states will need to commit significant resources towards the macroeconomic modeling agenda in the region, in an effort to resolve the highlighted challenges.

### 1.1 Purpose of the Macroeconomic Modeling and Forecasting Manual (MFM)

01 The overall aim of the MFM is to provide a modeling and forecasting tool to the Macroeconomic and Financial Management Institute of Eastern and Southern Africa (MEFMI) Member Countries, namely; Angola, Botswana, Kenya, Lesotho, Malawi, Mozambique, Namibia, Rwanda, Swaziland, Tanzania, Uganda, Zambia and Zimbabwe, which will assist in the designing, formulating and implementing of major economic policy measures and promote exchange of information and experiences through regional training. In short, the MFM intends to present a set of techniques that can be used for model development, policy analysis and forecasting across the member countries, taking into account their data and capacity constraints.

02 The main focus of MFM is therefore, to build capacity in the institutions responsible for macroeconomic policy-making in the MEFMI region. These institutions include: Ministries of Finance, Economic Planning and Development; Central Banks; Revenue Authorities; and Bureaus of Statistics; among others. The Manual will provide *a hands-on, step by step* process on how to carry out macroeconomic modeling, that is, the approach of "*learning by doing*". Its use will not only be restricted to training sessions by MEFMI, but it will be used in day-to-day economic modeling in these critical institutions.

03 The MFM secondary users include: Universities/Colleges/Research Centres which have expertise in macroeconomic modeling, and could be called upon to be part of the national modeling teams; Private Sector and non-governmental organizations (NGOs) which are custodians of key information necessary for macroeconomic modeling and forecasting; and Public corporations who are also custodians of critical information.

04 The *Manual* seeks to provide a user friendly prescription of how to handle macroeconomic data and use it for modeling as an input into evidence based policy making. The *Manual* will emphasize on the statistical description of key aggregates in a theoretically coherent approach using a wide range of casual observations that have been made on the economies. It will outline the methods and approaches to examining the relations in the movement of macroeconomic aggregates and their aggregate impact on the entire economy. The *Manual* will also bring to the fore the fact that whilst many of these theoretical points highlight the potential value of using models, the main practical consideration is whether the estimated models can provide policy relevant results.

05 Through the use of the *Manual*, MEFMI member countries will be able to:

- Improve the robustness of economic policy making as they apply the modeling and forecasting techniques presented in the *Manual* to real life situations;
- Design and develop their country specific macroeconomic modeling frameworks;
- Develop models of inflation, exchange rate, consumption, investment, monetary transmission mechanisms, fiscal deficit management and financing, revenue and tax modeling and forecasting, interest rates, economic growth and all other small macroeconomic models;
- Build fully-fledged macro models;
- Exchange information and experiences through regional training within the region with other MEFMI member states modeling experts and policy analysts.

06 Thus, the *Manual* will considerably strengthen MEFMI's current Modeling and Forecasting activities as part of the Macroeconomic Management Programme.

## 1.2 Scope of the Manual

07 Given the variation in modeling skill levels in both the primary and secondary users of the MFM, the *Manual* has been *designed in a pragmatic manner, keeping it simple but useful*, in order to meet the needs of users at all levels. In this regard, the *Manual* is a 'how to', step-by-step user guide in macroeconomic modeling and forecasting. The MFM is largely based on the methodology of macro-econometric modeling using time series data, which still dominates the macroeconomic modeling and forecasting agenda to date. The logical modeling framework with economic blocks is revisited; data handling issues are discussed, before the examples of estimating some of the equations of the blocks are

presented using the most common computer platforms being used in the MEFMI region, namely E-views and Excel. A simple macroeconomic prototype model from the region is presented as an example. In addition, simple, non-econometric forecasting methods are also discussed in the *Manual*.

08 However, to take care of the needs of some of the MEFMI region member states who are at a relatively more advanced stage in macroeconomic modeling, the MFM introduces economy-wide modeling using the simple input output (IO), social accounting matrices (SAMs) and simple computable general equilibrium (CGE) models. However, it is important to note that, this *Manual does not cover the complex methodologies* on dynamic stochastic general equilibrium models (DSGE) models and other more complex econometric modeling techniques.

### 1.3 Macro Models and Economic Policy Formulation

09 A model is basically a way of trying to represent relationships in a more concise way. In other words, **it is a simplification of reality**. In Economics, models are mainly used to inform policy makers as they try to come up with policy measures and their doses. The economic model provides a framework that is a representation of the real economy. There are several models that policy makers use, which differ in their types, scale and application. These assist by equipping them with ideas on how economic variables react to exogenous changes in economic variables. The results from the quantitative analysis of the economic variables help establish the responses of the economic aggregates under consideration to the various stimulants emanating from changes in the macroeconomic environment, and thus proffering a platform for different economic policies and policy measures to put the economy on a path of sustained growth and development.

10 Evidence based policy formulation requires **a critical analysis of the prevailing macroeconomic environment**. Models provide an important basis for a solid footing for anchoring policy on reality. Taking into consideration the results from the analysis and projections of economic models into the policy formulation processes may help policy makers come up with relevant and consistent policy interventions. Careful consideration needs to be taken of the current prevailing conditions in light of accumulated experience in relation to the behaviour of economic variables. Although the model may not really tell what the future will be, it gives solid possibilities that can be relied on in choosing policy options, and thus, optimizing the decision-making process.

11 By its nature, an economic model makes it possible **to clarify and organize rational decision making**. It summarizes the accumulated behaviour of the economy over different phases under different monetary and fiscal regimes. The quantitative nature of models makes it possible to quantify the effect of foreseeable future trends in these variables, taking into account the reaction of agents to eventual policy choices and decisions.

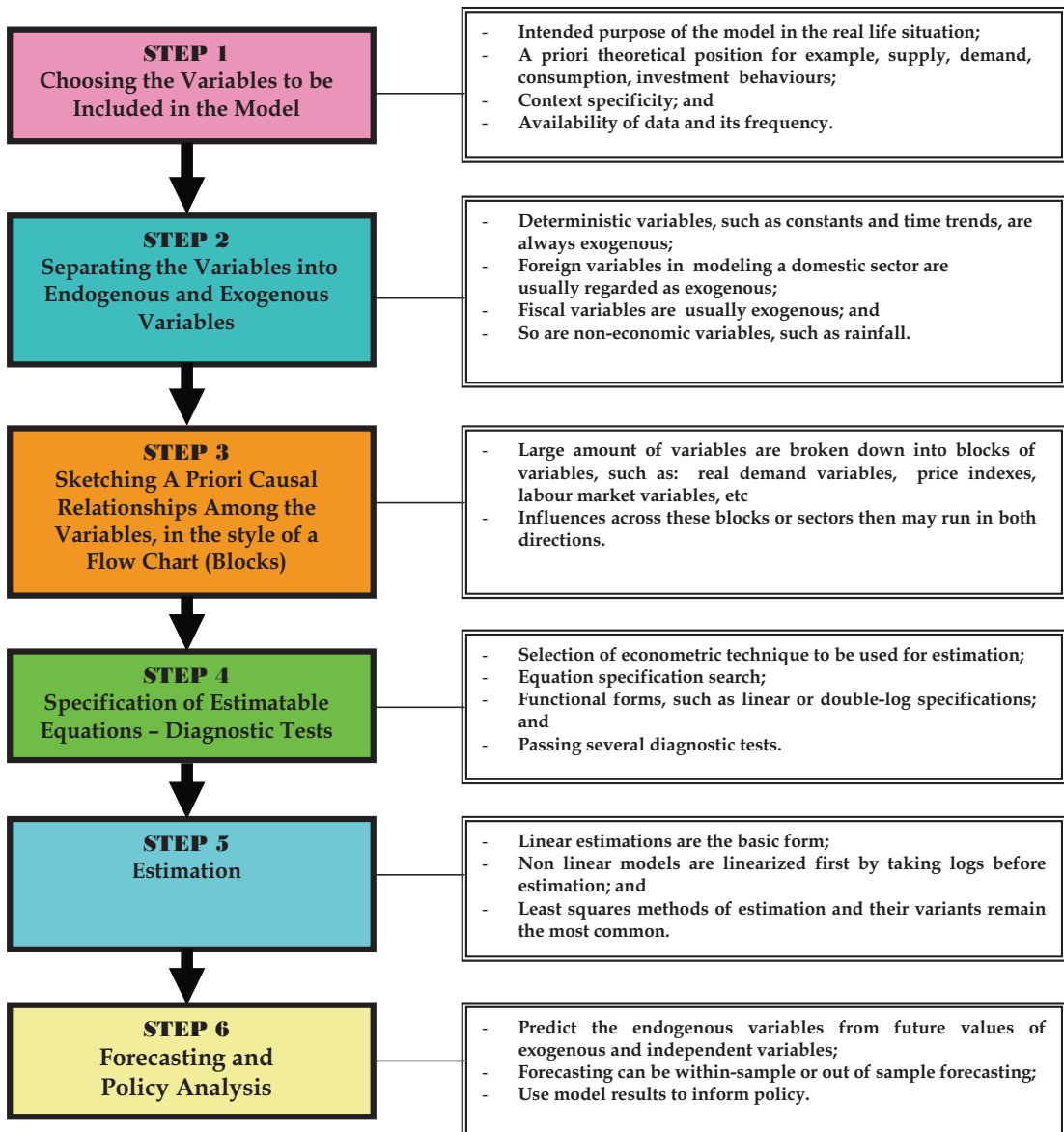
12 Generally, there are six basic steps in econometric model building as illustrated in Figure 1.1 on page 17. These will be elaborated on throughout the *Manual*.

13 It can be observed with confidence that the major obstacle for policy formation in most African countries is **limited knowledge of the way the macro-economy works**. It is therefore beyond doubt that both technocrats and bureaucrats should have solid understanding and knowledge of the behaviour of macroeconomic aggregates and the factors underlying such behaviour. This will give them an idea on how to assess the potential impact of proposed policy choices.

14 However, it has to be noted that these **models are context specific**. In other words, different economies cannot have the same model for purposes of policy formulation. The lack of local models has left most African countries with not much choice but to rely on models and policy choices prescribed by bi-lateral and multi-lateral institutions. This has seen the same modeling frameworks being prescribed for many different countries with different economic contexts. The net result has been the wrong policy doses, most notably the economic reform programmes of the late 1980s and early 1990s. In other words, a 'same-size-fits-all' type of model is fundamentally irrelevant in economic modeling. The knock-on effects from policies based on such models on the entire economy have been devastating since they have remained unexplained and have not been captured in the models.



**Figure 1.1: Basic Steps in Econometric Model Building**



15 It can be noted that many African countries do not have their policy making process supported by economic models, despite the growing world trend of an improvement in the construction and use of economic models. Economic models have also not only increased in their usage, but have also increased in complexity, mainly being driven by the diverse nature of economies in the modern era. These complexities have reduced even further, their usage in most African countries, mainly due to lack of capacity by these countries to both come up with relevant models and to use them more effectively. Many African countries are confronted with the lack of well-trained professionals capable of preparing consistent models that are relevant for comprehensive planning frameworks. This is further compounded by lack of requisite data for modeling purposes, leaving the option of building a model being solely based on data processed as estimates.

#### 1.4 Regional Developments in Economic Modeling

16 Most African countries face major constraints in economic modeling. They face both financial constraints to sustain economic modeling as well as the requisite skills to pursue economic modeling. This puts a severe constraint to economic modeling, especially on the selection of the ideal model to capture certain economic behaviour. There are two basic issues that arise from this, namely the best type of model among other types of models, and which model within that specific group of models. This arises from the fact that different models reflect differences in the theoretical and methodological paradigms. Controversies about model selection also reflect the debates about the competing explanations of how the economy works and the relevant methodology for statistical inference, (Soludo C. C. 2002).

17 However, as far back as 2002 it was reported that there was government CGE modeling capability for example in South Africa and Mozambique. Many other countries regularly used CGE models in policy analysis but relying on consultants and non-governmental research institutions.

18 Table 1.1 presents some of the modeling efforts in the MEFMI region.

<b>Lesotho</b>	<ol style="list-style-type: none"> <li>1. Ministry of Finance and Development Planning and the Central Bank of Lesotho have an <b>Econometric Model</b> using E-views, and</li> <li>2. A <b>Computable General Equilibrium (CGE) Model</b> using the General Algebraic Modeling System (GAMS).</li> </ol> <p>- The Database at the Ministry of Finance is linked with the Bureau of Statistics Database.</p>
<b>Zambia</b>	<ol style="list-style-type: none"> <li>1. Bank of Zambia has a <b>Bayesian VAR</b> using E-views and,</li> <li>2. An <b>ECM Inflation Model</b> incorporating inflation, petrol prices, exchange rate, money</li> </ol>

	<p>supply and output using Excel.</p> <p>3. The Ministry of Finance and National Planning (MFNP) has a <b>Macro Econometric Model (ZAMMOD)</b> which has well-organized data included in it and is estimated using Excel and other econometric computer software. The model is used for forecasting and performing simulations.</p> <p>4. The MFNP has a <b>Poverty module, which is attached to ZAMMOD</b>. This model is used to track macro-level outcomes on poverty. It is also used to forecast poverty and to target poverty reduction through macroeconomic interventions;</p> <ul style="list-style-type: none"> <li>• There is a time series data set of major macroeconomic variables located in the Bank of Zambia's Economics Department.</li> <li>• The Bank of Zambia is currently pursuing the possibility of using TSA for data storage and analysis.</li> <li>• ZAMMOD has a sheet (named source) which is essentially a time series data base of economic variables from 1994.</li> </ul>
<b>Swaziland</b>	<p>1. The Central Bank of Swaziland has a <b>Macro Econometric Model</b> using E-views.</p> <p>2. Economic Planning and Development has the <b>Threshold 21 (T21)-Long-Term Forecasting Model</b> using Vensim</p> <ul style="list-style-type: none"> <li>• Data is from the Central Statistics Office and the Swaziland Revenue Authority.</li> </ul>
<b>Botswana</b>	<p>1. Ministry of Finance and Development Planning (MFDP) had a <b>Small Macro Model - Econometric</b>, using E-Views and GAMS.</p> <p>2. MFDP had a Sectoral Model, the <b>Macro-Econometric Model for Botswana (MEMBOT)</b>, using E-Views and GAMS.</p> <p>3. MFDP has a Computable General Equilibrium Model (CGE), <b>BotsMod CGE Model</b>, using E-Views and GAMS.</p> <p>4. MFDP has <b>Inflation Targeting Frameworks (IT)</b>, using E-Views and GAMS.</p> <p>5. MFDP has <b>Financial Programming &amp; Policy Framework (FPP)</b>, using E-Views and GAMS.</p> <p>6. Bank of Botswana has the <b>Near Term Forecasting Model and Core Medium Quarterly Model (structural)</b>, using E-Views and GAMS.</p>
Source: MEFMI Member States, 2011	

19 Most of these models in African countries are of a short to medium term nature, normally matching the funding cycles of the development partners. This culture has continued to dominate most institutions responsible for modeling in many African countries to this day. Their over-reliance on donors for supporting the construction and maintenance of such models has also resulted in some African countries and institutions failing to sustain such modeling when the donors withdraw their funding. Highlighted below are some examples of key models currently being used in the MEFMI region.

20 Box 1.1 illustrates a prototypical example for a macro-econometric model of national income by Granger.

**Box 1.1: A Prototypical Example for a Macro-Econometric Model, the National Income Model by Granger**

$$C_t = a_1 + b_1 Y_t + e_{Ct}$$

$$I_t = a_2 + c_2 P_{t-1} + e_{It}$$

$$T_t = d_3 GDP_t + e_{Tt}$$

$$P_t = a_4 + b_4 Y_t + f_4 I_{t-1} + e_{Pt}$$

$$GDP_t = C_t + I_t + G_t$$

$$Y_t = GDP_t - T_t.$$

The endogenous variables are: consumption  $C$ , investment  $I$ , taxes  $T$ , profits  $P$ , gross domestic product  $GDP$ , disposable income  $Y$ . Apparently, the original Granger called  $Y$  the 'national disposable income' out of a misunderstanding, thus excluding the government sector from the national economy. Rather,  $Y$  appears to be a disposable income of the household and firm sectors. The exogenous variables are: government expenditure on goods and services  $G$  and the constant.

There are four structural or behavioural equations with error terms and two identities. The economy is closed, which is clear from the  $GDP$  definition. There are two simultaneous feedback cycles:

$C$  depends on  $Y$ ,  $Y$  depends on  $GDP$ , and  $GDP$  depends on  $C$ ;  $T$  depends on  $GDP$ ,  $GDP$  depends on  $C$ ,  $C$  depends on  $Y$ , and  $Y$  depends on  $T$ . Therefore, least-squares estimation will yield inconsistent estimates of all parameters (coefficients).

By substitution, for example  $GDP$  can be expressed by the predetermined variables:

$$GDP_t = A + \frac{G_t}{H} + c_2 \frac{P_{t-1}}{H} + e_{Gt}$$

where  $A = (a_1 + a_2) / H$ ,  $H = 1 - b_1 + b_1 d_3$ , and  $e_{Gt} = (e_{Ct} + e_{It} - b_1 e_{Tt}) / H$ .

Analogous substitution yields the reduced form of the system, which expresses all six endogenous variables by predetermined variables and error terms. Note that the error structure of the reduced form necessarily has a singular covariance matrix. The equations of the reduced form can be estimated consistently by least squares. Econometric textbooks consider the option of retrieving estimates for the original 'structural' coefficients 'backward' from these reduced form estimates by algebraic operations and call it indirect least squares. This method is rarely used in practice.

In order to forecast  $GDP_{t+1}$ , one may use:

$$GDP_t(1) = A + \frac{G_{t+1}}{H} + c_2 \frac{P_t}{H}$$

which assumes that approximations to the reduced-form coefficients  $A$ ,  $H^{-1}$ , and  $c_2/H$  are available in any case, either from least squares on the reduced form or from two-stage least squares on the structural model and simple 'forward' algebraic operations. While  $P_t$  is available at  $t$ ,  $G_{t+1}$  is not. Either one inserts a 'plausible value' for  $G_{t+1}$ , for example from the government's budget plan, or one uses an additional time-series or econometric equation for predicting  $G$ . The latter option 'essentially' completes the system by treating  $G$  as if it were endogenous' (Granger), that is, by endogenizing government policy. Granger suggests 'to use a model to forecast  $G$  and then to alter this forecast subjectively if relevant extra information is available to the forecaster. The success or lack thereof of such adjustments clearly depends on the quality of the information being utilized and the abilities of the forecaster who tries to use it'.

21 Box 1.2 illustrates recent progress in macroeconomic modeling and forecasting in the Banco De Mozambique.

**Box 1.2: Macroeconomic Modeling and Forecasting in the Banco De Mozambique, 2008 to 2012**

In 2008, the Banco de Moçambique asked for a second IMF Technical Assistance for further improvements in the monetary policy framework. Proportional to the size of the model to be developed, this process demanded additional effort towards the improvement of the database, by including other variables now available, and to the development and analysis of economic indicators, essentially, indicators of economic activity (the real sector). These efforts resulted in several meetings and agreements at different levels with the Instituto Nacional de Estatística (the National Statistical Office), and with different other institutions and companies considered as important sources of economic statistics and information. Notwithstanding the significant advances made in developing this range of indicators of economic activity and inflation, it is important to note that the availability of data is still a concern for further improvement of the model.

With this mission of Technical Assistance, the Banco de Moçambique has developed a small, semi-structural core macroeconomic model for Mozambique's economy that follows, according to Roger, Lima and Minella (2008)<sup>1</sup>, the basic structure of other central banks' models, but with some specific features of Mozambique's economy. Basically, as described by these authors, the model contains an IS curve for the output gap, a Phillips curve for CPI inflation, and a money demand equation, which is linked to the monetary instrument (base money) via the money multiplier. Formally, the model has:

Output gap equation (IS curve):

$$\tilde{y}_t = \beta_1 \tilde{y}_{t-1} + \beta_2 E_t \tilde{y}_{t+1} - \beta_3 (r_t - r_t^*) + \beta_4 (z_t - z_{t-1}^*) + \varepsilon_t^y,$$

Where:  $\tilde{y}_t$  = output gap

$r_t$  = real interest rate

$r_t^*$  = equilibrium real interest rate

$z_t$  = real exchange rate

$z_t^*$  = equilibrium real exchange rate and

$\varepsilon_t^y$  = shock to output gap.

This equation captures two transmission mechanism channels: the direct effect of the interest rate on economic activity and the effect of the real exchange rate on output via net exports.

Inflation equation (Phillips curve):

$$\pi_t = \alpha_1 \pi_{t-1} + (1 - \alpha_1) E_t \pi_{t+1} + \alpha_2 \tilde{y}_{t-1} + \alpha_3 (z_t - z_{t-1}) + \varepsilon_t^\pi,$$

Where:

$\pi_t$  = inflation rate

$y_t$  = output gap

<sup>1</sup> Roger, Lima e Minella (2008), Strengthening Monetary Policy Formulation and Implementation. IMF's Technical Assistance Report.

$z_t$  = real exchange rate and  
 $\varepsilon_t^\pi$  = shock to inflation.

The Phillips curve is the cornerstone of central banks' models. It captures the driving forces of inflation. The output gap term reflects the effects of the economy slackness on prices. The exchange rate term change captures the direct pass-through to tradables goods prices. This term could also be modeled as the nominal exchange rate change plus foreign prices change (with the corresponding change in the coefficients). The lagged inflation term reflects the stylised effect of the high degree of inflation persistence observed in the data. This sluggishness could reflect several factors, such as backward-looking behavior or institutional arrangements (such as indexation). The expected inflation term captures the forward-looking nature of pricing decisions. The presence of some price rigidity or costs to adjust prices turn firms extremely forward-looking when adjusting prices. Thus, it is necessary for them to make an assessment about future inflation values.

Considering that monetary policy of the Banco de Moçambique is based on a monetary targeting regime, there are two equations. The money demand equation:

$$(m_t^d - p_t) = \lambda_1 y_t - \lambda_2 i_t,$$

Where:  $(m_t^d - p_t)$  is the real demand for money

$y_t$  = output, and

$i_t$  = nominal interest rate

And the money supply equation:

$$m_t^s = \chi b_t^*$$

Where:  $m_t^s$  is money supply, and

$b_t^*$  is the monetary base target.

These equations in the system capture the first round in the transmission mechanism for monetary policy. Changes in money supply lead to changes in the interest rate, which in turn affects the other variables in the economy. In this specification, the central bank targets monetary base in order to affect money supply in the economy (via the money multiplier  $\chi$ ) and consequently, the interest rate.

The model was designed in an environment that allows conducting two kinds of monetary policy analysis, one where changes in the money base affect the interest rate and thus, the other macroeconomic variables; and the other, where the operational instrument is the interest rate.

The model has been in use by the Banco de Moçambique since December 2009, to produce forecasts for four quarters ahead, and since the second quarter of 2011 we have increased the horizon for up to eight (8) quarters. Our model has been performing accordingly since the deviations of our forecasts from the actual values have been very small.

Source: Banco de Moçambique, 2011

22 Box 1.3 presents Kenya's 'KIPPRA-Treasury Macroeconomic Model (KTMM)' as an example of some of the best practices in modeling in the region.

### **Box 1.3: KIPPRA-Treasury Macro Model (KTMM) - Kenya**

#### **Structure of the Model**

The KIPPRA-Treasury Macroeconomic Model (KTMM) is an aggregate supply/aggregate demand type of macro model of a market economy, and it describes the behaviour of market actors. It consists of equations not only from demand the side but also from supply. The model was developed by KIPPRA to analyze the effects of the economy on the Budget of the Government as well as the other way round: the effects of the Budget on the economy. It is a simultaneous model based on the national income accounting identity. The model is demand driven in the short-run, with multiplier effects through consumption and investment. An important assumption of the model is that any demand is met, that is, it is assumed that the price system ensures that there is always some excess capacity in the economy. High demand leads to high capacity utilisation rates of capital and low unemployment rates, however, which lead to wage and short-run price increases. With the important further assumption that the resulting inflation will decrease the competitiveness of exports, causing a reduction in exports and then lowering investments. In this way, the model has a tendency to return to equilibrium with 'normal' capacity utilisation and unemployment rates in the medium and long run.

These main feedback mechanisms in the real economy work through the wage-price spiral and also the real exchange rate. For instance, an increase in aggregate demand raises labour demand, reduces the unemployment rate, raises wages and starts a wage - price spiral. The resulting inflation causes a real appreciation, a reduction in competitiveness and a reduction in exports. The drop in exports and investment reduces demand again, until equilibrium is restored. An important point to note is that even though total demand may be stabilised in this way, the feedback mechanism may well change the composition of demand. For instance, if the original increase in demand came from an increase in government spending, the net result will be a shift from exports and investment to government spending, resulting in a government deficit and a current account deficit.

#### **Sheets of the Kenya Macro Model**

The KTMM is an Excel-based model. The Excel file with the Model consists of:

1. Sheet MANUAL: contains a detailed manual organised in blocks and details the design of the consistency framework, discusses every sheet of the model, calibration results of all semi-behavioural equations, explains the specification and coefficients of all behavioural equations, how to run a simulation, gives the sources and literature list, and discusses the activities performed during the construction stage of the model.
2. Sheet HELP: gives variants for a number of variables, ready to be used in all kinds of combinations of variables and years.
3. Sheet SOURCE: contains the information about macro time series data that were gathered from several sources. Column A gives the name of the time series, column B the unit and column C and D indicate the source. Column E and subsequent columns display calendar years, starting with 1971. These columns contain the values from the different sources. This sheet does not contain any formulas.
4. Sheet MODEL: its North-west quadrant contains cells with values of the so-called primary variables for the years in the past. All these values are calculated by formulas or references only using information from sheet SOURCE. The North-east quadrant contains formulas that calculate future values for the primary variables. The cells in the Southern part of the sheet Model contain help variables (also called auxiliary variables) their value is calculated by formulas using only information

from the primary variables. These help variables are used as explanatory variables in the (semi-) behavioural equations. So the time series in sheet MODEL contain all time series that are needed to estimate the behavioural equations and to calibrate the semi-behavioural equations. Sheet MODEL does not contain any input of values (other than calculated by formulas or references to sheet SOURCE. In column C of sheet MODEL, you can click in each row (variable) to go immediately to the relevant part of sheet Manual where you will find background information concerning that variable. The scheme of the KIPPRA-Treasury Kenya Macro Model (KTMM) is as shown in the Table below:

### Main worksheets of Kenya Model

5. Sheet HISTAN contains time series data and graphs to analyse the relation between variables.
6. Sheet OUTPUT contains the output of the key variables (in calendar years). All output variables are calculated using values from sheet MODEL and no other sheet.
7. Sheet FY contains the same information as in sheet OUTPUT for Calendar Years for the Central Government also for Fiscal Years.
8. Sheet THEORY contains the text of the theory behind the behavioural equations.
9. The sheet CALIBRATION contains for each (semi) behavioural equation a graph with a comparison of the actual figures of the historical time series, and the simulations based on the forecasting formula, using the realised, actual figures for the explanatory variables.
10. The sheet TAX contains the calculation of the average marginal wage tax rate, based on the tax rates per bracket and a division of wage earners to income brackets.
11. The sheet MICROEXPORT contains micro data for tea and coffee, to be used for the break-down of the export equation
12. The sheet ADD-FACTOR contains the add-factors that are in use in the (semi) behavioural equations in sheet MODEL.
13. Sheet sectors give a breakdown of forecasts to production growth of sectors of industry.
14. Sheet IMF OUTPUT gives some of the output variables in IMF format.
15. Sheet decomposition gives a breakdown of the forecasts for main variables to the contribution of each explanatory variable. This sheet makes analyses of forecasts and variants easier.
16. Sheet MONITOR gives monthly information concerning CPI etc. and a calculation scheme to judge the plausibility of the forecast of the running year, given the already available monthly figures and trends.
17. Sheet LOGBOOK is used to document all changes in the model that are brought into the most recent official version of the model.

*Source: website at: <http://www.micromacroconsultants.com/Engels/Downloads/Downloads/>; Micro-Macro Consultants, the main consultants who helped KIPPRA to develop the model.*

23 Box 1.4 presents the IMF Polak Model as an example of some of the international best practice in macroeconomic modeling.

#### Box 1.4: The IMF Model - the Polak Model

The Polak model was the main framework the IMF applied in its stabilisation program for developing countries. The model essentially consists of four equations: two identities, one for nominal money supply as a function of domestic credit and official foreign exchange reserves, the other for the relation of official reserves to current account and capital inflows; a behavioural equation for nominal demand for money as a function of changes in nominal income (ignoring other determinants of money balances such as interest rates and wealth); and an equilibrium condition, which assumes that the money market is in flow equilibrium and equates changes in nominal money supply to changes in the nominal demand for money. The main use of the Polak model was to assess the effects of changes in domestic credit on the



balance of payments. The model has a number of limitations. These include the assumption that changes in domestic credit have no effect on the determinants of money demand such as real income or domestic interest rates and the fact that it largely ignores the real side of the economy, concentrating on the external sector through its focus on balance of payments. The model is also limited in its application to economies with regimes of fixed exchange rate and from the inherent closure where exports and capital flows are treated as exogenous. The exogeneity of exports is realistic if the exchange rate is fixed, but where the exchange rate is flexible, exports have to be endogenous.

The Polak model has been extended to address some of the limitations inherent in the original formulation. One such extension is by Khan et al. (1990), which distinguishes explicitly between changes in real and nominal output and the sources of credit growth. Domestic credit consists of credit to the private sector and the government. Capital inflows are both private and public. Foreign direct investment intended for the private sector is distinguished from foreign borrowing mainly by the government. A significant extension to the original Polak model is the introduction of the government sector through a government budget constraint. This extension is appropriate in the sense that the balance of payments problems addressed by the original Polak model arise as a result of the difficulties encountered in financing the fiscal budget balance. The government budget constraint relates the budget deficit to foreign borrowing and changes in Central Bank credit. A limiting assumption in this extension is that the government cannot borrow directly from the domestic private sector through bonds. Despite the various extensions, the model has the acknowledged limitation of assuming that the economy produces one domestic good, which is used for both domestic consumption and exports. A more appropriate framework for developing economies would distinguish between exportables, non-tradables and importables.

*Source: Karingi and Ndungu (2000)*

24 Box 1.5 presents the World Bank Revised Minimum Standard Model (RMSM) as an example of some of the international best practice in macroeconomic modeling.

#### **Box 1.5: The World Bank Revised Minimum Standard Model (RMSM)**

The World Bank's revised minimum standard model (RMSM), is a real economy model whose objective is to make explicit the link between medium-term growth and its financing. The model has five relationships: an equation relating the desired level of investment to the change in real output; an imports equation; a private consumption equation; the balance of payments identity; and the national income identity. The target variables in the model are change in official foreign reserves and change in output. In the macroeconomic closure, exports are treated as exogenous. Key policy instruments are government expenditure, tax revenues and change in net foreign borrowing. This model is described as a trade gap model since the treatment of real output as a target variable essentially fixes imports and, given exogenous exports, the trade balance is fixed. The trade gap framework works well if policy-makers have sufficient control over capital inflows. Because of the foreign borrowing limits that economies invariably face, a savings gap framework of the same model is used. Where both the trade gap and the savings gap are binding constraints, RMSM becomes a two-gap model, and the solution depends on which constraint is more binding.

A main criticism of the two-gap RMSM is that in determining the financing requirements for alternative target rates of output growth, one cannot tell, a priori, whether the savings or the trade constraint will be binding. A second criticism, and one that is more serious in applying RMSM, is that it fails to incorporate the financial sector of the economy. The government sector is also minimal in the model while the exogeneity of prices is a major limitation. The original RMSM has been revised further to what is now called RMSM-X. It combines the extended Polak model with the Bank's RMSM model.

The improvement in RMSM-X is that the financial sector component is incorporated through the extended IMF model. Endogenising the prices in RMSM-X is a positive attribute. Nevertheless, RMSM-X still suffers from the limitations of the two models underlying it: the rudimentary nature of the financial programming model and its failure to include many important features such as bond financing of fiscal deficits; its failure to incorporate a labour market; and the use of an inflexible production function on the supply side. Different closure rules normally solve RMSM-X: normative, positive and policy closures. Whereas this makes the model versatile, it has a general weakness in that it is usually solved recursively, ignoring the simultaneous nature of key macroeconomic variables. The recursive solution to the model does not allow for explicit consideration of the relationships among variables.

*Source: Karingi and Ndungu (2000)*

## 1.5 Macroeconomic Modeling Challenges and Constraints Experienced by MEFMI Member States

25 MEFMI member states are at different stages of macroeconomic modeling, with efforts under way to build both small-to-medium, as well as large econometric and computable general equilibrium (CGE) models.

26 The most common computer platform being used are E- views, Excel, and General Algebraic Modeling System (GAMS), (particularly for Computable General Equilibrium Models).

27 Macroeconomic modeling efforts in the MEFMI region are encountering many challenges, which if not addressed by member states, constitute serious risks to the success of the MFM. Some of the challenges are generalised in the region, whilst others are country specific, and these include:

### **(i) Shortage of modeling staff**

- General lack of modeling capacity;
- High staff turnover of staff who are knowledgeable in modeling techniques;
- Modeling staff shortages, they are few in each modeling unit.

### **(ii) Computer software challenges**

- Computer software problems including limited econometric software packages;
- Need for more computer programme licenses;
- Inadequate knowledge of advanced Excel and other computer software packages required for modeling.

**(iii) Data challenges**

- Data series not long enough;
- Data gaps, including data unavailability;
- Unavailability of reliable, high frequency and timely data;
- It is difficult to accurately compare model forecasts and actual trends due to the use of out-of-date GDP bases, when the structure and composition of the economy has changed;
- Data bases are not centralised and in most cases they are kept by model builders on their personal computers, and when they leave employment, they go with the data and the model.

**(iv) Overall modeling skills challenges**

- Models not producing expected economically meaningful results;
- Lack of sustainability of the models developed;
- Dependency on external international consultants.

**(v) Institutional and coordination challenges**

- Team members based in different institutions and working on different models;
- Lack of coordination of the modeling efforts within each country; and
- Generally, coordinated modeling teams are non-existent in the countries, yet modeling requires serious monitoring to succeed.

28 If the MFM is to be useful in the region, MEFMI member states will need to commit significant resources towards the macroeconomic modeling agenda in the region, in an effort to resolve the highlighted challenges.

**1.6 Evolution of Macro - Models**

29 Economic modeling is not an entirely new phenomenon. It is an area that has for so many years received little attention, especially by policy makers. Modeling has an interesting history of more than half a century, having initially been a way to implement Keynes' General Theory, which was demand-driven emphasising the output, employment and wage trade-offs. As time went on, there were some modifications to try and introduce some alternative paradigms that were coming up such as the Monetarist view which emphasised money supply's role in inflation, New Keynesian approach which emphasised rational expectations, imperfect competition and sticky prices; and the New Classical supply-driven, rational expectations approaches (Bodkin, Klein, & Marwah, 1986a).

30 According to Bautista (1988) and Capros, Karadeloglou, and Mentzas (1990), there are basically two types of macroeconomic models: (i) macroeconometric models (MEMs), and (ii) computable general equilibrium (CGE) models.

31 The distinction between MEMs and CGE models can be related to the time horizon. The early CGE models involve comparative statics. In comparative statics, one starts from a base solution or scenario (also known as a counterfactual), then one of the exogenous variables or parameters is altered and the model calculates the new values for the endogenous variables. The comparison of the two sets of values of the endogenous variables suggests the estimated economic impact of the change in policy (intervention). The path taken by variables of interest to move from one equilibrium point to another is not of importance here, only the comparison of the two equilibria. Dynamic analysis, on the other hand, is concerned with looking at the path from one equilibrium to another. More recently, some CGE models, for example, Dixon Malakellis (1995), involve a dynamic adjustment process which can be used for short-term (3–5 years) and medium-term (5–7 years) analysis. MEMs provide information on the dynamics of the adjustment process, which is useful for short-term policy analysis.

32 Challen and Hagger (1983, pp. 2–22) classify MEMs into five categories: the Keynes–Klein (KK) model, the Phillips–Bergstrom (PB) model, the Walras–Johansen (WJ) model, the Walras–Leontief (WL) model, and finally the Muth–Sargent (MS) model. See Table 1.2 for the key features of each of these five categories of MEMs.

<b>Macroeconometric Model Category</b>	<b>Key Features</b>
1. The Keynes–Klein (KK) model	The main purpose of the KK model is to explain the Keynesian demand-oriented model of macroeconomic fluctuations. This model deals with the problems of short-run instability of output and employment, using mainly stabilisation policies. The basic Keynesian model has been criticised for not considering the supply side including production relations. Furthermore, this model does not give sufficient attention to the role of the money market, relative prices and expectations.
2. The Phillips–Bergstrom (PB) model	The second type of MEM, the PB, emerged in literature when Phillips (1954, 1957) used both Keynesian and Neoclassical theories within a dynamic and continuous time model to analyse stabilisation policy. The PB model is also a demand-oriented model. Generally speaking, the PB model mainly employs differential or difference equations to estimate the structural parameters of a stochastic model. Thus, in this approach the steady state and asymptotic properties of models are examined in a continuous time framework. However, this method is difficult to implement especially for large scale models.
3. The Walras–Johansen (WJ) model	The WJ, is mainly a multi-sector model. Walras (1954) postulates that the economy consists of various inter-dependent markets, which reach an equilibrium state by the profit maximising behaviour of producers and utility maximising actions of consumers in competitive markets. The

	various sectors in the WJ model are linked together via their purchases and sales from, and to, each other. The WJ model is highly non-linear and uses logarithmic differentiation.
4. The Walras–Leontief (WL) model	The WL model, has been more widely considered as the most appropriate MEM for developing countries (Challen and Hagger, 1983). The WL model is an example of a general equilibrium system. This type of modeling incorporates an input–output (IO) table into the Walrasian general equilibrium system. By using an IO table, given the values of the sectoral, or aggregate, final demand components, the sectoral output (or value added) can be obtained.
5. The Muth–Sargent (MS) model	The MS, is based on the evolution of the theory of rational expectations. The MS model is very similar to the KK model with regard to being dynamic, non-linear, stochastic and incorporating the use of discrete time. One of the distinguishing features of the MS model pertains to the formation of expectations. In this model, the formation of expectations is no longer a function of previous values of dependent variables. In fact, the expectation variables are not observed values, but can be obtained when the complete model is solved. By highlighting the role of the supply side and expectations in this model, the New Classical School aims at interpreting the inadequacy of demand management policies. On these grounds, Sargent (1976) devised a forward-looking MEM. Variants of this model indicate that there is no trade-off between inflation and unemployment in the short term, which is in contrast to both the Keynesian and Monetarist modeling perspective.
<i>Source: Constructed from Diebold Francis X, University of Pennsylvania and the National Bureau of Economic Research (NBER), 1997.</i>	

33 The economic impact of policy depends upon a number of interrelated factors: the specific policy and production technology employed; the sector’s embeddedness to the rest of the economy, through its demand for local resources; the extent to which new activity is created etc. These issues can be analysed using multi-sectoral economic models. Three well known economy-wide approaches are the Input-Output (IO), Social Accounting Matrix (SAM) and Computable General Equilibrium (CGE) modeling frameworks.

34 IO models were pioneered by Wassily W. Leontief in the early 1950s. However it has been noted that early antecedents of input-output models date back to the 18th century in the writings of Francoise Quesnay (Isard, 1960). Input-output models were extremely useful to the central planners of communist regimes and have become the staple for impact analyses throughout the world.

35 SAM's were originally developed at the “Cambridge Growth Project” in Cambridge, UK, which developed the first SAM in 1962 (Stone and Brown 1962). They were built as a matrix representation of the National Account, and came to the World Bank with Graham Pyatt in the 1960s.

36 CGE models are primarily based on neoclassical theory of general equilibrium, first formulated by Leon Walras in 1877 and later formalised by Arrow and Debreu (1954) and McKenzie (1954, 1959, 1981). The model seeks to explain production, consumption and prices in a whole economy in which agents respond to relative prices as a result of profit and utility maximising behaviours. Markets simultaneously adjust relative prices in order to reconcile endogenous supply and demand decisions, and thus, determining levels of production and consumption. Although most of the equations feature rigorous microeconomic foundations, the behaviour of economic agents must be consistent with the macroeconomic framework.

37. Diebold Francis X (University of Pennsylvania and the National Bureau of Economic Research (NBER, 1997) provide a good evolution of both structural and nonstructural macroeconomic forecasting. The starting point is tracing the rise and fall of the structural Keynesian system-of-equations paradigm, into assessing the long-running and on-going progress in the non-structural tradition, ending with an assessment of the rise of modern dynamic stochastic general equilibrium (DSGE) macroeconomic theory, its relationship to non-structural methods, and its implications for a new structural macroeconomic forecasting.

### 1.7 Criticisms of Macroeconometric Modeling and their Responses

38 From the 1970s onward, macroeconometric modeling was criticised predominantly on academic grounds. According to Pesaran (1995), the major criticisms of the traditional MEMs based on the Cowles Commission approach can be summarised in terms of six issues: (i) forecasting inadequacy; (ii) theoretical contrasts with rational expectations theory; (iii) structural instability (Lucas critique); (iv) arbitrary assumption of zero restrictions (i.e. causal ordering) or the endogenous–exogenous division of the model variables in order to pass the identification conditions; (v) the existence of the problem of unit roots and ignorance of co-integration and (vi) the time-series properties of the data.

39 However, the Lucas (1976) critique, inter alia, had a powerful influence in decreasing the application of MEMs for policy analysis. The Lucas critique led to a new area of research, which is referred to as analysis of “deep structural parameters” (Fair, 1987). It is mentioned that under alternative policy formulations, because all the economic agents base their decisions on the full information, “any change in policy will systematically alter the structure of econometric models” (Lucas, 1976, p. 41). Therefore, it is highly likely that the estimated coefficients of a MEM will vary as a result of agents anticipating and knowing policy measures. Consequently, Lucas rejects the use of MEMs for the policy analysis.

40 On the contrary, Intriligator et al. (1996) assert that MEMs are useful in structural analysis, forecasting and policy evaluation provided that they are subjected to some parametric tests prior to and after the release. The tests they recommend, which are of paramount importance in evaluating the validity of MEMs, can be classified into two major categories: testing for individual equations of the model and testing for the full model (system) as a whole.

41 The first category, which will be emphasised in this Manual consists of four sub-tests as follows:

- i. The standard t and F tests to check for the statistical significance of the estimated parameters;
- ii. Testing for the expected theoretical signs of the estimated parameters;
- iii. Diagnostic tests for checking various violations of the classical linear regression model such as normality, autocorrelation and serial correlation, heteroskedasticity, functional form, etc.; and
- iv. Checking for the stability of the estimated equation over time.

## 1.8 Lessons Learnt

42 After more than half a century of macro-modeling, the question arises of, what are the main lessons that can be learnt from past experience. This section discusses some of the lessons proposed by leading model-builders.

43 According to Bodkin and Marwah (1988) and Bodkin et al. (1991), the new generation of modelers should take advantage of the current developments to build large scale MEMs and run various econometric diagnostic tests. These current developments consist of improvements in computational capacity, new developments in econometric methods, new macroeconomic theories and advances in the quality and availability of the required data.

44 Since the analysis of the economy will be more difficult when there are numerous equations in the model, advocates of small scale modeling conclude that the small model can explain the economy in a better way. They argue that one needs to Keep It Sophisticatedly Simple (KISS). It is "much easier to see the forest when the trees are fewer" (Bodkin & Marwah, 1988, p. 301). This Manual is based on this principle<sup>2</sup>.

---

<sup>2</sup> However, Klein (1999) argues that small models may fail capture the complex nature of an economy and this may result in misleading policy conclusions.

45 The main lessons from economy-wide analysis (IO, SAM, CGE) according to Devarajan and Robinson (2002) that would be useful for policy analysis is that economic models should have the following desirable features:

- i. **Policy relevance:** The models should link values of policy variables to economic outcomes of interest to policy makers and useful in policy debates
- ii. **Transparency:** The links between policy variables and outcomes should be easy to trace and explain.
- iii. **Timeliness:** Policy models must be based on relevant data, which implies that they must be implemented with recent data if they are to be used in ongoing policy debates.
- iv. **Validation and estimation:** Estimated model parameters and model behaviour need to be validated for the domain of application of the model. That is, the model must be determined to achieve accurate results for the domain of potential policy choices under consideration in the policy debate.
- v. **Diversity of approaches:** Validating results from policy models is greatly strengthened by analysis using a variety of models and at different levels of aggregation. Such diversity tests the robustness of the results and the importance of assumptions made in the various approaches.

## 1.9 Structure of the Manual

46 The MFM is divided into two major Parts. These two major parts are presented after the preface and the introduction to the *Manual*, which cover: background to the *Manual*; purpose and scope of the *Manual*; macro models and economic policy formulation including the modeling challenges in the MEFMI region; evolution of macro-models and lessons learnt; regional developments in economic modeling including examples of some of the models in use in the region; ending with the structure of *Manual*.

47 Part I of the MFM is divided into 4 Chapters as follows: Chapter 2 covering the Theoretical/Logical Framework and Chapter 3 on the Estimation of the Equations of the Blocks, (mainly step-by-step practical illustrations in E-views and Excel). The same chapter also covers issues on data requirements including data types and frequency, data evaluation, data gaps, inconsistencies and remedies and basic database (a “how to” guide). Chapter 4 is on the Building of a Prototype Model, (mainly step-by-step practical illustrations in E-views and Excel). Chapter 5 covering Other Forecasting Methods brings Part I of the *Manual* to a conclusion.

48 Part II of the MFM consists of Chapter 6 on Introduction to Economy-Wide Modeling, which basically introduces computable general equilibrium (CGE) models starting with input and output (IO) analysis and social accounting matrix (SAM) analysis, with examples illustrated in both GAMS and Excel.



# **PART I**

# **MACRO-ECONOMETRIC MODELING USING TIME SERIES DATA**

---

---

## CHAPTER 2

### THEORETICAL / LOGICAL FRAMEWORK

In this chapter, we take the reader through

- Key Interactions between the Macro Blocks
- Description of the blocks
  - Real sector Block
  - Fiscal Block
  - Monetary Block
  - External Block
- Closure Rules
- Consistency Framework in the Blocks
- Possible Challenges when modeling the blocks

#### 2.1 Need for Theory in Macroeconomic Modeling

49 In this chapter, we describe the general theoretical underpinning of structural models. This can easily be extended to general equilibrium models. The Chapter describes in detail the different blocks of the model including the real sector block, the monetary sector block, the external sector block, fiscal sector block and the government sector. The closure rules and possible challenges one would encounter when modeling these blocks are also discussed.

50 Recent attempts to analyze macroeconomic activity have gone deeper to seek the micro foundations of macro behavior. For instance, modelers have attempted to establish the behavior of individual households and firms and how they interact in the goods market through supply and demand forces of goods and their interaction in the labour markets and other markets. The need for a deeper understanding of these interactions have necessitated the emphasis on the theory that guides those relationships. Structural modeling therefore tends to emphasize more theory to try and understand the inter-linkages between the variables under study.

51 Most macroeconomic models outline the theories upon which they are based. These are commonly described as blocks. The blocks represent the main economic agents in the country and the interaction between the blocks/agents defines the outcome in the economy including unemployment levels, economic growth rates, inflation rates, etc.

---

---

## 2.2 Choosing the Right Model to Build

52 The first task that a macro modeler would do before starting to build a model would be to identify the kind of model that will deliver what they want. There are several kinds of models; partial equilibrium models vs general equilibrium models, long-term vs short-term models, quarterly vs annual vs monthly models, etc.

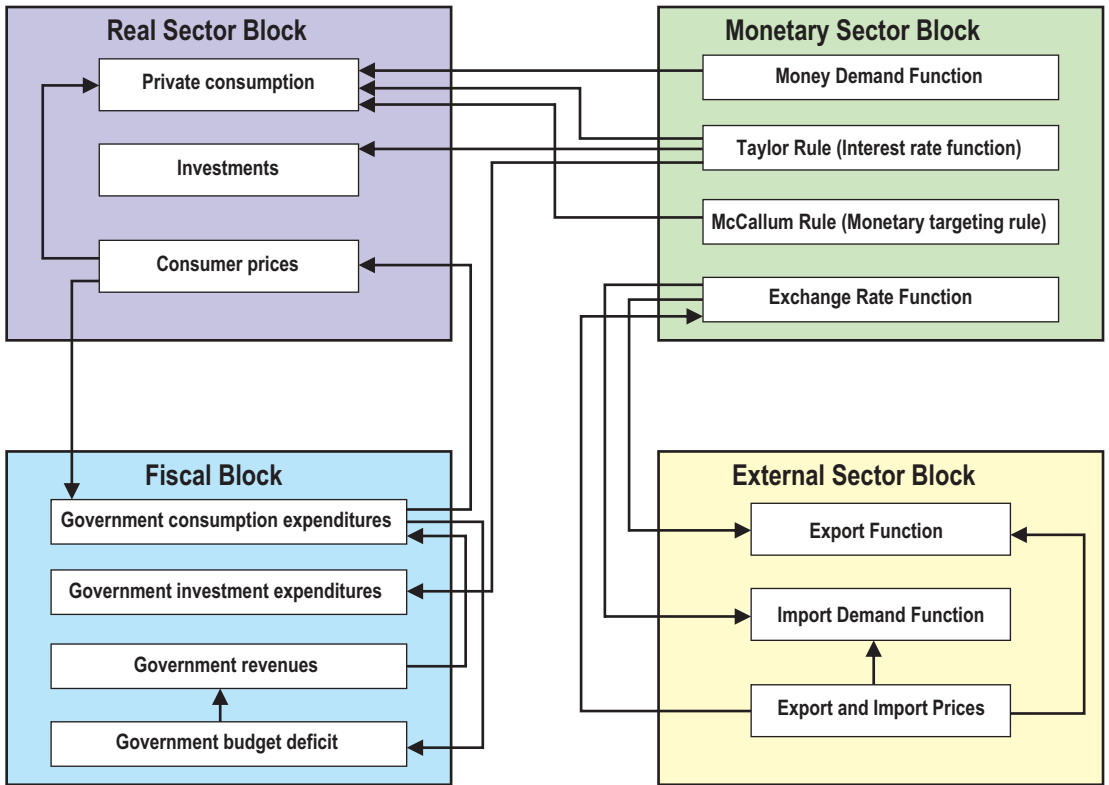
53 Several factors come into consideration when choosing a model. The first and most important is the kind of policy question that the modeler wants to answer. If the modeler wants to use the model to determine how inflation responds to changes in money supply or the Central Bank Rate, a short-run model would be more appropriate compared to a case where the main interest is to determine the response of government expenditure on economic growth. Central bankers, in particular, are mostly concerned with day-to-day liquidity forecasting and therefore may be more interested in a shorter term model. Officials from the ministry of planning and finance on the other hand, may be more interested in long term development planning. They may therefore be more interested in a long-term model. The second important consideration would be the type of data available. You may want to build a quarterly model but you do not have data on some or most of the variables in the equations of the model. As is well known, high frequency data like quarterly and monthly data for most variables including GDP are not available in most developing countries. It will therefore be difficult to build a quarterly model where some of the most important variables are not available in quarterly series. Other than data issues, you will have to choose the theoretical basis of your model, is it Keynesian, Neo-Keynesian, demand side or supply side etc.

## 2.3 Key Interactions between Macro Blocks

54 The starting point in model building is to figure out the interactions between the various economic agents in the economy. The interactions are normally identified in line with the theoretical basis upon which your model is based. The interactions are normally captured in blocks, see Figure 2.1.

55 Modelers specify interactions between blocks differently from one model to another. This depends on the assumptions behind each block and the number and kind of behavioral equations in each block. The figure below illustrates some of the key interactions in a hypothetical macro model. It is important to note that the interactions as given in the figure below may not be very exhaustive and the modeler must think about other interactions that may be relevant for their economy.

Figure 2.1: Key Interactions in a Macro Model



## 2.4 Description of Blocks

56 Most macroeconomics models consist of several blocks including production, expenditure, balance of payments, government, monetary and prices blocks. They could be also categorized into real sector block, monetary sector block, external sector block and the fiscal sector block. Other than that, there are several behavioral equations and identities in each block. The next sub-section gives a description of some of the possible blocks in a macro-model. It is noteworthy that the blocks must not necessarily be as they are given here. Some modelers merge some blocks while others disaggregate them further.

### 2.4.1 Real Sector Block

The real sector block may consist of consumption, and investments and their prices including goods prices, wages and interest rates.

#### (A) Private Consumption function

57 There are several consumption theories that can be used to model consumption. Some of these theories include;

**(a) The Keynesian Consumption Function**

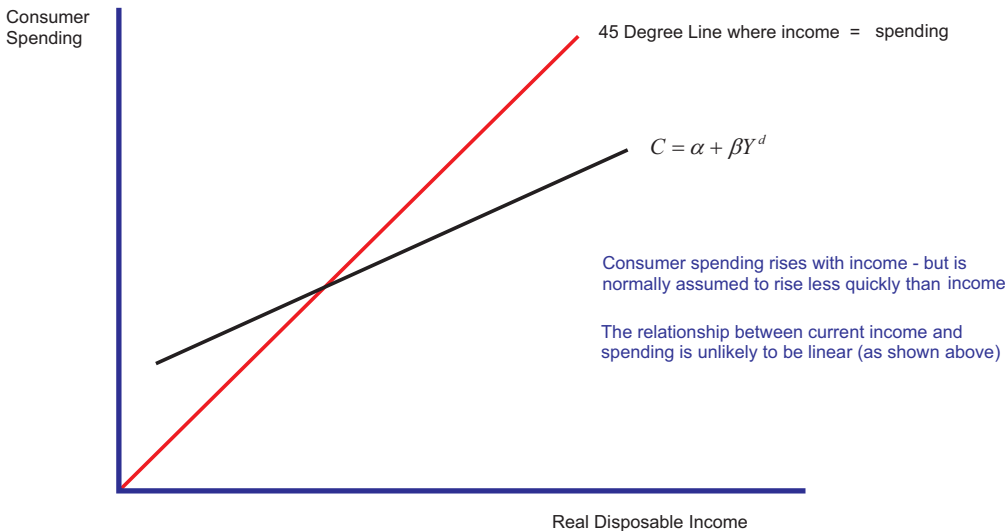
58 Keynes argued that 'The fundamental psychological law is that men are disposed, as a rule and on the average, to increase their consumption as their income increases, but not by as much as the increase in their income'. Keynesian consumption theory therefore states that consumption is solely a function of disposable income. That is:

$$C = \alpha + \beta Y^d$$

where,

$C$  = Consumer expenditure, and  $\alpha$  is the autonomous consumption. This is the level of consumption when the person earns no income i.e. when income was zero.  $\beta$  is the marginal propensity to consume (mpc). This is percentage of each additional income earned that is spent. If you spend 50 percent of the increased earnings, then mpc equals to 0.5.

**Figure 2.2: 45 Degree Line (Income = Spending)**



59 Disposable income  $Y^d$  is the total gross income minus deductions from direct taxation and benefits.

60 Empirically, the Keynesian consumption function is specified as;

$$C_t = \alpha + \beta Y_t^d + \varepsilon_t$$

where  $\varepsilon_t$  is the error term.

**(b) Inter-temporal Choice Consumption Model**

61 This was introduced by Irving Fisher. It assumes that the consumer is forward-looking and chooses consumption for the present and future periods to maximise lifetime satisfaction. Consumer's choices are subject to an inter-temporal budget constraint, a measure of the total resources available for present and future consumption. In this model, consumption is divided into different periods, period 1 for instance (present period) where the person works, earns income, consuming part of it and saving part of it. In the second period (future), the person consumes what he/she saved. Consumption in period one is therefore equal to the income in period one less savings done in period one; that is;  $S = Y_1 - C_1$ . The budget constraint facing the consumer in period two therefore is given by;

$$C_2 = Y_2 + (1+r)S$$

$$\Rightarrow C_2 = Y_2 + (1+r)(Y_1 - C_1)$$

Rearranging the terms gives:

$$(1+r)C_1 + C_2 = Y_2 + (1+r)Y_1$$

Dividing through by  $(1+r)$  we get the inter-temporal budget constraint as:

$$C_1 + \frac{C_2}{1+r} = Y_1 + \frac{Y_2}{1+r}$$

62 Fisher's inter-temporal consumption theory therefore states that the current level of consumption depends only on the present value of lifetime income and interest rates.

**(c) Relative Income Hypothesis**

63 Duesenberry (1947) noted that in 1935 dissaving grew as a percentage of income. In particular, dissaving was greater in 1935 than in the relatively prosperous year 1941. He theorized that households must sacrifice saving to maintain their highest level of consumption. Thus:

$$\frac{S_t}{Y_t} = F \left[ \frac{Y_t}{Y_{peak}} \right]$$

For simplicity, he assumes linearity, so that.

$$\frac{S_t}{Y_t} = a + b \left[ \frac{Y_t}{Y_{peak}} \right]$$

$$Y_t = C_t + S_t \quad \text{so that} \quad 1 = \frac{C_t}{Y_t} + \frac{S_t}{Y_t} \quad \text{and} \quad \frac{S_t}{Y_t} = 1 - \frac{C_t}{Y_t}$$

$$\text{But} \quad \frac{C_t}{Y_t} = (1-a) - b \left[ \frac{Y_t}{Y_{peak}} \right] \quad \text{so that}$$

$$C_t = \left\{ 1 - a - b \left[ \frac{Y_t}{Y_{peak}} \right] \right\} Y_t$$

If  $Y$  grows steadily, then

$$\frac{Y_t}{Y_{peak}} = \frac{Y_t}{Y_{t-1}} = k, \text{ and}$$

$$C_t = (1 - a - bk)Y_t, \text{ or}$$

$$C_t = \lambda Y_t, \text{ with } \lambda = (1 - a - bk)$$

#### (d) The Permanent Income Hypothesis

64 This theory, introduced by Milton Friedman, postulates that people maximise utility based on their permanent (expected life-time) income. They allocate their income inter-temporally. The budget constraint is:

$$c_1 + \frac{1}{1+r}c_2 = Y_1 + \frac{1}{1+r}Y_2 = w_e$$

Where  $q_1, q_2$  are planned real receipts from human and nonhuman wealth in each period, and  $w_e$  is expected wealth. Optimization yields:

$$c_1 = i_A(r)w_e$$

$$c_2 = i_B(r)w_e$$

This implies that consumption depends upon income only if the income affects expected wealth. Therefore, changes in income need not affect current consumption. We may view  $yp = rwe$  as permanent income – what really affects the consumption decision.

$$c = cP + cT$$

$$y = yP + yT$$

Both consumption and income has permanent (p) and transitory (T) components. For each individual agent:

$$cP = k(r, v, x)yP$$

where  $r$  is the rate of return,  $v$  is the ratio of human to nonhuman wealth (permanent income) and  $x$  are the demographics. By summing across individuals, we have:

$$CP = \kappa YP,$$

This implies that permanent consumption is proportional to permanent income. When reduced to an estimable form, this becomes:

$$c_t = \kappa \alpha_0 Y_t + \lambda c_{t-1} + \theta_t$$

### (e) Life Cycle Hypothesis (LCH)

65 The life Cycle consumption Hypothesis was developed by Franco Modigliani, Albert Ando, and Richard Bloomberg. It assumes that each representative agent will die, and knows when he/she will die, how many periods  $T$  he/she will live, and how much his/her life-time income will be. The consumer smoothens consumption expenditure over his/her life, spending  $1/T$  of his/her life-time income each period. The consumption function implied by this logic is:

$$C_t = \frac{1}{T} [Y_t^1 + (N-1)\bar{Y}^{1e} + A_t]$$

with the aggregate estimable consumption function looking like this:

$$C_t = b_1 Y_t^1 + b_2 \bar{Y}^{1e} + b_3 A_t$$

Ando and Modigliani argue that expected future labor income is proportional to current income, so that the function can be reduced to:

$$C_t = (b_1 + b_2 \beta) Y_t^1 + b_3 A_t$$

### (B) Investment Function

66 Investment demand can be divided into private investment and public investment. Private investment can be classified by sectors if sectoral models are included or taken as an aggregate. Investments may be considered as a function of capital stock in the previous period and the user cost of capital (interest rates). Corporate profits may also be important determinants of investments since retained profits may be used to finance investment. In addition, the capacity utilization rate may play a role as a direct indicator of the difference between optimal and actual capacity. Adding these elements, we get:

$$I_t = \theta_0 + \theta_1 K_{t-1} + \theta_2 r_t + \theta_3 \text{Pr} + \varepsilon_t$$

where Pr denotes retained profits.

67 Other possible theoretical models of investments that could be used include the Accelerator Model. The Accelerator Model attempts to capture some measure of current business conditions (growth of the economy or lack of it), and use that to explain the level of investment. The desired capital stock is proportional to the level of output:

$$K_t^d = \alpha Y_t$$

Investment is the process of moving from the current level of capital to a desired level:

$$I_{n,t} = K_t^d - K_{t-1}$$



It assumes that whatever the capital stock ended up being last period was the level of capital that businesses actually wanted:

$$K_{t-1} = K_{t-1}^d = \alpha Y_{t-1}$$

This allows us to rewrite

$$I_{n,t} = K_t^d - K_{t-1}$$

As

$$I_{n,t} = K_t^d - K_{t-1} = \alpha Y_t - \alpha Y_{t-1} = \alpha(Y_t - Y_{t-1})$$

$$I_{n,t} = \alpha \Delta Y_t$$

### (a) Cost of Capital Approach

68 Keynes argued that firms compare the expected revenue stream from the new capital to the user cost of capital. The user cost of capital is the total cost to the firm of employing an additional unit of capital for one period. The new capital might be funded by borrowing, selling stock shares, retained earnings, etc. This suggests that investment is a function of income, the cost of capital and the capital in the previous period.

$$I_{n,t} = I(Y_t, CC_t, K_{t-1})$$

where;  $CC_t$  is the user cost of capital in the current period.

If a firm invested its retained earnings or monies raised by selling stock shares, it could earn the current interest rate. Investment will depend on the real interest rate  $\phi$ , where:

$$\phi = r - p^e$$

where  $r$  is the nominal interest rate and  $p^e$  is the expected rate of inflation. Part of the user cost of capital is the depreciation rate  $\delta$ .

$$CC = \phi + \delta = r - \dot{p}^e + \delta.$$

If the government gives a subsidy equivalent to  $\tau$ , then the effective cost of capital to the firm is

$$CC = (1 - \tau)(r - \dot{p}^e + \delta).$$

The investment function therefore is given as:

$$I_{n,t} = I(Y_t, r_t, \dot{p}_t^e, \tau_t, K_{t-1})$$

69 Thus, investment is related to the rate of change in output. If the economy is growing rapidly, then investment grows rapidly and if the economy is not growing, then investment slows, and net investment (after depreciation) may actually be negative. As a result of adjustment costs and practical time-to-build considerations, the entire adjustment to the desired capital stock may not be done in one period. The firms may only finance a partial adjustment. Let  $\lambda$  be the fraction of the gap between the desired and

actual capital stock that the firms pursue. This leads to:

Or equivalently,

$$I_{n,t} = \lambda (Y_t - K_{t-1})$$

This is referred to as the flexible accelerator model of investment.

### (C) Real Sector Prices

70 A number of prices obtain in the real sector. These include the consumer prices, wage rates and even interest rates as the price of investment goods. Several models have different specifications for the price equations, depending on the assumed structure of the domestic economy. Generally, consumer prices are determined mainly by real income levels (real GDP), wage rates are determined by a number of factors including employment levels, labour productivity, real income among other factors. We will look at the determination of interest rates from the monetary sector.

#### 2.4.2 Fiscal Block

71 Under the fiscal block, we have government revenues and government expenditures.

##### (a) Government Expenditure Functions

72 Government expenditure is divided into consumption expenditure and investment expenditure. Consumption expenditure includes both recurrent expenditures and transfer payments. Expenditures depend on nominal GDP and other relevant variables. Some regressions include a time trend to capture the fact that government expenditures will not in most cases follow its determinants, but may be pre-determined, sometimes influenced by political populism.

##### (b) Government Revenue Functions

73 Government revenue originates from tax and non-tax sources. Because of substantial dependence of tax on imports, taxes are divided into import duties, other trade taxes and internal taxes. Import duties and trade taxes are related to imports. Internal taxes are specified as a function of nominal GDP. Similarly, non-tax revenue is a function of nominal GDP and other relevant factors. In the government block therefore, the only endogenous item is the tax revenues. Other sources of finance including donor funds could also be included in the model, but these are exogenous. The main tax heads include excise taxes, value added taxes, corporate taxes, income taxes including pay as you earn (PAYE) and import duty.

74 Tax estimations generally come up with tax elasticities obtained as the coefficient  $\beta$  in the functional relationship:

$$T = \alpha B^\beta \varepsilon$$

where  $T$  is the tax revenue,  $B$  is the tax base (GDP or value of imports depending on which tax head is being estimated), while  $\beta$  is the coefficient of interest- tax elasticity and  $\varepsilon$  is a multiplicative error term. Taking logs, we have:

$$\ln T = \alpha + \beta \ln B + \ln \varepsilon$$

The tax elasticity  $\beta$  is defined as the responsiveness of tax revenues to movements in the base.

75 The proportional adjustment (PA) method is normally used to isolate the discretionary effects from the revenue effects. Other relevant factors may be incorporated in the models for each tax head. Also, the base may change, depending on which of the tax heads is being analysed. For instance, if we wanted to determine the elasticities of domestic excise taxes, the base would no longer be GDP but the value of domestic excise taxes while the regression would include, among other things, inflation rate and private consumption. The regression would therefore be of the form;

$$Excise\_domestic = \eta_0 + \eta_1 excise\_value\_domestic + \eta_2 GDP + \eta_3 inflation + \eta_4 private\_consumption + \varepsilon$$

### 2.4.3 Monetary Block

76 Monetary block consists of mainly the demand for money and the supply for money functions. Other equations in the monetary block may include interest rates and exchange rates, inter-bank rate, reserve equations, etc.

#### (a) The Money Demand Function

77 Most models start from the quantity theory of money, where money stock depends on income (real GDP) as a transaction motive and the price level. Some models extend the quantity theory to include nominal interest rate as the opportunity cost of holding money balances. In this case, the money demand equation would be of the form:

$$M^d = \alpha Y - \beta \Delta i + \gamma P$$

where  $M^d$  denotes demand for money  $Y$  is the real GDP and  $i$  the nominal interest rates,  $P$  the domestic price level.

#### (b) The Taylor Rule (Interest rate Function)

78 According to the Taylor rule, the nominal interest rate should respond to deviations of actual inflation rates from *target* inflation rates and of actual Gross Domestic Product (GDP) from *potential* GDP (Taylor, 1993). It is mainly useful for the Central Banks that target interest rates to influence monetary policy. It is specified as;

$$i_t = \pi_t + r_t^* + \alpha(\pi_t - \pi_t^*) + \beta(y_t - \bar{y}_t)$$

In this equation,  $i_t$  is the target short-term nominal interest rate,  $\pi_t$  is the rate of inflation,  $\pi_t^*$  is the desired rate of inflation,  $r_t^*$  is the equilibrium real interest rate,  $y_t$  is the logarithm of real GDP, and  $\bar{y}_t$  is the logarithm of potential output. Potential GDP can be determined by a linear trend or Hodrick Prescott Filter.

### (c) The McCallum Rule

79 **McCallum rule** specifies a target for the monetary base, which could be used by a central bank. It is an alternative to the well-known Taylor rule and is useful for those Central Banks that use monetary aggregate targeting instead of interest rate targeting. The rule gives a target for the monetary base in the next quarter (McCallum, 1994). The target is:

$$m_{t+1} = m_t - \Delta \bar{v}_{t-16} + 1.5(\Delta \pi^d + \Delta \bar{Y}) - 0.5 \Delta x_{t-1}$$

Where  $m_{t+1}$  is the natural logarithm of monetary base at time  $t$  (in quarters);  $\Delta \bar{v}_{t-16}$  is the average quarterly increase of the velocity of the monetary base over a four year period from  $t-16$  to  $t$ ;  $\Delta \pi^d$  is desired rate of inflation, i.e. the desired quarterly increase in the natural logarithm of the price level;  $\Delta \bar{Y}$  is the long-run average quarterly increase of the natural logarithm of the real GDP; and  $\Delta x_{t-1}$  is the quarterly increase of the natural logarithm of the nominal GDP from  $t-1$  to  $t$ .

80 It is important to note that the Taylor Rule and the McCallum Rules cannot be both used in the same model.

### (d) Exchange Rate Function

81 The exchange rate in a flexible exchange rate regime is an endogenous variable. It can appear in the external block or in the monetary block depending on the theoretical basis of the model.

82 Some modelers have followed the classical theorists and argued that the exchange rate is the purchasing power of the domestic currency in terms of the foreign currency. This relationship (which by no means is not the only way to model exchange rates) can be expressed as:

$$E = \frac{P^d}{P^f}$$

where  $E$  is the nominal exchange rate,  $P^d$  is the nominal domestic price level and  $P^f$  are the nominal foreign prices. Taking the logarithms of above equation gives:

$$e = p^d - p^f$$

where  $e = \log(E)$ ;  $p^d = \log(P^d)$ , and  $p^f = \log(P^f)$ . This relationship is what Keynes called the Purchasing Power Parity (PPP). Other authors have used the Dornbusch overshooting model to derive exchange rate equations from the domestic and the foreign money demand functions represented by:

$$m^d = p^d + y^d - r^d - \pi^d$$

$$m^f = p^f + y^f - r^f - \pi^f$$

where  $m^f$  represents foreign money demand and  $m^d$  represents domestic money demand,  $y^d$  represents the domestic income levels,  $r^f$  represents the foreign interest rates while  $r^d$  represents the domestic interest rates.  $\pi^f$  and  $\pi^d$  are the foreign and domestic inflation rates respectively. From equations above,  $p^d = p^f + e$ . Substituting the domestic money demand equation into this we have;

$$e = (m^d - y^d + r^d + \pi^d) - (m^f - y^f + r^f + \pi^f)$$

After collecting like terms, the equation above becomes:

$$e = (m^d - m^f) + (y^f - y^d) + (r^d - r^f) + (\pi^d - \pi^f)$$

83 The equation above gives the classical exchange rate determination model, which is basically a monetary side model because it does not consider the supply side of the economy. Based on the interpretations of Mussa and Rugoff (1983) that trade and current account balances are terms that allow for changes in the long run exchange rate, several authors have incorporated the supply side to this equation using current account balances and net capital flows.

### (e) Other Important Monetary Policy Transmission Mechanisms

84 It is important for the modeler to determine which of the numerous monetary policy transmission mechanisms work in their economy. Of particular importance to test will include (but not limited to) the following;

#### (i) The Interest Rate Channel

85 The interest rate channel assumes that monetary policy changes affect output through its effects on interest rates and investments. A tight policy reduces interest rates acting as an incentive to investments and thereby increasing output. The channel can be characterized schematically as;

$$M \downarrow \Rightarrow i \uparrow \Rightarrow I \downarrow \Rightarrow Y \downarrow$$

where  $M$  is real money supply,  $i$  is real interest rate,  $I$  is real investment and  $Y$  is real GDP.

86 This causal structure can be estimated empirically to determine whether the interest rate channel is effective in this economy.

### (.) The Credit Channel

87 This can work through several ways;

#### - The Balance-Sheet Channel

88 A tight monetary policy directly weakens borrowers' balance sheets by increasing interest expenses, reducing net cash flows and weakening the borrower's financial position. Rising interest rates are also associated with declining asset prices, which shrink the value of the borrower's collateral. Lower net worth means that lenders in effect have less collateral for their loans. A decline in net worth, thus leads to decreased lending to finance investment spending. The balance-sheet channel can be schematically represented as;

$M \downarrow \Rightarrow \text{Interest expenses} \uparrow \Rightarrow \text{Net Cash flows} \downarrow \Rightarrow \text{consumption and investments} \downarrow \Rightarrow \text{GDP} \downarrow$   
or

$M \downarrow \Rightarrow \text{Asset prices} \downarrow \Rightarrow \text{net worth} \downarrow \Rightarrow \text{loans} \downarrow \Rightarrow \text{investments} \downarrow \Rightarrow \text{GDP} \downarrow$

#### - The Bank-Lending Channel

89 Here, monetary policy changes affect commercial bank liquidity positions reducing the amount of money available to lend as loans. This reduces investments and GDP.

$M \downarrow \Rightarrow \text{bank liquidity} \downarrow \Rightarrow \text{loans} \downarrow \Rightarrow \text{investments} \downarrow \Rightarrow \text{GDP} \downarrow$

#### - Exchange Rate Channel

90 Under a flexible exchange rate, tight monetary policy lead to a rise in domestic interest rates vis-à-vis foreign interest rates. This appreciates the exchange rates. The appreciation causes a fall in net exports (NX) and hence, aggregate output (Y).

$M \downarrow \Rightarrow (i - i^*) \uparrow \Rightarrow e \downarrow \Rightarrow \text{NX} \downarrow \Rightarrow Y \downarrow$

#### - Other Asset Prices Channel

91 A tight monetary policy leaves less money in the hands of the public and therefore spending falls including spending in the stock market and in the real estate. This decreases the demand for equities and real estate and consequently lowering their investment in equities and real estates. This lowers aggregate income.

$M \downarrow \Rightarrow i \uparrow \Rightarrow \text{Equity demand and real estate demand} \downarrow \Rightarrow \text{investments} \downarrow \Rightarrow \text{output} \downarrow$

### 2.4.4 External Sector Block

92 The balance of payments block may consist of several equations. This will involve both current account side and capital account sides. The question is, how does the current account come to equilibrium and how does the capital account come to equilibrium? In the current account, import and export functions are the most used. These are the equations that will determine the volume of exports and imports.

#### (a) Export Functions

93 Normally, exports are determined by the interaction of foreign demand, which is a function of relative export price and the level income of the trading partners  $Y^f$ .

$$x = Y^f \alpha \left( \frac{e p_x}{p_d} \right)^\beta$$

where  $e$  is the real exchange rate. Other relevant factors that drive exports could be added to the export function and the function linearized by taking logs to get:

$$\ln x = \beta_1 \ln Y^f + \beta_2 \ln(p_x + e - p_d) + \phi Z$$

where  $Z$  is the vector of other variables

#### (b) Import Demand Function

94 Similarly, the demand for imports is normally given as a function of the real exchange rate, import prices, which are in most cases exogenously determined and domestic income, normally represented by GDP. This is given as:

$$\text{Im} = \left( \frac{P_{im} e}{WPI} \right)^{\delta_1} GDP^{\delta_2}$$

$$\ln \text{Im} = \delta_0 + \delta_1 \ln \left( \frac{P_{im} e}{WPI} \right) + \delta_2 \ln GDP$$

Where WPI is the world price index and  $P_{im}$  is the price of imports and  $e$  is the nominal exchange rates.

#### (c) External Sector Prices

95 Most models also have import prices and export prices endogenously determined. Import prices are determined by, among others, the domestic real income while export prices are determined by foreign income.

### 2.4.5 Closure Rules

96 Each of the blocks mentioned above have their own closure rules. These are discussed below:

#### (a) Closure Rule for the Real Sector Block

97 In the real sector, when using the expenditure approach, the closure rule is the usual national income accounting identity given as  $Y = C + I + G + (X - M)$

#### (b) Closure Rule for the Fiscal Sector Block

98 In the fiscal sector, the closure rule is the budget deficit equation where budget deficits (def) equals to government revenues (rev) less government expenditures (exp) i.e  $def = rev - exp$

#### (c) Closure Rule for the Monetary Sector Block

99 The closure rule in the monetary sector is money supply equation which states (on the asset side) that money supply equals to the net foreign assets plus net domestic assets. This must include the assets of both the central bank and the commercial banks, i.e  $M^s = NFA + NDA$

100 On the liabilities side, the closure rule is that money supply equals to the currency outside banks plus total deposits.

101 It is important to ensure that the liabilities and the asset side balance. The reconciliation is normally done by adding the value of "other assets" to the assets side.

#### (d) Closure Rule for the External Sector Block

102 The closure rule in the external sector is the balance of payments identity. It says that balance of payments equals the capital account plus current account minus the financial account plus or minus the balancing item (statistical discrepancy) i.e.

$$BOP = Current\ Account + Capital\ Account - Financial\ Account \pm Balancing\ Item$$

### 2.4.6 Consistency Framework in the Blocks

103 The modeler must always check that the data used is consistent with the way the data in the country is collected and used.

104 In the real sector, the consistency framework is that of the system of National Accounting (SNA) which the country is using at the time of modeling.



105 In the monetary sector, the data must be consistent with the monetary survey or the banking survey data as collected by the monetary authorities.

106 In the fiscal sector, the data must be consistent with government financial statistics (GFS).

107 In the external sector, data must be consistent with the BOP manual.

#### **2.4.7 Possible Challenges when Modeling the Blocks**

108 Each of the blocks present their own unique challenges when modeling them. Some of these are highlighted below:

##### **(a) Possible Challenges when Modeling the Real Sector Block**

109 There are several challenges that are likely to be encountered in the real sector block and which the modeler must take cognisance of from the beginning and think about how to solve them. These include:

- In most countries, GDP by expenditure is different from GDP by production. There is, therefore, need to reconcile the production approach and the expenditure approach. This is normally done by adding the value of statistical discrepancy in the GDP by expenditure series used.

- When modeling using low frequency data, one runs into the problem that data on several variables may not be available. This includes data on wage rates, capacity utilisation, etc.

- The available data may be under different SNAs. For instance, data up to the year 2000 may have been collected under SNA 1993 and thereafter under SNA 2000. Reconciling this needs the input of the data collection body. It is not just enough to reconcile by changing the base because some components of GDP may have changed with the change in SNA.

##### **(b) Possible Challenges when Modeling the Fiscal Sector Block**

In the fiscal sector block, the following could be the possible challenges and points to note when modeling the sector:

- Expenditure is normally difficult to model since it is not determined entirely by the fundamentals but largely by political decisions.

- When modeling for a federal government like Uganda, laws that govern revenue allocation to counties may complicate the modeling. The modeler must therefore take cognizance of such issues.

- The level of data aggregation of the revenue side may sometimes not allow the modeler to estimate specific tax heads. For instance, if data is collected only as import duties, it may not be possible to separate ordinary imports from oil imports. However, it may not make much sense to estimate very small components of tax revenues, for instance, if they do not form a significant portion of the country's tax revenues. In this case, some level of aggregation would be desirable.

**(c) Possible Challenges when Modeling the Monetary Sector Block**

- It is sometimes not possible to know all the transmission channels operating in the country and which ones could be strong. However, this could be known by testing them inside the model. The only thing that the modeler will need to be careful about is to ensure that they identify all the possible transmission channels that could exist in the country and test them in the model.

- The modeler needs to know the monetary policy framework that the country is operating in and the monetary targets.

- Other issues which the modeler must note are; the stability of the demand for money and the money multipliers, which normally are not known in advance and must be determined.

**(d) Possible Challenges when Modeling the External Sector Block**

- In most countries, the exports data from the BOP accounts and the SNA are not always the same. This is also true for import data. The data from the BOP accounts and the SNA must therefore be reconciled before use.

---

---

## CHAPTER 3

### ESTIMATION OF THE EQUATIONS IN THE BLOCKS

#### Objectives of this Chapter

At the end of this chapter the reader should be familiar with the following:

- In-putting data into E-views;
- Preliminary Analysis of the data;
- Diagnostics in time Series data – Testing for Stationarity
  - Unit roots testing
  - Correcting for unit roots
- Estimation of time series data:
  - Estimation when all data is stationary and there are no feedback effects (single equation models)
  - Estimation when data is non-stationary and co-integrated but no feedback effects (error correction models)
  - Estimation when data is non-stationary and co-integrated and there are feedback effects (vector error correction models)

#### 3.1 Pre-Estimation Issues

Before estimating the equations, several factors need to be taken into consideration:

##### 3.1.1 Identifying the variables to be included in the Model

110 Often, one is guided by the theory behind the equation of interest. If it is consumption, you could start from any of the many consumption theories that you think is applicable to your case. From the theoretical model specification, you will be able to identify the variables that you must have in the model.

111 However, there could be some other variables that determine your variable of interest and which are not explicitly mentioned by the theory you are applying. This is where literature review comes in. From the literature review, one will be able to identify the other variables that they could possibly include in the model.

112 The theoretical model is then transformed into an empirical model by including a disturbance term.

113 Once you have your variables (generally), you may first want to look at your data before you start the estimations just to ensure that it is clean.

## 3.2 Data

114 Before forecasting, several questions need to be asked as far as your data is concerned. Are the data “dirty”? For example, are there “ragged edges”? That is, do the starting and ending dates of relevant series differ? Are there missing observations? Are there aberrant observations, called outliers, perhaps due to measurement error? Is the data in a format that inhibits computer analysis?

### 3.2.1 Types of Macroeconomic Data

115 Macroeconomic data can be classified in terms of its frequency, either high or low. High frequency data consists of data that is available more frequently or more regularly including the daily, weekly, monthly and quarterly data. Low frequency data, on the other hand, is available less regularly and these include the semi-annual and the annual data.

### 3.2.2 Getting Data into E-views

116 There are several ways to get data into excel One method would be to copy directly from excel and paste in E-views. Assume you have a time series data in excel in the format below;

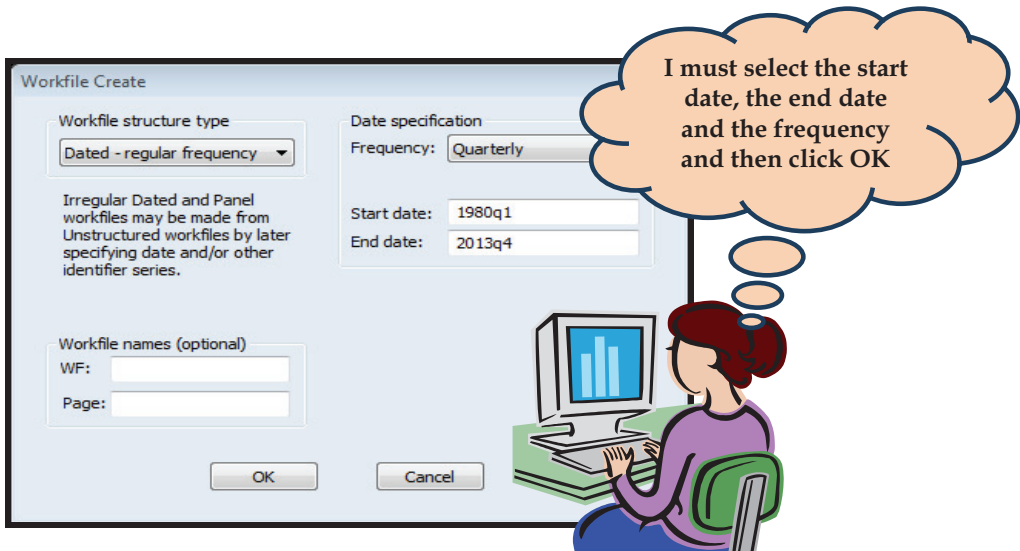
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	obs	CPISA	MRRAND	WAGES	RGDP	DATE															
2	2000Q1	64.19375	2.44218	16277634	4.393205																
3	2000Q2	63.90625	2.322493	12285293	697.869																
4	2000Q3	63.99625	2.237627	14229781	1291.767																
5	2000Q4	64.34375	2.197581	16099997	1787.366																
6	2001Q1	66.06875	2.172356	17905243	2186.086																
7	2001Q2	66.13125	2.191952	19644217	2489.979																
8	2001Q3	67.53125	2.246368	21316020	2656.678																
9	2001Q4	69.26875	2.336604	22920652	2809.111																
10	2002Q1	73.4375	2.501554	24809913	2994.208																
11	2002Q2	75.0125	2.643676	26138943	2785.026																
12	2002Q3	76.0875	2.803861	27261341	2547.667																
13	2002Q4	76.6625	2.902109	28176207	2175.543																
14	2003Q1	75.14375	3.368392	27889674	1042.729																
15	2003Q2	75.35625	3.50678	26787025	648.8914																
16	2003Q3	75.70625	3.597244	29874392	361.63																
17	2003Q4	76.19375	3.669783	31151776	177.4199																
18	2004Q1	77.11563	3.338617	33423439	200.8149																
19	2004Q2	77.75938	3.33962	34769150	160.5648																
20	2004Q3	78.42188	3.377012	35963171	187.3246																
21	2004Q4	79.10313	3.450791	37035603	255.8769																
22	2005Q1	79.65313	3.37353	37192633	384.4031																
23	2005Q2	80.37188	3.813457	38314993	547.5928																
24	2005Q3	81.30938	3.865144	39619068	756.9646																
25	2005Q4	82.36663	3.802658	41104860	1015.917																
26	2006Q1	83.36875	3.786258	43648691	1667.482																
27	2006Q2	84.73125	3.733037	45087987	1886.315																
28	2006Q3	85.28125	3.693389	46493770	2019.511																
29	2006Q4	88.01875	3.647316	47776940	2068.614																
30	2007Q1	90.14688	3.667813	47894163	1868.677																

117 From the excel sheet;

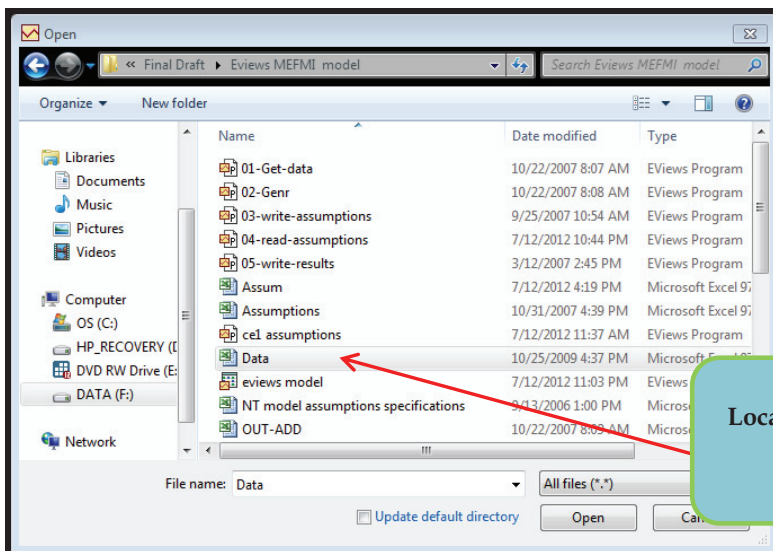
- Note location of upper leftmost observation. e.g., B2 in the case above.
- Note the variable names in the order in which they appear in the excel sheet.
- Note the first and last value of the first variable.
- Note the count (or dates for a time series data set - in this case, 1980q1-2004q4).
- Note where the excel file is saved, whether desktop or C drive, etc. and the file name.

118 To move the data into E-views, you first start E-VIEWS. There are two steps to moving in the data; first is to create a work-file and secondly, import the data.

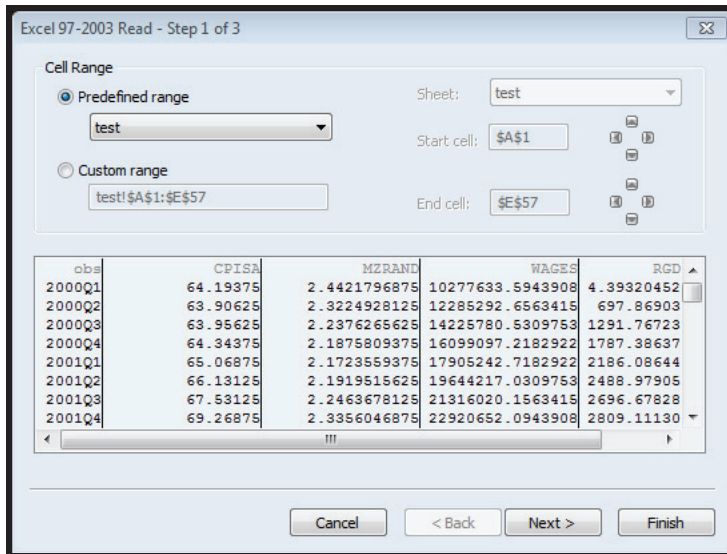
119 To create the work-file click: File→ New → Work-file, then fill in appropriately for your data:



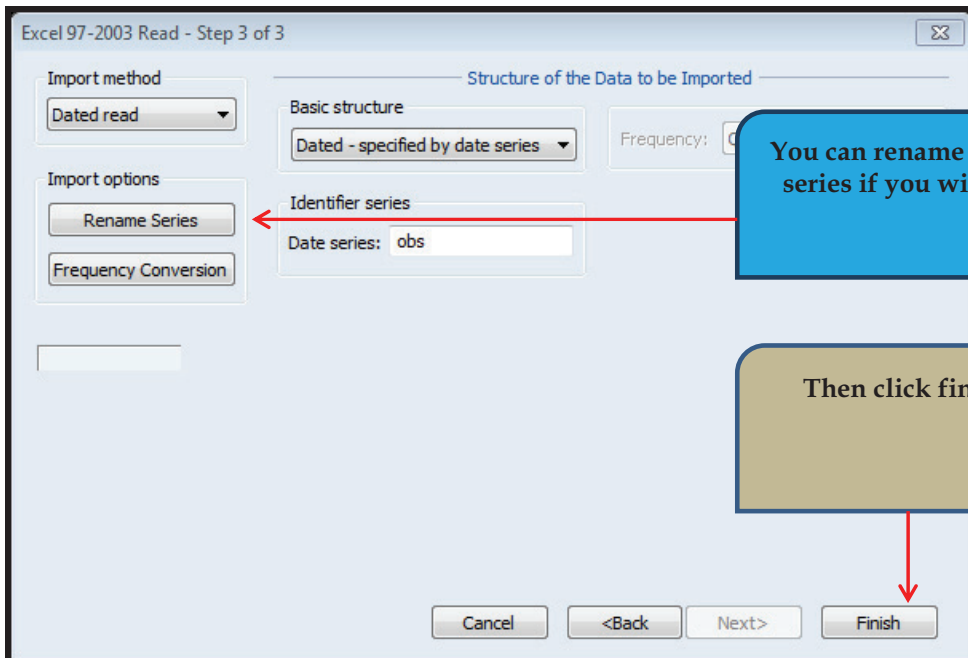
120 Click OK, then click Procs → Import → Import from file and a dialog box appears;



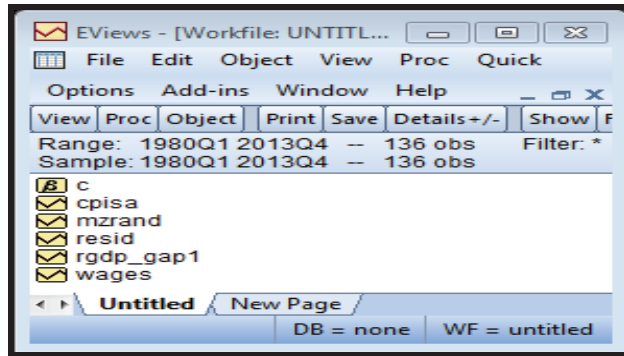
121 At this point you will be prompted to locate the excel sheet from where you saved it. The following dialog box will then appear;



122 Just click next. Then the next dialog box appears asking you to name the variable in the first column. This is the column of the year and we name it Year.



123 After this step, click OK and your data will be imported to E-views.



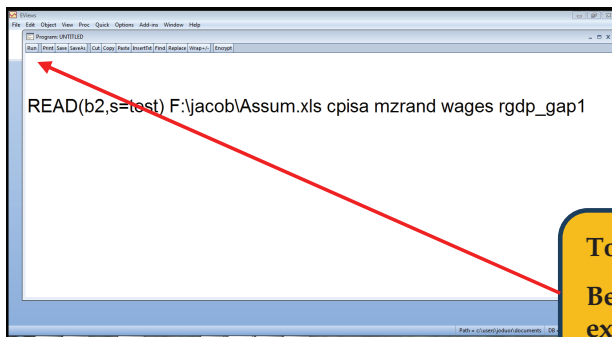
124 In addition to the imported series, there is also “c”, a vector for the constant, and “resid”, a place holder for residuals after regressions are run. It is a good idea to ensure that the data is imported correctly. You can ctrl-click each of the imported series and right click, then open as a group to show the values of all the series as:

125 Once you “show” your series, you can click VIEW from within the GROUP window

126 The other option would be to write a program in E-views to read the data directly from excel.

127 To write a program, go to File from your command window, then to New Object, Program, and a blank prompt window will appear. You can now type your commands here instructing E-views to read your data from excel. In our case we type;

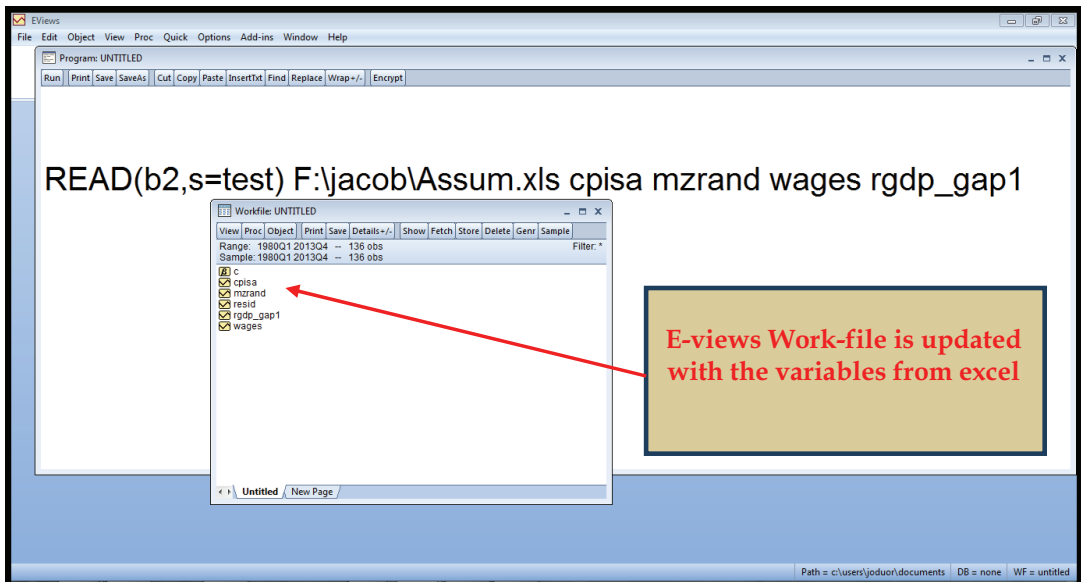
```
READ(b2,s=test) F:\jacob\Assum.xls cpisa mzrand wages rgdp_gap1
```



To read the data click “Run”  
Before running, ensure the  
excel sheet is closed

128 The above command instructs E-views to read the data from a folder named jacob in the F drive. The name of the file is Assum, saved in the 2003/2007 version of excel (that is why we use .xls, otherwise we would have used .xl if the file was saved in a higher version of excel). In addition, the command instructs E-views to start reading the data from cell b2, in the sheet named test. And it instructs E-views to read the values of cpisa mzrand wages and rgdp\_gap1

129 When you run E-views, work-file will be updated with the new variables as shown below:



### 3.2.3 Data Cleaning and Transformations

130 In most cases, data that is used in macro modeling is normally data collected over long periods of time, for different purposes that are not specifically modeling and also may have been collected using very rigid data collection techniques. Thus, much of that data is normally not controlled to suit the requirements of any form of modeling. As mentioned earlier, data must be reconciled to conform to a single SNA, which is currently being used in the country. In addition, it is wise to convert all the data to have one base.

131 Data transformation may also be necessary, for instance, by taking logs if the functional form of the model was non-linear and you want to transform the data into a linear form for use by OLS.

132 It is also important at this point to manually go through the data and plot it in graphs to identify any outliers which could be as a result of mistakes in data entry. These



outliers could be counter-checked to ensure that no mistakes were made during data entry.

133 The main graphical techniques used in exploratory data analysis are, a box plot, histogram, multi-vari chart, run chart, pareto chart, scatter plot, stem-and-leaf plot, odds ratio and the chi-square. Other descriptive statistics could also be used to see how the data looks like.

134 Before starting estimations, and after looking at the data, you need to remind yourself of the question that you intend to answer. This can also help you to further transform the data to be able to answer the questions. For instance, if you are interested in elasticities, you will need to ensure that your data is in log form.

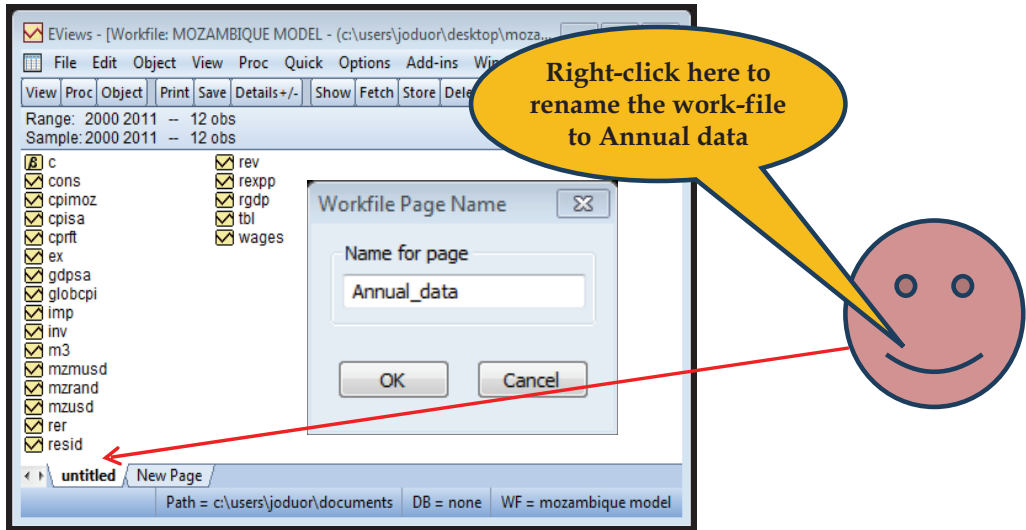
135 You must ensure that data is available for the kind of questions you are proposing to study. For example, the database may not allow you to estimate regional income functions if you are interested in regional models if you do not have regional income data like is the case in most developing countries. In the same way, you may not formulate questions to study the determinants of sectoral income like agriculture, fishing, manufacturing, and services, if you do not have sectoral GDP data like is the case in most countries in Africa and elsewhere.

136 Before estimations, and after ensuring that the data is clean, you will need to run an initial estimation and check whether that model you have is not mis-specified. The model could have missed some important explanatory variables, it could have included unnecessary variables or you could have adopted the wrong functional form.

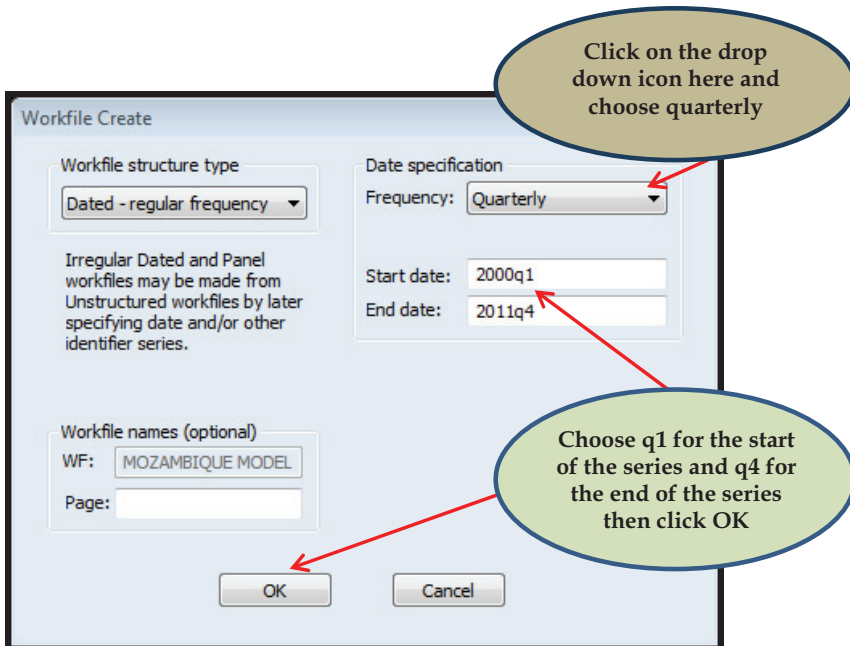
### **3.2.4 Data Frequency Conversions**

137 Frequency conversion allows you to convert a series of data from one work-file frequency to another. For example, you might have quarterly data that you want to use together with annual data, requiring you to convert it from quarterly to annual. In E-Views, this means that you copy a series from one page of a work-file to another page with a different frequency. Although E-Views allows you to copy from one work-file to another work-file, it often makes sense to perform the copy between pages of the same work-file.

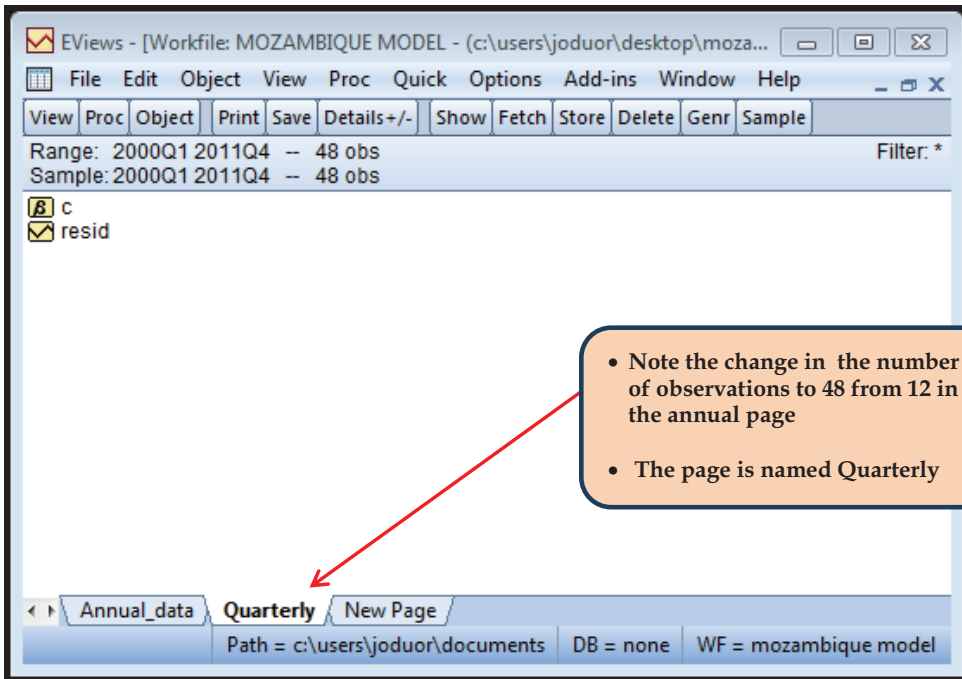
138 Assuming you already have your data in an annual E-views work-file as shown below, you can create a new page within the same work-file and rename both pages. In this example, we use the following work-file containing Mozambique annual data from 2000 to 2011.



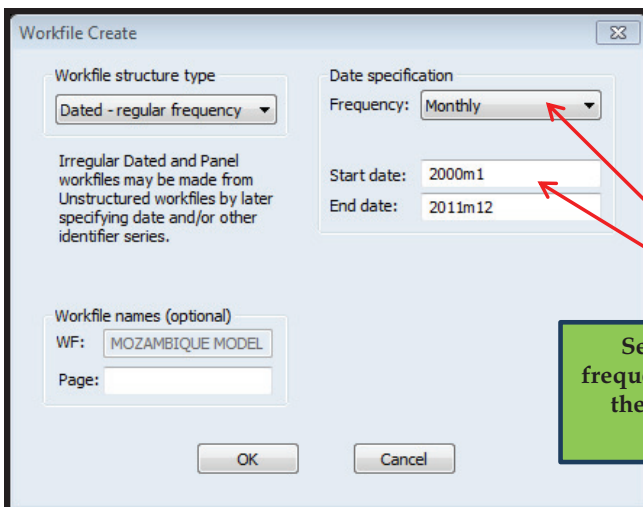
139 Then click on the “New page” and E-views will ask you to “Specify by Frequency/Range”. When you click on that, a new dialog box pops up, asking you to specify the frequency of the new page.



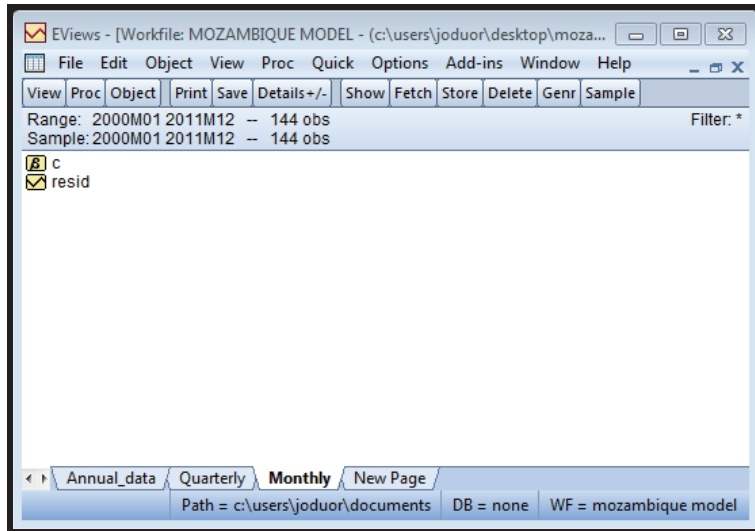
140 And you have your quarterly page, which you can rename in the same way that you did with the annual page.



141 You can click again on the “New Page” and create a monthly work-file

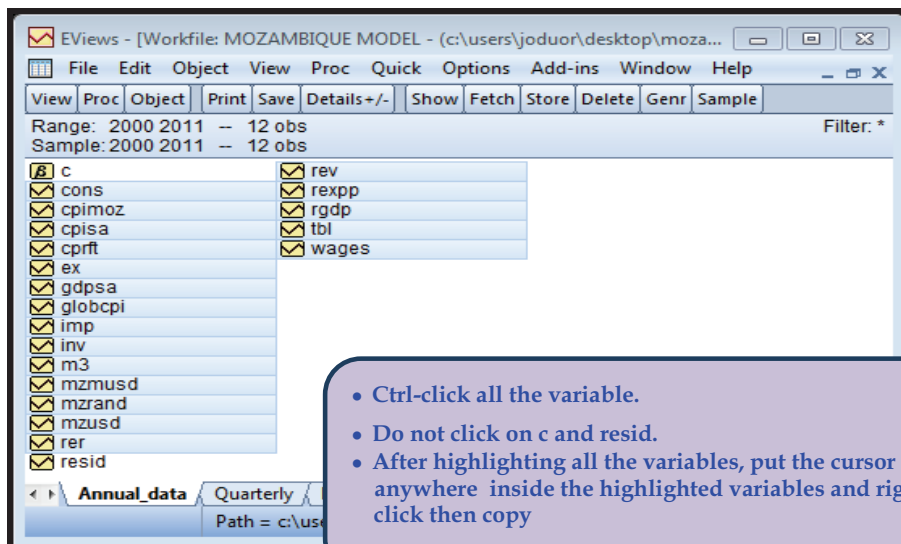


142 When you click, OK you get:



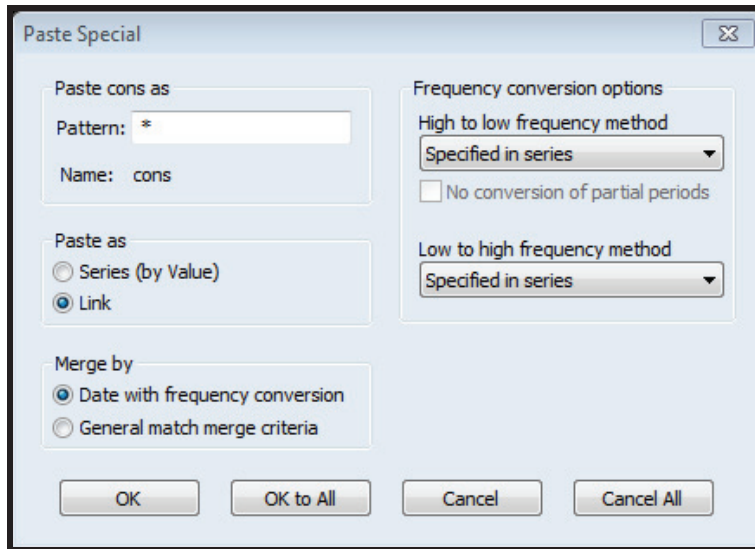
143 The monthly work-file has 144 observations up from 48 in quarterly form.

144 The work-file now has three pages - an Annual page, a Quarterly page and a Monthly page. The annual page contains annual data. We only obtained the annual data and want to convert it to quarterly and monthly. To do this, we select the Annual page, Ctrl-click the all the variables and select copy:



- Ctrl-click all the variable.
- Do not click on c and resid.
- After highlighting all the variables, put the cursor anywhere inside the highlighted variables and right click then copy

145 We then select the quarterly page, right click in some of the white space, and select Paste Special. This brings up the paste special dialogue:

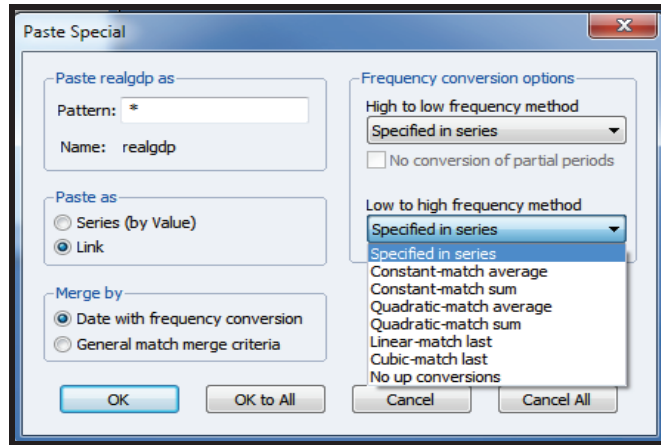


146 The "Pattern" box allows you to specify the name of the variables that will be used in the destination page. If you want to keep the same name as the original series, just leave that box as it is (i.e. with a "\*" in it).

147 Following that, you have a choice whether to paste as a "Series (by Value)" or as a "Link". If you choose to paste as a Series, then the copy will be a one-off event. The Quarterly page will have a snapshot of the annual data at the time the copy was done. If you choose to paste as a Link, the data in the Quarterly page will permanently be linked to the data in the Annual page - that is whenever the data in the Annual page changes, the data in the Quarterly page will also change.

148 The right hand side of the dialog allows you to specify the frequency conversion options. There are two sets of options - "High to low" and "Low to high". If your source page has a higher date frequency than your destination page, you should set the "high to low" options. If your source page has a lower date frequency than your destination page, you should set the "low to high" options.

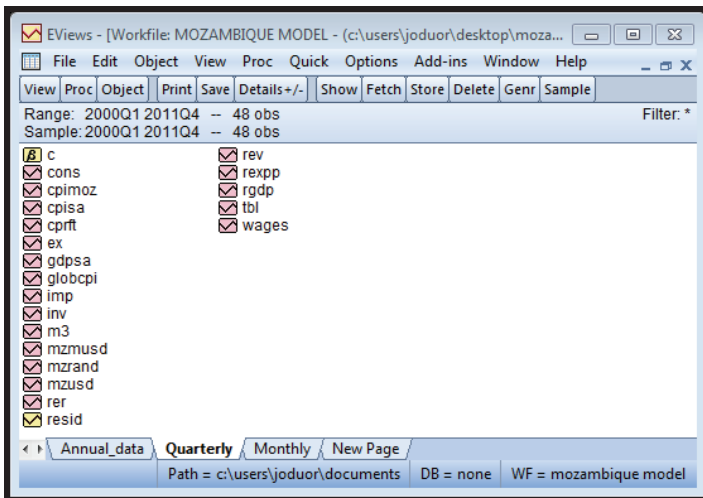
149 Since we are moving from annual to quarterly, we are moving from a lower frequency to a higher one, and so should set the "Low to high" options:



- **Constant-match average:** This simply repeats the low value for each of the high observations. i.e. each quarter in 2000 will have the same value as the annual value for 2000.
- **Constant-match sum:** This simply repeats the low value divided by the number of observations in the high page. i.e. each quarter in 2000 will have the annual value for 2000 divided by 4 (since there are 4 quarters in a year).
- **Quadratic-match average:** This performs a proprietary local quadratic interpolation of the low frequency data to fill in the high observations.
- **Quadratic-match sum:** This performs the same proprietary local quadratic interpolation, but this time the low frequency data is divided by the number of observations first.
- **Linear-match last:** This inserts the low observation value into the last period of the high frequency data, then performs linear interpolation on the missing values. i.e. 2000Q4 will be given the annual 2000 value, 2001Q4 will be given the annual 2001 value, then linear interpolation will fill in 2001Q1, 2001Q2, 2001Q3.
- **Cubic-match last:** This is the same as linear-match last, but rather than linear interpolation, a cubic spline is used for the interpolation.

150 We can choose “quadratic-Match average’ in our case for illustration purposes. But ideally, one should copy each variable at a time and chose the appropriate interpolation method for that variable. For instance, you can choose Linear-match last for GDP which means that you are assuming that real GDP grows at a linear rate throughout the year.

151 We then paste a link. Once we hit OK, you can see the series has been copied to the Quarterly page:



152 And the frequency conversion is complete. If you open the variables in the quarterly page as a group you can see how the frequency conversion worked:

The screenshot shows the EViews software interface with a group of variables selected. The main window title is "EViews - [Group: UNTITLED - Workfile: MOZAMBIQUE MODEL: Quarterly]". The menu bar includes File, Edit, Object, View, Proc, Quick, Options, Add-ins, Window, and Help. The toolbar contains buttons for View, Proc, Object, Print, Save, Details+/-, Show, Fetch, Store, Delete, Genr, and Sample. Below the toolbar, the workfile information is displayed: Range: 2000Q1 2011Q4 -- 48 obs, Sample: 2000Q1 2011Q4 -- 48 obs, Filter: \*. On the left, a list of variables is shown with checkboxes, including c, cons, cpimoz, cpisa, cprft, ex, gdpsa, globpci, imp, inv, m3, mzmusd, mzrand, mzusd, rer, and resid. On the right, a list of variables is shown with checkboxes, including rev, rexpp, rgdp, tbi, and wages. At the bottom, the page selection is set to "Quarterly", and the path is shown as "Path = c:\users\joduro\documents" and "DB = none" and "WF = mozambique model".

153 You can follow the same procedure to convert into monthly data.

### 3.2.5 Data Gaps, Inconsistencies and Remedies

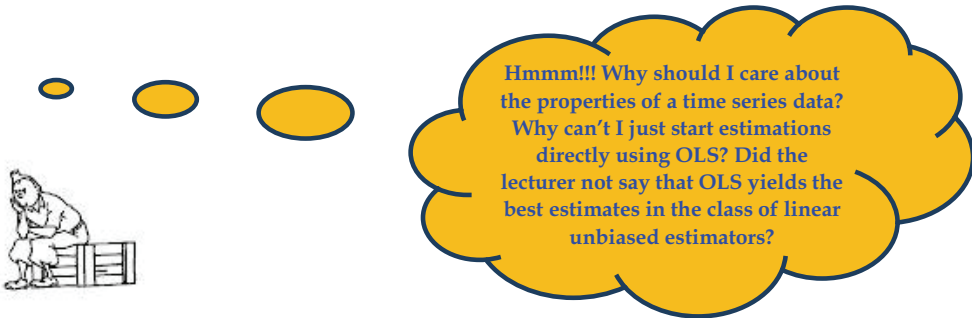
154 Sometimes, macroeconomic data are produced by non-analytical staff. This causes data, be subject to both methodological and compilation discrepancies or errors. It is not uncommon to find breaks and or unexplained outlier figures in macroeconomic time series. It is important, therefore, by the modeler to make adjustments to data so as to

minimize compilation errors. Sudden bumps in data series should trigger the modeler to follow-up with data compilers to make corrections and or find explanation. Sometimes explanations made by compilers may lead to reclassification of some of the items included in the “misbehaving” time series.

155 In most cases, some of the series needed to develop a model may not be available in the routinely supplied macroeconomic statistics. For instance, the modeler might want to have government revenue disaggregated into tax revenue, sale of government goods and services and government property income, while the data that is routinely supplied is aggregated into tax and non-tax revenue only. Under this situation, the modeler would need to seek the missing data details from the compilers and also push for inclusion of that level of detail in future submission of SGO statistics. As long as the modeling work is for facilitation of macroeconomic policy formulation it is justifiable for the modeler to ask for the level of detail recommended by the macroeconomic data compilation manuals.

### 3.3 Diagnostics of Time Series Data – Testing for Stationarity

156 Remember we are using time series data (data collected over a period of time) and we may get into trouble if we just start estimations without looking at some of the properties of the time series data.



157 Conventional asymptotic theory for ordinary least squares (e.g. standard prove of consistency and asymptotic normality of OLS estimators) assume stationarity of the explanatory variables.

158 However, not all economic time series, are stationary. In many cases, economic theory posit causal relationships between economic series that are non-stationary e.g. aggregate consumption, national income, exchange rates etc. If variables are non-stationary we may have nonsensical regressions named so by Yule (1926) or spurious regression in the terminology of Granger and Newbold (1974).

159 Yule (1926) shows that spurious correlation may persist in large samples despite the absence of any connection between the underlying series. The estimation of




the spurious regressions by OLS is shown to lead to results that cannot be interpreted within conventional testing procedure due to; first; conventional calculated t-statistics on the coefficients do not have the standard t-distribution; secondly, the null of no relationship between the series will tend to be rejected more frequently in large samples.

160 If the two series are stationary, we would fail to reject the null of no relationship. Even the joint significance test will face the same problems. The F-statistic diverges, leading to rejections growing with the sample size  $T$ , despite the lack of any relationship. However, despite the problems with t-test and F-test, some information can be gleaned from residual autocorrelation tests about the misspecification of the model.

161 A quick way of detecting spurious regressions is the high  $R^2$ , significant t-ratios but very big standard errors.

162 It is therefore important that estimations are only done with stationary data. If a series is non-stationary, then the non-stationarity must be detected and the variable made stationary before estimations are conducted with it. But what exactly is stationarity?



.....but what exactly is stationarity?..  
and why is it that only time series  
data is mostly non-stationary? Why  
not cross sectional data?

### 3.3.1 Meaning of Unit Roots and Stationarity

163 A stochastic process is said to be weakly, covariance or second-order stationary if its *mean* and *variance* are constant over time and the value of covariance between two periods depend only on the distance or gap between the two time periods and not the actual time at which the covariance is computed.

### 3.3.2 Requirements for Weak Stationarity

164 If you think about a stochastic process  $X_t$ , the series will be stationary if;

- i. The mean is constant:  $E(X_t) = \mu$
- ii. The variance is constant:  $Var(X_t) = E(X_t - \mu)^2 = \sigma^2$
- iii. The covariance of the series in different periods is zero. This requirement can also be stated differently; that the covariance depends on the difference between the lags and not when it is observed

$$\gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)] = Cov(x_t, x_{t-k} = 0)$$

### 3.3.3 Why are Time Series mostly non-stationarity?

165 The simplest, purely statistical time series model is the first order autoregression, or AR(1) process given as;

$$X_t = \rho X_{t-1} + \varepsilon_t$$

166 Because most time series depend on their lags, the requirement that  $Cov(x_t, x_{t-k} = 0)$  is normally mostly violated (not always). If that requirement is violated, then the series will be non-stationary.

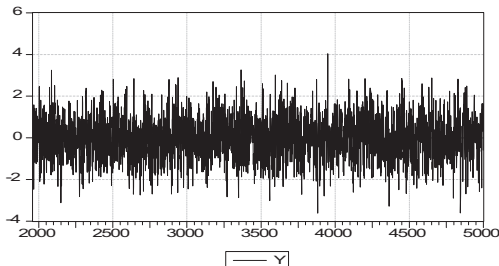
167 To determine therefore whether the series is non-stationary or not, we must test.

### 3.3.4 Types of Time Series Processes

As mentioned above, not all time series will be non-stationary.

168 A stationary process is given by:

Figure 3.1: A Stationary Process



169 Notice that the above series meets the requirements of;

- Constant mean
- Constant variance
- The covariance depends on the difference between the lags and not when it is observed e.g.  $Cov(y_t, y_{t-1}) = Cov(y_{t-6}, y_{t-7})$  and  $Cov(y_{t-2}, y_{t-4}) = Cov(y_t, y_{t-3})$  i.e. covariance between two variables with the same lag difference are the same irrespective of when it is observed

170 There are three possibilities in which nature could generate time series variables

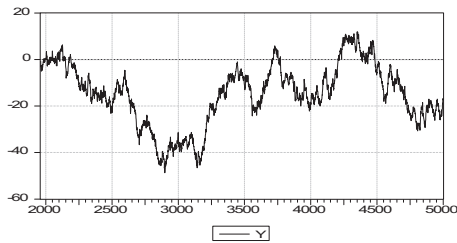
- i. Pure random walk
- ii. Random walk with drift
- iii. Random walk with drift and time trend

### 3.3.5 Pure Random Walk

171 For a pure random walk we have the observed values today depend on the values in the previous period plus a white noise process. A white noise process is a process with zero mean, constant variance and zero covariance. A white noise process is therefore a special case of a stationary process.

$$y_t = y_{t-1} + \varepsilon_t$$

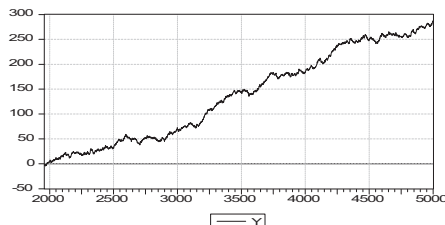
Figure 3.2: A Pure Random Walk Process



172 Random walk with drift but initial value is still zero.

$$y_t = \alpha + y_{t-1} + \varepsilon_t$$

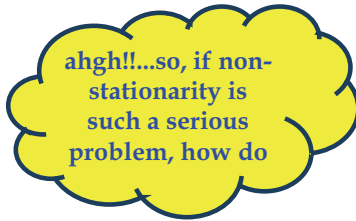
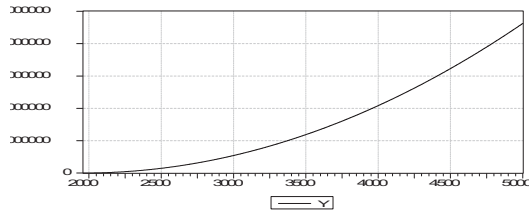
Figure 3.3: Random Walk with a Drift



(.) Random walk with drift and time trend.

$$y_t = \alpha + y_{t-1} + \beta t + \varepsilon_t$$

Figure 3.4: Random Walk with Drift and Time Trend

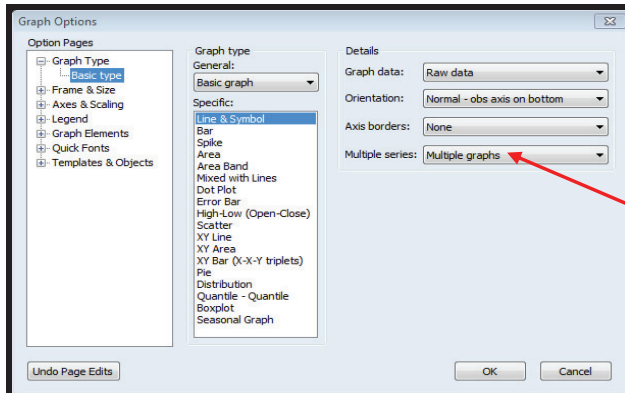


### 3.4 Testing for Unit Roots

#### (a) Visual inspection

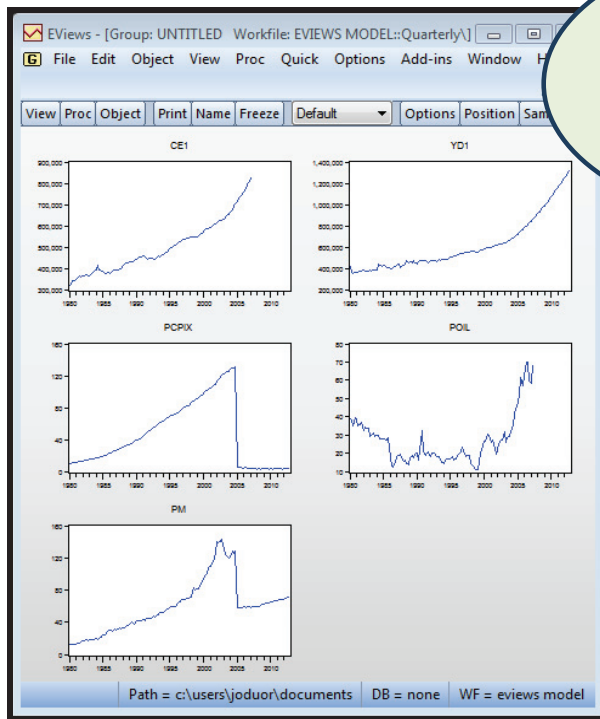
173 The first task to do when testing for stationarity is to plot the variables in graphs and see whether they could be stationary or not.

174 Let us plot a few variables from our database and visualize the graphs. To do this, you can Ctrl-click all the variable you want to see their graphs and open as a group. Then View, Graph, Multiple graphs. Let us see the graphs of **ce1 yd1 pcpix poil pm**



Choose multiple graphs

175 We get;



From these graphs, which ones seem to be stationary and which ones are non-stationary? What should you look at? Are the means constant? What about the variances?



### (b) Formal unit root tests

176 Several formal tests have been proposed to tests for unit roots. The simplest of them is the Dickey Fuller Test.

177 Given a time series process generated as  $y_t = \alpha y_{t-1} + \varepsilon_t$  with  $\varepsilon_t$  as a white noise process,  $y_t$  is non-stationary if the root of this simple difference equation is unit (one). The root of this equation is  $\alpha$ . When  $\alpha = 1$ , then the process is said to have a unit root (a root that is equal to one) and the process is non-stationary. It is stationary if the root is less than one in absolute terms. The process of formally testing whether  $y_t$  is non-stationary or not therefore reduces to testing whether  $\alpha = 1$

178 If  $\alpha = 1$  then we say that the process is a unit root process. It is common to hear people say that they are testing for unit roots. This is the same as saying that you are testing whether the series is non-stationary or whether  $\alpha = 1$

179 Dickey and Fuller developed a formal procedure to test whether  $(\alpha - 1) = 0$  using the following transformation.

Given  $y_t = \alpha y_{t-1} + \varepsilon_t$ , subtracting  $y_{t-1}$  from both sides, you get.

$$y_t - y_{t-1} = \alpha y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = (\alpha - 1)y_{t-1} + \varepsilon_t$$

Letting  $\rho = \alpha - 1$  implies that when  $\alpha = 1$ ,  $\rho = 0$

180 This implies that presence of unit root is equivalent to  $\rho = 0$

$$\Delta y_t = \rho y_{t-1} + \varepsilon_t$$

181 This is the Dickey Fuller test equation with the null hypothesis that  $H_0 : \rho = 0$ , against the alternative that  $H_a : \rho < 0$  or  $\rho_a < 1$

182. The DF Unit Root Test is based on three regression forms, ie without constant and trend, with constant, and then with constant and trend:

1.  $\Delta y_t = \rho y_{t-1} + \mu_t$  Without Constant and Trend
2.  $\Delta y_t = \alpha + \rho y_{t-1} + \mu_t$  With Constant
3.  $\Delta y_t = \alpha + \rho y_{t-1} + \beta T + \mu_t$  With Constant and Trend

### (c) Other Unit Root Tests

183 The other common unit root tests include:

- The Augmented-Dickey-Fuller Test
- Phillip-Perron Tests
- Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
- The Lagrange Multiplier (Lm) KPSS Test for Stationarity
- Ng and Perron Tests (NP)
- Dickey Fuller GLS Test (DFGLS)
- Elliot, Rothenberg, and Stock Point-Optimal ERS Test

## 3.5 Testing for Unit Roots in E-views

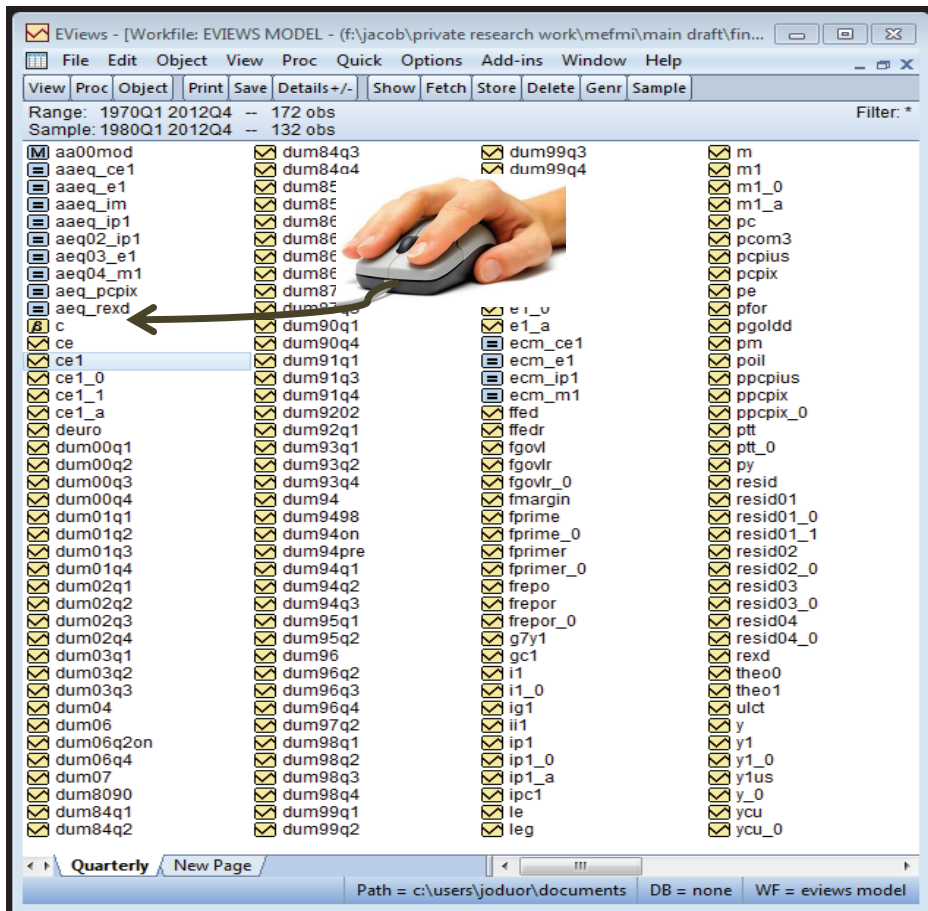
184 Assume you have all your data already in E-views.

185 From the work-file, pick any one of the series that you want to test for stationarity on.

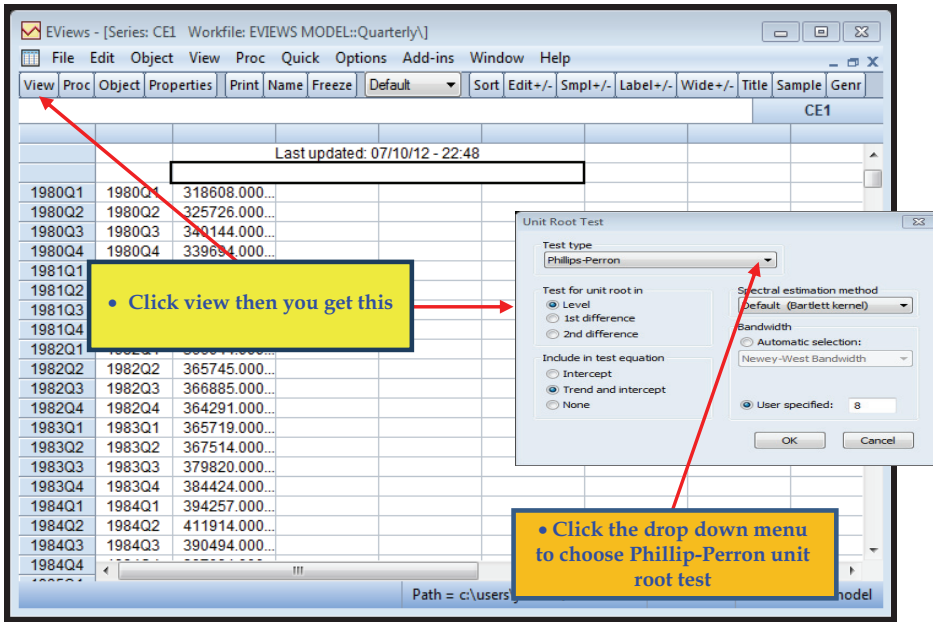
186 Let us use the consumption (ce1) from our work-file given in the CD to

illustrate the points.

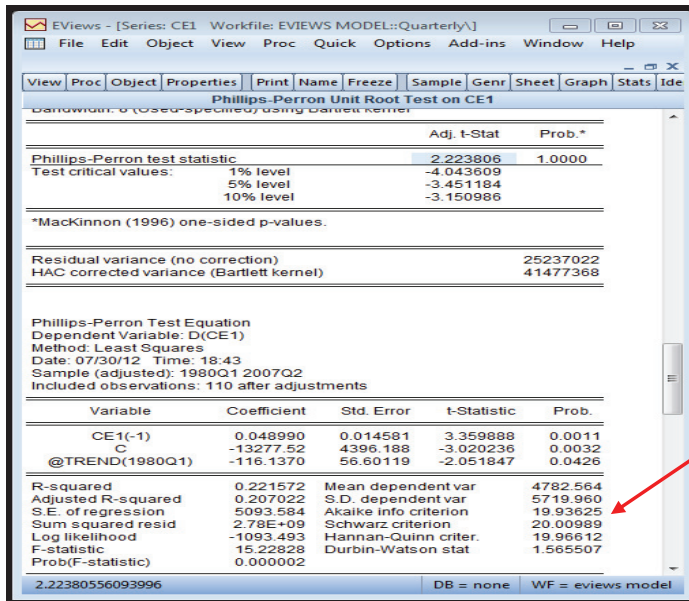
187 To access the data, open the CD, open the folder called E-views MEFMI Model and then open the work-file called E-views Model. Double click on the ce1 variable to open.



188 After double clicking ce1, you get the sheet containing the values of ce1. When you are there, click on View, Unit root Test and the next window appears.



189 Let us use the "Phillip-Perron", at "Level". The results from this unit root test is given as:



190 Before we interpret the results, we must choose the optimal lag length. We choose the lags that give the minimum information criteria (Schwartz - SIC, Akaike - AIC



and Hannan-Quinn - HIC). Remember, we started with eight lags. We can now try seven lags and see whether the SIC and AIC reduces or not. Note the values of SIC and AIC in the test above. They are 19.93, 20.00 and 19.96 for SIC, AIC and HIC

191 With seven lags we have;

Choose seven lags

192 And the results with seven lags;

Phillips-Perron Unit Root Test on CE1  
Bandwidth: 7 (User-specified) using Bartlett kernel

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	2.217123	1.0000
Test critical values:		
1% level	-4.043609	
5% level	-3.451184	
10% level	-3.150986	

\*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	25237022
HAC corrected variance (Bartlett kernel)	41608001

Phillips-Perron Test Equation  
Dependent Variable: D(CE1)  
Method: Least Squares  
Date: 07/30/12 Time: 18:46  
Sample (adjusted): 1980Q1 2007Q2  
Included observations: 110 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CE1(-1)	0.048990	0.014581	-3.359888	0.0011
C	-13277.52	4396.188	-3.020236	0.0032
@TREND(1980Q1)	-116.1370	56.60119	-2.051847	0.0426

R-squared	0.221572	Mean dependent var	4782.564
Adjusted R-squared	0.207022	S.D. dependent var	5719.960
S.E. of regression	5093.584	Akaike info criterion	19.93625
Sum squared resid	2.78E+09	Schwarz criterion	20.00989
Log likelihood	-1093.493	Hannan-Quinn criter.	19.96612
F-statistic	15.22628	Durbin-Watson stat	1.565507
Prob(F-statistic)	0.000002		

Path = c:\users\joduo\documents DB = none WF = eview

AIC = 19.93  
SIC = 20.00  
HIC = 19.966

SIC, AIC and HIC are not very different at seven lags than they are at eight lags.

We continue with this process until we get the lag that gives us the minimum SIC, AIC and HIC. If they contradict, use the lag that gives the minimum SIC.

In our case, the information criteria are the same for Phillips-Perrn at all lags. So, we can go ahead and interpret these results.

193 From the results we see that that the test statistic is 2.21 and the critical value at 5 percent significance level is -3.45. Since the test statistic of 2.21 is less than the critical value of -3.45 in absolute terms, then we conclude that consumption for this hypothetical country is non-stationary at levels (before differencing).

194 You need to conduct the tests using the other unit test approaches and see whether the conclusion changes.

### 3.6 Making Non - Stationary Variables to be Stationary

195 A variable can either be trend stationary or difference stationary. A trend stationary variable is a non-stationary variable which exhibits a deterministic trend over time and becomes stationary after removing the trend. A difference stationary variable on the other hand contains stochastic components and may or may not contain a deterministic trend and it becomes stationary after differencing (Oduor, 2008).

196 Making non-stationary variables to be stationary depends on whether the variable is trend stationary or difference stationary.

197 A trend stationary variable is made stationary by detrending, while a difference stationary variable is made stationary by differencing.

### 3.7 Detrending a Trend Stationary Variable

198 Suppose a series  $y_t$  always changes by a constant amount from one period to the next plus some stationary component, the deviations of these series are only temporary departures from the trend. As such, the long term forecast will converge to the trend. This type of model is called trend stationary. If we had a model without any stochastic component given as:

$$y_t = \alpha + y_{t-1}$$

The change in  $y_t$  is

$$\Delta y_t = y_t - y_{t-1} = \alpha$$

199 The solution to this difference equation is given as:

$$y_t = y_{t-1} + \alpha^t$$

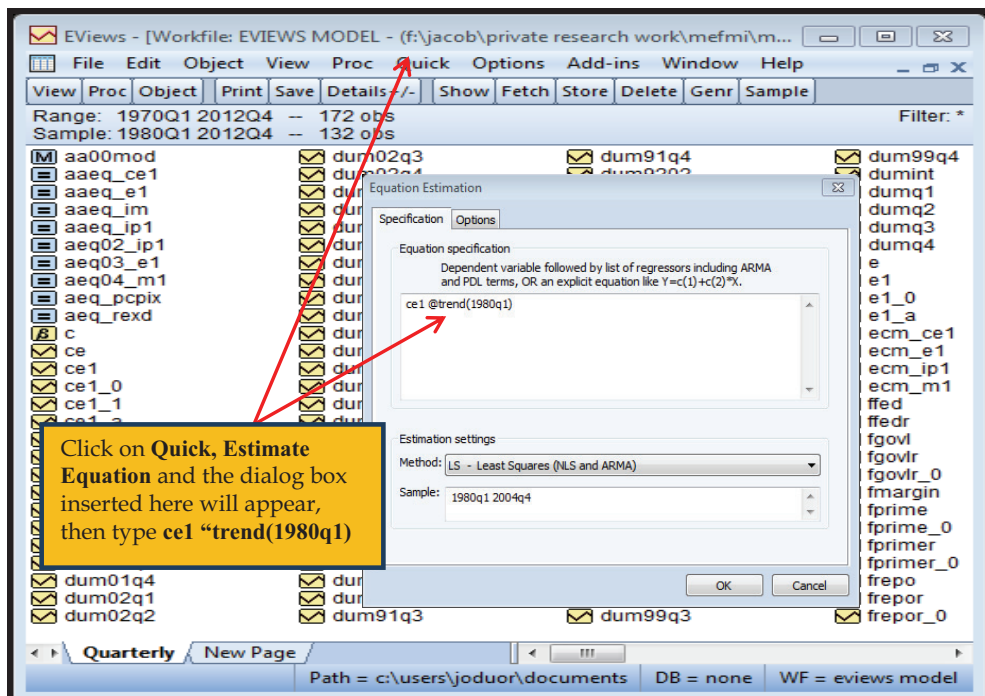
where  $y_{t-1}$  is the initial condition at period zero.

200 It is clear from above that the solution to this difference equation is just a linear time trend and  $\alpha$  is the slope. It will be sufficient to make the series  $y_t$  stationary by removing the trend. This is called detrending. This variable is called a trend stationary variable.

201 To determine whether a non-stationary series which exhibits a trend is trend stationary or difference stationary, we need first to attempt to remove the trend and see whether the variable becomes stationary. If it does become stationary after detrending, it is trend stationary, if it does not, then it is difference stationary. Detrending a variable involves regression of the variable against time.

202 We illustrate the detrending procedure using the example of consumption in our EViews Model work-file given in the CD. Remember, when testing for unit roots, we found out that consumption is non-stationary. What we do not know is whether it is trend stationary or difference stationary.

203 We first regress consumption against a time trend using the following command. Go to **Quick, Estimate Equation** and a new dialog box will appear. Inside the dialog box, type **ce1 “trend(1980q1)**



204 We illustrate, in more detail, how to conduct estimations in E-views in the next section. The results are given as;

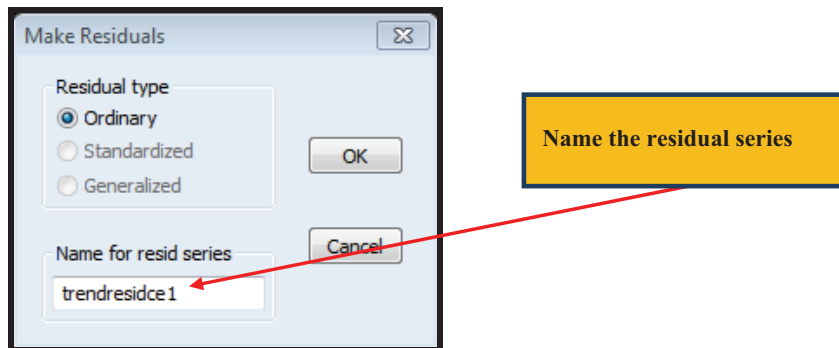
Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND(1980Q1)	8006.916	279.0949	28.68887	0.0000

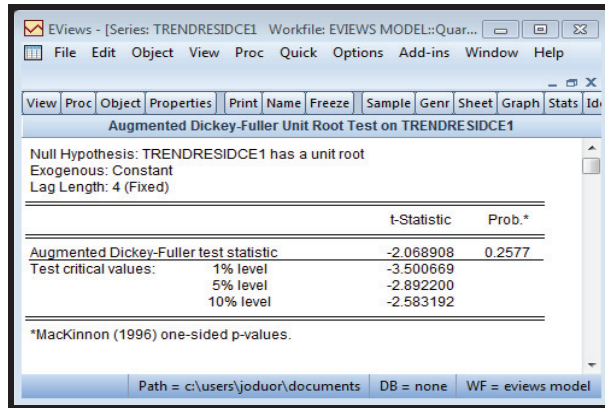
R-squared	-1.696707	Mean dependent var	475855.8
Adjusted R-squared	-1.696707	S.D. dependent var	97387.55
S.E. of regression	159926.5	Akaike info criterion	26.81277
Sum squared resid	2.53E+12	Schwarz criterion	26.83882
Log likelihood	-1339.638	Hannan-Quinn criter.	26.82331
Durbin-Watson stat	0.001743		

Path = c:\users\joduor\documents DB = none WF = eviews model

205 From the above equation, we generate the residuals and test whether the residuals are stationary or not. To generate the residuals, we go to Proc, Make Residual Series. A dialog window will then appear looking like this;



206 Test for unit root on the residuals in the same way as illustrated earlier in the unit root testing section. The results of the unit roots tests on the residuals are given as;

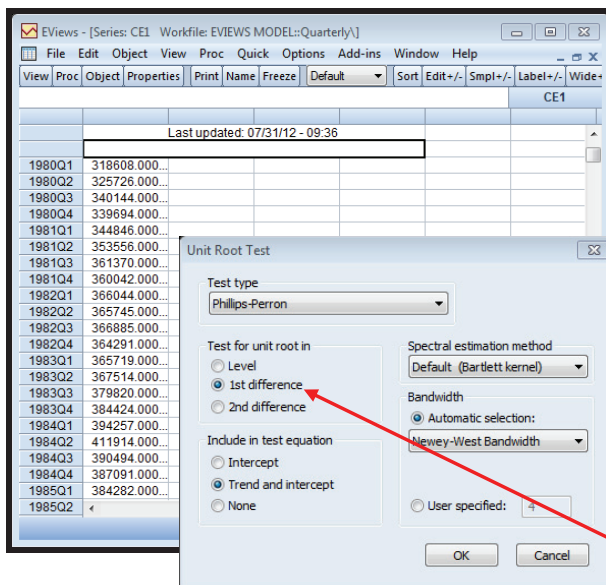


207 From the results above, we cannot reject the null hypothesis that the residuals are non-stationary. This implies that consumption is not trend stationary.

### 3.8 Correcting for non-Stationarity by Differencing

208 We have found out that consumption is not trend stationary, meaning that it is difference stationary. We therefore need to difference to make it stationary.

209 To do this in E-views, we go back to our unit root testing window by double clicking on consumption from the work-file and testing for unit roots now at first difference.



Notice that we are now using the first difference and not

210 The final results after the process of selecting the optimal lag length is given as;

211 The results are given in the chart below. The Phillip-Perron test Statistic is -8.011822 against a critical value of -3.456319 at the 5 percent level. The associated p-value is 0.000.



212 The results show that consumption is stationary at first difference.

### 3.9 Order of Integration

213 In the example above, consumption becomes stationary after differencing once. It is therefore said to be integrated of order one, denoted by  $I(1)$ .

214 Some variables will become stationary after differencing two and they are said to be integrated of order two, denoted by  $I(2)$ .

215 A series with no deterministic component, which has a stationary invertible ARMA representation after differencing  $d$  times, is said to be integrated of order  $d$  denoted  $x_t \sim I(d)$  (Engle and Granger 1987: 252). A series is therefore integrated of order one denoted as  $x_t \sim I(1)$  if it is non-stationary at levels but becomes stationary after differencing once. It is integrated of order two denoted as  $x_t \sim I(2)$  if it becomes stationary after differencing twice and so on. A series that is stationary at levels (that will not require any differencing) is said to be integrated of order zero i.e.  $x_t \sim I(0)$

### 3.10 Estimation of Time Series Data

216 There are several types of econometric estimation techniques that can be used when dealing with time series data.

217 The choice first depends on the kind of data that you have. If all the variables are stationary and has no feedback relationships, then you can estimate a single equation with all the variables in their levels.

218 If the variables are non-stationary and cointegrated but has no feedback relations, then you need to estimate an error correction model.

219 If the variables are cointegrated and have feedback relations (contemporaneous relations), then we estimate a vector error correction model.

220 In this section, we will illustrate the estimation procedure in these three situations; when all variables are stationary and have no feedback effects (single equation approach); when the variables are non-stationary and cointegrated (error correction models - ECM) and when the variables are cointegrated and have feedback effects (vector error correction models - VECM)

#### 3.10.1 Estimation with Stationary Variables - Single Equation Approach

221 For illustration purposes, assume that we have tested for unit roots and found that the variables in the consumption equation are all stationary at levels and there are no feedback effects among the variables, then we will have to use the single equation approach.

222 We will now illustrate how to conduct a single-equation estimation assuming all the variables in the equation are all stationary at levels. We will also illustrate some of the diagnostic tests that one must be careful to do when conducting the tests.

223 We also take the opportunity to take readers through the process of conducting a simple single equation estimation in E-views.

224 You must remember that you must first test for unit roots in the variables and determine that they are all stationary at levels before you can decide to estimate a single equation model in levels, other than an error correction model for instance.

225 The single equation models are normally used when the data is all stationary and therefore the short run coefficients are the same as the long-run coefficients. In addition, it is used when there are no feedback effects among the variables in the model. That is, if  $y$  is the dependent variable and,  $x$  is the independent variable, then it is not the case that  $x$  also depends on  $y$ . One of the estimation techniques normally adopted for estimation of single equation models is the ordinary least squares, which is used when all the Gauss-Markov assumptions are met. Variables that are non-stationary and have no long-run relationship (not cointegrated) can also be represented in a single equation framework in their differences. This is will only be a short-run model.

226 First step is to choose your variables. Let us use the variables in a hypothetical consumption equation (for illustration purposes) from the data in the CD provided with this manual. To access the data, go to the CD and open the folder called 'E-views MEFMI Manual'. Inside that folder is an E-views work-file called 'E-views Model'. Double click on it to open.

227 In this example, consumption (ce1) is hypothesized to depend on the autonomous consumption part (c), disposable income (yd1), domestic prices (pcpix) oil price (poil) and import prices(pm). The variables therefore are; **ce1 yd1 i1 pcpix poil pm**

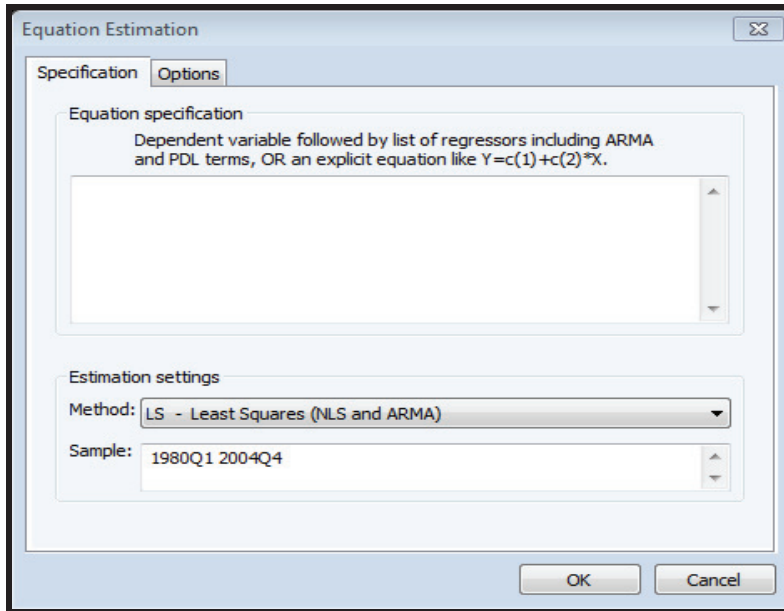
228 Step Two is to test for unit roots on all the variables. We have assumed for the sake of illustration that all the variables in this equation are stationary at levels. This is actually not true because we have already found that consumption is I(1). BUT, we proceed.

229 To run a simple linear regression in E-views, when you are in your work-file, click **Object** → **New Object** → **Equation**

230 Type the name of the dependent variable, followed by the other variables (this is not case sensitive meaning that you can write in capital letters the names of the variables that are saved in small letters). Separate each variable by spaces. Ensure to type "c" for the constant if you want to include a constant in your estimation.



231 E-views will bring a prompt window which looks like this;



232 There are a few things to note here; type the name of the variables in the big space. In the estimation settings part, the first part brings a drop-down menu with the estimation method. The default is the ordinary least squares (LS), which you will use unless you want to use another estimation approach like the ARCH. In the next space, you have the option of adjusting the sample range that you would like to use. In the above case, the default is given as 1980q1 to 2004q4. If you only wanted to use part of the sample, say up to 1990q4, for the estimation, you can change the sample range to read 1980q1 1990q4.

233 The command that we type in the open space is therefore

**ce1 c yd1 i1 pcpix poil pm**

234 Ensure that you include only the names of variables that are already in the database. Also ensure that you type the names of the variables as they are saved in the E-views work-file without any spelling mistakes.

Equation Estimation

Specification Options

Equation specification  
Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like  $Y=c(1)+c(2)*X$ .

ce1 c yd1 pcpix poil pm

Is the ordering or the variables here important?

Estimation settings  
Method: LS - Least Squares (NLS and ARMA)  
Sample: 1980Q1 2004Q4

Yes, the first must be the dependent variable; the other variables can be typed in any order

235 After typing your variables, you are now ready to click “OK” and see the results. The results window will look like this;

EViews - [Equation: UNTITLED Workfile: EIEWS MODEL::Q...

File Edit Object View Proc Quick Options Add-ins Window

Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: CE1  
Method: Least Squares  
Date: 07/30/12 Time: 09:06  
Sample: 1980Q1 2004Q4  
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	111267.6	23243.87	4.786965	0.0000
YD1	0.599910	0.066198	9.062325	0.0000
PCPIX	1028.081	192.7730	5.333117	0.0000
POIL	167.9236	141.0900	1.190188	0.2369
PM	107.1790	134.1687	0.798838	0.4264

R-squared 0.989075 Mean dependent var 475855.8  
Adjusted R-squared 0.988615 S.D. dependent var 97387.55  
S.E. of regression 10391.44 Akaike info criterion 21.38406  
Sum squared resid 1.03E+10 Schwarz criterion 21.51432  
Log likelihood -1064.203 Hannan-Quinn criter. 21.43678  
F-statistic 2150.104 Durbin-Watson stat 0.500291  
Prob(F-statistic) 0.000000

Path = c:\users\joduor\documents DB = none WF = eviews model

236 The results show the coefficient, standard error, t-statistic and the probability values. As a rule of thumb, the coefficient is statistically different from zero at the 5 percent significance level if the t-statistic is equal to or greater than 2 in absolute value

and the number of degrees of freedom is 20 or more (Gujarati, 1995 pg 129). In this case, we say that the variable is important in explaining the dependent variable. If the t-statistic is less than 2, then the independent variable is not important in explaining the dependent variable.

237 It, also possible and easier to judge using the p-values. If the p-value is less than 0.05, then the variable is important in explaining the dependent variable. As can be seen above, except for the prices of oil, all the other variables in the above equation are important in explaining consumption.

238 The other statistics that are reported include the R-squared, which tells the proportion of the variations in the dependent variable that are explained by the explanatory variables. In the above example, the R-squared of 0.99 reported means that the explanatory variables that are included in the model explains 99 percent of the variations in the dependent variables.

239 To save these results, we click on Name and a window will come that asks you to type the desired name. The results will then be saved as an equation.

### **A. Diagnostic Tests**

240 Before making any inferences out of the initial regression results, it is important to run a number of diagnostic tests to determine whether any of the ordinary least squares (OLS) assumptions have been violated. If any of the assumptions are violated, then remedial measures must be taken before any inferences are made.

241 The first important set of diagnostic tests is the residual tests. To see how the residuals from the estimated equation behave. We can go to View → Actual, Fitted, Residual → Actual, Fitted, Residual Graph. This will show the actual, the fitted and the residuals. From the residuals, one can identify periods with sharp residuals values and which can be re-looked at. They could be outlier values which could be removed in the subsequent estimations.

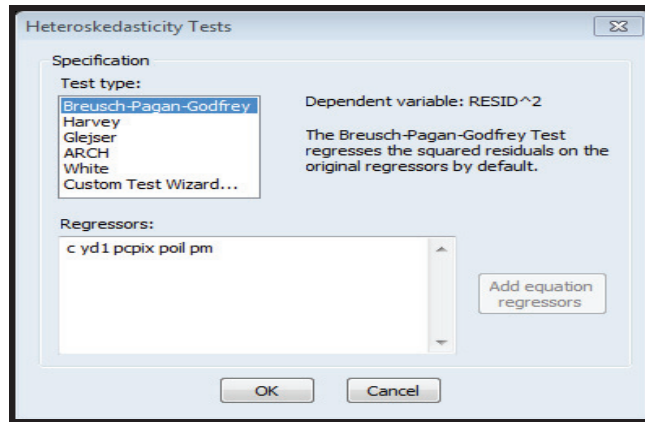
#### **(a) Testing for Heteroskedasticity**

242 Heteroscedasticity is a term used to describe the situation when the variance of the residuals from a model is not constant. When the variance of the residual is constant, we call it homoskedasticity. OLS assumes that the variance of the residuals is constant. If the variance is not constant (if there exists heteroskedasticity) then OLS is not an appropriate method to estimate the coefficients.

243 There are many tests that can be used to test for heteroskedasticity. One of them is the Breusch - Pagan - Godfrey Test. The Breusch - Pagan - Godfrey Test tests the null

hypothesis that the variance of residual ( $u$ ) is constant against the alternative that the variance of residual ( $u$ ) is not constant.

244 To run the Breusch-Pagan-Godfrey Test in E-views; from Equation View, click View → Residual Diagnostics → Heteroskedasticity Tests... A window will then appear and you click Breusch-Pagan-Godfrey.



245 The results that you get may look like this;

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.76E+09	7.77E+08	-2.267190	0.0256
YD1	4930.031	2213.933	2.226820	0.0283
PCPIX	-9654017.	6447094.	-1.497422	0.1376
POIL	10432718	4718611.	2.210972	0.0294
PM	-4129330.	4487134.	-0.920260	0.3598

246 The p-value shows that we reject null. Normally, we reject the null when the p-value is less than 0.05. Remember, the null we are testing is that the residuals ( $u$ ) are Homoskedastic against the alternative that the residuals ( $u$ ) are heteroskedastic. So the results show that the residuals are not homoskedastic. We therefore have a problem of heteroskedasticity.

247 If the residuals are heteroskedastic, then the OLS procedure is not the appropriate estimator.

248 The appropriate estimator is the generalised least squares (GLS) procedure where the variables in the equation are transformed using the heteroskedastic variances. If the variances are not known, then they need to be estimated.

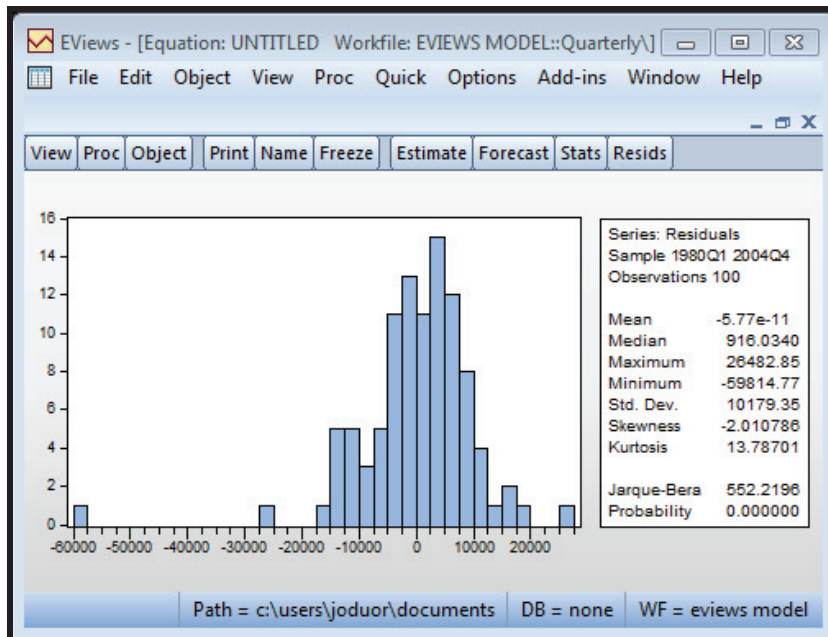
### (b) Testing for Normality

249 One of the main assumptions of the classical normal linear regression model (CNLRM) is that the residuals are normally distributed. The hypothesis tests on the coefficients obtained by OLS are based on this assumption.

250 To detect whether the residuals are normally distributed or not, we use the Jarque-Bera Statistic. The null hypothesis is that the residuals are normally distributed against the alternative that the residuals are not normally distributed.

251 Again, if the p-value of Jarque-Bera statistics is less than 5 percent (0.05) we can reject null and accept the alternative, that is residuals are not normally distributed.

252 To test in E-views, go to View → Residual Diagnostics → Histogram-Normality Test and look at the p-values and the result of the test would look like this;



253 Jarque Berra statistics is 552.2 and the corresponding p value is 0.000. Since p value is less than 5 percent we reject null meaning that population residual ( $u$ ) is not normally distributed.

254 If you find that the residuals are not normally distributed like in the case above, you can fix by increasing the number of observations.

### (c) Testing for Multicollineality

255 Multicollineality is the dependence between two explanatory variables.

256 To detect multi-collinearity, we inspect the Correlation Matrix. High correlation among the independent variables is a sign of multicollineality.

257 Another symptom of multicollineality is very high  $R^2$  but very few significant t-ratios.

258 To create a correlation matrix in E-views, select the group of independent variables by highlighting all of them from the command window right click and open as a group. Alternatively go to the command window and type `show ce1 c yd1 pcpx pm`



The screenshot shows the EViews software interface. A callout bubble points to the command window with the text "Type show ce1 c yd1 pcpx pm". The command window shows a list of variables: ce1, c, yd1, pcpx, and pm. The main window displays a list of variables with checkboxes for selection.

259 Then go to View → Covariance Analysis...Select "Correlations" (unselect "Covariance"). Off-diagonal values indicate the correlation between the variables. These correlations should be ideally less than 0.5. But this is rarely the case.

Covariance Analysis: Ordinary  
Date: 07/30/12 Time: 11:27  
Sample (adjusted): 1980Q1 2007Q2  
Included observations: 110 after adjustments  
Balanced sample (listwise missing value deletion)

Covariance Correlation	CE1	C	YD1	PCPIX	POIL	PM
CE1	1.58E+10 1.000000					
C	0.000000 NA	0.000000 NA				
YD1	1.53E+10 0.994776	0.000000 NA	1.49E+10 1.000000			
PCPIX	1946407. 0.396478	0.000000 NA	1707418. 0.358485	1521.604 1.000000		
POIL	890144.1 0.536312	0.000000 NA	899609.2 0.558673	-186.2073 -0.361965	173.9233 1.000000	
PM	3228406. 0.718315	0.000000 NA	3004581. 0.689060	1241.846 0.891472	1507220 0.003200	1275.319 1.000000

Path = c:\users\jodur\documents DB = none WF = eviews model

260 The results above show a very high correlation between ce1 and yd1, which may require that we use another proxy for yd1 instead of yd1.

261 Some of the ways to solve the problem is:

- i. Increase the number of observations in the sample
- ii. Use non-sample information such as imposing restrictions on the coefficients/parameters.
- iii. Drop the variable that is causing multicollineality and use a proxy

#### (d) Testing for Autocorrelation or Serial Correlation Test

262 To test for auto-correlation in the residual, we can use the Breusch-Godfrey serial correlation LM test : BG test.

263 Serial correlation is a statistical term used to describe the situation when the residual is correlated with lagged values of itself. In other words, if residuals are correlated, we call this situation serial correlation. Serial correlation can occur due to incorrect model specification; omitted variables; use of incorrect functional form and incorrectly transformed data.

264 When using the BG test, the null hypothesis is that there is no serial correlation (no correlation between residuals) against the alternative that there is serial correlation in the residuals. To test for serial correlation in E-views, from the equation view, **View, Residual Diagnostics, Serial Correlation LM Test...**E-views will then ask you for the number of lags to use. Let us use 5 lags. The results would like this:

Breusch-Godfrey Serial Correlation LM Test

F-statistic	31.22657	Prob. F(2,93)	0.0000
Obs*R-squared	40.17490	Prob. Chi-Square(2)	0.0000

Test Equation:  
 Dependent Variable: RESID  
 Method: Least Squares  
 Date: 07/30/12 Time: 12:02  
 Sample: 1980Q1 2004Q4  
 Included observations: 100  
 Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	70305.92	21078.46	3.335439	0.0012
YD1	-0.205573	0.060468	-3.399699	0.0010
PCPIX	462.6006	166.7566	2.774106	0.0067
POIL	75.73776	110.9286	0.682762	0.4965
PM	34.17963	104.9833	0.325572	0.7455
RESID(-1)	0.633717	0.099785	6.350803	0.0000
RESID(-2)	0.132186	0.105540	1.252474	0.2135

R-squared	0.401749	Mean dependent var	-5.77E-11
Adjusted R-squared	0.363152	S.D. dependent var	10179.35
S.E. of regression	8123.401	Akaike info criterion	20.91031
Sum squared resid	6.14E+09	Schwarz criterion	21.09268
Log likelihood	-1038.516	Hannan-Quinn criter.	20.98412
F-statistic	10.40886	Durbin-Watson stat	1.315215
Prob(F-statistic)	0.000000		

Path = c:\users\joduo\documents DB = none WF = eviews model

265 Since the  $p$ -value ( 0.000) of **Obs\*R-squared** is less than 5 percent ( $p > 0.05$ ), we reject the null hypothesis meaning that residuals ( $u$ ) are serially correlated.

### (e) Linearity Test (Adopting the Wrong Functional Form)

266 One of the assumptions of the OLS is that the model is linear. We can test for non-linearity using the Ramsey Regression Equation Specification Error Test (RESET) test.

267 RESET tests whether non-linear combinations of the estimated values help explain the endogenous variable. The intuition behind the test is that, if non-linear combinations of the explanatory variables have any power in explaining the endogenous variable, then a linear model is mis-specified, meaning that the specification of the model should be non-linear.

268 To test for non-linearity in E-views (from the equation view), select **View/Stability Tests/Ramsey RESET Test...**, you can enter 3 in the Number of fitted Terms (this denotes the order of the polynomial. If you think the correct specification of a quadratic form, then enter the number 2. However, it is better to start at higher orders like 3 or 4 and if higher order polynomials are rejected, move to lower order. It means that if you enter 3 and the results show that the model is not a polynomial of order 3, then you can try 2).

269 The null hypothesis that the coefficients on the added variables are jointly zero.



270 The results would look like;

Ramsey RESET Test  
Equation: UNTITLED  
Specification: CE1 C YD1 PCPIX POIL PM  
Omitted Variables: Powers of fitted values from 2 to 4

	Value	df	Probability
F-statistic	2.456646	(3, 92)	0.0679
Likelihood ratio	7.706105	3	0.0525

F-test summary:

	Sum of Sq	df	Mean Squares
Test SSR	7.61E+08	3	2.54E+08
Restricted SSR	1.03E+10	95	1.08E+08
Unrestricted SSR	9.50E+09	92	1.03E+08
Unrestricted SSR	9.50E+09	92	1.03E+08

LR test summary:

	Value	df
Restricted LogL	-1064.203	95
Unrestricted LogL	-1060.350	92

Unrestricted Test Equation:  
Dependent Variable: CE1  
Method: Least Squares  
Date: 07/30/12 Time: 12:08  
Sample: 1980Q1 2004Q4  
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	391816.4	126523.6	3.096785	0.0026
YD1	-12.70459	7.136816	-1.780148	0.0784
PCPIX	-22155.84	12465.03	-1.777424	0.0788
POIL	-4060.207	2088.395	-1.944176	0.0549
PM	-2149.536	1139.637	-1.886158	0.0624
FITTED^2	6.63E-05	3.62E-05	1.829907	0.0705
FITTED^3	-8.67E-11	4.81E-11	-1.802451	0.0747
FITTED^4	4.21E-17	2.36E-17	1.785515	0.0775

Path = c:\users\joduor\documents DB = none WF = eviews model

271 Since the p-value is 0.06 which is greater than 0.05 but less than 0.1, we fail to reject the null hypothesis at 5 percent but reject at 10 percent. At the 5 percent level therefore, the results show that the added coefficients are jointly equal to zero and therefore the model is linear. At the 10 percent level, we reject the null the coefficients are not jointly equal to zero (they are important in the model) and therefore the true model is one that has the additional coefficients (the non-linear model).

272 If you find that the model is non-linear, it can be made linear by taking the logs of the variables before they are used in the estimations.

### (f) Omission of Relevant Variables Test

273 We have noted in the previous section that a model should be kept as simple as possible. However, a researcher may omit variables that should be included in the model in the process of trying to make the model simple. It is worth noting that there is no amount of simplicity that can be traded for the omission of relevant variables. If an important explanatory variable is omitted, the model is said to suffer from omitted variable bias and is said to be underfitted. Let us consider the problems of this with an example.

274 To test for omitted variables in E-views (from the Equation view), select

**View/Coefficient Tests/Omitted Variables-Likelihood Ratio...** In the dialog that opens, list the names of the test variables, each separated by at least one space. Suppose, for example, that the initial regression is:

**Ls ce1 c yd1 pcpix poil pm gc1 m1**

275 If you enter the list: **gc1 m1** (government consumption and money demand) in the dialog that appears, then E-views reports the results of the unrestricted regression containing the two additional explanatory variables, and displays statistics testing the **null hypothesis that the coefficients on the new variables are jointly zero**. The top part of the output depicts the test results:

	Value	df	Probability
F-statistic	7.990447	(2, 93)	0.0005
Likelihood ratio	15.87382	2	0.0004

	Sum of Sq	df	Mean Squares
Test SSR	1.51E+09	2	7.53E+08
Restricted SSR	1.03E+10	95	1.08E+08
Unrestricted SSR	8.75E+09	93	94113770
Unrestricted SSR	8.75E+09	93	94113770

	Value	df
Restricted LogL	-1064.203	95
Unrestricted LogL	-1056.266	93

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	80853.15	23854.76	3.389392	0.0010
YD1	0.486234	0.070714	6.876111	0.0000
PCPIX	-439.2234	232.6770	-1.887596	0.0622
POIL	-4.363760	154.0454	-0.028328	0.9775
PM	394.9799	144.4546	2.734277	0.0075
GC1	0.410022	0.146185	2.804817	0.0051
M1	0.255495	0.065192	3.919112	0.0002

276 In this example, the tests reject the null hypothesis that the coefficients on the new variables are jointly equal to zero. This means that the two series, in-fact, belong to the equation at a 5 percent significance level. This implies that the additional set of variables gc1 and m1 are jointly significant (necessary) in the model. The rule is always to reject the null hypothesis when the probability value is less or equal to 0.05 if you are using 5 percent significance level.

277 The consequences of omitting the relevant variable are as follows:

- Omitting a relevant variable lumps the effects of the relevant variables together with the error term and therefore the expected value of the error term will no longer be equal to zero. This violates the Gauss - Markov Assumptions about the mean of the errors.
- More importantly, in the case where the omitted variable, let us call it  $x$ , is correlated with the included variable, let us call it  $y$ , then  $Cov(x^*, y) \neq 0$ . As a

result, Gauss-Markov theorem does not apply. In general, OLS estimate of the regression coefficient is biased, ie,

$$E(\hat{\beta}^*) \neq \beta$$

- The bias will only be zero when the coefficient of the omitted variable is significantly equal to zero or when the two variables are not correlated.
- In addition, the standard errors on the estimated coefficients will be biased. In the misspecified model:

$$Var(\hat{\beta}_1^*) = \frac{\sigma^2}{\sum s_i^2}$$

- But variance of the 'true' estimator is:

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum s_i^2 (1 - r_{12}^2)}$$

where  $r_{12}$  is the correlation coefficient between  $x$  and  $y$ . This means that:

$$\text{If } r_{12}^2 > 0, \text{ then } Var(\hat{\beta}^*) < Var(\hat{\beta})$$

- The variance of estimated coefficient is also *biased*. We are placing 'too much' confidence in our coefficient estimates. The result is that the  $t$  test will be misleading (this is true *even if*  $r_{12}=0$ , because our estimate of  $\sigma^2$  will also be *biased*.)

278 Running a model with the above shortcomings will result in the entire hypothesis testing results and confidence intervals losing any meaning and thus, conclusions drawn from such results cannot be relied on. Therefore, developing a model requires a lot of care. One should pay close attention to the theory underlying the phenomenon under studying order to ensure that all theoretically relevant variables are included in the model.

279 The remedial measure is easy if we know or have determined through the above test which variable has been omitted. If the omitted variable is available include it in the model. If the omitted variable is not available, try to find a proxy variable that is closely related to this missing variable.

### (g) Inclusion of Unnecessary Variables (Redundant Variables Test)

280 In some cases, there is a high temptation of running away from under-fitting a model by including as many variables as one can think of. This can be done without noticing that one is also running into another problem, over-fitting or over specifying, which results from including too many variables. Some researchers want to include all sorts of variables, sometimes even including those that have no theoretical basis. In most cases, one is not sure about the role of such variables in the model, especially when the theory underlying the phenomenon under study is not well developed. Over-fitting a model has the main effect on increasing the R-squared, showing as if the predictive power of the model has increased.

281 To test for redundant variables in E-views, select **View/Coefficient Tests/Redundant Variables-Likelihood Ratio...** In the dialog that appears, list the names of each of the test variables, separated by at least one space. Suppose, for example, that we included exports (e1) and investments (i1) as part of the variables that explain consumption in the initial regression as:

**ce1 c yd1 pcpix poil pm e1 i1**

282 If you type the list: **e1 i1** in the dialog, then E-Views reports the results of the restricted regression dropping the two regressors, followed by the statistics associated with the test of the null hypothesis that the coefficients on the two variables are jointly zero (that the two variables are not necessary in the model). In our case we have:

Redundant Variables Test  
Equation: UNTITLED  
Specification: CE1 C YD1 PCPIX POIL PM E1 I1  
Redundant Variables: E1 I1

	Value	df	Probability
F-statistic	11.88421	(2, 93)	0.0000
Likelihood ratio	22.75931	2	0.0000

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	2.09E+09	2	1.04E+09
Restricted SSR	1.03E+10	95	1.08E+08
Unrestricted SSR	8.17E+09	93	87851643
Unrestricted SSR	8.17E+09	93	87851643

LR test summary:

	Value	df
Restricted LogL	-1064.203	95
Unrestricted LogL	-1052.823	93

Restricted Test Equation:  
Dependent Variable: CE1  
Method: Least Squares  
Date: 07/30/12 Time: 12:48  
Sample: 1980Q1 2004Q4  
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	111267.6	23243.87	4.786965	0.0000
YD1	0.599910	0.066198	9.062325	0.0000
PCPIX	1028.081	192.7730	5.333117	0.0000
POIL	167.9236	141.0900	1.190188	0.2369
PM	107.1790	134.1687	0.798838	0.4264

R-squared 0.989075 Mean dependent var 475855.8  
Adjusted R-squared 0.988615 S.D. dependent var 97387.55  
S.E. of regression 10391.44 Akaike info criterion 2138406

Path = c:\users\joduo\documents DB = none WF = eviews model

283 The p-values associated with the F-statistic and the log-likelihood ratio are both less than 0.05 (at 5 percent significance level) implying that we reject the null hypothesis that the two variables are not necessary in the model.

284 The consequences of over-fitting a model are as follows:

- The OLS estimators are all unbiased and consistent. This is because if the included variable is not relevant in explaining the phenomenon, its coefficient is expected to be zero, thus leaving the relevant variables estimating their true values.
- The variance from the estimated model is correctly estimated and thus the hypothesis testing and confidence interval analysis remain valid.
- However the estimated coefficients from the estimated model are inefficient, with their values and their standard errors larger than those estimated from the true model. This will invalidate the confidence intervals based on the standard errors of the coefficients
- This makes the t-ratio smaller than it should be, and makes it more likely that we won't be able to reject the null hypothesis (that the coefficients are statistically not different from zero) when we should. In other words, we are likely to conclude from the resulting t-statistics that the variables are not important explanatory variables when in actual sense they are.

285 By looking at the two specification errors, it can be seen that it is better to include irrelevant variables if one cannot correctly specify a model, although this philosophy is not encouraged. The coefficients, as has been highlighted before, are still unbiased and consistent, the variance is still correctly estimated and the standard hypothesis testing procedures apply. The only penalty we suffer is that the inclusion of such superfluous variables tends to result in our probability inferences about the true parameters being less precise because the confidence intervals are wider.

### **3.10.2 Estimation with Non-Stationary, but Cointegrated variables with No Feedback Effects - Error Correction Models**

286 We mentioned above that after testing for unit roots and finding that all your variables are stationary at levels (no feedback effects), then you use a single equation of the variables at their levels.

287 We also mentioned that if we test for unit roots and find that the variables are non-stationary and cointegrated then we use an error correction model. We will at the end of this section mention what to do when you encounter a situation of non-stationary variables that are not cointegrated.

288 We also mentioned that you can correct for non-stationarity by differencing if the variables are difference stationary.

289 Differencing however, results in the loss of some valuable long-run information in the data. You cannot therefore use coefficients from a regression using only differenced variables to make inferences about the long-run relationships, but only short-run relationships.

290 But we need to know something about the long-run relationship even among the variables that are non-stationary at levels. This inference is only possible if the non-stationary variables are cointegrated.

**But what is cointegration and how do you test for it?**

291 The notion of long run equilibrium in econometrics implies that two or more variables may wander away from each other in the short run but move together in the long run.

292 When the variables wander away from each other, they are called random walk processes. In the long run however, it may be possible that these variables move in the same direction (have a long run relationship). In this case, there may be a linear combination of these random walk processes, that is, a white noise (stationary) and the variables are said to be cointegrated. The process by which the random walk processes come back towards (attract) each other is called an error correction mechanism. We will come back to error correction mechanisms later.

293 Several approaches have been proposed in the literature to test for cointegration among variables. The two most popular however, are the Engle-Granger two Step Method and the Johansen's test. We will illustrate how to test for cointegration using only these two approaches.

**A. Engle-Granger Two Step Cointegration Method**

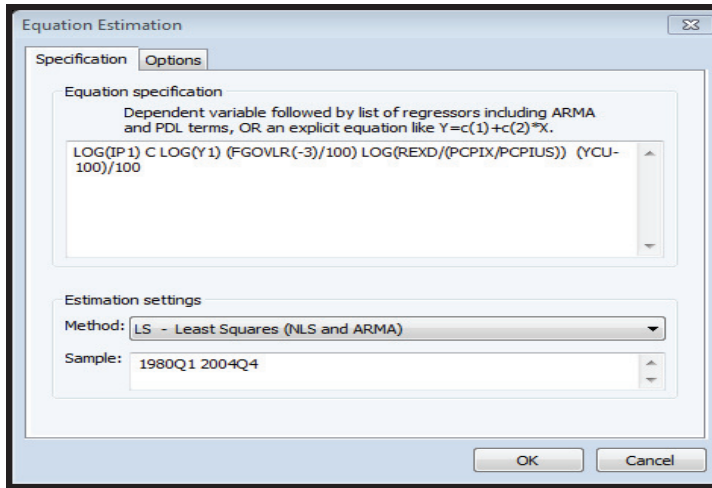
294 Let us use the investments equation (IP1) in our database to test for cointegration among the variables in that model

295 The variables in the investment equation are; Private Investment (IP1); FGOVLR-real bond rate; YCU-capacity utilisation and Y1-real GDP.

296 There are two steps to be followed;

**(a) Step One-Estimate the Long-run Equation**

297 The long run equation to be estimated is given as:  $\text{LOG}(\text{IP1}) \text{ C } \text{LOG}(\text{Y1})$   
 $(\text{FGOVLR}(-3)/100) \text{ LOG}(\text{REXD}/(\text{PCPIX}/\text{PCPIUS})) (\text{YCU}-100)/100$



298 The results are given as:

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9.267427	1.321822	-7.011101	0.0000
LOG(Y1)	1.569875	0.108177	14.51210	0.0000
FGOVL R(-3)/100	0.180970	0.327577	0.552451	0.5820
LOG(REXD/(PCPIX/PCPIUS))	-0.111929	0.057701	-1.939804	0.0555
(YCU-100)/100	1.821002	0.590567	3.083482	0.0027

R-squared	0.806847	Mean dependent var	11.31802
Adjusted R-squared	0.798449	S.D. dependent var	0.213483
S.E. of regression	0.095842	Akaike info criterion	-1.802064
Sum squared resid	0.845081	Schwarz criterion	-1.669347
Log likelihood	92.40009	Hannan-Quinn criter.	-1.748400
F-statistic	96.07640	Durbin-Watson stat	0.130687
Prob(F-statistic)	0.000000		

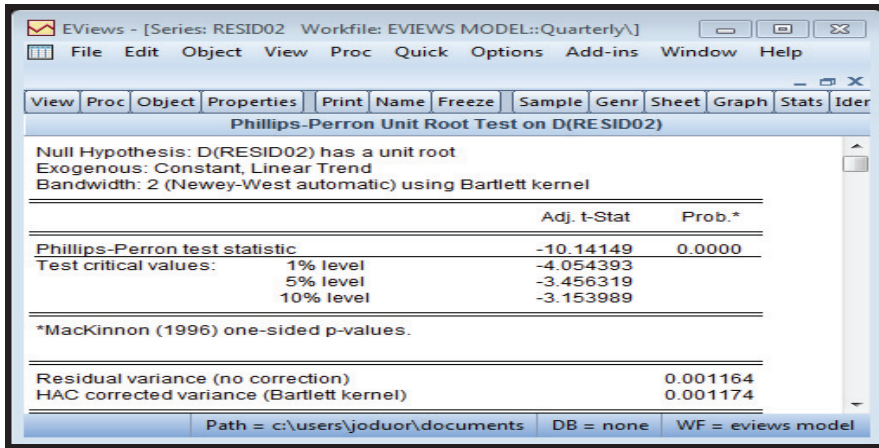
### (b) Step Two – Test for Stationarity of the Residuals

299 From the estimation of the long run equation, generate the residuals and test whether the residuals are stationary or not.

300 If the residuals are stationary, then the variables are cointegrated; if the residuals are not stationary, then the variables are not cointegrated.

301 To generate the residuals, click **Proc, Make Residual Series** and name the residuals. We have named the residuals as **resid02**

302 We can then test for the unit roots in the residuals (resid02) in the usual way. The unit root test results on the residuals are given as;



303 Using Phillip-Perron test, the results show that the residuals are stationary meaning that the variables in the investment equation are cointegrated.

**We will illustrate how to test for cointegration using the Johansen approach under the Vector Error Correction Models.**

## B. Error Correction Models

304 Granger's Representation Theorem (Granger 1983) has shown that whenever cointegration exists between non-stationary series, there must exist an error correction mechanism maintaining it.

305 Non-stationary variables that are cointegrated must therefore be represented by an error correction model.

306 Assume we have two variables  $y_t$  and  $Z_t$  that are cointegrated, an error correction model can be formulated in the form;

$$\Delta y_t = \delta_1 + \phi_y [y_{t-1} - \alpha_1 z_{t-1}] + \sum_{i=1}^p \gamma_{1i} \Delta y_{t-i} + \sum_{i=1}^p \lambda_{1i} \Delta z_{t-i} + \varepsilon_{yt}$$



$$\Delta z_t = \delta_2 + \varphi_z [y_{t-1} - \alpha_1 z_{t-1}] + \sum_{i=1}^p \gamma_{2i} \Delta y_{t-i} + \sum_{i=1}^p \lambda_{2i} \Delta z_{t-i} + \varepsilon_{zt}$$

where  $i = 1, \dots, p$  are the number of lags,  $\varphi_y$  and  $\varphi_z$  are the adjustment parameters. It is possible to use the deviation series in period  $(t-1)$  estimated in step two, that is  $\{\tilde{\varepsilon}_{t-1}\}$  as an instrument for the explanatory variables  $y_{t-1} - \alpha_1 z_{t-1}$  to avoid cross-equation restrictions. The above error correction models then become;

$$\Delta y_t = \delta_1 + \varphi_y \tilde{\varepsilon}_{t-1} + \sum_{i=1}^p \gamma_{1i} \Delta y_{t-i} + \sum_{i=1}^p \lambda_{1i} \Delta z_{t-i} + \varepsilon_{yt}$$

$$\Delta z_t = \delta_2 + \varphi_z \tilde{\varepsilon}_{t-1} + \sum_{i=1}^p \gamma_{2i} \Delta y_{t-i} + \sum_{i=1}^p \lambda_{2i} \Delta z_{t-i} + \varepsilon_{zt}$$

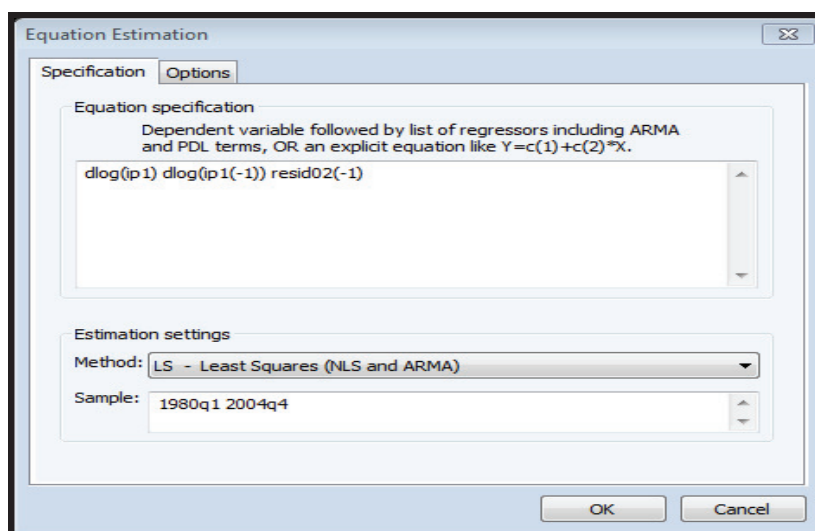
$\{\tilde{\varepsilon}_{t-1}\}$  is the error correction term. The other variables are in their first differences,  $\Delta$  is the difference operator. This model can then be estimated using OLS.

### C. Formulating an Error Correction Model in E-views

307 We have just tested for cointegration using the Engle and Granger two step method.

308 From that process, we generated the residuals from the long-run model and named then resid02.

309 From the formulation of the ECM above, the ECM is basically first differences of the non-stationary variables and the error term, which is lagged once as shown in the example of the investment equation below ;



310 The results are given as;

Dependent Variable: DLOG(IP1)  
Method: Least Squares  
Date: 07/31/12 Time: 12:33  
Sample (adjusted): 1980Q3 2004Q4  
Included observations: 98 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(IP1(-1))	0.120885	0.098692	1.224879	0.2236
RESID02(-1)	-0.075087	0.039252	-1.912931	0.0587

R-squared 0.005847 Mean dependent var 0.007579  
Adjusted R-squared -0.004509 S.D. dependent var 0.035998  
S.E. of regression 0.036079 Akaike info criterion -3.785998  
Sum squared resid 0.124965 Schwarz criterion -3.733244  
Log likelihood 187.5139 Hannan-Quinn criter. -3.764660  
Durbin-Watson stat 2.077666

Path = c:\users\joduor\documents DB = none WF = eviews model

311 The coefficient of resid02 is the error correction term. It should ideally be negative and less than one. A negative sign means that the short-run shocks adjust towards the long-run path.

312 The results show that the error correction term in the above ECM is 0.07 or 7 percent. This is interpreted as 'seven percent of the short-run shocks will be adjusted back to the long run path in one period (in our case in one quarter, because we are using quarterly data). It means that it takes about 14 quarters for a temporary shock on investments to adjust fully back to the long-run path.

313 The coefficients of the ECM as given in the output above are the short-run coefficients. When the variables are cointegrated, the coefficients of the estimation of the variables in their levels are no longer spurious. Stock (1989) shows that the coefficients are super-consistent when the variables are cointegrated. That is the reason why you can be able to interpret these coefficients as the long-run coefficients while the coefficients from the estimation (in levels) of variables that are non-stationary but are not cointegrated are spurious.

### 3.10.3 Estimation with non-Stationary, but Cointegrated Variables with Feedback Effects - Vector Error Correction Models (VECM)

314 When there are feedback effects, we use the vector autoregressions (VAR).

315 VAR econometrics analysis entails estimating regression equations in which the current value of each variable is expressed as a function of lagged values of itself and of each of the selected variables. No variable is assumed to be exogenous a priori and no variable is excluded from the autoregressive equation for any of the variables in the system. VAR sidesteps the need for theoretical underpinning of the models. If you have variables that have no known theoretical relations and the variables are suspected to have feedback effects, then the best modeling approach is the VAR.

316 Just like in the single equation approach, when equations in a VAR are cointegrated, then they must be represented by an error correction model. In this case, the model is called a Vector Error Correction Model (VECM).

**See the Appendix C for the technical details of a VAR and a VECM**

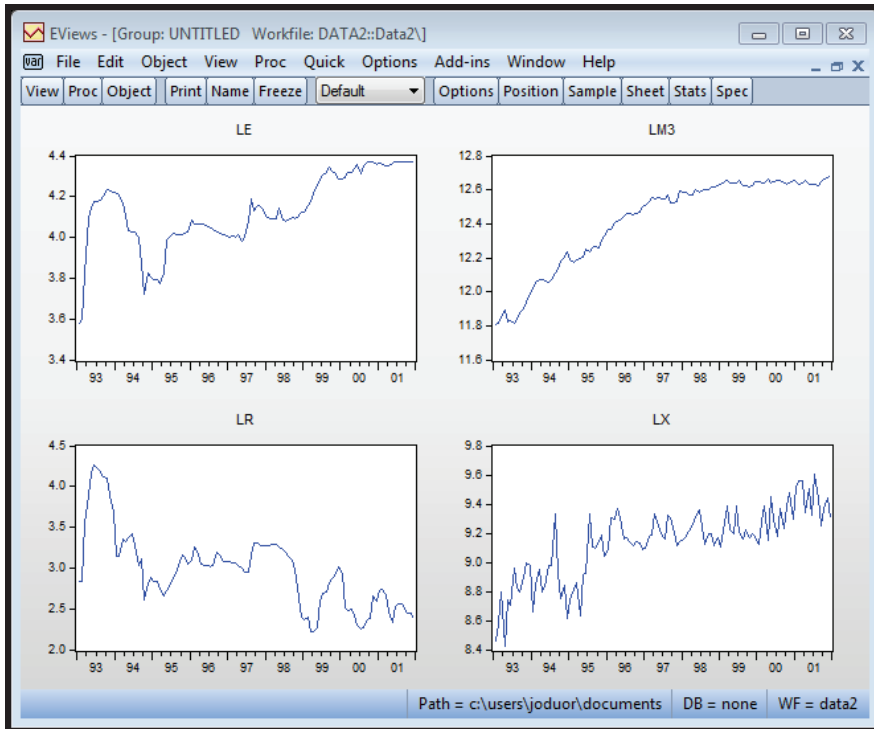
#### A. Estimating a VECM in E-views

317 To illustrate how to conduct Johansen cointegration test and thereafter estimate a VECM, we use monthly Kenyan data on the log of exchange rates (le), log of money supply (lm3), log of interest rates (lr) and log of exports (lx) for the period January 1993 to December 2001 (108 observations). The data is also provided in the CD accompanying this manual and E-views work-file is named 'Kenyan Data' in the 'E-views MEFMI Manual' folder.

##### **STEP 1: Determine the order of integration of variables**

318 The first step in the exercise involves the determination of order of integration. This is done using the formal unit root tests as was illustrated earlier.

319 It is always a good idea to first plot the variables to see how they look like.



320 A casual inspection of the plots of the variables above shows that the variables could be non-stationary. Formal tests which we will not report here show that the four variables are  $I(1)$ . Those interested in how to conduct the tests can refer back to the section on unit root tests.

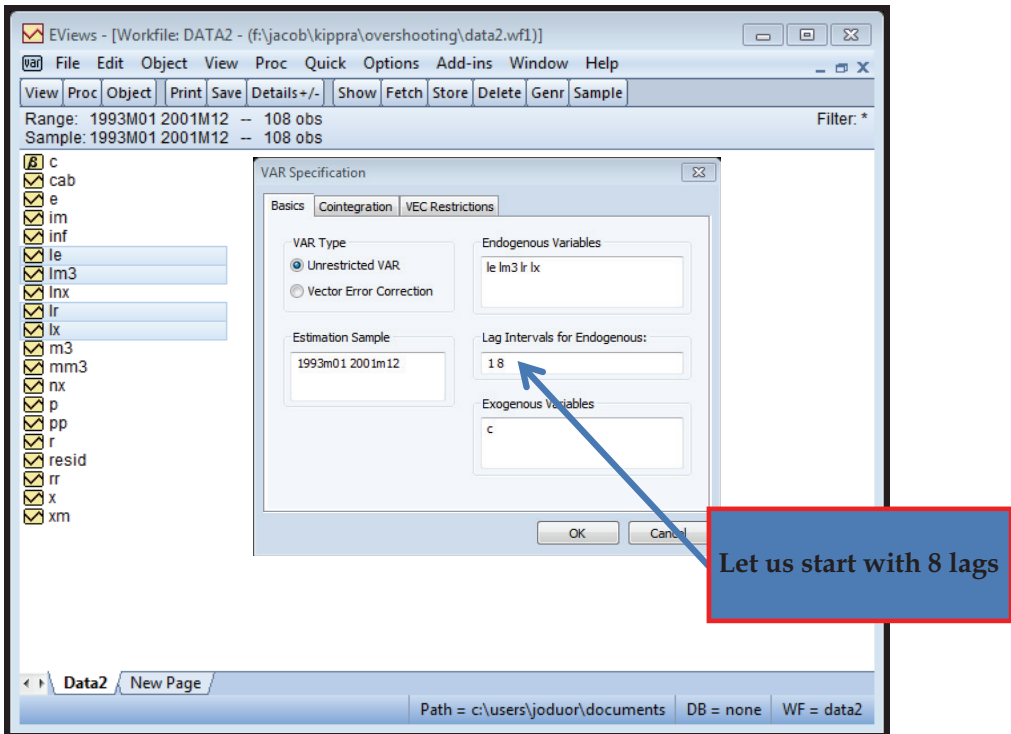
### STEP 2: Determine the number of lags for the VAR

321 Before estimating a VAR, you must think about the appropriate lags to use. Using too many lags will reduce the degrees of freedom while too few lags may leave out a lot of information in the data. One option to use in determining the optimal lag length are the various information criteria. These include the Akaike information criterion (AIC), Schwarz information criterion (SC), Hannan-Quinn information (HQ) criterion, Final prediction error (FPE) and sequential modified LR test statistic.

322 To choose the optimal lag length, you will need to estimate the VAR first. Highlight the variables and right click on **...Open, as Var...** to open the variables as VAR.

323 This opens a menu, which has **Basics/Cointegration/VEC restrictions**. Select **unrestricted VAR**. In the lag length you can put between 6 or 12 lags depending on the number of observations you have. Maybe start with 10 lags; if you find that the estimation cannot be done due to few observations reduce to 8 lags or so. In our case, we

start with 8 lags. To put 8 lags, change to 18 in the space written *lag intervals for endogenous variables*.



324 The program will estimate the reduced form VAR with the following output.

The screenshot shows the EViews Vector Autoregression Estimates output window. The output is a table of parameter estimates for the VAR(5) model. The table includes the following data:

	LE	LM3	LR	LX
LE(-1)	1.131887 (0.12401) [ 9.12752]	-0.064449 (0.05403) [-1.19277]	-0.036275 (0.43250) [-0.08387]	0.756101 (0.39170) [ 1.93030]
LE(-2)	-0.205907 (0.18408) [-1.11858]	0.029325 (0.08021) [ 0.36561]	0.645301 (0.64200) [ 1.00514]	0.636628 (0.58145) [ 1.09490]
LE(-3)	0.017786 (0.18902) [ 0.09410]	0.008994 (0.08236) [ 0.10921]	-0.640106 (0.65923) [-0.97099]	-1.797826 (0.59705) [-3.01119]
LE(-4)	0.177378 (0.19845) [ 0.89384]	0.099701 (0.08647) [ 1.15305]	0.392036 (0.69211) [ 0.56644]	0.434660 (0.62682) [ 0.69343]
LE(-5)	-0.137184 (0.17866) [-0.76784]	-0.158759 (0.07785) [-2.03938]	0.087786 (0.62311) [ 0.14088]	-0.502459 (0.56433) [-0.89036]

325 After this estimation, go to **View/lag structure/lag length criteria**. In the lag specification, select larger lag. For instance, 12 lags. Higher versions of E-views uses several information criteria to select the lags (We are using E-views 7 for the tests):

Click View/lag structure/lag length criteria. Choose 8 lags and click OK

	LE	LM3	LR	LX
LE(-1)	1.131887 (0.12401) [9.12752]	-0.064449 (0.05403) [-1.19277]	-0.036275 (0.43250) [-0.08387]	0.756101 (0.39170) [1.93030]
LE(-2)	-0.205907 (0.18408) [-1.11858]	0.029325 (0.08021) [0.36561]	0.645301 (0.64200) [1.00514]	0.636628 (0.58145) [1.09490]
LE(-3)	0.017786 (0.18902) [0.09410]	0.008994 (0.08236) [0.10921]	-0.640106 (0.65923) [-0.97099]	-1.797826 (0.59705) [-3.01119]
LE(-4)	0.177378	0.099701	0.392036	0.434660 (0.62682) [0.69243]
L				-0.502459 (0.56433) [-0.89036]

Lag Specification dialog box: Lags to include: 8

326 The results of the lag length test are given as;

VAR Lag Order Selection Criteria  
Endogenous variables: LE LM3 LR LX  
Exogenous variables: C  
Date: 07/31/12 Time: 16:02  
Sample: 1993M01 2001M12  
Included observations: 100

Lag	LogL	LR	FPE	AIC	SC	HQ
0	106.3951	NA	1.52e-06	-2.047901	-1.943694	-2.005727
1	601.9891	941.6287	1.04e-10	-11.63978	-11.11875*	-11.42891*
2	614.8922	23.48363	1.10e-10	-11.57784	-10.63998	-11.19827
3	636.9154	38.32030*	9.82e-11*	-11.69831*	-10.34362	-11.15004
4	649.8888	21.53588	1.05e-10	-11.63778	-9.866260	-10.92081
5	660.3453	16.52133	1.19e-10	-11.52691	-9.338564	-10.64125
6	673.6360	19.93595	1.28e-10	-11.47272	-8.867549	-10.41836
7	686.8239	18.72689	1.39e-10	-11.41648	-8.394481	-10.19342
8	702.5257	21.04040	1.44e-10	-11.41051	-7.971690	-10.01876

\* indicates lag order selected by the criterion  
LR: sequential modified LR test statistic (each test at 5% level)  
FPE: Final prediction error  
AIC: Akaike information criterion  
SC: Schwarz information criterion  
HQ: Hannan-Quinn information criterion

-11.15004 Path = c:\users\jodur\documents DB = none WF = data2

327 It can be observed from the table above that the majority of the criteria suggest that we use 3 lags (see the row in bold). You always go with what the majority suggest.

### STEP 3: Cointegration test

328 To do cointegration test, go to **View/cointegration test**. But now you choose one lag because you have established that three lags are optimal

329 You will also need to determine the trend characteristics of the VECM. The variables may have either deterministic or stochastic trends. Similarly, the cointegrating equations may have intercepts and deterministic trends. Without going into too many technical details about this, we will assume that there is no trend but that there are intercepts and choose option number three. Readers who are interested in the technical details can consult econometrics text like Enders (2004).

Change to three lags

330 After selecting the trend characteristics, we test for cointegration. From the estimated output window, and choose option 2 (intercept but no trend) and click OK. Remember to change the lag length, so that you use three lags as was determined above. We therefore put 1 3 in the space for lags. This means that we are using lags one to three in the estimation. The results are given as:

Unrestricted Cointegration Rank Test (Trace)				
Hypothesized		Trace	0.05	
No. of CE(s)	Eigenvalue	Statistic	Critical Value	Prob.**
None *	0.262758	62.67086	54.07904	0.0071
At most 1	0.142197	30.96765	35.19275	0.1331
At most 2	0.100725	15.01610	20.26184	0.2254
At most 3	0.037498	3.974749	9.164546	0.4159
Trace test indicates 1 cointegrating eqn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
**MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Cointegration Rank Test (Maximum Eigenvalue)				
Hypothesized		Max-Eigen	0.05	
No. of CE(s)	Eigenvalue	Statistic	Critical Value	Prob.**
None *	0.262758	31.70320	28.58808	0.0193
At most 1	0.142197	15.95155	22.29962	0.3017
At most 2	0.100725	11.04135	15.89210	0.2488
At most 3	0.037498	3.974749	9.164546	0.4159
Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
**MacKinnon-Haug-Michelis (1999) p-values				

331 Both the Trace Test and the Maximum eigenvalue test indicate that there is only one cointegrating equation.

#### STEP 4: Test long-run restrictions

332 After finding that there is only one cointegrating equation, we can now estimate the VECM which includes the long-run theoretical restrictions.

333 The formulation of a VECM is given in Appendix C.

334 It is now time to estimate the VECM. Before we do that, we need to formulate the VECM with our trend assumptions, to see how it looks like. Since we included the intercept in the cointegrating space, there are at most 4 cointegrating vectors and loading matrices as follows.



$$\begin{bmatrix} \Delta \ln e_t \\ \Delta m_t \\ \Delta r_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \\ d_{41} & d_{42} & d_{43} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{22} & \gamma_{22} & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} \end{bmatrix} \begin{bmatrix} \Delta \ln e_{t-1} \\ \Delta i \ln m_{t-1} \\ \Delta \ln r_{t-1} \\ \Delta x_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \\ \alpha_{51} & \alpha_{52} & \alpha_{53} & \alpha_{54} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} & \beta_{51} \\ \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} & \beta_{52} \\ \beta_{13} & \beta_{23} & \beta_{33} & \beta_{43} & \beta_{53} \\ \beta_{14} & \beta_{24} & \beta_{34} & \beta_{44} & \beta_{54} \end{bmatrix} \begin{bmatrix} \ln e_{t-1} \\ \ln m_{t-1} \\ \ln r_{t-1} \\ x_{t-1} \\ C \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix}$$

335 Remember, we found out that there is only one cointegrating relation among the variables. With only one cointegrating relations, our VECM can take several forms like;

$$\begin{bmatrix} \Delta \ln e_t \\ \Delta m_t \\ \Delta r_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \\ d_{41} & d_{42} & d_{43} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{22} & \gamma_{22} & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} \end{bmatrix} \begin{bmatrix} \Delta \ln e_{t-1} \\ \Delta i \ln m_{t-1} \\ \Delta \ln r_{t-1} \\ \Delta x_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \alpha_{41} \\ \alpha_{51} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} & \beta_{51} \end{bmatrix} \begin{bmatrix} \ln e_{t-1} \\ \ln m_{t-1} \\ \ln r_{t-1} \\ x_{t-1} \\ C \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix}$$

336 If the cointegration relation is among the variables in the second equation, we will have;

$$\begin{bmatrix} \Delta \ln e_t \\ \Delta m_t \\ \Delta r_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \\ d_{41} & d_{42} & d_{43} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{22} & \gamma_{22} & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} \end{bmatrix} \begin{bmatrix} \Delta \ln e_{t-1} \\ \Delta i \ln m_{t-1} \\ \Delta \ln r_{t-1} \\ \Delta x_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{12} \\ \alpha_{22} \\ \alpha_{32} \\ \alpha_{42} \\ \alpha_{52} \end{bmatrix} \begin{bmatrix} \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} & \beta_{52} \end{bmatrix} \begin{bmatrix} \ln e_{t-1} \\ \ln m_{t-1} \\ \ln r_{t-1} \\ x_{t-1} \\ C \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix}$$

337 There are four of such possibilities because we have four possible cointegrating relations. We have found, however, that there is only one cointegrating relation. What we do not know is which of the four possible relations is the one we got. The task then is to determine which of the possible four relations is binding here. This is done by imposing long-run restrictions in the VAR.

338 There are four possibilities. If the cointegration relation is among the variables in the first equation then the one **cointegrating vector** is given as:

$$\beta_{11} \ln e_{t-1} + \beta_{21} m_{t-1} + \beta_{31} \ln r_{t-1} + \beta_{41} x_{t-1} + \beta_{51} C$$

339 This cointegrating vector feeds into the 4 different equations as follows;

$$\begin{aligned} \Delta \ln e_t &= \alpha_{11} [\beta_{11} \ln e_{t-1} + \beta_{21} \ln m_{t-1} + \beta_{31} \ln r_{t-1} + \beta_{41} x_{t-1} + \beta_{51} C] \\ \Delta \ln m_t &= \alpha_{21} [\beta_{11} \ln e_{t-1} + \beta_{21} \ln m_{t-1} + \beta_{31} \ln r_{t-1} + \beta_{41} x_{t-1} + \beta_{51} C] \\ \Delta \ln r_t &= \alpha_{31} [\beta_{11} \ln e_{t-1} + \beta_{21} \ln m_{t-1} + \beta_{31} \ln r_{t-1} + \beta_{41} x_{t-1} + \beta_{51} C] \\ \Delta x_t &= \alpha_{41} [\beta_{11} \ln e_{t-1} + \beta_{21} \ln m_{t-1} + \beta_{31} \ln r_{t-1} + \beta_{41} x_{t-1} + \beta_{51} C] \end{aligned}$$

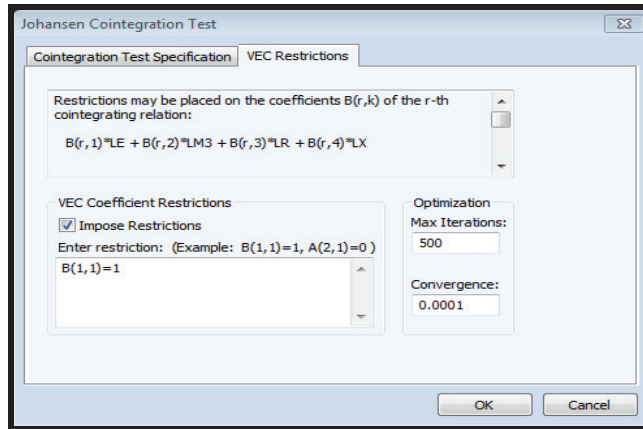
340 The reason why we cannot just assume that the relationship obtained is on the first equation is that, all the variables in the system have unique theory behind them which could be binding in this data or not. What we do not know is which theory is binding and which is not, but we know that out of the four theories governing the relations, one is binding in the long run. This is now where theory comes in. You need to know the theoretical relationships among the variables in the model and test each of them to see which theoretical relations are represented by the one cointegrating relation.

341 We can first test whether the theoretical relationship governing exchange rates and the other variables is supported by this model. If not, then we proceed to test maybe the money demand theory represented by the second equation in the system, or the Taylor rule represented by the interest rate equation in the third row. We will do this until we find which long-run relationship is represented by the one cointegrating relation that we found. To start, let us long run relationship with exchange rates so that the restricted VECM is given as:

$$\begin{bmatrix} \Delta \ln e_t \\ \Delta m_t \\ \Delta r_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \\ d_{41} & d_{42} & d_{43} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{22} & \gamma_{22} & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} \end{bmatrix} \begin{bmatrix} \Delta \ln e_{t-1} \\ \Delta \ln m_{t-1} \\ \Delta \ln r_{t-1} \\ \Delta x_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \alpha_{41} \\ \alpha_{51} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} & \beta_{51} \end{bmatrix} \begin{bmatrix} \ln e_{t-1} \\ \ln m_{t-1} \\ \ln r_{t-1} \\ x_{t-1} \\ C \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix}$$

342 Observe here that the coefficient of exchange rates in the exchange rate equation is normalised to one. To impose the restriction that this coefficient is one and all the others in the system are non-zero, go to **view/cointegration test**. Select **VEC restrictions**.

343 In the VEC coefficient restrictions, click on **impose restrictions**. A blank space to type the restrictions is then activated. In our case we type B(1,1)=1.



344 Note that the ordering of the variables is important here because  $\mathbf{B}(1,1)=1$  instructs E-views to test whether truly the coefficient  $\beta_{11}$  equals to one. In this case, exchange rates must be the first variable in the order. If exchange rates were the second variable, this restriction test would be on something else but not on the coefficient of exchange rates. After imposing the restriction, we get the following results.

Tests of cointegration restrictions:

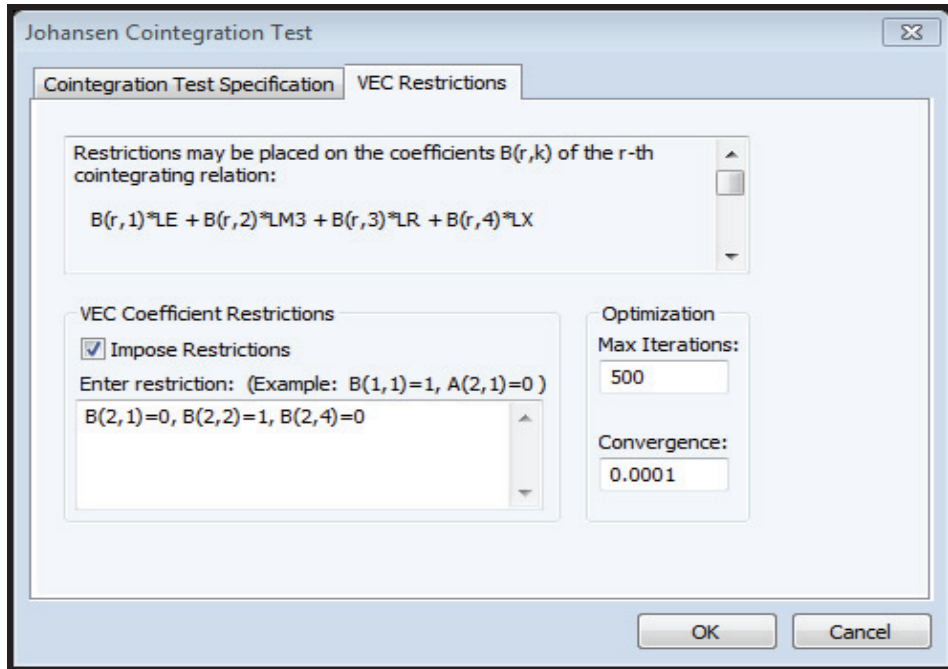
Hypothesized No. of CE(s)	Restricted Log-likelihood	LR Statistic	Degrees of Freedom	Probability
1	658.7748	NA	NA	NA
2	666.4983	NA	NA	NA
3	669.7710	NA	NA	NA

NA indicates restriction not binding.

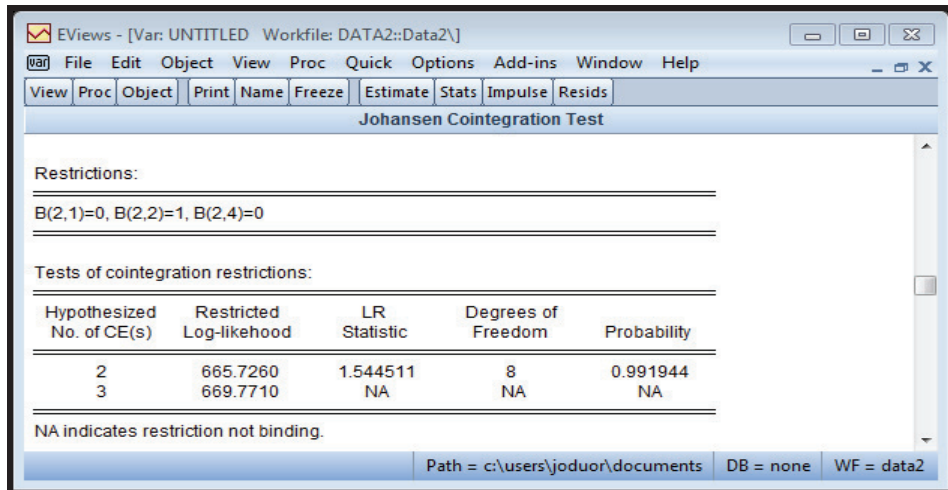
Path = c:\users\joduor\documents DB = none WF = data2

345 The results show that the restrictions were invalid, implying that there is no long run relationship in the exchange rate equation with all those variables as we have them.

346 We then try another set of restrictions. Let us try the money demand function and put the restrictions,  $\mathbf{B}(2,1)=0$ ,  $\mathbf{B}(2,2)=1$ ,  $\mathbf{B}(2,4)=0$ . These restrictions suppose that money demand is not a function of exchange rates or exports but a function of interest rates only.



347 The results are given as;



348 The hypothesis is not rejected, implying that data is consistent with a traditional demand for money equation specified in the VAR.

349 The estimated long run and VECM model is given as:

Restricted cointegrating coefficients (not all coefficients are identified)				
LE	LM3	LR	LX	C
-5.423911	-13.10727	-49.24875	13.72985	185.6976
0.000000	1.000000	36.88201	0.000000	-105.6951
0.644471	18.12708	12.40849	-23.46497	-44.22526
Adjustment coefficients (standard error in parentheses)				
D(LE)	0.007908	0.008246	0.005717	
	(0.00343)	(0.00402)	(0.00188)	
D(LM3)	0.004602	0.006295	9.01E-05	
	(0.00165)	(0.00194)	(0.00091)	
D(LR)	0.019778	0.021163	0.007431	
	(0.01202)	(0.01412)	(0.00661)	
D(LX)	-0.011411	-0.017144	0.009103	
	(0.01224)	(0.01438)	(0.00674)	

350 From the estimated coefficients, the long-run demand for money function for Kenya can be written as;

$$\ln m_t = 36.88r - 105.69$$

351 It is important to observe the elasticities obtained here may not make a lot of sense as we suspect that the model is wrongly specified. A traditional money demand function should have income as one of the explanatory variables. We did not have it in our model. So we do not labour to make sense out of the results. Just to illustrate the process.

352 One important point that comes out from the preceding discussion is the importance of selecting the right variables to include in the model. Theory must always guide you from the beginning, because as can be seen above, we did not include income variable in the model and therefore, we cannot conclusively infer about money demand function for Kenya.

---

---

## CHAPTER 4

### BUILDING A PROTOTYPE MODEL

#### Objectives of this Chapter

At the end of this chapter, the reader should be familiar with the how-to of building a typical econometric model in E-views. In particular, the reader should be familiar with;

- Estimate Equations in E-views;
- Create E-views model objects;
- Create indentities;
- Do forecasts: In-Sample & Out of Sample;
- Understand the Use of Add Factors;
- Impulse responses;
- Use the model to do policy simulations.

#### 4.1 Introduction

353 Using the single equation Engle Granger approach described in the previous chapter (paragraph 220, page 67), we will show how to build a simple macroeconomic model in E-views.

354

#### Why E-views?

There is variety of software packages that can be used in developing macro economic models. In this case, we have chosen E-views as it is generally the preferred software for this type of exercise. It is also the software used in the MEFMI macroeconomic modeling course and thus, it is natural to extend its use to this manual.

Other types of economic software packages are Matlab, GAMS, etc.

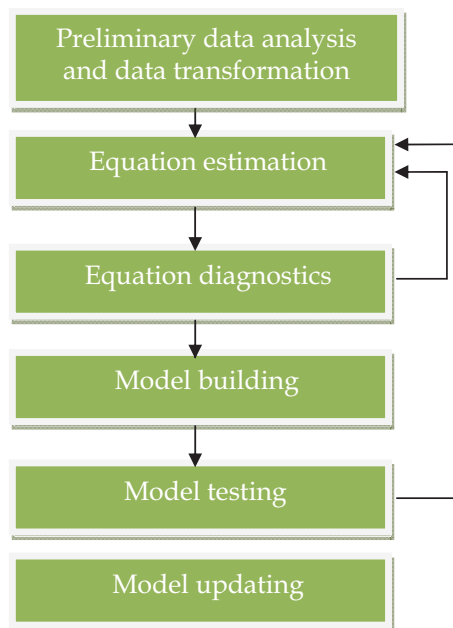
### Why macro model based on two-step Engle-Granger equations?

There are different techniques and types of models available to economists. Presenting all of these in this manual will create the risk of overwhelming and confusing the reader, while at the same time not addressing the prime objective of this manual, which is to create a simple step-by-step guide to macro model building. The selected technique is relatively simple and easily implementable in member countries based on the availability of data and capacity. The VAR, VECM techniques presented in the previous section and the ARIMA methodology presented in the next section aim to complement this macro model.

## 4.2 A Simple Model of a Hypothetical Economy in E-views Platform

355 The critical steps involved in building a model are illustrated in the figure below and explained thereafter;

Figure 4.1: Critical Steps in Model Building



356 After the identification of the theory and the variables to use for each equation in the model, and after transferring the data from excel to E-views, the first step would be

to conduct preliminary data analysis and data transformations. Preliminary data analysis involves plotting the data to see whether there could be some abnormal data points that may need closer scrutiny. Such abnormal data points are called outliers. Outliers can result from mistakes in data entry or just shocks to the variables during a particular period. Data transformation involves getting the logs of variables where necessary, growth rates and ratios.

357 The next step is estimation of the equations. After ensuring that the data is clean (has no mistakes in the entry and has been transformed, the modeler starts the equation estimation process. For time series data that you will most likely be using, the first step is to test for unit roots. In the preceding step, you have plotted the variables and you already have an idea from the graphs whether the variables are stationary or not. But casual inspection is not enough; you must formally test for unit roots. If you find that the data is non-stationary, you test for cointegration. Existence of cointegration (from the Granger Representation Theorem) necessitates the use of the error correction model.

358 After estimating the equation that you feel best represents the reality in your economy, you need to conduct model diagnostics, testing whether there are any violations of OLS assumptions including the existence or not, of heteroskedasticity, autocorrelation, normality or errors and multi-collinearity.

359 After this, the modeler brings all the equations into the model one at a time. It is advisable to bring in one equation, solve it in-sample and out-of-sample, use it for simulating some policy and test its stability by generating the impulse response functions before you bring in the next equation into the model.

360 Finally, all the variables are brought into the model and the overall model stability is tested. The model can be updated as new information comes and new solutions derived.

### 4.3 Bringing the Estimated Coefficients into the Model

361 After you have read the data and done the preliminary data analysis, you can start estimating your equations. This was illustrated in Chapter 4. Here, we will move directly to equation estimation. **Remember you always have to do the preliminary data analysis!**

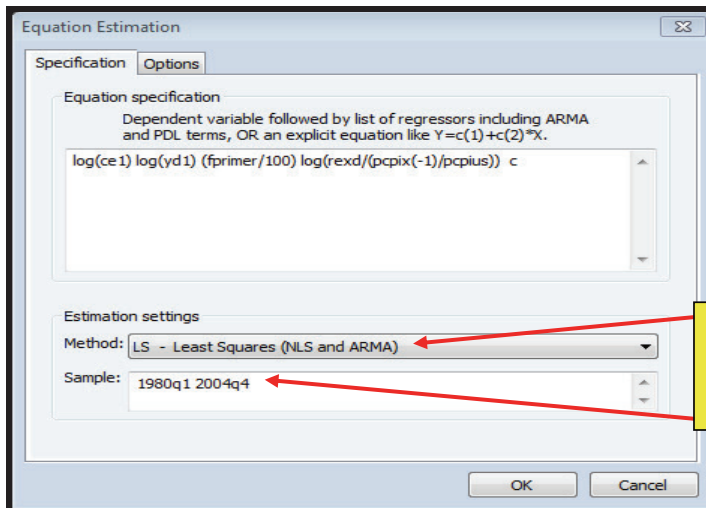
362 We will estimate the equation in two steps; first estimate the long run equation and generate the residuals from there and then use the residuals to the error correction model. The long run consumption equation is given in the box below:



$\log(\text{ce1}) \log(\text{yd1}) (\text{fprimer}/100) \log(\text{rexd}/(\text{pcpix}(-1)/\text{pcpius})) \text{ c}$   
 where ce1-real household  
 consumption yd1-real disposable  
 income frprimer-the real interest  
 rate rexd-the rand-dollar  
 exchange rate pcpix-cpi  
 inflation pcpius-the US  
 inflation rate

Remember that the long  
run equation has no  
differenced variables

363 We had illustrated earlier how to estimate an equation in E-views. To run the estimation, just type the commands given in the box above into the equation window. When you are in the "Model" work-file, go to "Quick", and then to 'Equation Estimation". The data is given in the E-views work-file in the accompanying CD. The work-file is named "E-views model". The long run equation is named "aaeq\_ce1". You could run the same equation on your own and see the kind of results that you get. The output from the above estimation is given below;



Remember to specify the  
correct method and time  
period. In this case 1980:1  
2004:4

364 The output from this equation is given by;

The screenshot shows the EViews software interface with the following data:

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(YD1)	1.056421	0.025817	40.91918	0.0000
FPRIMER/100	0.174481	0.000000	-2.484227	0.0155
LOG(REXD/(PCPIK(-1)*PCPIUS))	0.002729	0.017662	0.154320	0.8777
C	-0.801917	0.283844	-2.825203	0.0057

Other statistics shown in the screenshot include:

- R-squared: 0.975045
- Adjusted R-squared: 0.974205
- S.E. of regression: 0.032267
- Sum squared resid: 0.099949
- Log likelihood: 203.5196
- F-statistic: 1250.301
- Prob(F-statistic): 0.000000

A yellow callout box with a red arrow pointing to the coefficient of FPRIMER/100 contains the text: "Do these coefficients make economic sense? Are they in line with theory?"

365 From the long-run equation we generate the residuals. When you are in the equation window, go to "proc", "make residual series" then click 'OK'.

The 'Make Residuals' dialog box is shown with the following settings:

- Residual type:  Ordinary,  Standardized,  Generalized
- Name for resid series: resid01

The 'OK' button is highlighted, and a red arrow points from the 'Name for resid series' field to the 'OK' button.

Remember to name the residuals and click OK

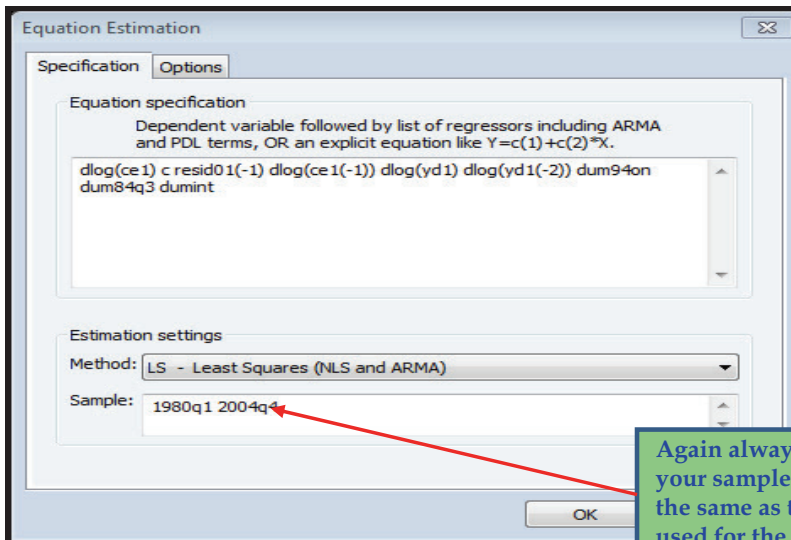
366 After this, we estimate the short run equation with the residuals lagged once (this is the error correction model). The first lag of the residuals gives the speed of adjustment or the error correction term. Again, go to “Quick”, “Estimate Equation” and type the following;

```
dlog(ce1) c resid01(-1) dlog(ce1(-1)) dlog(yd1) dlog(yd1(-2)) dum94on
dum84q3 dumint
```

The variables are in their first differences since they are non-stationary at levels.

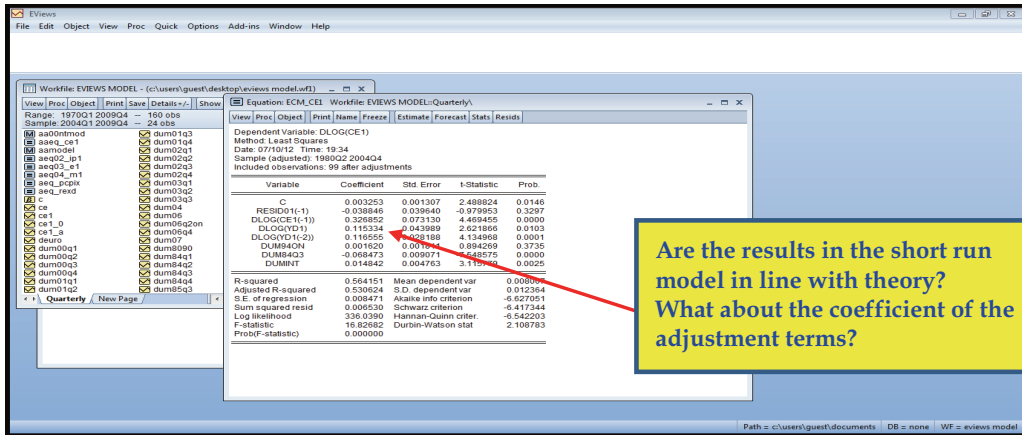
Remember that the short run model has differenced variables and also includes the new variable we just generates resid01

367 When you click Ok, a prompt window comes;



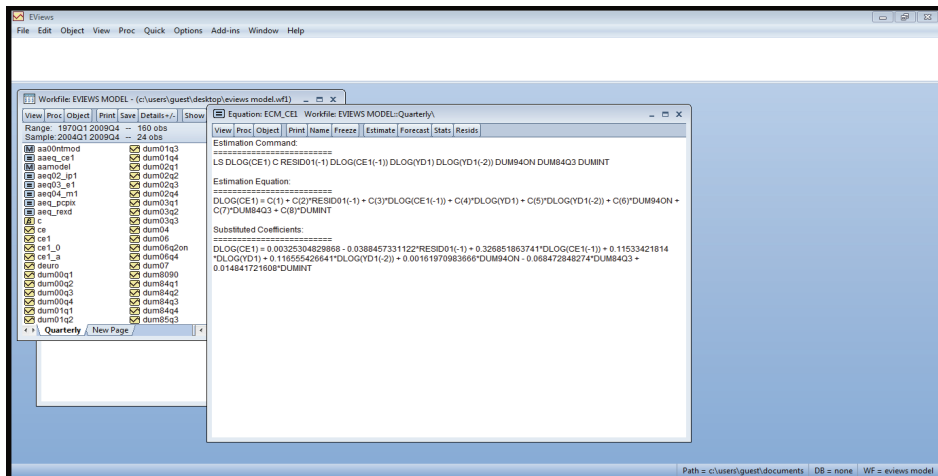
Again always ensure that you set your sample period and it must be the same as the sample period you used for the long run model

368 These are the results from the short run model.



369 Please note that we have not made any efforts here to ensure that the long run or the short-run is in line with theory, but that is something one must make sure they do before proceeding. Always ensure that your model is in line with theory in terms of the sign and size of the coefficients.

370 From this estimation, you will now need to generate the representations.



371 You can click on the "Name" in the above window and name your short run equation as 'ecm\_ce1' or any other name you may choose.

372 The main equation in the model will be the short-run model with the error correction term, but you need the long run to help you generate the long-term errors to be used for contrasting the error correction term.

373 Once you are convinced that your equation is reasonably good you have to do the post estimation test illustrated in Chapter 4. Based on the outcomes from these tests you may have to re-estimate your equation.

Remember that your equation needs to satisfy three criterion:

1. Theoretical fundamentals
2. Statistical properties
3. Work well with other equations in the model.

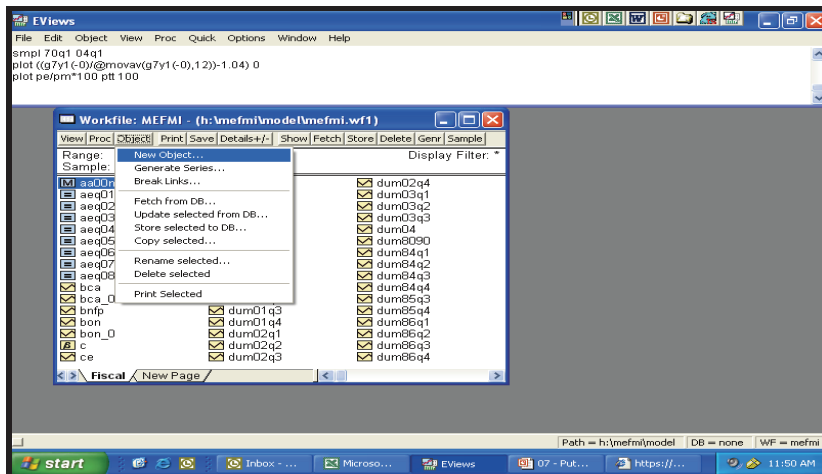
You will not be able to always satisfy all three conditions and hence, there will be a need for some trade offs and flexibility. This is part of the art of model building.

374 We assume that our equation is reasonably good and proceed to the next step.

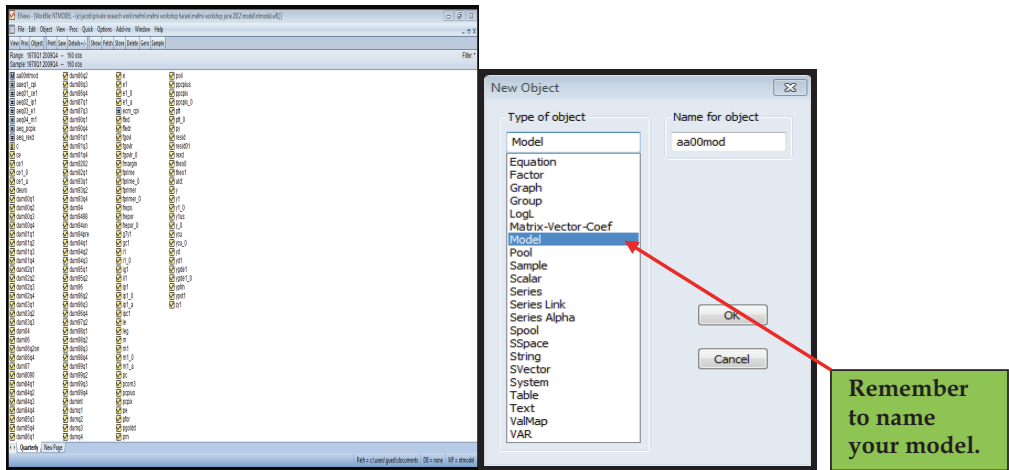
## 4.4 Model Building

### 4.4.1 Creating the Model Object

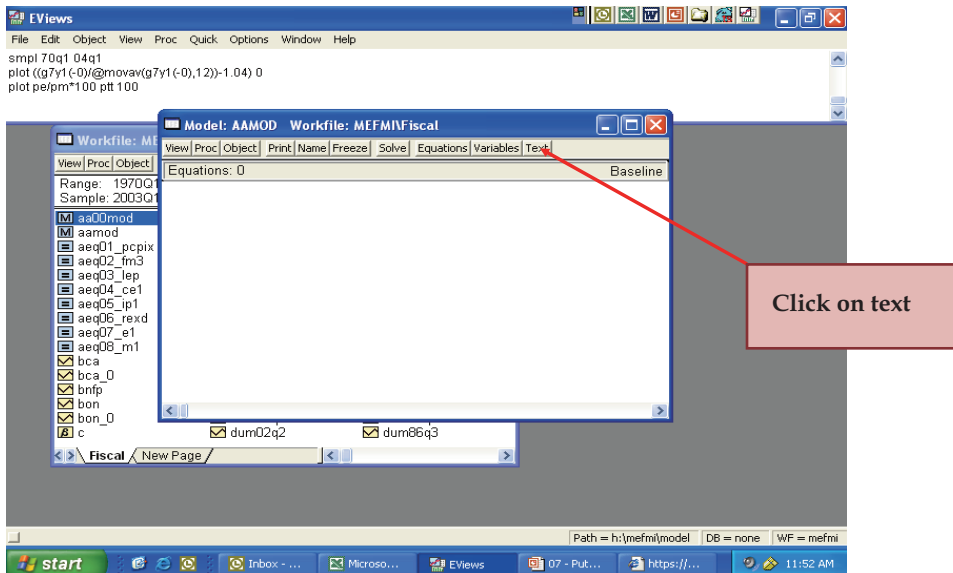
375 Once you have an estimated equation and you are happy with its properties you can proceed to putting into a model object. This is as indicated below. To do that, go to Objects, then to New Object and select "Model



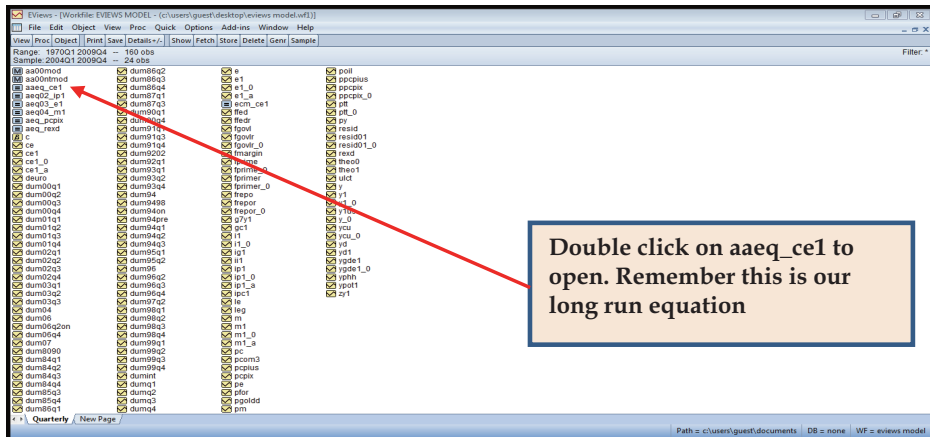
376 This will prompt another window asking you to give a name to the Model and then click OK. We name our model as *aa00mod*. We start with *aa* so that the model can appear on top of the other items in the work-file to make it easy for us to get it when we need it. E-views lists items in the work-file in an alphabetical order.



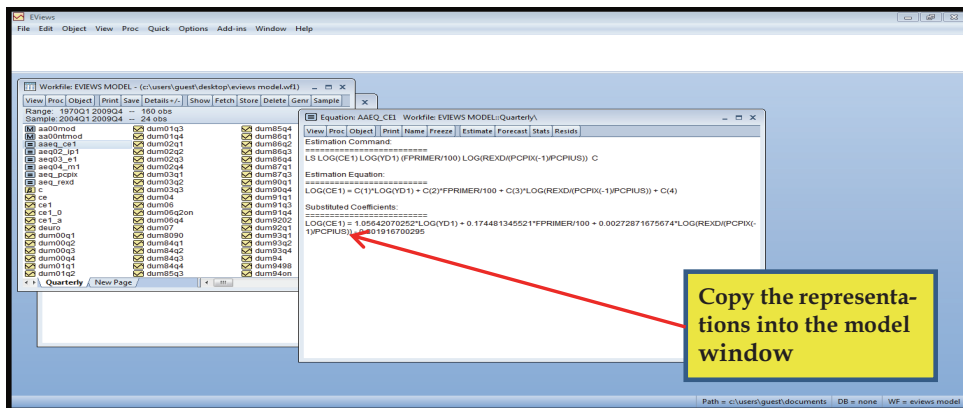
377 You will need to click on text to be able to paste text in the clear space in the model window.



378 After that you can now copy equation into work-file. This is done by first double clicking the estimated equations to open.



379 When the equation window is open, go to View, then Representations and copy the substituted coefficients to paste in the model window.



380 Before you paste into the model window, remember that we want to generate the residuals using the long-run coefficients. Remember that;

$$\text{LOG}(CE1) = C(1)*\text{LOG}(YD1) + C(2)*\text{FPRIMER}/100 + C(3)*\text{LOG}(\text{REXD}/(\text{PCPIX}(-1)/\text{PCPIUS})) + C(4) + \text{resid01}$$

381 This implies that;  
 $\text{resid01} = \text{LOG}(CE1) - (C(1)*\text{LOG}(YD1) + C(2)*\text{FPRIMER}/100 + C(3)*\text{LOG}(\text{REXD}/(\text{PCPIX}(-1)/\text{PCPIUS})) + C(4))$

382 The error term is therefore alternatively given by everything in the right hand side as above.

383 The long-run equation with the substituted coefficients is given as;

$$\text{LOG}(\text{CE1}) = 1.05642070252 * \text{LOG}(\text{YD1}) + 0.174481345521 * \text{FPRIMER} / 100 + 0.00272871675674 * \text{LOG}(\text{REXD} / (\text{PCPIX}(-1) / \text{PCPIUS})) - 0.801916700295$$

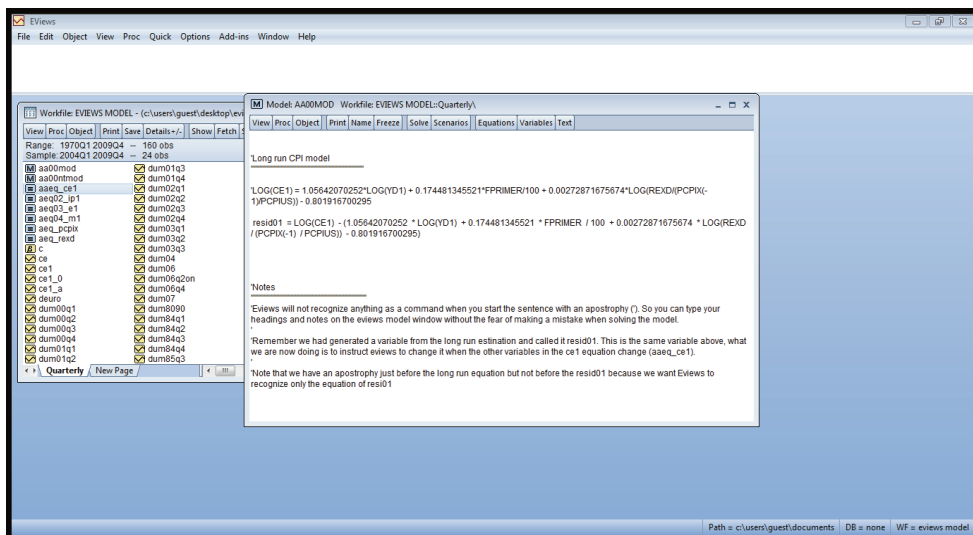
384 The resid01 that we generated earlier is therefore given as;

$$\text{resid01} = \text{LOG}(\text{CE1}) - (1.05642070252 * \text{LOG}(\text{YD1}) + 0.174481345521 * \text{FPRIMER} / 100 + 0.00272871675674 * \text{LOG}(\text{REXD} / (\text{PCPIX}(-1) / \text{PCPIUS})) - 0.801916700295)$$

385 Remember we had generated a variable from the long run estimation and called it resid01. This is the same variable above. What we are now doing is to instruct E-views to change it when the other variables in the ce1 equation change.

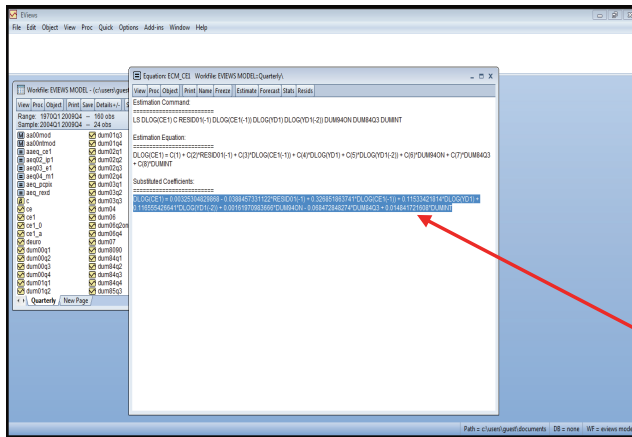
386 This is the first variable we take to the model. What we do is to copy the long run equation into the model and write resid01 on the left of everything, write an equal sign (=) just after that and change the equal sign after log(CE1) to a minus sign(-) and put a bracket after the minus sign and another bracket at the end of the line (after 0.801916700295)

387 To paste, you can right click on the model window (after clicking on text) and a prompt will appear enabling you to paste, click on paste and your equation will appear in the model window;



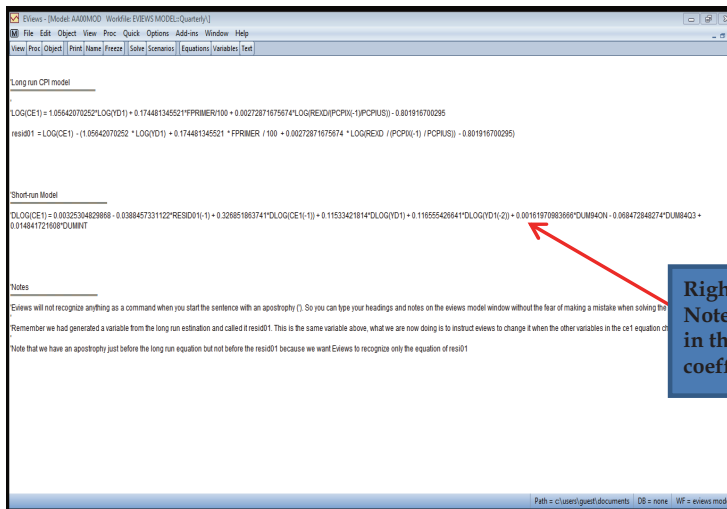
388 The second step involves bringing the short-run equation into the model. We follow the same steps by double-clicking the ecm\_ce1 equation from the work-file in the same way we did for the long run model. When the equation window opens, we go to View, Representations and copy the equation with the substituted coefficients as;





Highlight, right click on the highlighted text and then copy

389 We paste this into the model window;



Right click and paste here. Note that we have the resid01 in the short run model. Its coefficient is the speed of

390 Before proceeding you need to identify the residuals as an identity. You do this by writing @identity just before the resid01 equation. What you get will look like what is in the visual aid below after the next paragraph.

391 Secondly, you want to get the values of ce1 from the short-run equation but on the right hand side you have the first difference of ce1 (dce1). Remember that  $dlog(ce1) = dlog(ce1) - log(ce1(-1))$ . What this means is that  $log(ce1) = dlog(ce1) + log(ce1(-1))$ . To remove the logs we have  $ce1 = exp(dlog(ce1) + log(ce1(-1)))$ . But we know and can substitute the estimated coefficients of  $dlog(ce1)$  from the short run model into the equation above. So what we have is;

$$CE1 = \exp(0.00325304829868 - 0.0388457331122*RESID01(-1) + 0.326851863741*DLOG(CE1(-1)) + 0.11533421814*DLOG(YD1) + 0.116555426641*DLOG(YD1(-2)) + 0.00161970983666*DUM94ON - 0.068472848274*DUM84Q3 + 0.014841721608*DUMINT + \log(ce1(-1)))$$

392 This is the equation we will have active in the model. So the model window will look like this;

The logs are removed with the exp command.  
The dlog(ce1) equation is now disabled  
using the apostrophy at the beginning

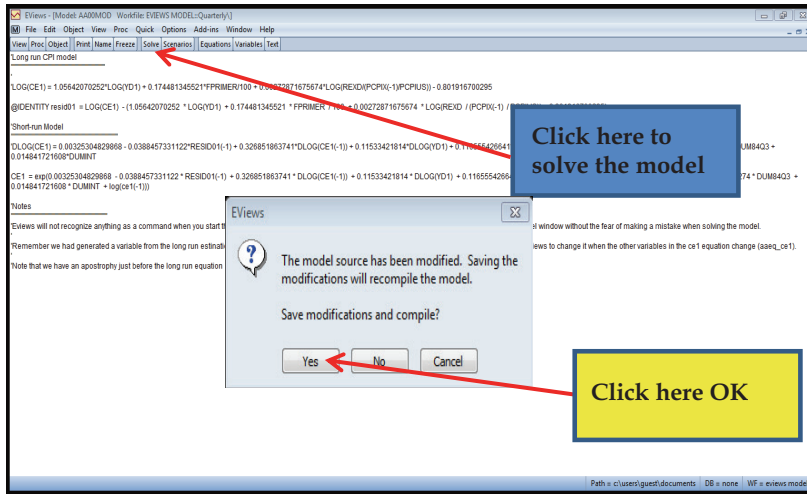
@identity is added here

393 Now you have one complete equation in the model.

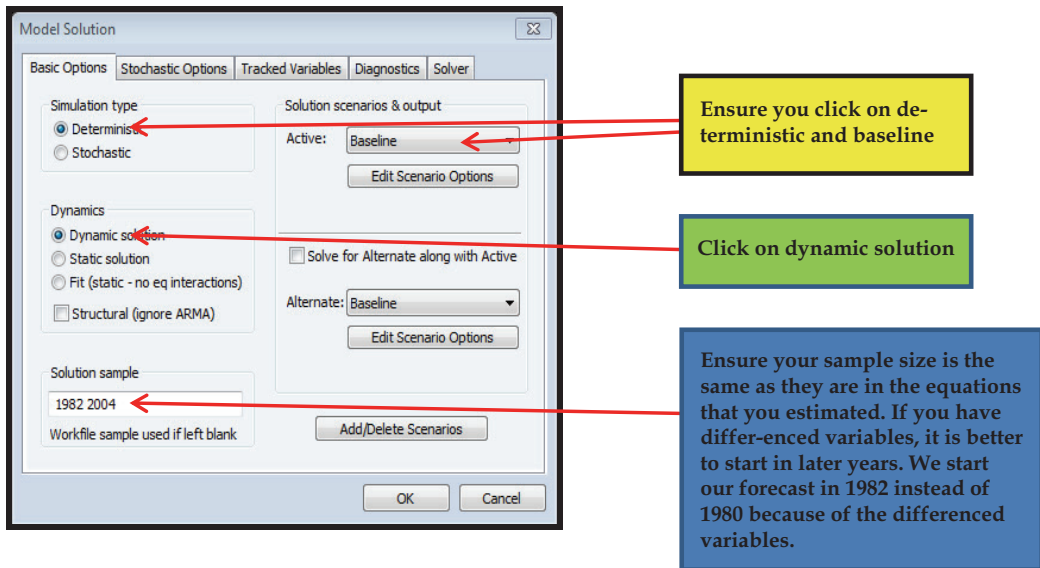
#### 4.4.2 In-Sample Forecasting with a Single Equation

394 You can run first forecast in-sample with this one equation and see how best your forecasts track the actuals within the sample (in-sample).

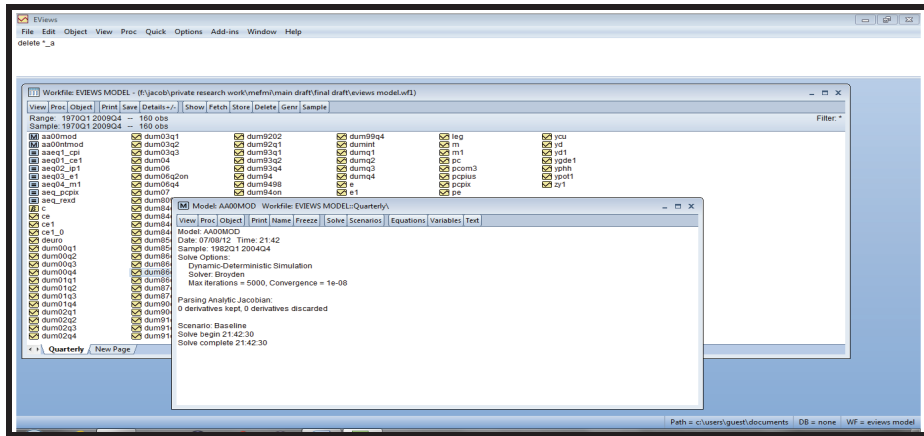
395 When you are in your model window, click solve as shown below. E-views will inform you that the model source has been modified and whether you want to save the modifications and recompile the model. You click OK.



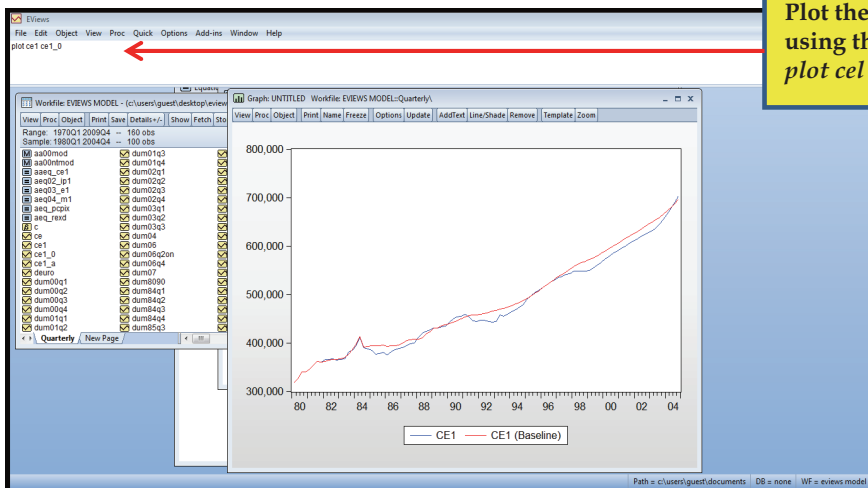
396 A prompt then appears asking you to specify the simulation type, the dynamics, solution scenario and the sample period. The simulation type would be deterministic with a dynamic solution and a baseline solution scenario. After specifying these, click “OK”.



397 If the model solves without errors, then this is the window you will see.



398 You can now check how best your forecasts track your actuals in-sample. This is the in-sample forecast. To do that, go to your command window and type `plot ce1 ce1_0`. E-views will have generated a variable which ends with `_0`. In this case, you will see a new variable named `ce1_0` in your work-file. Compare this with the actuals. To compare, hold the control (CTRL) key on the keyboard, and then click on `ce1` and then click on `ce1_0` while still holding the CTRL key. Put the cursor on the highlighted variables and right click, then open as a group. After the sheet has opened, you can view the variables in a graph form by clicking on View, graphs. Click on the single graphs option and you get;



399 As you can see the fitted values of `ce1` track the actuals very well until about 1985 when the deviations start to widen.

400 We can improve the performance of our in-sample forecasts by ensuring that our model is well specified and all econometric problems are sorted.

#### 4.4.3 Adding the Add-Factors

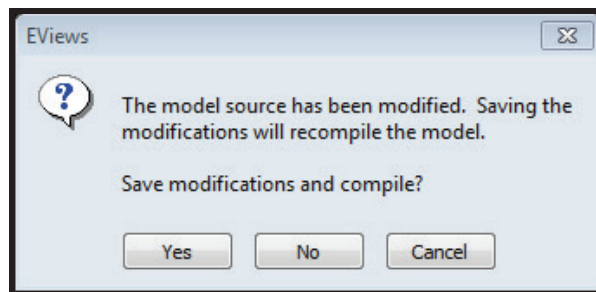
401 The next thing is to forecast out of sample (beyond the period of the current data into the future)

402 Forecasting out of sample, however, requires that we start our forecasts at the latest (current) value of the actuals. But you can see from the graph above that our forecasts are different from the actuals at the latest date of data (2004). What we therefore need is that our actuals be the same as our forecasts at the beginning of out-of-sample forecast.

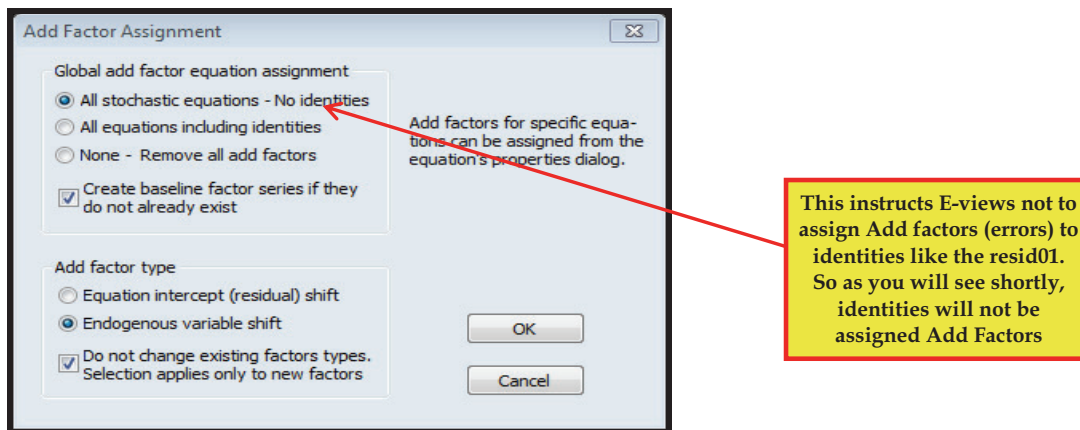
403 For our forecasts to be the same as the actuals at the current date, we must 'return' the errors (residual terms) back to the fitted line so that it can start forecasts from the actuals. This is done by adding the add factors.

404 To add the Add-factors, go to Proc from the Model window, then Click Add Factors then Equation Assignment;

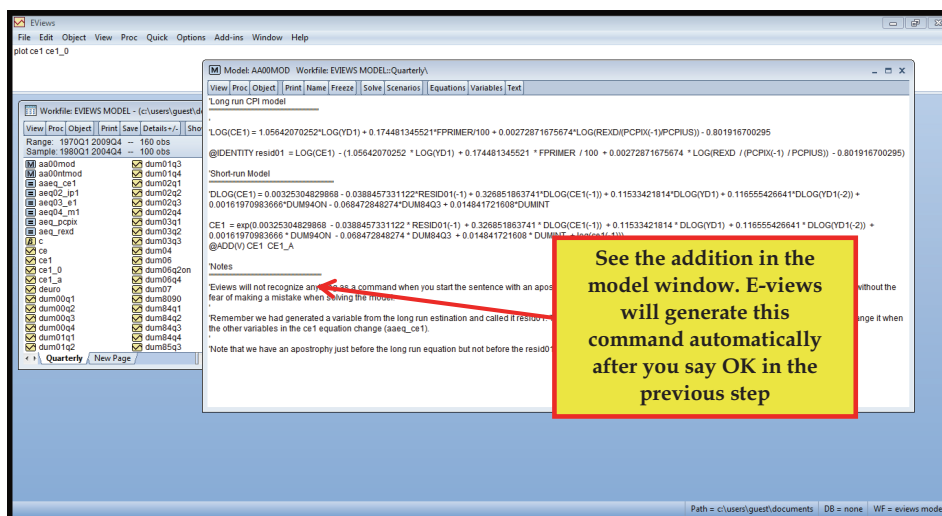
405 A prompt window will appear requiring you to save the modifications and compile. You click OK to save the modifications



406 Another prompt then appears immediately asking you to choose the Add-Factor type and Add-Factor Assignment. In the assignment you assign "All stochastic equation - No identities". This is the first option on top and then choose endogenous variable shift as the type of Add-Factor.



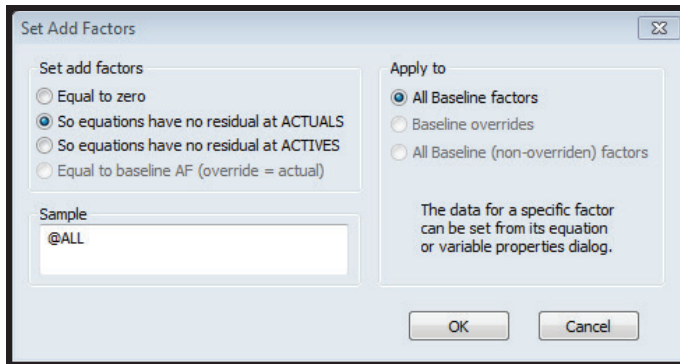
407 After you click OK, what you get would look like this;



408 Notice the additional commands @ADD(V) CE1 CE1\_A highlighted above.

409 After assignment of the Add-Factors, you can then calculate the values of the Add-Factors.

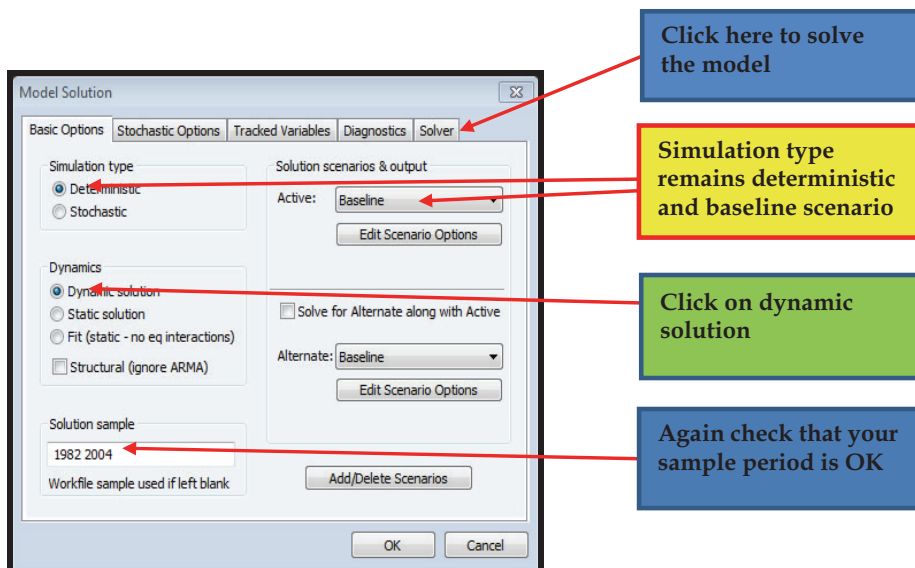
410 To set the values of the Add-factors from E-views, go to Procs from the Model window,, then Add-Factors and then Set Values.



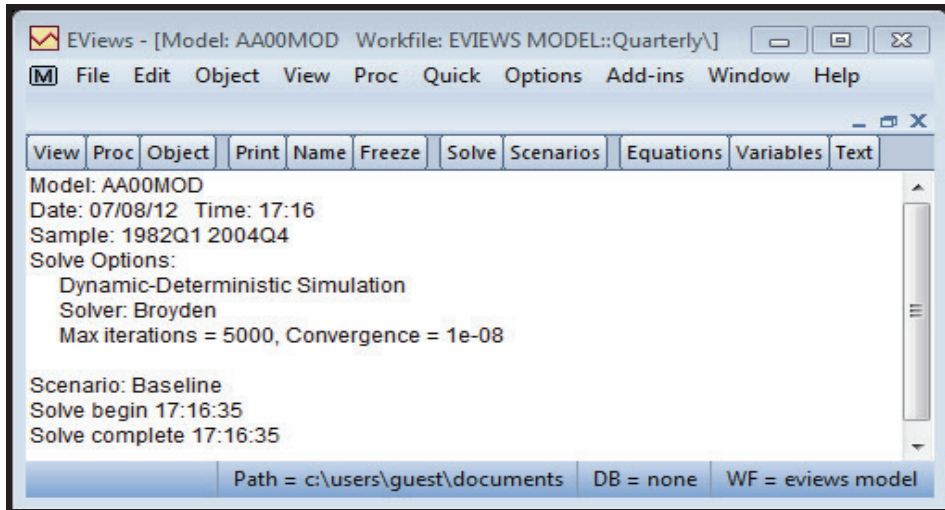
411 You need to set the Add-factors so that equations have no residuals at actuals and then click OK.

412 You are now ready to solve the model again in-sample with the Add-Factors just to be sure that the residuals have been added back to the fitted values and so the forecast in-sample are the same as the actuals in-sample. To solve click "Solve" from the model window.

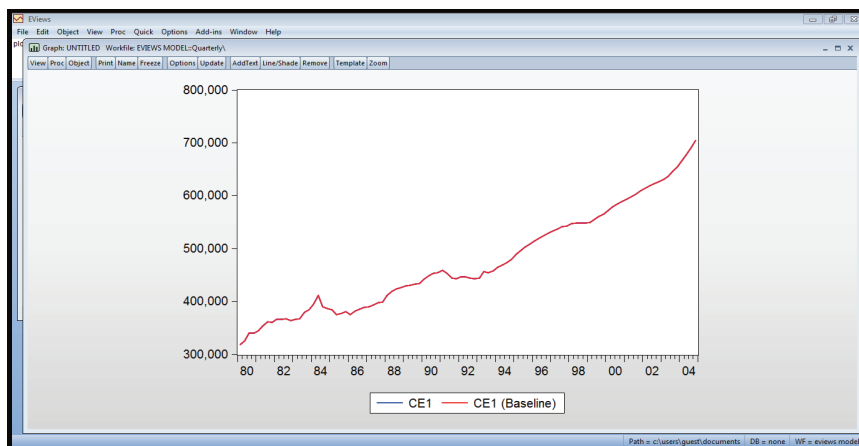
413 Just like before, a prompt appears asking you to specify the simulation type, the dynamics, solution scenario and the sample period. We use the same settings as before. Our intention is to overwrite the baseline. Remember, what we wanted by adding the Add-Factors is to make our fitted the same as the actuals. The simulation type would therefore remain deterministic with a dynamic solution and a baseline solution scenario. After specifying these, click "OK"



414 If the model solves without errors, then this is the window you will see.



415 You can now check that the actuals and the forecast values are the same in-sample. To do that, go back to your work-file. E-views will have generated a variable which ends with `_0`. This is the in-sample forecast. Compare this with the actuals. Hold the control (CTRL) key on the key board, and then click on `ce1` and then click on `ce_0` while still holding the CTRL key. Put the cursor on the highlighted variables and right click, then open as a group. After the sheet has opened, you can also view the variables in a graph form by clicking on view, graphs. Click on the single graphs option and you get;



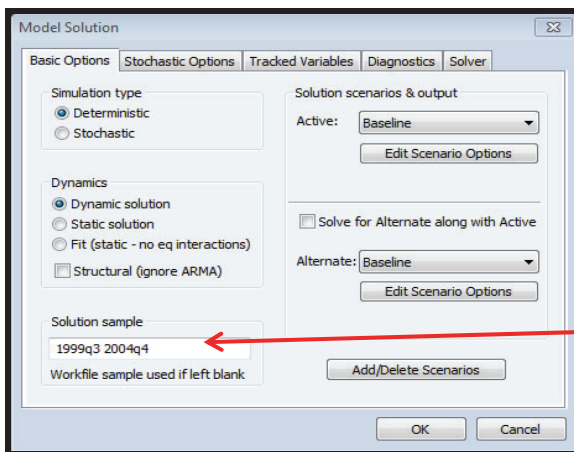
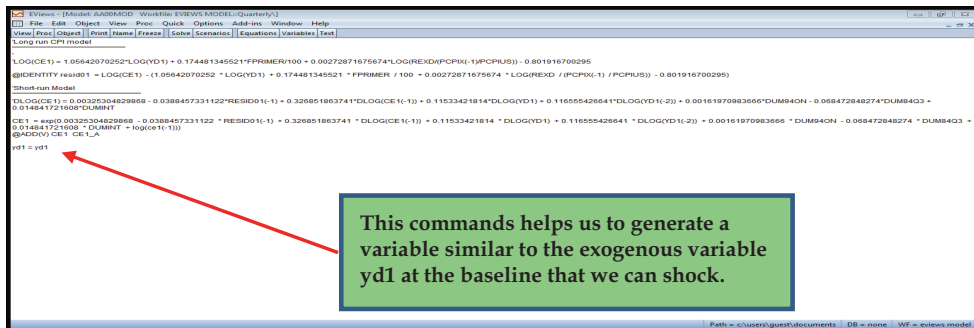
416 We now see that our actuals and the baseline are exactly the same.



#### 4.4.4 Impulse Responses

417 Before out-of-sample forecasting however, it is also important to generate the impulse responses of your equation to ensure that they converge to the long-run. The impulse responses reflect the adjustment process in the dependent variable in response to a shock in one of the exogenous long-term variables.

418 For example, in your model prompt, type the command  $yd1=yd1$  (as indicated below) and solve under baseline. You period should be the last 20 observations.



419 What you should notice is that EViews has created a variable called  $yd1_0$ , which is equal to the actual value  $yd1$ . Now let's shock this baseline variable by typing  $yd1 = yd1_0 * 1.01$

Model: AA00MOD Workfile: EViews MODEL=Quarterly\

Long run CPI model

LOG(CE1) = 1.05642070252\*LOG(YD1) + 0.174481345521\*FPRIMER/100 + 0.00272871675674\*LOG(REXD/(PCPI\*(1/PCPIUS))) - 0.801916700295

@IDENTITY resid01 = LOG(CE1) - (1.05642070252 \* LOG(YD1) + 0.174481345521 \* FPRIMER / 100 + 0.00272871675674 \* LOG(REXD / (PCPI\*(1) / PCPIUS))) - 0.801916700295

Short-run Model

DLOG(CE1) = 0.00325304829868 - 0.0388457331122\*RESID01(-1) + 0.326851863741\*DLOG(CE1(-1)) + 0.11533421814\*DLOG(YD1) + 0.116555426641\*DLOG(YD1(-2)) + 0.00161970983666\*DUM94ON - 0.068472848274\*DUM84Q3 + 0.014841721608 \*DUMMIT

CE1 = exp(0.00325304829868 - 0.0388457331122 \* RESID01(-1) + 0.326851863741 \* DLOG(CE1(-1)) + 0.11533421814 \* DLOG(YD1) + 0.116555426641 \* DLOG(YD1(-2)) + 0.00161970983666 \* DUM94ON - 0.068472848274 \* DUM84Q3 + 0.014841721608 \* DUMMIT + log(ce1(-1)))

@ADD(V) CE1 CE1\_A

yd1 = yd1\_0\*1.01

We increase disposable income by 1 per cent in the baseline. Note that we are using yd1\_0.

420 Then click Solve in the model window. E-views will ask you whether you want to accept the changes and you will say OK. Then click solve and a new prompt will come up. You will keep everything the same as in the baseline except Solution Scenario and Output Option. You will change this to Scenario 1.

Model Solution

Simulation type

Deterministic

Stochastic

Dynamics

Dynamic solution

Static solution

Fit (static - no eq interactions)

Structural (ignore ARMA)

Solution sample

1999q3 2004q4

Workfile sample used if left blank

Solution scenarios & output

Active: Scenario 1

Edit Scenario Options

Solve for Alternate along with Active

Alternate: Baseline

Edit Scenario Options

Add/Delete Scenarios

OK Cancel

Change to Scenario 1 and click OK. We want to compare how household consumption in the baseline compares to an alternative scenario where disposable income is 1 per cent higher.

Notice that the sample period remains the same as in the previous step

421 You will see that Eview has created two variables yd1\_1 and ce1\_1. Whenever you solve a model under a specific scenario, E-views creates exogenous variables with suffix reflecting the scenario. If you solved the model under scenario 2, then E-views would have generated yd1\_2 and ce1\_2.

422 In the command window type plot (ce1\_1/ce1\_0-1)\*100 (yd1\_1/yd1\_0-1)\*100

Model: A400MOD Workfile: EIEWS MODEL:Quarterly

Long run CPI model

Range: 1970Q1 2009Q4 -- 160 d  
Sample: 1999Q1 2004Q4 -- 24 obs

LOG(CE1) = 1.05642070252\*LOG(YD1) + 0.174481345521\*FPRIMER/100 + 0.00212871675674\*LOG(REXD/(PCPIX(-1)/PCPIUS)) - 0.801916700295

@IDENTITY resid01 = LOG(CE1) - (1.05642070252 \* LOG(YD1) + 0.174481345521 \* FPRIMER / 100 + 0.00212871675674 \* LOG(REXD / (PCPIX(-1) / PCPIUS))) - 0.801916700295

Short-run Model

DLOG(CE1) = 0.00325304829868 - 0.0388457331122\*RESID01(-1) + 0.326851863741\*DLOG(CE1(-1)) + 0.11533421814\*DLOG(YD1) + 0.116555426641\*DLOG(YD1(-2)) + 0.00161970983666\*DUM94ON - 0.068472848274\*DUM84Q3 + 0.014841721608 \*DUMINT

CE1 = exp(0.00325304829868 - 0.0388457331122 \* RESID01(-1) + 0.326851863741 \* DLOG(CE1(-1)) + 0.11533421814 \* DLOG(YD1) + 0.116555426641 \* DLOG(YD1(-2)) + 0.00161970983666 \* DUM94ON - 0.068472848274 \* DUM84Q3 + 0.014841721608 \* DUMINT + log(ce1(-1)))

@ADD(V) CE1 CE1\_A

yd1 = yd1\_0 \* 1.01

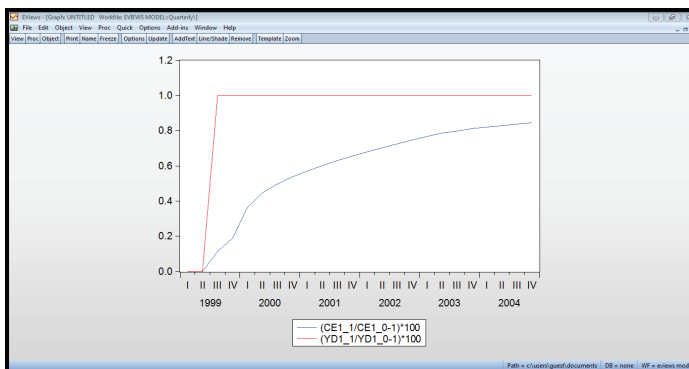
Path = c:\user1\guest\documents D6 = none WF = eviews.model

Notice two things here;

1. The commands are types in this space
2. The command `smpl 1999q1 2004q4` instructs E-views to plot the graph beginning two years before 1999q3 which was the beginning of our sample. After typing this command just press the Enter key in the keyboard

- The command in the prompt window requests that E-views plots the percentage change between consumption in the baseline and scenario 1. The same for disposable income.

423 The impulse response graph is plotted by E-views as;



- The red line shows that the shock in disposable income is correctly specified at 1% while the blue line shows the response in consumption.
- From the graph consumption increases following an increase in income by 1% converging at 0.8%.

424 It is important to remember that the sample over which the graph is plotted should start two periods before the shock and finish in the last period of the shock!

425 You can try shocking disposable income by more than 1 per cent and evaluating the impact on household consumption.

Repeat the same shock for the prime interest rate and the exchange rate. A few things you have to remember:

- Do each shock individually hence remember to switch off any previous shocks.
- To delete any previous shock type on the command window; Delete \*\_0 (delete space underscore zero). This deletes the baselines. If you type delete \*\_1, this deletes all the scenario 1s and so on.
- When you shock rates by 1 per cent, you do NOT multiply the baseline rate by 1.01 but rather you add 1 i.e +1.

426 The same methodology can be used to study policy questions asked by the Minister of Finance or the Governor of the central Bank. For example:

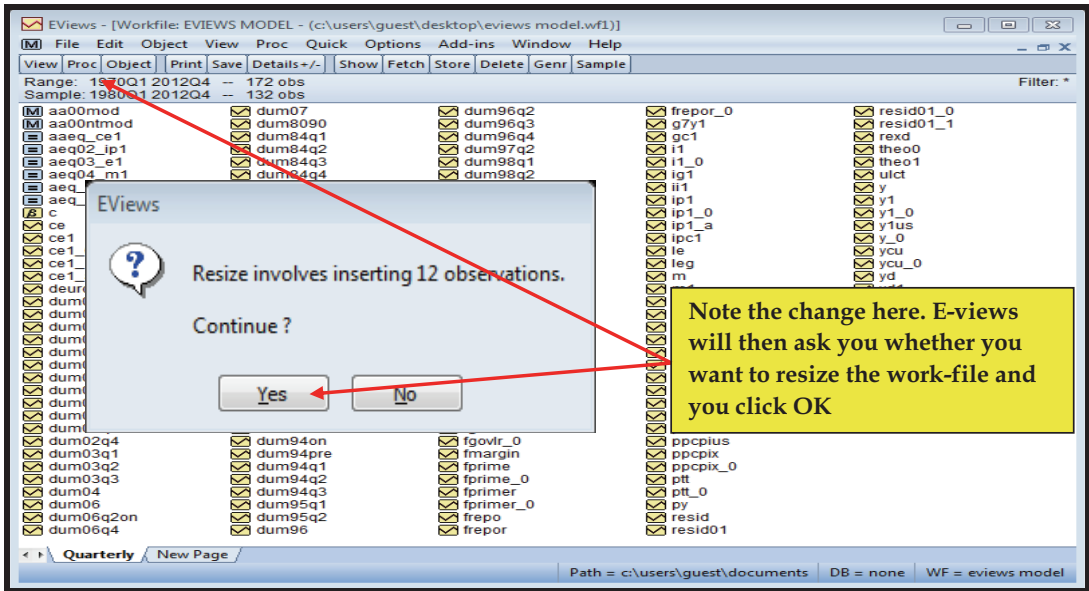
- a. By how much does consumption rises if disposable income is increased by X per cent?
- b. What is the impact of interest rates on household consumption?
- c. The exchange rate has depreciated by X per cent, what should we expect to be the impact on consumption?

427 You can also use this simple one equation model to forecast. We will illustrate how you do this with the single equation model and extend it to a model with multiple equations.

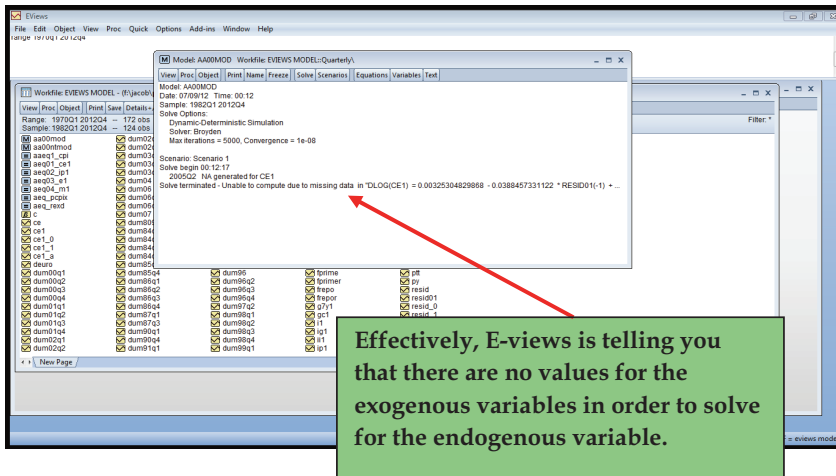
#### 4.4.5 Making Assumptions

428 In order to forecast the dependent variable, you need to provide assumptions for your exogenous variables and ensure that the range and sample of your Work-file include the future period you want to forecast.

429 For example, the last data point in our illustrative example is 2009q4. In order to add another 3 years to the range, we type *range 1970 2012*. Then, you also need to change the sample by typing *smpl 1980q1 2012q4*. This is the period we used for our in-sample forecast. If you don't know what the difference is between the sample and the range consult the E-views Guide.



430 Go back to your model and click the Solve button. Under Solution sample extend the period to 2012 and click OK. E-views will generate the following message, indicating that it is unable to solve due to missing data.



431 Type in the command window the following command `show yd1 fprimer rexd pcpix pcpius dum94on dum84q3 dumint`. We are asking E-views to give us a list of all the exogenous variables. You will be presented with something as the screen below. Remember that the screen will be equation and data specific.

The screenshot shows the EViews software interface with a data table. A yellow callout box with a red arrow points to the 'edit' button in the table header. The table contains columns for 'Year', 'Y1', 'Y2', 'Y3', 'Y4', 'Y5', 'Y6', 'Y7', 'Y8', 'Y9', 'Y10', 'Y11', 'Y12', 'Y13', 'Y14', 'Y15', 'Y16', 'Y17', 'Y18', 'Y19', 'Y20', 'Y21', 'Y22', 'Y23', 'Y24', 'Y25', 'Y26', 'Y27', 'Y28', 'Y29', 'Y30', 'Y31', 'Y32', 'Y33', 'Y34', 'Y35', 'Y36', 'Y37', 'Y38', 'Y39', 'Y40', 'Y41', 'Y42', 'Y43', 'Y44', 'Y45', 'Y46', 'Y47', 'Y48', 'Y49', 'Y50'. The rows represent different variables, with some cells containing numerical values and others containing 'NA' (Not Available).

432 You can enter values in the sheet by clicking on the *edit* button. The values that you enter are called assumptions and should generally represent the views of your institution i.e they should be agreed as part of a consultation process.

433 Alternatively, we can highlight the exogenous variables above, copy and paste in a new excel file and make assumptions from excel and read the data back into E-views. It may be easier to enter the assumptions in excel because we can easily use formulas in excel and then transfer the complete data to E-views. In our case, we use this approach.

434 We make the following assumptions (remember we have copied the data into excel. We copied all the data from 1982 so we will read all the data from 1982 to 2012 into E-views);

- Let us assume that  $yd1$  will increase by 2% every quarter. We therefore get the value of  $yd1$  for 2005q1 as  $2004q1 \times 1.02$
- We assume that  $fprimer$  will increase by 1%. We get the value for 2005q1 onwards as above
- We assume that  $cpus$  will increase by 1%
- We assume that  $cpix$  will increase by 4%
- We assume that  $tharedx$  will increase by 6%
- The dummy variables remain as before

435 We copy the formulas downwards from 2005q1 to 2012q4 to get the excel work-file below;

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
97	2006Q4	782286.8	7.481154	764.8801	153.6414	198.4779	1	0	0												
98	2006Q1	778551.5	7.684989	810.7729	159.7871	200.4627	1	0	0												
99	2006Q2	784122.5	7.915539	859.4193	160.1788	202.4673	1	0	0												
100	2006Q3	810095	8.153095	910.9844	172.8257	204.4919	1	0	0												
101	2006Q4	826205.1	8.397595	965.6435	179.7387	206.5369	1	0	0												
102	2007Q1	842729.2	8.649523	1023.582	186.9283	208.6022	1	0	0												
103	2007Q2	859583.7	8.909089	1084.597	184.4054	210.6893	1	0	0												
104	2007Q3	876775.4	9.176279	1150.097	202.1816	212.7951	1	0	0												
105	2007Q4	884310.9	9.451567	1219.103	210.2689	214.9231	1	0	0												
106	2008Q1	912187.1	9.725114	1292.249	218.6787	217.9723	1	0	0												
107	2008Q2	930441.1	10.02717	1369.784	227.4268	219.243	1	0	0												
108	2008Q3	949049.9	10.32798	1451.971	236.5239	221.4355	1	0	0												
109	2008Q4	969039.9	10.63782	1539.089	245.9649	223.6498	1	0	0												
110	2009Q1	987391.5	10.95696	1631.434	255.8243	225.8863	1	0	0												
111	2009Q2	1007139	11.28567	1729.32	266.0572	228.1452	1	0	0												
112	2009Q3	1027292	11.62404	1833.18	276.6995	230.4295	1	0	0												
113	2009Q4	1047828	11.97296	1943.064	287.7675	232.7309	1	0	0												
114	2010Q1	1068784	12.33215	2059.648	299.2762	235.0582	1	0	0												
115	2010Q2	1090169	12.70212	2183.227	311.2403	237.4088	1	0	0												
116	2010Q3	1111963	13.08318	2314.221	323.6993	239.7829	1	0	0												
117	2010Q4	1134202	13.47567	2453.074	336.6473	242.1807	1	0	0												
118	2011Q1	1156887	13.87965	2600.268	350.1132	244.6026	1	0	0												
119	2011Q2	1180024	14.29634	2756.274	364.1177	247.0486	1	0	0												
120	2011Q3	1203625	14.72523	2921.65	378.6824	249.519	1	0	0												
121	2011Q4	1227937	15.16689	3096.949	393.8207	252.1142	1	0	0												
122	2012Q1	1252251	15.622	3282.766	409.5829	254.5344	1	0	0												
123	2012Q2	1277296	16.09068	3479.732	426.9662	257.0797	1	0	0												
124	2012Q3	1302842	16.57338	3688.516	443.0049	259.6096	1	0	0												
125	2012Q4	1328899	17.07058	3909.827	460.725	262.247	1	0	0												

436 We can save the work-file in a folder where we can easily get it. We save it in drive F. It is always easy to save it in drive c if c is accessible to you. Ensure that you note the version of excel you are saving. Name the file and the sheet where the data is. In our case we save the file as Assum, the sheet where the data is we call Assumptions and we save as excel 2003-2007 version. This is going to be important when reading the data into E-views.

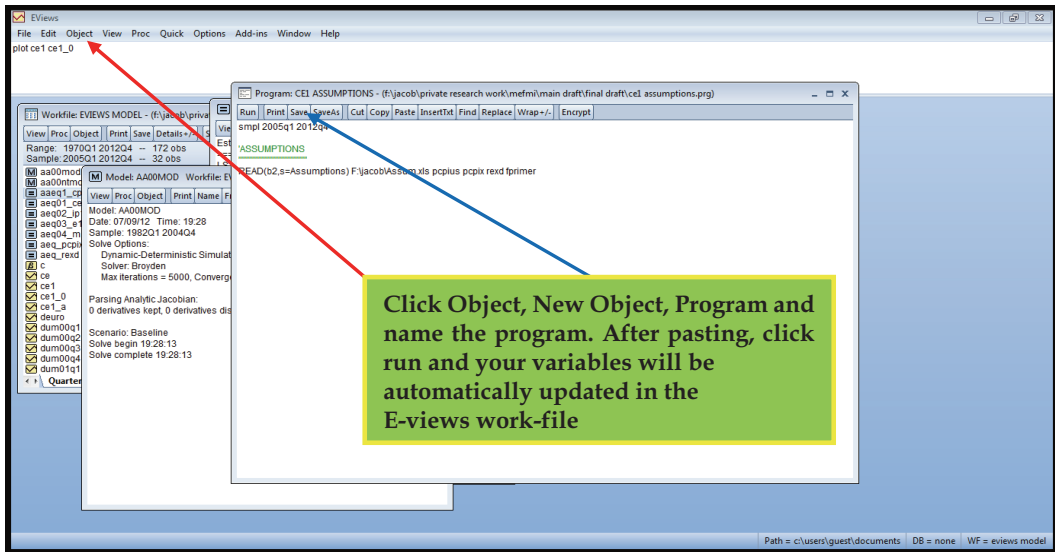
437 We are now ready to read the data into E-views.

438 To read the data into E-views, you need to create a program to read the data from excel into E-views. To do that, go to File from your command window, then to New Object, Program, and a blank prompt window will appear. You can now type your commands here instructing E-views to read your data from excel. In our case we typed;

#### ASSUMPTIONS

```
READ(b2,s=Assumptions) F:\jacob\Assum.xls yd1 pcpius pcpix rexd fprimer dumon
dum84q3 dumint dum94on
```

439 The above command instructs E-views to read the data from a folder named Jacob in the F drive. The name of the file is Assum, saved in the 2003/2007 version of excel (that is why we use .xls, otherwise we would have used .xl if the file was saved in a higher version of excel). In addition, the command instructs E-views to start reading the data from cell b2 which is 1982q1, in the sheet named Assumptions.



440 We can view the data now. In your command window, you can type *show pcpius pcpix rexd fprimer dumon dum84q3 dumint*

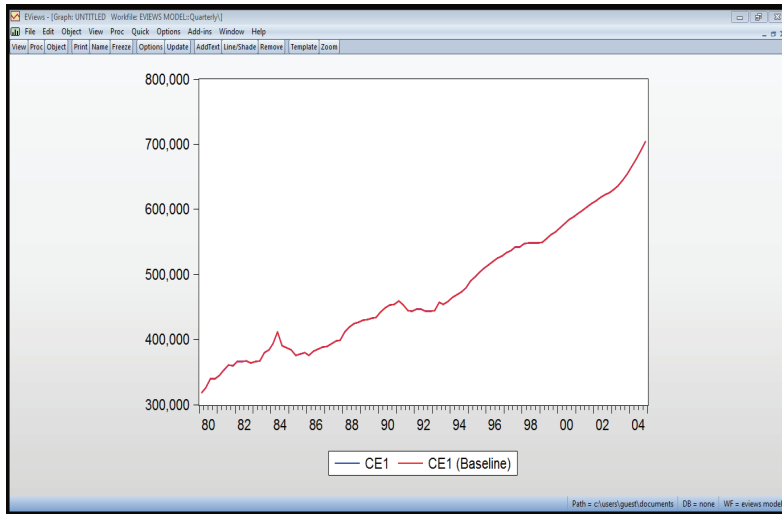
The screenshot shows the EViews software interface displaying a data table. The table has the following columns: obs, PCPIUS, PCPIX, REXD, FPRIMER, DUMON, DUM84Q3, and DUMINT. The data is organized into rows corresponding to different observations (years and quarters).

obs	PCPIUS	PCPIX	REXD	FPRIMER	DUMON	DUM84Q3	DUMINT
2003Q4	65072.245	7.204541828	675.610000	128.823333	184.6000	1.000000000	0.000000000
2004Q1	670043.644	7.043478291	677.310000	128.133333	185.2667	1.000000000	0.000000000
2004Q2	681371.992	6.894589695	659.030000	129.466666	188.9333	1.000000000	0.000000000
2004Q3	693302.578	7.119209681	637.770000	130.190000	189.6000	1.000000000	0.000000000
2004Q4	705158.061	6.620919073	605.856666	131.333333	190.7233	1.000000000	0.000000000
2005Q1	718261.223	6.828012245	642.280000	135.586666	192.4607	1.000000000	0.000000000
2005Q2	733946.447	7.028293642	680.740000	142.050133	194.5671	1.000000000	0.000000000
2005Q3	748319.376	7.243839252	721.584893	147.732139	195.5127	1.000000000	0.000000000
2005Q4	763285.784	7.681154429	784.880002	153.641424	198.4779	1.000000000	0.000000000
2006Q1	778551.479	7.684898952	810.772887	159.787081	200.4527	1.000000000	0.000000000
2006Q2	794122.508	7.915538734	859.419290	165.179564	202.4673	1.000000000	0.000000000
2006Q3	81004.498	8.353084898	819.984416	172.827027	204.4818	1.000000000	0.000000000
2006Q4	826205.058	8.397959503	865.643481	179.738735	206.5389	1.000000000	0.000000000
2007Q1	842729.159	8.649522894	1023.26209	186.926284	208.6022	1.000000000	0.000000000
2007Q2	856583.142	8.909086981	1054.49701	194.465416	210.6683	1.000000000	0.000000000
2007Q3	876775.417	9.176278838	1150.08683	202.181632	212.7951	1.000000000	0.000000000
2007Q4	894310.925	9.451687294	1219.10254	212.289898	214.9821	1.000000000	0.000000000
2008Q1	912197.144	9.735114420	1292.24880	218.679653	217.0723	1.000000000	0.000000000
2008Q2	930441.087	10.027167648	1369.78373	227.426840	219.2430	1.000000000	0.000000000
2008Q3	948049.908	10.27982078	1451.97075	236.523913	221.4355	1.000000000	0.000000000
2008Q4	968030.907	10.637822158	1539.08900	245.984870	223.6488	1.000000000	0.000000000
2009Q1	987391.625	10.956995621	1631.43434	255.824295	225.8863	1.000000000	0.000000000
2009Q2	1007138.35	11.286469251	1728.32040	266.987235	228.1452	1.000000000	0.000000000
2009Q3	1027282.14	11.624235491	1833.07862	278.699525	230.4286	1.000000000	0.000000000
2009Q4	1047827.78	11.972962550	1943.04440	291.787506	232.7309	1.000000000	0.000000000
2010Q1	1068784.34	12.321514420	2058.64827	299.278206	235.0582	1.000000000	0.000000000
2010Q2	1090160.02	12.702115975	2183.22716	311.249334	237.4088	1.000000000	0.000000000
2010Q3	1111963.22	13.083179453	2314.20779	323.699308	239.7829	1.000000000	0.000000000
2010Q4	1134202.49	13.475674838	2453.07404	336.647200	242.1807	1.000000000	0.000000000
2011Q1	1156886.54	13.873945083	2600.25848	350.113171	244.6025	1.000000000	0.000000000
2011Q2	1180024.27	14.296343430	2762.27399	364.117698	247.0488	1.000000000	0.000000000
2011Q3	1203624.75	14.725253739	2921.65043	378.682406	249.5190	1.000000000	0.000000000
2011Q4	1227697.25	15.166980791	3096.94846	393.829702	252.0142	1.000000000	0.000000000
2012Q1	1252251.19	15.622004714	3282.74543	409.982890	254.6344	1.000000000	0.000000000
2012Q2	1277296.22	16.090980488	3479.73241	425.986206	257.0797	1.000000000	0.000000000
2012Q3	1302824.14	16.573380200	3685.81938	443.048854	259.6505	1.000000000	0.000000000
2012Q4	1328898.99	17.070581712	3909.82734	460.725048	262.2470	NA	NA

441 All your exogenous variables are now updated until 2012. You are now ready to forecast out of sample.

442 Just to be sure that your actuals and fitted have not been interfered with in the process of getting the assumptions, you can plot *ce1* and *ce1\_0* (the actuals and the fitted values of *ce1*) to confirm. Remember we have to start our out-of-sample forecast from the last actual value. To do this, go to your command window and type *plot ce1 ce1\_0* and you get;

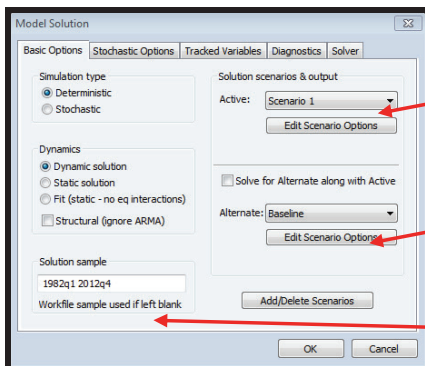




443 We are now sure that our actuals and fitted values have not been changed. We can proceed;

#### 4.4.6 Out-of-Sample Forecasting with a Single Equation

444 To forecast out of sample, go to the your model window and click solve. The procedure is the same as what we have illustrated before. What you need to do now is to change the scenario output to scenario 1. You remember you had run the models in-sample and had obtained the fitted values, which was the same as the actuals upto the last actual value which was 2004q1. We now want to forecast from 2005q1 to 2012q4. You need to ensure that there are no other variables in the work-file which ends with `_1`. You can delete them using the command `delete *_1` from your command window. After typing this command press the Enter button on the keyboard and accept to delete in the next prompt window that will appear.

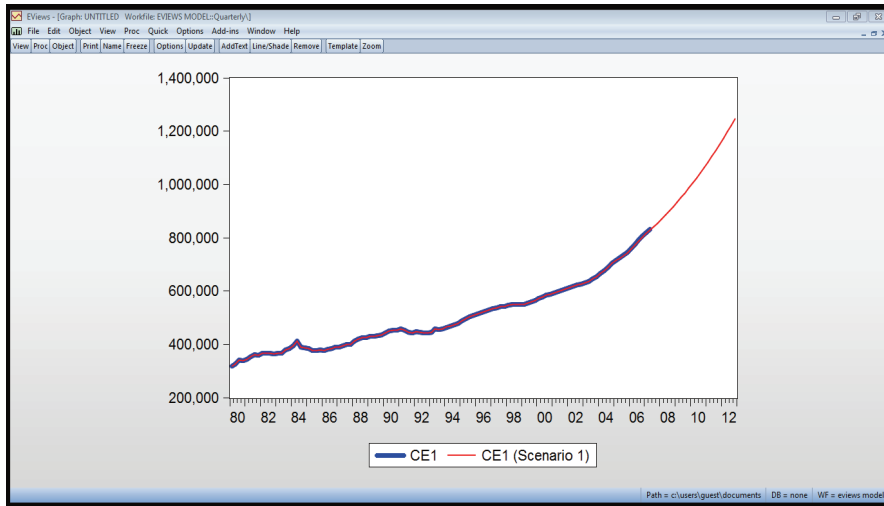


Ensure you change the scenario to 1

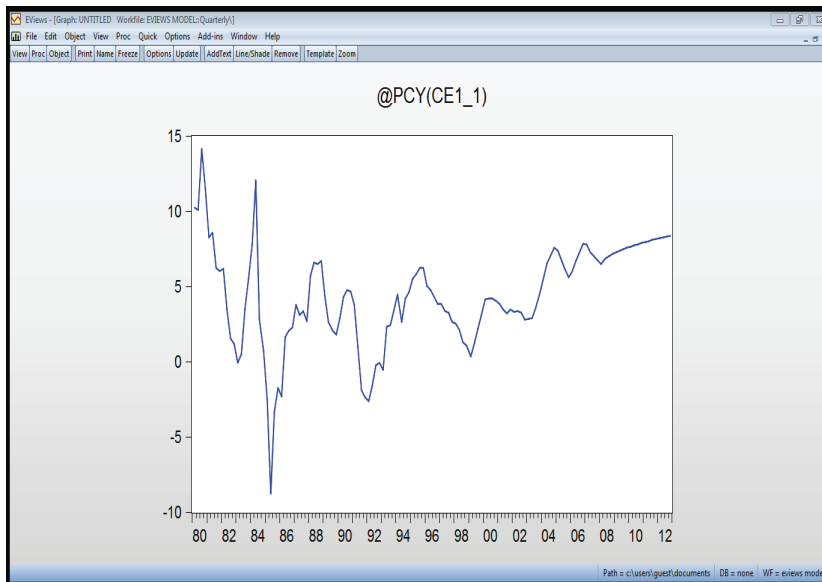
This remains as the baseline

The sample period must be changed  
1982 to 2012

445 After the model is solved successfully plot the `ce1` and `ce_1` to see how the out-of-sample forecast looks like. In the command window type `plot ce1 ce_1` and you get;



446 y-o-y % change in real household consumption. Use the command `plot @pcy(ce1_1)`. Note that we are plotting `CE1_1` and not `CE1`. What is the expected growth in consumption based on the assumptions for 2012?



447 We have gone through the process of building a one equation model and using it for policy analysis and forecasting. We will use the same methodology to create a model made of 4 equations.

448 Before we proceed, you will have to estimate on your own three additional equations following the steps outlined above and in chapter 4. Our estimated equations are given in the “E-views model” work-file in the CD. We give for each equation the model that we estimated and you can try if you can replicate our results. We must again warn that we did not go through diagnostic testing of these equations and so the results may be a little different from yours if you go through all the diagnostic tests for each equation.

- **Investment**

- Long run investment equation**

$$\log(ip1) \text{ c } \log(y1) \text{ (fgovlr(-3)/100) } \log(rexd/(pcpix/pcpius)) \text{ (ycu-100)/100}$$

- Short-run investment equation**

$$DLOG(IP1) = C(1)*DLOG(IP1(-1)) + C(2)*RESID02(-1)$$

- **Exports**

- Long run exports equation**

$$LOG(E1) = C(1)*LOG(G7Y1) + C(2) + C(3)*LOG((PE)/PY)$$

- Short-run exports equation**

$$DLOG(E1) = C(1) + C(2)*RESID03(-1) + C(3)*DLOG(G7Y1(-1)) + C(4)*LOG(PE(-2)/PY(-2)) + C(5)*DUM94Q2 + C(6)*DUM96Q3$$

- **Imports**

- Long run imports equation**

$$LOG(M1) = C(1)*LOG(YGDE1) + C(2)*LOG((PM)/PY) + C(3)$$

- Short-run imports equation**

$$DLOG(M1) = C(1)*RESID04(-1) + C(2)*DLOG(YGDE1) + C(3)*DLOG(PM(-4)/PY(-4)) + C(4)*DUM86Q4 + C(5)*DUM87Q1$$

449 The equations that you estimate depend on the type of model you want to build. In this particular example, we require equations for investment, exports and imports. In your case you may also require exchange rate, employment, interest rates. The policy question at hand and the data availability will determine the structure of your model.

#### 4.4.7 Putting all the Equations Together

450 In the windows below, we present the equation that we will be using as part of this example.

Some questions for you to think as you go through the equations:

- Do they make sense in terms of theory?
- What are the long-run elasticities?
- What other variables can be used to explain the dependent variables?

Note the use of PM/PY and PE/PY. These are measures of competitiveness. If import prices rise faster than domestic prices then imports are losing competitiveness. If export prices rise faster than domestic prices then we gain competitiveness, hence we export more. PE/PM is the terms of trade.

previous shocks.

- When you shock rates by 1 per cent, you do NOT multiply the baseline rate by 1.01 but rather you add 1 i.e +1.
- Does household consumption converge to the respective long-run elasticity?

451 For each of the other three equations, we follow the same steps as we did in the one equation example, estimating the long-run and the short-run and solving to ensure that there are no mistakes as we add more equations into the model. After we put all the equations in the model, the model window will look like this;

```

EViews - (Model: A400MOD  Workfile: EViews MODEL--Quarterly)
File Edit Object View Proc Quick Options Add-ins Window Help
View Proc Object Print Name Freeze Solve Scenarios Equations Variables Text

HOUSEHOLD CONSUMPTION
.
Long run consumption model
LOG(CE1) = 1.05642070252*LOG(YD1) + 0.174481345521*FPRIMER/100 + 0.00272871675674*LOG(REXD/(PCPIX/(PCPIUS))) - 0.801916700295
@IDENTITY resid01 = LOG(CE1) - (1.05642070252 * LOG(YD1) + 0.174481345521 * FPRIMER / 100 + 0.00272871675674 * LOG(REXD / (PCPIX/(PCPIUS))) - 0.801916700295)

Short-run consumption Model
DLOG(CE1) = 0.00325304829868 - 0.0388457331122*RESID01(-1) + 0.326851863741*DLOG(CE1(-1)) + 0.11533421814*DLOG(YD1) + 0.116555426641*DLOG(YD1(-2)) + 0.00161970983666*DUM94N - 0.068472848274*DUM84Q3 + 0.014841721608*DUMINT
CE1 = exp(0.00325304829868 - 0.0388457331122 * RESID01(-1) + 0.326851863741 * DLOG(CE1(-1)) + 0.11533421814 * DLOG(YD1) + 0.116555426641 * DLOG(YD1(-2)) + 0.00161970983666 * DUM94N - 0.068472848274 * DUM84Q3 + 0.014841721608 * DUMINT + log(ce1(-1)))
@ADD(V) CE1 CE1_A

INVESTMENT
Estimate LOG(IP1) C LOG(Y1) (FGOVL(R(-3)/100) LOG(REXD/(PCPIX/PCPIUS))) (YCU-100)/100
Substituted coefficients are given as LOG(IP1) = -9.26625308398 + 1.57150280244*LOG(Y1) + 0.147768507761*FGOVL(R(-3)/100 - 0.115103473241*LOG(REXD/(PCPIX/PCPIUS))) + 1.8675286979*(YCU-100)/100
We have saved the residuals are resid02
@IDENTITY resid02 = LOG(IP1) - (-9.26625308398 + 1.57150280244 * LOG(Y1) + 0.147768507761 * FGOVL(R(-3) / 100 - 0.115103473241 * LOG(REXD / (PCPIX / PCPIUS))) + 1.8675286979 * (YCU - 100) / 100)

short-run investment model
DLOG(IP1) c DLOG(IP1(-1)) resid02(-1)
DLOG(IP1) = 0.146744547768*DLOG(IP1(-1)) - 0.0755785183517*RESID02(-1)
IP1 = exp(0.146744547768 * DLOG(IP1(-1)) - 0.0755785183517 * RESID02(-1) + log(ip1(-1)))
@ADD(V) IP1 IP1_A

@IDENTITY I1 = (ip1 * ipc1 + lg1)

```

```

Exports
-----
Long-run exports
-----
LOG(E1) LOG(G7Y1) C LOG(PE)PY)
*Substituted coefficients LOG(E1) = 1.45552177003*LOG(G7Y1) + 5.70240248403 + 0.0781552274096*LOG((PE)/PY)
@IDENTITY resid03 = LOG(E1) - (1.45552177003 * LOG(G7Y1) + 5.70240248403 + 0.0781552274096 * LOG((PE) / PY))

Short-run model
-----
DLOG(E1) LOG(E1(-1)) LOG(G7Y1(-1)) C DLOG(PE(-2))PY(-2)) DUM84Q2 DUM96Q3
DLOG(E1) = 0.00249150688114 + 0.296212112618*RESID03 + 1.00902413076*DLOG(G7Y1(-1)) - 0.0266261330729*LOG(PE(-2))PY(-2)) - 0.12147954974*DUM84Q2 + 0.070592257901*DUM96Q3
E1 = exp(0.0194813461411 - 0.418033274881 * RESID03(-1) - 1.19850159891 * DLOG(G7Y1(-1)) - 0.0283303515352 * LOG(PE(-2) / PY(-2)) - 0.144877431411 * DUM84Q2 + 0.123954978033 * DUM96Q3 + log(e1(-1)))
@ADD(V) E1 E1_A

Imports
-----
M1-real imports; YGDE-gross domestic expenditure; PM-import price deflator; PY-GDP deflator; E1-real exports; G7Y1-G7 growth; PE-export price deflator.
Longrun-equation
-----
Estimate LOG(M1) LOG(YGDE1) LOG(PM)PY) C
*Substituted coefficients LOG(M1) = 2.01985603852*LOG(YGDE1) - 0.361803325374*LOG((PM)PY) - 15.385052219
@IDENTITY resid04 = LOG(M1) - (2.01985603852 * LOG(YGDE1) - 0.361803325374 * LOG((PM) / PY) - 15.385052219)

Short-run model
-----
DLOG(M1) resid04(-1) DLOG(YGDE1) DLOG(PM(-4))PY(-4)) DUM86Q4 DUM87Q1
DLOG(M1) = -0.208145770851*RESID04(-1) + 1.65541332931*DLOG(YGDE1) + 0.214048824633*DLOG(PM(-4))PY(-4)) - 0.227948831051*DUM86Q4 + 0.129137453196*DUM87Q1
M1 = exp(-0.208145770851 * RESID04(-1) + 1.65541332931 * DLOG(YGDE1) + 0.214048824633 * DLOG(PM(-4) / PY(-4)) - 0.227948831051 * DUM86Q4 + 0.129137453196 * DUM87Q1 + log(m1(-1)))
@ADD(V) M1 M1_A

```

452 Notice that for each equation we have an Add-factor. The role of the Add-Factor and how it is added was explained in the one-equation model earlier.

#### 4.4.8 Adding Identities

453 Also, notice that at the bottom of the model window, there are identities. As mentioned earlier, the residuals are identities. Other identities include the variables that are calculated as ratio and which are Show how identities are added and what type of identities

```

IDENTITIES
-----
Gross Domestic Product (GDP)
-----
@IDENTITY y1 = ygde1 + e1 - m1
@IDENTITY y = y1 * py / 100)

Capacity utilization
-----
@IDENTITY ycu = y1 / ypot1 * 100)

Interest rates
-----
@IDENTITY ppcpix = (ppcpi / pcpic(-4) - 1) * 100
@IDENTITY fprime = (frepo - fmargin)
@IDENTITY fprime = (fprime - ppcpix)
@IDENTITY frepo = (frepo - ppcpix)
@IDENTITY fgovt = (fgovt - ppcpix)

Terms of Trade
-----
@IDENTITY pt = (pe / pm) * 100

```

#### 4.4.9 Forecasting In and Out of Sample with All the Equations

454 After this step, you can now solve the in-sample in the same way we illustrated with the one equation example. Solve first at the baseline and the values of all the equations will be calculated in sample. The values of the forecasts and the actuals will be the same because you have added the Add-factors already.

455 After solving in-sample and testing for the stability of the equations, again in the same way like we did in the one-equation model, you increase the sample size to 2012 and make assumptions for all the exogenous variables in each equation. Remember, earlier on we had created an excel file of assumptions from where we read the assumptions for the exogenous variables in the one-equation model. What you need to do now is to identify the exogenous variables in the other three equations take their values to the same excel file from where we can read their assumed values.

456 Check every one of the three equations and identify all the exogenous variables. To visualize the values of the other variables, go to the command window and type;

*show fgoal fmargin frepo g7y1 gc1 ig1 iil ipc1 pcpus pcpix pe pm py rexd yd1 ypot1 zy1*  
and you get

The screenshot shows the EViews software interface with a data table. The table has columns for variables: GDI, IG1, IPC1, PCPIUS, PCPIX, PE, PM, PY, REXD, YD1, YPOT1, and ZY1. The rows represent quarterly data from 2004Q2 to 2012Q4. The values are numerical, representing different economic indicators. The interface also shows the menu bar (File, Edit, Object, View, Proc, Quick, Options, Add-ins, Window, Help) and the command window at the bottom.

457 You can see from the E-views work-file which of the exogenous variables have missing values beyond 2009q4

458 Copy the equations that are still not in your excel assumptions file into excel.

You could also create a program to write them directly into excel as;

459 After this step, my full excel file will look like this;

460 For the sake of simplicity in this case, we grow all the variables with missing values by 5% up to 2012. So the excel file will look like this;

461 Now we have all the values of the exogenous variables up to 2012. We can read then into E-views. To read them, go to the assumptions program file you created and run. Remember we called this program file as 'cel\_assumptions'. Also, remember to note the path of the excel file (where you saved the file because you need to instruct E-views to get the excel file through that path).

462 Now solve with scenario1. Then, compare baseline with scenario1 to see how the model performs out-of-sample. This is one way to establish the performance of the model.

463 Scrutinize forecast results. It is possible to transfer your output to excel. You need to write a program in E-views that will write the results into excel;

```

'WRITE RESULTS.prg
'-----
'This program write the model results to:
'Tables-FORECASTS - for clients to see model results

smpl 2005q1 2012q4

WRITE(t) c:\modelout-RES.xls ce1_0 e1_0 fgovl fgovlr_0 fmargin fprime_0
fprimer_0 frepo frepor_0 g7y1 gc1 i1_0 ig1 ii1 ip1_0 ipc1 m1_0 pcplus pcpix
pe pm pppcix_0 py rexd y1_0 ycu_0 yd1 ygd1_0 ypot1 zy1
  
```

464 The excel work-file will then look like this;

Code	Variable name	Estimate 2006				FORECAST 2007				FORECAST 2008				ACTUAL		FORECAST	
		q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	2005	2006	2007	2008
CE1_0	Real Final household consumption	761,662	776,048	790,458	805,358	819,822	830,792	843,888	857,645	872,305	886,164	900,472	914,866	729,963	783,239	838,037	893,452
	% change y-o-y	6.5%	7.1%	7.5%	8.0%	7.7%	7.1%	6.8%	6.5%	6.4%	6.7%	6.7%	6.7%	6.6%	7.3%	7.0%	6.6%
	% change q-o-q (base)	0.3%	0.1%	7.0%	7.0%	7.4%	6.5%	6.5%	6.7%	7.0%	6.5%	6.6%	6.5%				
IG_0	Real gross fixed capital formation:	209,514	218,898	223,901	231,683	242,418	251,688	258,307	248,297	248,278	248,498	249,798	252,244	195,200	220,040	248,981	249,696
	% change y-o-y	11.4%	12.2%	12.8%	14.2%	16.2%	16.5%	12.5%	-2.2%	2.0%	-1.2%	6.5%	3.6%	9.0%	12.7%	13.0%	0.5%
	% change q-o-q (base)	13.9%	12.8%	14.0%	16.4%	21.8%	14.2%	-1.0%	-4.1%	6.0%	6.3%	2.1%	3.9%				
GC1	Real Gov consumption	220,586	228,682	226,139	228,773	236,337	234,589	237,217	238,967	242,728	245,531	248,366	251,235	214,440	226,025	237,005	246,965
	% change y-o-y	5.6%	8.0%	5.3%	2.0%	7.2%	-2.5%	-4.9%	-4.9%	2.7%	4.7%	4.7%	4.7%	5.2%	5.4%	4.9%	4.2%
	% change q-o-q (base)	-3.3%	10.7%	-4.4%	4.7%	13.9%	-3.1%	-6.7%	-6.7%	4.7%	4.7%	4.7%	4.7%				
PI_0	Real fixed capital formation: Total Private	152,877	157,178	162,838	167,458	172,958	178,398	175,473	176,501	168,078	165,808	164,548	164,323	142,784	159,913	174,332	165,688
	% change y-o-y	11.1%	12.0%	12.0%	12.7%	13.1%	13.5%	8.3%	1.8%	-2.6%	-2.7%	-6.2%	-3.6%	8.3%	12.8%	9.0%	-5.0%
	% change q-o-q (base)	12.5%	11.4%	13.6%	14.1%	13.8%	13.2%	-6.4%	-10.9%	-6.8%	-6.3%	-3.9%	-6.5%				
GI1	Real fixed capital formation: Government	31,310	32,482	33,584	34,806	35,141	37,078	38,569	39,589	40,630	41,708	42,812	43,844	29,500	33,940	37,968	42,275
	% change y-o-y	7.7%	10.3%	12.8%	16.0%	15.4%	15.6%	14.8%	13.7%	12.4%	11.9%	11.9%	11.9%	6.5%	11.7%	14.9%	11.3%
	% change q-o-q (base)	10.5%	16.0%	14.1%	15.4%	16.2%	16.9%	11.0%	11.0%	11.0%	11.9%	11.9%	11.9%				
PC1	Real fixed capital formation: Public corporations	25,227	26,220	27,444	29,419	34,311	35,828	36,995	38,307	39,661	40,972	42,428	43,838	22,894	27,071	36,260	41,728
	% change y-o-y	18.3%	17.5%	17.1%	21.4%	36.0%	35.8%	34.4%	29.9%	18.3%	16.9%	15.9%	15.9%	16.0%	18.3%	33.9%	15.1%

465 If you are not satisfied with model forecasts, you can adjust assumptions and/or Add-Factors after making sure that nothing is wrong with model and programs etc. After adjusting the assumptions and the Add-factors, you can just read the assumptions back into your E-views work-file. Once you are satisfied with the forecast, you now have a baseline forecast, you can now do alternative simulations.

466 You need to maintain a separate excel sheet with the original data from where you can always enter new data. For instance, assume we get the data for quarter two of 2004, we can enter this data into the original data sheet in excel and run the E-views program that we developed to read it from excel.



### 4.4.10 Policy Simulations with All the Equations

467 You could use the model to answer all policy questions in the same way we did in the one-equation model. Assume how a one percentage change in the G7 economic growth will affect the other variables in the model. You can write an E-views program for this and then write the simulations into excel. You will get the screen below. The file is named 'shock 01 - g7y1' under the folder 'E-views MEFMI model' in the attached CD.

```

EViews - [Program SHOCK01 - G7Y1 - If(jacob/private/research/woof/mefmi/main/draft/final/draft/eviews/mefmi/model/shock01 - g7y1.prg)]
File Edit Object View Proc Quick Options Add-ins Window Help
Run [Find] [Save] [Cut] [Copy] [Paste] [Insert] [Find] [Replace] [Wrap +/-] [Encrypt]
'Shock01 - g7y1.prg
-----
'Scenario 01
-----
G7 output is 1% higher than in baseline from 2005q1
'delete * _1
'delete g7y1t

SMPL 1970Q1 2012q4
series g7y1T = g7y1

SMPL 2006Q1 2012q4
series g7y1T = g7y1*1.01

smpl 2005q1 2012q4
show g7y1T (g7y1T/g7y1-1)*100

SOLVE aa01modG7

plot (y1_1/y1_0-1)*100 (ppcpix_1-ppcpix_0) (py1_1-py1_0)
plot (rwd_1/rwd_0-1)*100 0
plot (ycu_1-ycu_0) 0

STOP

smpl 2003q1 2007q4
WRITE(t) c:\model\out-SEN.XLS ((y1_1/y1_0-1)*100) (py1_1-py1_0) ((wrp_1/wrp_0-1)*100) ((wrpr_1/wrpr_0-1)*100) ((lep_1/lep_0-1)*100) ((dct_1/dct_0-1)*100) ((fm3_1/fm3_0-1)*100) ((ppcpix_1/ppcpix_0-1)*100) ((ppcpix_1-ppcpix_0) (frepo_1-frepo_0) (frepo_1-frepo_0) ((cex1_1/cex1_0-1)*100) ((t1_1/t1_0-1)*100) ((gpt1_1/gpt1_0-1)*100) ((ygdg1_1/ygdg1_0-1)*100) ((e1_1/e1_0-1)*100) ((m1_1/m1_0-1)*100) ((rwd_1/rwd_0-1)*100) ((pe_1/pe_0-1)*100) ((pm_1/pm_0-1)*100) ((pdt_1/pdt_0-1)*100) (pbca_1-pbca_0) (py_1/py_0-1)*100)

```

468 The excel file will then look like this.

	2003	2004	2005	2006	2007	2008
1	Scenario 01					
2	Economic G7 growth shock. Level of G7					
3	output is 1% higher from 2004q2					
4						
8						
9	Real wage levels (average wage per worker)	0.00	0.00	0.07	-0.08	-0.25
10	Non wage levels (average wage per worker)	0.00	-0.01	0.05	-0.04	-0.11
11	Total employment levels	0.00	0.00	0.01	-0.02	-0.02
12	Unit Labour Cost levels	0.01	0.00	-0.07	-0.14	-0.25
13	M3 money supply levels	0.48	0.40	-0.38	-0.97	-1.89
14	CPIH index levels	0.00	0.01	0.01	-0.22	-0.34
15	CPIF inflation rate (in %)	0.00	0.01	0.07	-0.22	-0.34
16	Nominal repo rate	1.00	1.00	1.00	0.00	0.00
17	Real repo rate	1.00	0.99	0.93	0.68	0.57
18	Household consumption levels	0.00	-0.01	-0.04	-0.07	-0.09
19	Government levels	-0.08	-0.07	-0.04	0.16	-0.15
20	Gov consumption levels	-0.08	-0.08	0.05	0.25	0.48
21	GCE levels	-0.02	-0.03	-0.02	0.02	0.01
22	Export levels	0.00	0.04	0.03	-0.04	-0.23
23	Import levels	-0.07	-0.08	-0.05	0.10	0.31
24	Real per US\$ levels	0.26	0.26	-0.08	-1.12	-2.01
25	Export inflation levels	0.25	0.24	-0.11	-1.08	-1.93

469 As explained in the case of results, we can in the same way write separate programs in E-views that will transfer the output of these simulations into excel so that we can easily report them to policy makers. In the model folder, we have included the programs and the excel sheets with different scenarios. You can go through them for your practice.

#### 4.4.11 Assessing Forecasting Accuracy

Allowing for the likelihood that economic outturn will be different from those projected, there are some basic yardsticks that can be used to measure forecast performance. One useful measure of accuracy has the average size of forecast error. We can also test whether or not forecasts are biased, i.e. whether or not they tend to consistently under-predict or over-predict and finally, we can test our forecasts against those made by external Forecasters.

##### (a) Using out of Sample Mean Square of Errors

Forecasters generally agree that forecasting methods should be assessed for accuracy using out-of-sample tests rather than goodness of fit to past data (in-sample tests). In most cases, in-sample errors are likely to understate forecasting errors. Method selection and estimation are designed to calibrate a forecasting procedure to the historical data. But the nuances of past history are unlikely to persist into the future, and the nuances of the future may not have revealed themselves in the past. Over-fitting and structural changes may further aggravate the divergence between in-sample and post-sample performance. Makridakis et al., (1982) shows that forecasting errors generally exceed in-sample errors, even at reasonably short horizons.

470 Moreover, common extrapolative forecasting methods, such as exponential smoothing, are based on *updating* procedures, in which one makes each forecast as if one were standing in the immediate prior period. For updating methods, the traditional measurement of goodness-of-fit is based on *one step-ahead errors* - errors made in estimating the next time period from the current time period. However, research shows (e.g., Schnaars, 1986) that errors in forecasting into the more distant future will be larger than those made in forecasting one step ahead. The second aspect to the argument is that methods selected by best in-sample fit may not best predict post-sample data.

471 One way to ascertain post-sample forecasting performance is to wait and see in real time. Real time assessment, however, has practical limitations for forecasting practitioners, since a long wait may be necessary before a reliable picture of a forecasting track record will materialize. As a result, tests based on *holdout* samples have become commonplace. The *fit period* is used to identify and estimate a model (or method) while the *test period* is reserved to assess the model's forecasting accuracy. If the forecaster withholds all data about events occurring after the end of the fit period, the forecast-accuracy evaluation is structurally identical to the real-world-forecasting environment, in which we stand in the present and forecast the future. However, 'peeking' at the held-out data while selecting the forecasting method pollutes the evaluation environment.

472 A popular measure of forecast accuracy is out-of-sample mean squared error,  $MSE = E(e^2)$ . The more accurate the forecast, the smaller is MSE. MSE is equal to the

sum of the variance of the error and the square of the mean error.

$$\text{var}(e) = E(e^2) - (E(e))^2 = \text{MSE} - (E(e))^2.$$

The elementary decomposition of the variance,

$$\text{MSE} = \text{var}(e) + (E(e))^2.$$

473 A forecast is therefore *unbiased* if the mean forecast error is zero. Why might unbiased forecasts be desirable? Are they *necessarily* desirable? Other things constant, a reduction in bias (mean error) is good, as it will clearly reduce MSE. However, other things may not be constant; in particular, an increase in bias may be more than offset by a large reduction in variance, resulting in a decrease in MSE. Suppose that  $(\log y)_{t+h,t}$  is an unbiased forecast of  $(\log y)_{t+h}$ . Then  $\exp((\log y)_{t+h,t})$  is a *biased* forecast of  $y_{t+h}$ . More generally, if  $(f(y))_{t+h,t}$  is an unbiased forecast of  $(f(y))_{t+h}$ , then  $f^{-1}((f(y))_{t+h,t})$  is a biased forecast of  $y_{t+h}$ , for the arbitrary nonlinear function  $f$ . The expected value of a nonlinear function of a random variable is *not* the same as the nonlinear function of the expected value, a result usually called "Jensen's inequality." Thus unbiasedness is not preserved under nonlinear transformations. Various "corrections" for the bias in  $\exp((\log y)_{t+h,t})$  have been proposed. In practice, however, bias corrections may increase the variance of the forecast error even if they succeed in reducing bias because the bias corrections involve estimated parameters, the very act of implementing them introduces a fresh source of error.

### (b) Using Theil's U-Statistic

474 Sometimes it's informative to compare the accuracy of a forecast to that of a "naive" competitor. A simple and popular such comparison is achieved by the U statistic, which is the ratio of the 1-step-ahead MSE for a given forecast relative to that of a random walk forecast  $y_{t+1,t} = y_t$ ; that is,

$$U = \frac{\sum_{t=1}^T (y_{t+1} - y_{t+1,t})^2}{\sum_{t=1}^T (y_{t+1} - y_t)^2}.$$

475 One must remember, of course, that the random walk is not necessarily a naive competitor, particularly for many economic and financial variables, so that values of U near one are not necessarily "bad." It is important to emphasize that for macroeconomic and financial series that can be well-approximated by random walks, a U-statistics near one do not necessarily indicate poor forecast performance.

(c) **Analysis using Historical Forecast Errors**

476 A run of errors in the same direction is sometimes taken as evidence of bias or inefficiency, as the forecaster does not appear to have learnt from past errors. But when considering persistence in errors it is important to distinguish between one-step ahead projections and multi-step ahead projections: It is only for one-step ahead forecasts that the previous forecast error can be observed before the next forecast is made.

(d) **Comparing your Forecasts with Other External Forecasts**

477 Another test of your projections is to compare performance with equivalent projections made by other forecasters.

#### 4.4.12 **Improving Forecasting Accuracy**

478 There are several ways to improve forecast accuracy, these include;

(a) **Forecast Combination**

479 One common way to improve forecasting accuracy has been to combine forecasts. Forecasters like using several different forecasting techniques and then comparing the forecasts and improving on them.

(b) **Consensus Forecasts**

480 In a policy making environment where there could be different institutions with their own forecasting tools, it would be important to come together and discuss the different forecasts and agree on some consensus forecasts. Normally, the mean or mode or median of all the forecasts are taken as the final forecast and reported. The consensus forecasts often perform very well relative to the individual forecasts.

(c) **Judgemental Forecasts**

481 Judgement is used throughout the modeling and forecasting process to modify quantitative forecasts. Judgement can be formally used to modify quantitative forecasts through forecast combination, when, for example, an “expert opinion” is combined with a model-based forecast

(d) **Delphi method for Combining Experts’ Forecasts**

482 The “Delphi method” is a structured judgmental forecasting technique that sometimes proves useful in very difficult forecasting situations not amenable to quantification, such as new-technology forecasting. The basic idea is to survey a panel of experts anonymously, reveal the distribution of opinions to the experts so they can revise

their opinions, repeat the survey, and so on. Typically the diversity of opinion is reduced as the iterations proceed.

483 Delphi and related techniques are fraught with difficulties and pitfalls. There is no guarantee that the iterations will converge, and even if they do, it's not clear why we should have confidence in the final forecast. At the same time, it's not at all clear that we should dispense with such techniques; they may be of real value because they at least provide a structured framework for attempting to reach consensus on difficult matters.

---

---

## CHAPTER 5

### OTHER FORECASTING METHODS

#### Objectives of this Chapter

At the end of this chapter, the reader should be familiar with the other forecasting methods available including:

- Forecasting from ARIMA models
- Forecasting from Exponential smoothing methods
- Forecasting from single equations

#### 5.1 Introduction

484 In the previous chapter we have illustrated how forecasting can be done from a structural econometric model. In this chapter we focus on a few other forecasting approaches that may be of interest to the user without necessarily having to build a very detailed model. Most policy analysts in the Central Banks and government ministries frequently use ARIMA models, exponential smoothing and single equations for forecasting. So we will dwell mainly on how to forecast from these three plus other simple forecasting approaches. The data used in this section is contained in the CD provided with this manual. To open the data, go to the CD and the folder named 'E-views MEFMI Model' from the CD. From the folder open the E-views file named 'E-views Model'.

#### 5.2 Other Forecasting Methods

485 Several forecasting methods are frequently used. These include:

- Unconditional Mean;
- Exponential Smoothing;
- Auto regressive Moving Average (ARIMA);

There are several other regression Based Forecasting Approaches which are beyond the scope of this manual (see page 158, paragraph 542)

##### 5.2.1 Unconditional Mean

486 In this model, the forecast equal to the unconditional mean over the recent past, that is:

$$y_t = \alpha + \varepsilon_t$$

where  $y_t$  is the variable of interest. The forecast from this model is simply

$$E(y_{t+h|t}) = \alpha.$$

In this case,  $E(y_{t+h|t}) = E(y_{t+h}|y_t)$  is the  $h$ -step ahead forecast.

487 This very simple method is appropriate when the process generating the  $\{y\}$  sequence is stationary. That is the mean of the series is constant over time, the variance is constant over time and the covariance is also constant over time. The main

impediment in using the unconditional mean to forecast next period's values is that most time series are non-stationary and therefore, an additional observation next period will likely change the mean of the series.

### 5.2.2 Exponential Smoothing

488 Exponential smoothing is a simple method of adaptive forecasting. It is an effective way of forecasting when you have only a few observations on which to base your forecast. Unlike forecasts from regression models which use fixed coefficients, forecasts from exponential smoothing methods adjust based upon past forecast errors. For additional discussion, see Bowerman and O'Connell (1979).

489 There are several exponential smoothing approaches. These include the simple exponential smoothing, weighted exponential smoothing, Holt's Linear Method (Double Exponential Smoothing) and Holt-Winters method (Triple Exponential Smoothing).

490 Holt's Linear Method is appropriate for series with a linear time trend and additive seasonal variation.

491 Holt-Winters method is appropriate for series with a linear time trend and no seasonal variation. This method is similar to the double smoothing method in that both generate forecasts with a linear trend and no seasonal component.

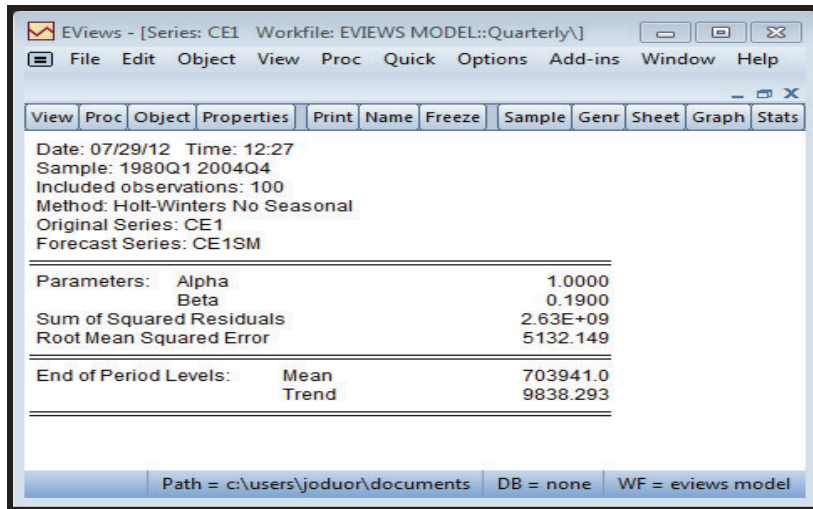
### 5.2.3 Exponential Smoothing in E-views

492 Open the ce1sm series from the 'E-views Model' work-file by double clicking on it. Then select Proc/Exponential Smoothing....

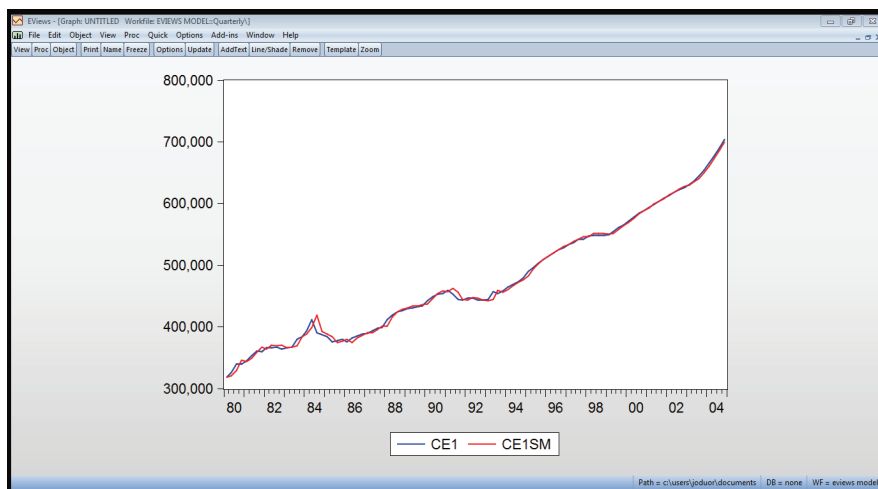
Name the forecasts



493 We use the Holt-Winters – No seasonal method since there are no seasonal patterns in the data. Let us name the smoothed forecasts as ce1sm, and estimate all parameters over the period 1980q1-2004q4. Leave the remaining settings at their default values. When you click OK, E-views displays the results of the smoothing procedure. The first part displays the estimated (or specified) parameter values, the sum of squared residuals, the root mean squared error of the forecast.



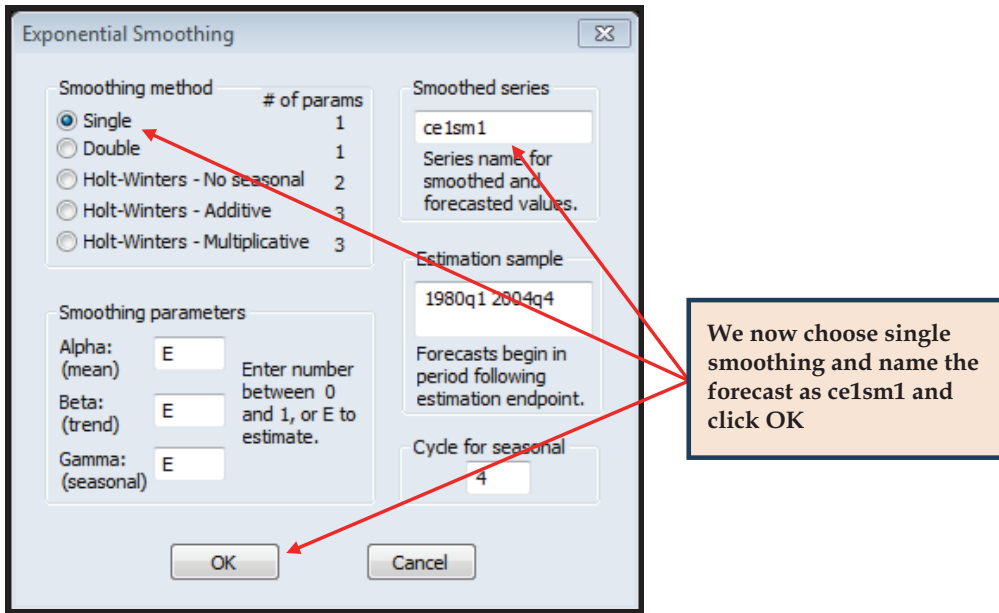
494 We can plot the actual values (ce1) and the smoothed forecasts (ce1sm) on a single graph to see how the forecasts track the actuals. We get:





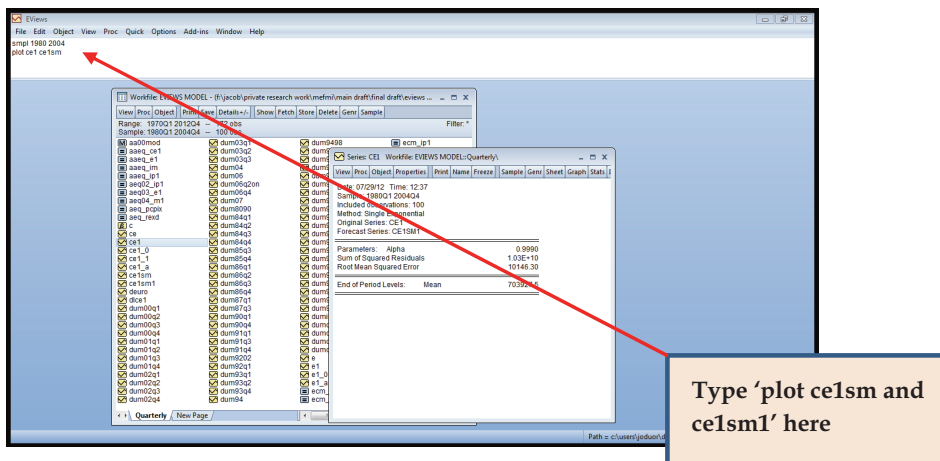
495 The forecasts are track closely the actuals.

496 We can also use a different smoothing approach, to see how different the forecasts would be. Let us now choose Holt-winters single smoothing. We have;

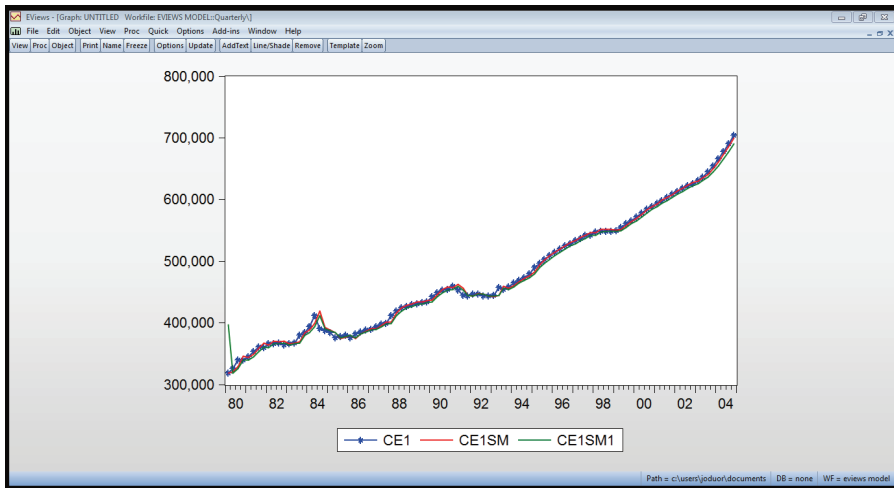


497 E-views will store the new forecasts as ce1sm1 in the work-file. We can compare the actual (ce1) with the two forecasts ce1sm and ce1sm1.

498 To do that, type in the command window, plot ce1sm and ce1sm1.



499 And you get;



500 The two forecasting methods do not give very different results.

501 We can use any of the approaches to forecast the values of consumption out-of-sample (into the future). Let us extend the sample to 2012 from 2004 and use Holt-Winters - No Seasonal method. When you click Proc/Exponential Smoothing, you get;

Exponential Smoothing

Smoothing method # of params

- Single 1
- Double 1
- Holt-Winters - No seasonal 2
- Holt-Winters - Additive 3
- Holt-Winters - Multiplicative 3

Smoothing parameters

Alpha: (mean)  Enter number between 0 and 1, or E to estimate.

Beta: (trend)

Gamma: (seasonal)

Smoothed series

ce1sm2  
Series name for smoothed and forecasted values.

Estimation sample

1980q1 2012e4

Forecasts begin in period following estimation endpoint.

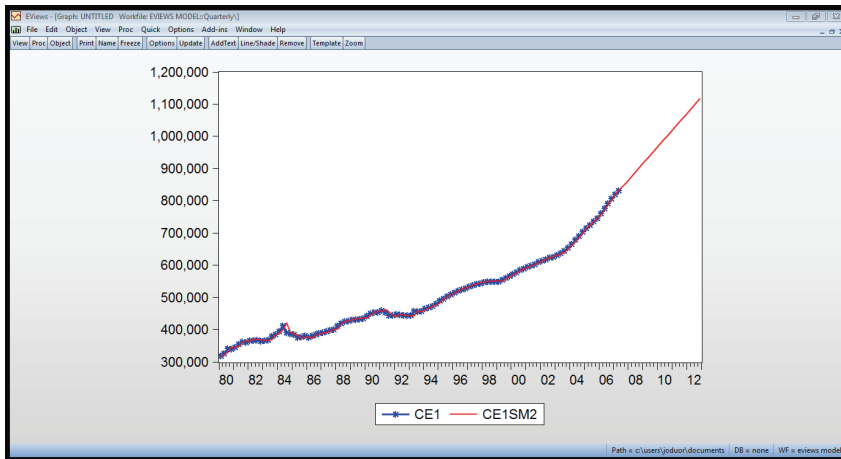
Cycle for seasonal

4

OK Cancel

Change the sample period to 2012

502 When we plot the out of sample forecasts against the actuals we have.



503 The thick line with symbols is the actual while the plain line in red is the forecast of consumption into 2012.

## 5.2.4 Autoregressive Moving Average (ARIMA)

504 This is a more complex class of moving average models that are capable of reflecting autocorrelations inherent in data. It can outperform exponential smoothing when the historical data period is long and data are nonvolatile. In the class of ARMA are several models. These include:

### (a) Random Walk models

505 Random walk models are special cases of ARIMA (0,1,0). They have been found to be robust to common forms of structural change including intercept shifts (Clements and Hendry (2002)).

The form of this model is given by;

$$y_t = y_{t-1} + \varepsilon_t$$

where  $y_t$  is the variable of interest. The  $h$ -step ahead forecast from this model is simply

$$E(y_{t+h|t}) = y_t$$

where  $E(y_{t+h|t}) = E(y_{t+h} | y_t, y_{t-1}, \dots)$  is the  $h$ -step ahead forecast. This model also requires that the  $y_t$  process is stationary.

## (b) Autoregressive (AR) Model

506 Autoregressive (AR) model describes a stochastic process that can be represented by a weighted sum of its previous values and a white noise error. In practice, univariate representations can often be captured by low-order systems. The general form of an AR model is given by;

$$y_t = \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t$$

where  $y_t$  is the variable of interest. The lag order  $p$  is chosen by the different information criteria like Akaike and Schwartz. The forecast from this model is

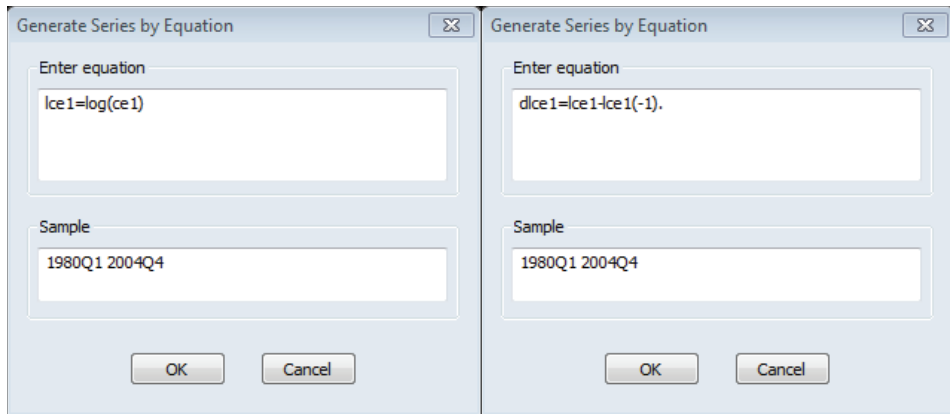
$$E(y_{t+h|t}) = \alpha_0 + \alpha' y_t \text{ and } E(y_{t+h|t}) = E(y_{t+h} | y_t, y_{t-1}, \dots)$$

## Choosing the right ARIMA Model and forecasting from an ARIMA Model in E-views

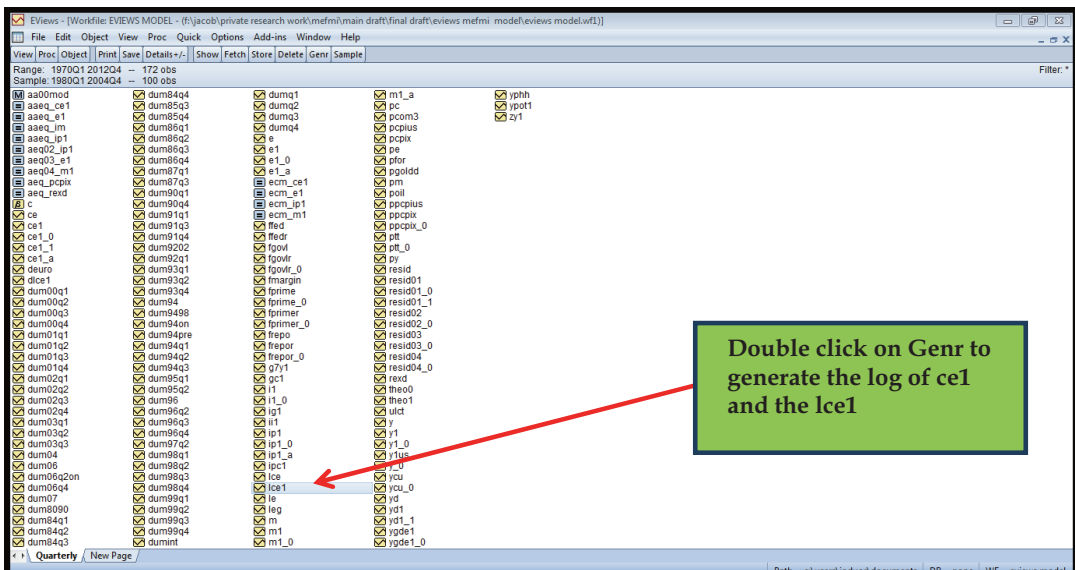
507 To demonstrate how to choose the right ARIMA model, we use the data contained in the CD that was provided with this manual. To access the data, open the folder named 'E-views MEFMI Model from the CD'. Click on the 'Genr' to generate the log of  $ce1$ .

The screenshot shows the EViews software interface. The 'Genr' menu option is highlighted in the top menu bar. A yellow speech bubble with a blue border points to the 'Genr' button, containing the text: "Click on 'Genr' to generate the log of ce1 and the first difference of the log of ce1". A cartoon character of a man in a white lab coat and glasses is pointing towards the speech bubble. The background shows the EViews main window with a list of variables and their properties.

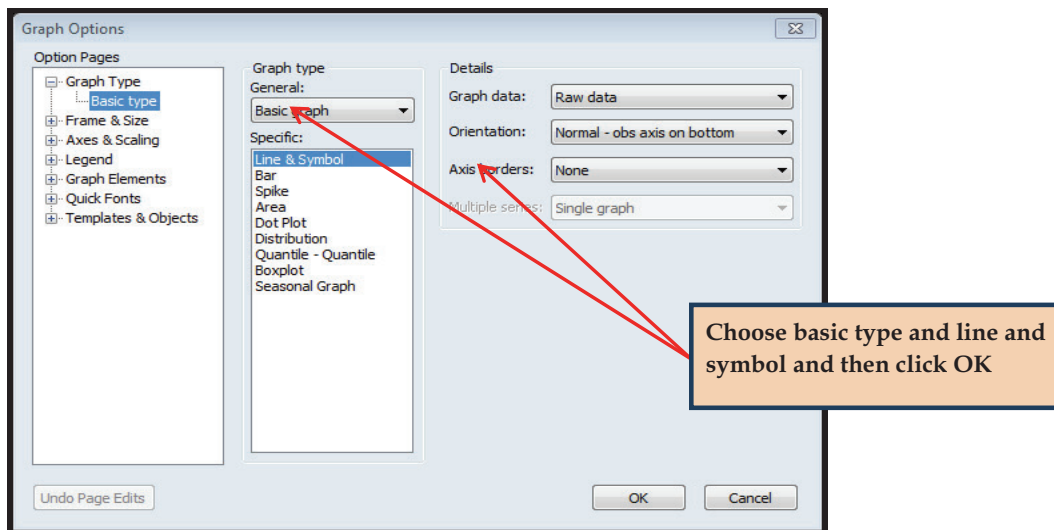
508 A dialog box come and you type the command  $lce1=log(ce1)$  and click OK. Then generate the first difference of  $lce1$  by clicking “Genr”, and typing  $dlce1=lce1-lce1(-1)$ .



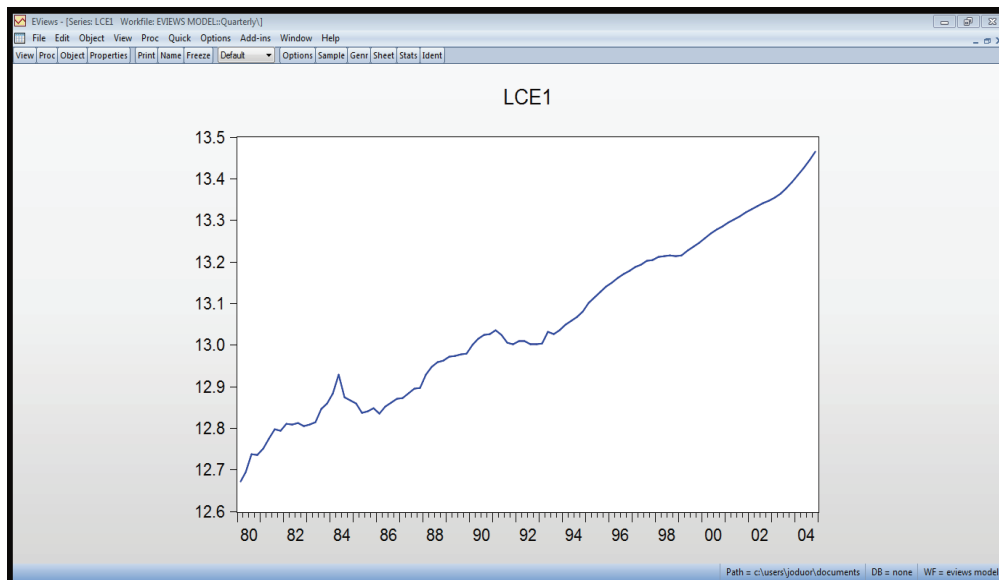
509 After that double click on the variable  $lce1$  (the consumption series) to open. Remember we are using E-views version 7. Lower versions may have differnt dialog boxes from the one show below.



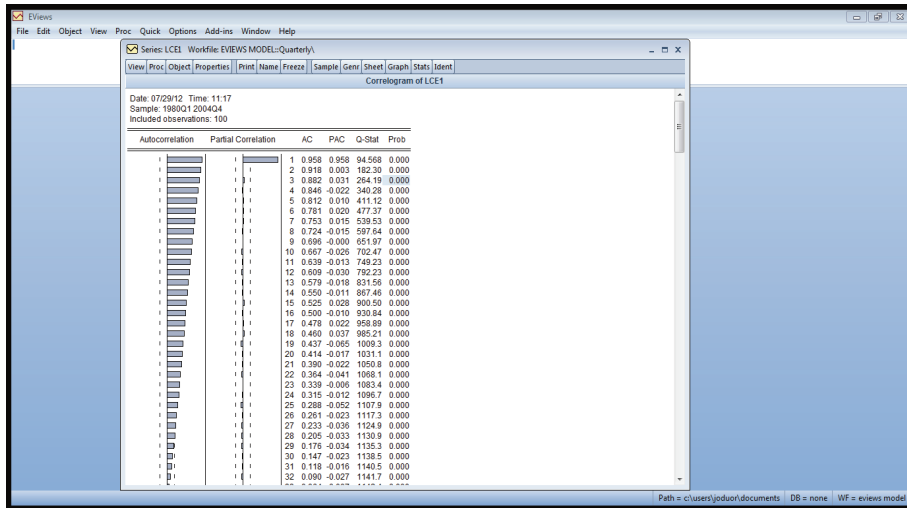
510 After opening the series, you can check its data to ensure that it is consistent with the Excel file and choose "View", "Graph" to get a general idea about the time series is stationary or not.



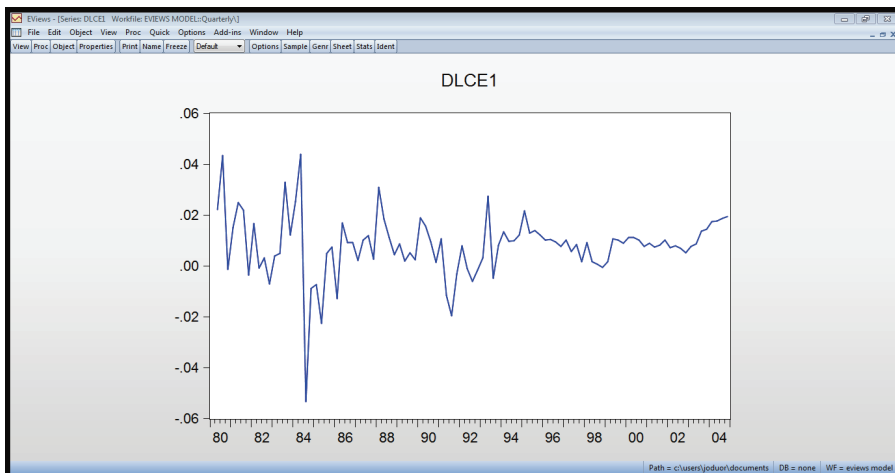
511 And you get graph that look like this.



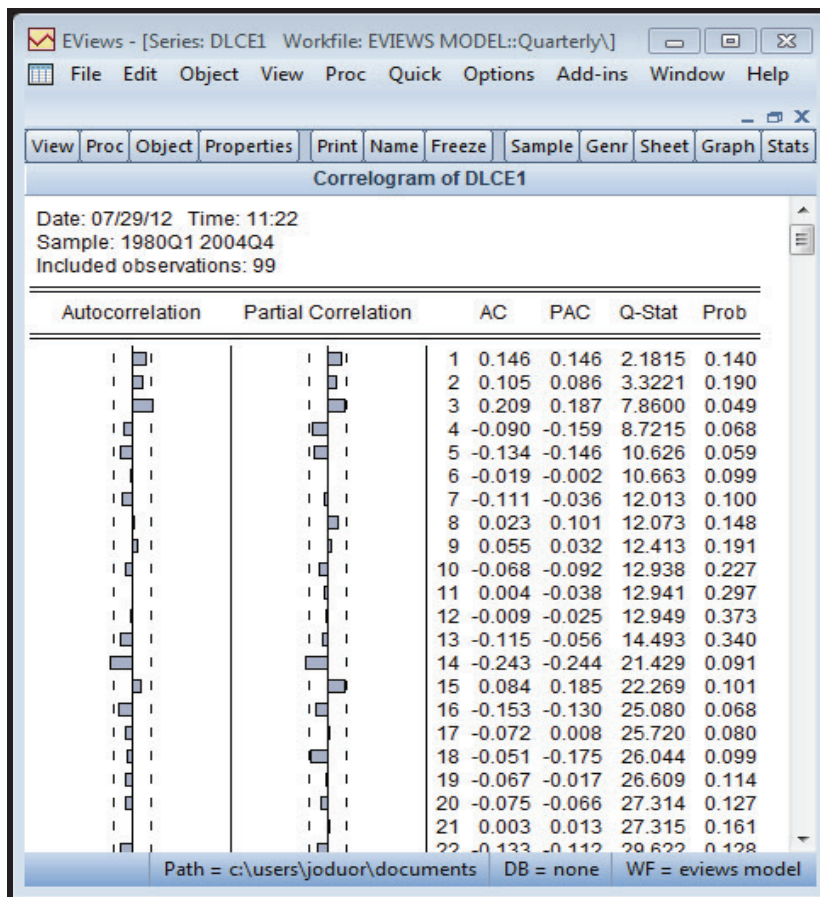
512 Next go to "View", "Correlogram" to get the tentatively identify patterns and model components (i.e. the degree of  $p$ ,  $d$ ,  $q$  of ARIMA). A small prompt window will appear asking you to specify the number of lags and whether you want to use levels or first difference. We choose 36 lags and levels and get;



513 From the above graphs, you can see that there is only one significant spike for PACFs. The graph of correlogram suggests that ARIMA(1,0,0) may be an appropriate model. Then, double click on the first-difference of "lce1" which we have already generated as 'dlce1' to see whether the time series becomes stationary at first difference before further finding AR( $p$ ) and MA( $q$ ). Maybe you can start by looking at the graph of dlce1 by and you get.



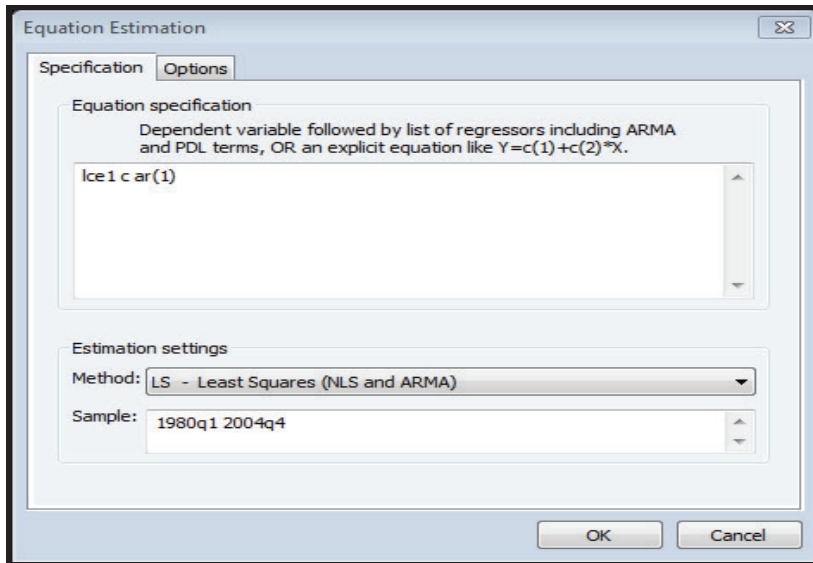
514 We can now view the correlogram of the dlce1 in the same way as done above.



515 The first-difference series "dlce1" shows no significant patterns in the graph of correlogram. This supports the that the ARIMA(1,0,0) is suitable for the time series. Then, we can estimate the ARIMA(1,0,0) model as following steps:



516 Step 1. Choosing "Quick", "Estimate Equation", then specify the mode and type "lce1 c ar(1)"



517 Click "OK", the result is:

Dependent Variable: LCE1  
Method: Least Squares  
Date: 07/29/12 Time: 11:30  
Sample (adjusted): 1980Q2 2004Q4  
Included observations: 99 after adjustments  
Convergence achieved after 5 iterations

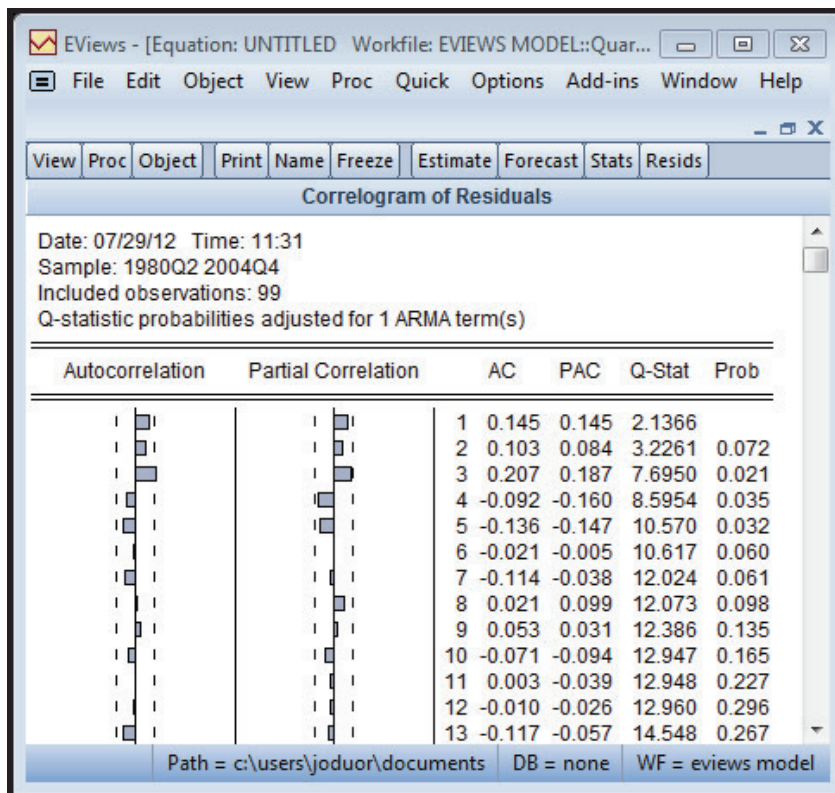
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.742344	31.53023	0.213837	0.8311
AR(1)	1.001270	0.006346	157.7871	0.0000

R-squared	0.996119	Mean dependent var	13.05650
Adjusted R-squared	0.996079	S.D. dependent var	0.198426
S.E. of regression	0.012425	Akaike info criterion	-5.918217
Sum squared resid	0.014975	Schwarz criterion	-5.865790
Log likelihood	294.9517	Hannan-Quinn criter.	-5.897005
F-statistic	24896.77	Durbin-Watson stat	1.688592
Prob(F-statistic)	0.000000		

Inverted AR Roots      1.00  
Estimated AR process is nonstationary

Path = c:\users\joduor\documents    DB = none    WF = eviews model

518 Step 2. choosing "View", "Residual diagnostics", "Correlogram-Q-Statistic" the result is:



519 [Since there is no significant spikes of ACFs and PACFs, it means that the residuals of this selected ARIMA model are white noise, so that there is no other significant patterns left in the time series, then we can stop at here and don't need to further consider another AR(p) and MA(q)].

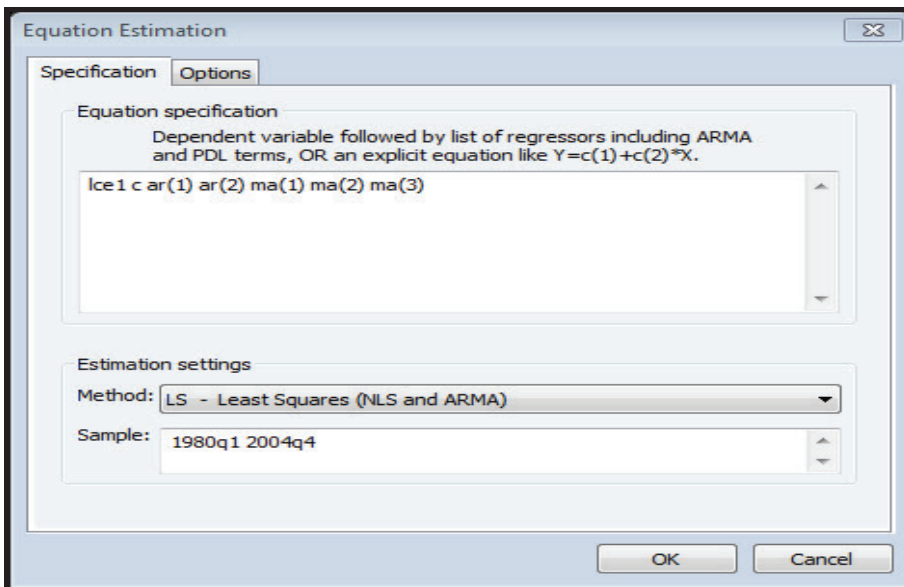
520 **The criteria to judge for the best model are as follows:**

- Relatively small of Schwarz and Akaike information criterion
- Relatively small of SEE
- Relatively high adjust  $R^2$
- Q- statistics and correlogram show that there is no significant pattern left in the ACFs and PACFs of the residuals, it means the residuals of the selected model are white noise.

521 You may try other ARIMAs with the `ce1` variable and compare the statistical results and fill in the table below. WE have already filled the statistics from the ARIMA (1,0,0) that we estimated above.

ARIMA MODEL	BIC	AKAIKE	ADJUSTED R2	SEE
(1, 0, 0)	-5.87	-5.92	0.996	
(1, 0, 1)				
(2, 0, 0)				
(0, 0, 1)				
(0, 0, 2)				
(1, 1, 0)				
(0, 1, 0)				

522 To estimate an ARIMA(2, 1, 3) go to Quick, Estimate and type `lce1 c ar(1) ar(2) ma(1) ma(2) ma(3)`.



### 5.2.5 Forecasting from an ARIMA Model

523 To forecast from ARIMA(1,0,0) model, click “Forecast” when you are in the results window and you will get;

Forecast

Forecast of  
Equation: UNTITLED      Series: LCE1

Series names  
Forecast name:   
S.E. (optional):   
GARCH(optional):

Method  
 Dynamic forecast  
 Static forecast  
 Structural (ignore ARMA)  
 Coef uncertainty in S.E. calc

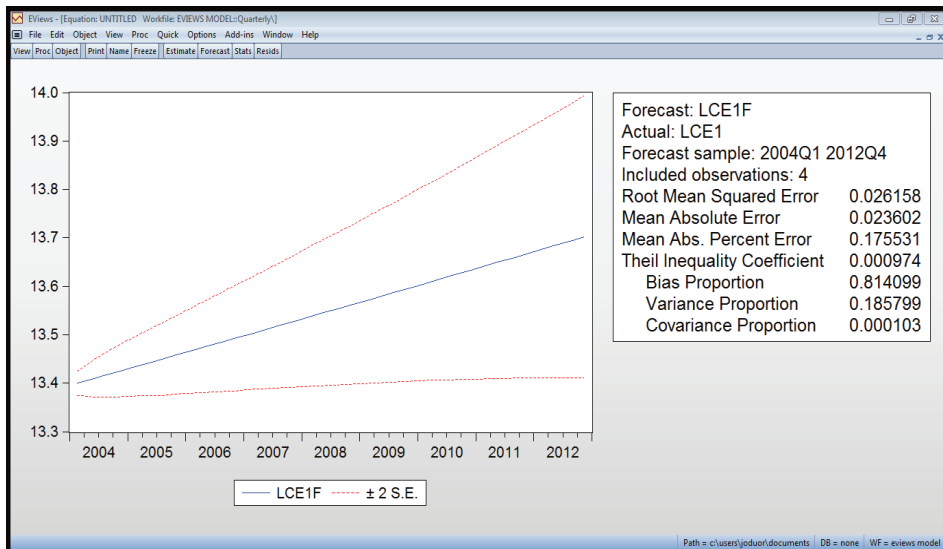
Forecast sample

Output  
 Forecast graph  
 Forecast evaluation

Insert actuals for out-of-sample observations

524 When you click 'OK' you will get;



### 5.2.6 Forecasting from an Estimated Equation

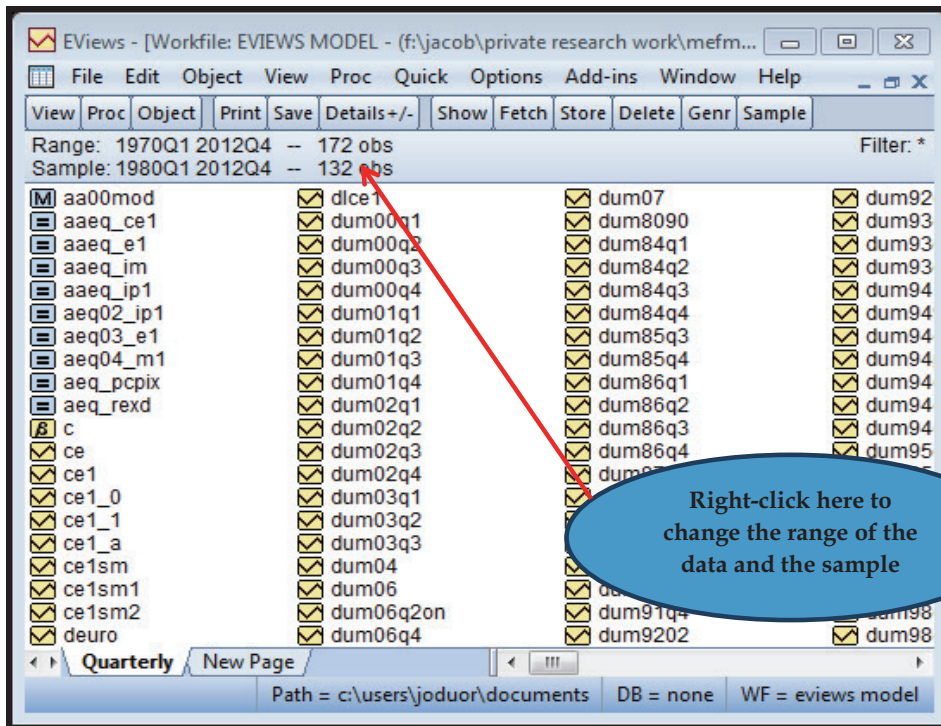
525 We estimate the consumption equation as;

LS LOG(CE1) LOG(YD1) (FPRIMER/100) LOG(REXD/(PCPIX(-1)/PCPIUS)) C

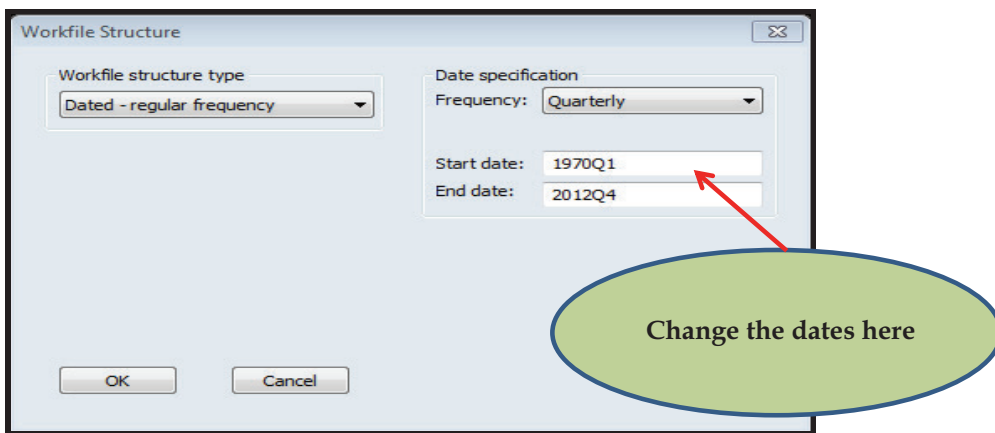
**Step One: Expand the sample range and enter new data for the new time periods:**

526 You can forecast first within the sample and then out of the sample. To be able to forecast out-of-sample, we need to expand the range of the data beyond the end of the available data. To do this, go to:

527 Open the work-file, right-click on "RANGE".



528 A new dialog box will appear. Change the ending date from 2004 to 2012 which is the end of our forecast horizon.

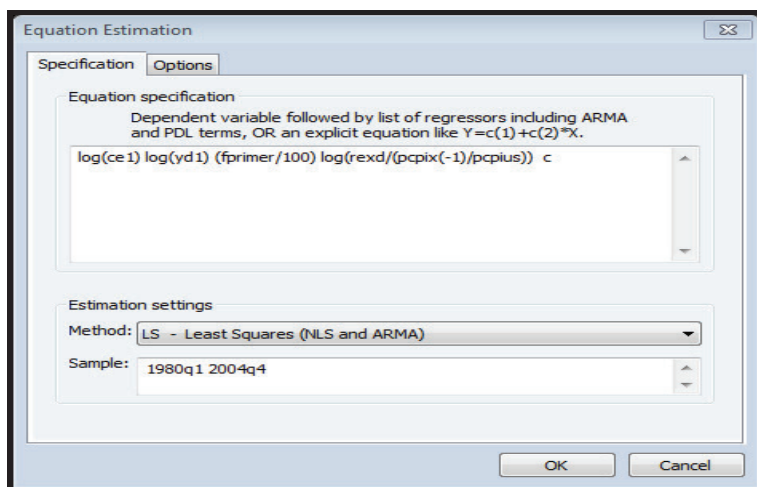


529 To forecast out of sample you will need to know the values of the exogenous variables beyond 2004q4. We get these values by imposing assumption on how the exogenous variables will evolve beyond 2004q4. Our sample data reaches 2004q4, so we assume that we do not have the values of all the variables beyond 2004. We have to dwell on how to make assumptions in the previous Chapter (Chapter Four) and so we will not repeat the process here. Let us assume that we now have the values of the exogenous variables up to 2012.

530 We are now ready to estimate our equation and forecast.

## Step 2: Estimating the equations:

531 To estimate the equation, go to; QUICK/ESTIMATE EQUATION and a dialog box appears. Type your commands in the window;



532 The consumption equation is given in the “E-views Model Work-file” contained in the CD. The estimation results are named as aaeq\_ce1. The results are given as:

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(YD1)	1.056421	0.025817	40.91918	0.0000
FPRIMER/100	0.174481	0.070806	2.464227	0.0155
LOG(REXD/(PCPIX(-1)/PCPIUS))	0.002729	0.017682	0.154320	0.8777
C	-0.801917	0.283844	-2.825203	0.0057

R-squared	0.975045	Mean dependent var	13.05265
Adjusted R-squared	0.974265	S.D. dependent var	0.201136
S.E. of regression	0.032267	Akaike info criterion	-3.990393
Sum squared resid	0.099949	Schwarz criterion	-3.886186
Log likelihood	203.5196	Hannan-Quinn criter.	-3.948218
F-statistic	1250.301	Durbin-Watson stat	0.800131
Prob(F-statistic)	0.000000		

Path = c:\users\joduor\documents DB = none WF = eviews model

### Step 3: Forecasting from the estimated equation:

533 To forecast from the estimated equation, the following steps are followed;

534 From the regression results window click on FORECAST.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(YD1)	1.056421	0.025817	40.91918	0.0000
FPRIMER/100	0.174481	0.070806	2.464227	0.0155
LOG(REXD/(PCPIX(-1)/PCPIUS))	0.002729	0.017682	0.154320	0.8777
C	-0.801917	0.283844	-2.825203	0.0057

R-squared	0.975045	Mean dependent var	13.05265
Adjusted R-squared	0.974265	S.D. dependent var	0.201136
S.E. of regression	0.032267	Akaike info criterion	-3.990393
Sum squared resid	0.099949	Schwarz criterion	-3.886186
Log likelihood	203.5196	Hannan-Quinn criter.	-3.948218
F-statistic	1250.301	Durbin-Watson stat	0.800131
Prob(F-statistic)	0.000000		

Path = c:\users\joduor\documents DB = none WF = eviews model

Click “Forecast”

535 The forecast dialog box then appears.

Forecast

Forecast equation  
AAEQ\_CE1

Series to forecast  
 CE1  LOG(CE1)

Series names  
Forecast name: ce1f  
S.E. (optional):  
GARCH(optional):

Method  
Static forecast  
(no dynamics in equation)  
 Structural (ignore ARMA)  
 Coef uncertainty in S.E. calc

Forecast sample  
1980q1 2012q4

Output  
 Forecast graph  
 Forecast evaluation

Insert actuals for out-of-sample observations

OK Cancel

1. Name the forecast – We call the forecast ce1f
2. Then click OK

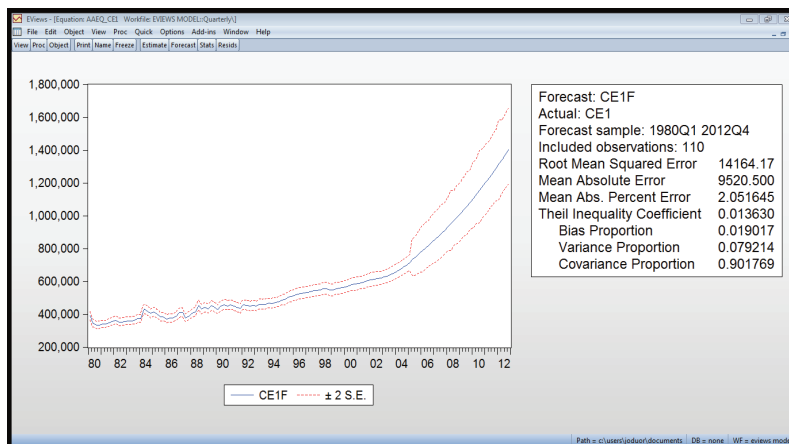
536 Give a name to the forecasts and to the standard error of the forecast. We name the forecasts ce1f.

537 Click/Check Forecast Graph and/or Forecast Evaluation to get forecast evaluation information.

538 Enter the desired Sample range for the forecast.

539 Check “Insert Actuals for out-of-sample”.

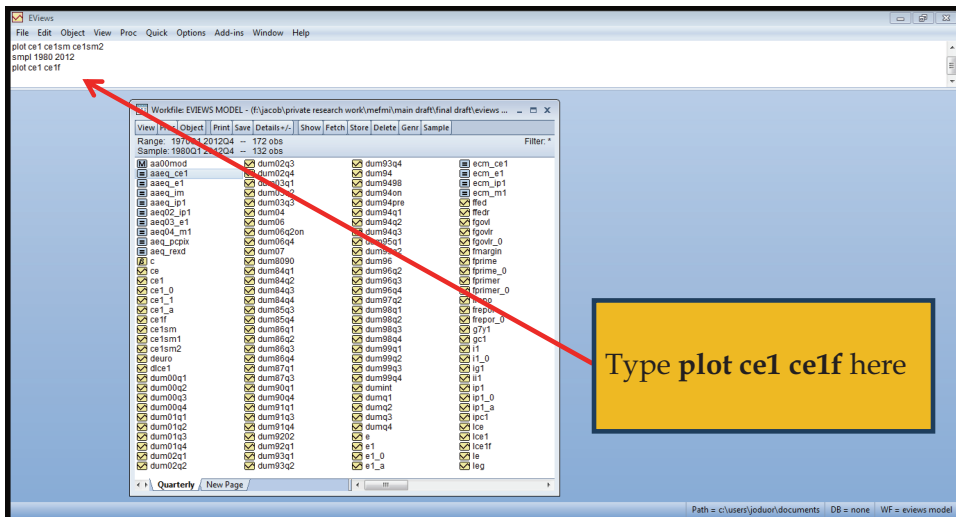
540 Click OK and you will see the Forecast Evaluation results.



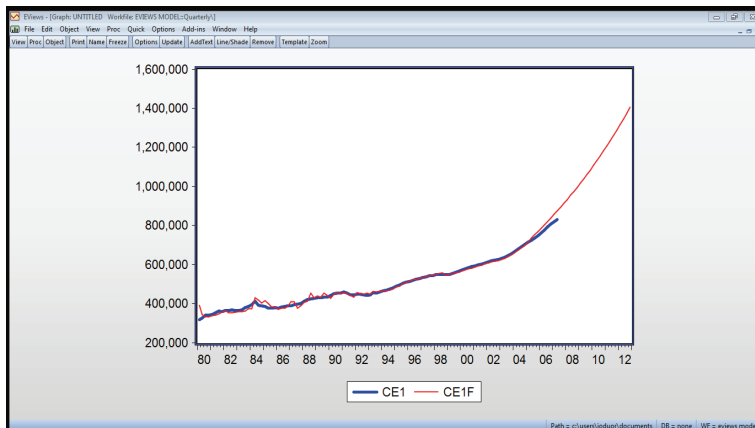


541 The results show that our forecasts are reasonable, given that the covariance proportion accounts around 90.2 percent of the variation. The bias and variance proportions also reported in the statistics above show how far the variation of the forecast is from the variation of the actual series respectively. The covariance proportion measures the remaining unsystematic forecast errors. If the forecast is good, the bias and the variance proportions should be small and thus most of the bias would be unsystematic. After doing the forecasts;

542 We can plot the actual `ce1` and the forecasts to see how best the forecast tracks the actuals.



543 The plot of the actual and the forecast is given as;



### 5.2.7 Other regression-based forecasting approach

544 There are several other regression models that can be used for forecasting and which could be used in different circumstances which are beyond the scope of this manual. These include:

- Vector autoregression (VAR)
  - Reduced form VAR
  - Bayesian VAR
  - Recursive VAR
  - Structural VAR
- Non-Linear Regression models
  - Markov-switching models (these are also univariate)
  - Smooth-transition autoregression (STAR) - (are also univariate)
  - Time-Varying parameter models
- Dynamic Stochastic General Equilibrium (DSGE) model

545 Readers who are interested in greater details of these other modeling and forecasting approaches are advised to seek guidance from relevant econometric text books.

**PART II**

**ECONOMY-WIDE  
MODELING**

---

---

## CHAPTER 6

### INTRODUCTION TO ECONOMY - WIDE MODELING

This chapter is aimed at practitioners interested in assessing the overall impact of policies on the economic system by developing models of economy wide interaction. The aim is to:

- \* Provide skills required to acquire relevant data, choose suitable techniques and use appropriate software packages to address substantive policy reform issues, equip practitioners with the necessary insights and skills to develop and criticize models for economy wide simulations;

At the end of this chapter, readers should have knowledge of;

- The three methods for making simulation studies, i.e. input output modeling, social accounting matrix modeling and computable general equilibrium modeling.
  - The software options available for estimating economy wide models.
  - GAMS with a view to an intuitive understanding of how algebraic model formulations can be expressed and presented under GAMS
- \* The chapter shows how the IO and SAM models can be solved using the widely used Excel software.
  - \* The chapter ends by showing how to use GAMS to solve the CGE models presented.

#### 6.1 Background

##### 6.1.1 Introduction

546 This part of the manual is aimed at a practitioner who is interested in assessing the overall impact of policies on the economic system by developing models of economy-wide interaction. The aim is to provide the skills required to acquire relevant data, choose the appropriate technique and use the appropriate software package to address substantive policy reform issues. A related aim is to equip the practitioner with the necessary insights and skills not only to develop but also to criticize models for economy-wide simulations.

547 The chapter presents three methods for making simulation studies, i.e. input output modeling, social accounting matrix modeling and computable general equilibrium modeling. For pedagogic reasons, we start with a look at *Input-Output* (IO) models in which quantities produced are determined given technology and demand levels.

548 The theoretical assumptions necessary to move from input-output accounts to an economy-wide economic model will be covered as will some of the more interesting

---

applications of input-output analysis to real world problems. We will also show how the models can be solved using the widely used Excel software.

549 The extension of input-output accounts to consider income flows from primary factors to households and government transfers to households allows construction of input-output/social accounts. The *Social Accounting Matrix* (SAM) provides the data necessary for the generation of computable general equilibrium models, but at the same time, given the necessary theoretical assumptions, lends itself to the construction of another class of fixed price models – the SAM model. SAM models permit a broader range of policy variables to be considered than IO models.

550 We then move to a *General Equilibrium Model* in which prices and quantities are determined simultaneously, given technology, preferences and endowments. We will introduce basic general equilibrium models and discuss how to introduce sophisticated features to these models. SAM accounts provide the basis for the construction of *computable general equilibrium* (CGE) models. The relationship between neoclassical economic theory and the structure of the standard CGE model will be discussed at length. Methods of moving from SAM accounts to model construction and solution will be reviewed. The CGE model has proven to be a useful economic framework for analysing many issues involving economic policy. We will review some of that literature as well as the emerging literature where CGE models have been used to focus on issues on financial flows and natural resource management in developing countries.

551 Finally, the options of software for estimating these economy-wide models are introduced. While the dominant software packages amongst researchers have been General Algebraic Modeling System (GAMS) and General Equilibrium Package (GEMPACK), people are increasingly writing their own programs in GAUSS, MATLAB or EXCEL. At the end of this chapter we will present a quick overview to Excel and GAMS which are the programming systems (languages) used for numerical implementation and solution of the economic models discussed in this section. The aim is to provide the user with a quick and intuitive understanding of how algebraic model formulations can be expressed under Excel and GAMS. The chapter ends by showing how to use Excel and GAMS to solve the models presented.

552 While the chapter deals with the empirical economy-wide approaches to policy modeling, it does not cover all such approaches. The chapter does not focus on linear optimisation models, game-theoretic models, neo-classical growth models, neo-Keynesian models and overlapping generations models. Several handbooks on macroeconomics *e.g.* Blanchard and Fischer (1989)<sup>3</sup> describe each of these model types, many references on economic modeling can be found in Van den Bergh (1999). An introduction into the mathematical methods used in most of these model types is given in

---

<sup>3</sup> They label neo-Keynesian models “new Keynesian models”; this is just one example of the lack of a common terminology encountered in the literature.

Chiang (1984). The rest of this chapter introduces some common elements of economy-wide modeling and then, each of the model types used is explained in more detail in separate sections.

### 6.1.2 Different Varieties of Economy-Wide Models

553 Within each modeling structure, there is room for variation in the specification of the model.

Basic characteristics that have to be chosen include:

- **Theoretical (methodological) versus applied (empirical):** does the model describe a hypothetical economy, allowing maximum focus on the mechanisms of the model (methodological approach), or does it provide statements about the size of the policy effects in a real economy (empirical approach). The neo-Keynesian models do not lend themselves easily for methodological analysis, as these models are based on extrapolating historical trends.
- **Static versus dynamic:** the number of periods in a model can be one (static model), multiple but finite (dynamic model), or infinite (some theoretical neo-classical growth models and overlapping generations models); another distinction often encountered in the literature is models describing the past, the present, the near future, or the far future.
- **For dynamic models only: myopic versus forward-looking:** agents can base their behaviour purely on the current allocation of resources (myopic behaviour) or can take a long term view for their decisions (forward-looking behaviour). The knowledge about future prices can be perfect (perfect foresight) or incomplete. If all agents are myopic, the model can be solved in a recursive-dynamic procedure, solving for one period at a time, while forward-looking models are solved for all periods simultaneously.
- **Deterministic versus stochastic:** a model may or may not contain uncertainty with respect to the expectations of agents, the model structure and the data (stochastics).
- **Calibrated versus estimated:** the parameters of the model may be chosen by the modeler based on existing literature and expert judgment, or they may be econometrically estimated, using historical time series data. Most neo-classical models are calibrated, while most neo-Keynesian models are econometrically estimated.
- **Geographical scale:** models exist that describe one individual firm or household (micro-economic models), while others describe the whole world; other commonly used aggregation levels are sectoral, regional, national and fluvial.
- **Different parts of the model can be fully integrated or be specified in separate sub-models,** which can have different characteristics; the feedback between the modules can be one-way (e.g. the economic module can affect the natural module but not *vice versa*) or both ways.

- The **specification of technological progress** can be absent, exogenous or endogenous. Next we introduce IO and SAM models<sup>4</sup>

## 6.2 Input-Output (IO) Tables

### 6.2.1 Background

554 IO matrices describe industry-level inputs and outputs. Most producers may require primary or intermediate inputs. They may produce intermediate or final output or both. An input-output table describes these inter-industry transactions.

555 Supply and use (SU) tables include transactions taking place in goods and services in an economy for a specific year, and are recorded in a matrix form. SU Tables are sometimes called rectangular IO tables or make and use tables. They are used in the construction of IO matrices.

556 Supply tables describe the sources available to the economy: local output and imports. Thus, supply tables present the value of different products at basic prices in each characteristic production industry and the total supply of products (of import and domestic production) at basic and purchaser prices. The supply table presents data on the total resources available in the economy in a given year, both from import and from domestic outputs. The table presents the resources by groups of characteristic products manufactured in each industry. *Uses tables* present the interrelations among different industries in the economy (intermediate uses), and among these industries and the final uses. Thus, *Use tables* mark the value of production in each industry and the use of each product at basic prices and purchaser prices. The two tables are used to construct the input output matrix/table

557 Note some differences between the IO and the National accounts: The National Accounts measure the net domestic product, meaning, the value added of all the industries while IO Tables present the detailed connections between resources and uses in each industry, meaning, the distribution of each industry's output by uses (consumption, capital formation and export). These connections are emphasised in the IO Tables. Because it involves much work to compile an input-output table, only a few governments produce these regularly.

558 IO tables are constructed based on certain principles, which are important for the understanding of the different components of the matrix. These are as follows:

- The tables are constructed on the basis that there is general equilibrium in the economy. Thus, the figures are based on a fundamental identity—i.e. supply equals demand. This means that supply or sectoral output  $X$ , must equal  $Y$  (final demand) +  $Z$  (intermediate demand).

<sup>4</sup> These examples are an adaptation from IFPRI (2004) and TIPS (2005).

- ii. All the data is reflected in monetary terms.
- iii. It is easiest, although not necessary, to assume that each sector or activity produces only one output.
- iv. Inputs are required in fixed proportions to output in each sector – this implies constant returns to scale.

559 IO tables offer a snapshot of the structure of the economy in terms of different sectors. We can thus, determine which industries (sectors) are important for which production sectors. We can also see the differences in sectors in terms of their contribution to the production of inputs, and final goods in the economy. We are able to learn and compare the differences in the primary factor requirements of the different sectors. We can tell which sectors are export oriented and which are not. We can tell the extent of relative dependence on imported inputs. As a result we can make informed guesses of what the likely impact of changes in the exchange rate would be on different sectors, for example. In addition IO models can be used for planning purposes and thus assist policy makers.

560 The rows in the tables detail the sectoral sales by products, and the columns display the sectoral input purchases by products.

561 An example of an IO table is given in Table 6.1:

		Intermediate Demand			Final Demand (D)				
To  From		1. Agriculture	2. Industry	3. Services	4. Consumption	5. Investment	6. Government	7. Exports	8. Total Output
1. Agriculture		O11	O12	O13	C1	I1	G1	E1	X1
2. Industry		O21	O22	O23	C2	I2	G2	E2	X2
3. Services		O31	O32	O33	C3	I3	G3	E3	X3
4. Imports		M1	M2	M3	Mc	Mi	Mg		M
5. Gross value added		gva1	gva2	gva3					GVA
..5a. Depreciation		depr1	depr2	depr3					depr
..5b. Wages		w1	w2	w3					w
..5c. Profit		prof1	prof2	prof3					prof
<b>6. Total input</b>		<b>X1</b>	<b>X2</b>	<b>X3</b>	<b>C</b>	<b>I</b>	<b>G</b>	<b>E</b>	

562 Table 6.1 above depicts three production sectors as follows: Agriculture, Industry and Services. These sectors are producing goods and services and they are selling their products and services to other sectors (indicated as *intermediate demand*) and to *final demand categories*, such as consumption, investment, government expenditure and exports. Sectors will also sell some of their products to other firms within the same sector.



These deliveries are called *internal deliveries (sales)*. All these aspects are shown in row numbered 1-3 of the table.

563 As seen in the table, each of the sectors appears twice in the table; as a column and as a row. As a column, a sector plays the role of purchasing goods and services. The sector then uses the goods and services as inputs in the production process. Hence the columns contain input data. For example reading downwards under the column for agriculture, the agricultural sector is found to purchase O11 units from itself, O21 from industry and O31 from services. It also uses M1 of imports.

564 As a row, a sector plays the role of producing and selling the goods and services to itself and the other sectors in the economy. The row therefore contains the sales or output data. By reading along the row, we are able to determine the sectoral contribution to output. Row numbered 4 of the table indicates that some inputs are imported from abroad. These imports may include intermediate demand or final demand categories, such as consumption, investment or government expenditure. In row numbered 5 gross value added is shown as the difference between total output ( $X_1$ ) and the inputs of a sector.

565 Note that total input in agriculture ( $X_1$ ) equals to the total output in the same sector ( $X_1$ ). Gross value added is subdivided into depreciation (to compensate for the wear and tear of capital), the remuneration of labor (wages) and the remuneration of capital (profits). Depreciation, wages and profits are called '*primary inputs*'.

### 6.2.2 Input-Output Modeling

566 IO models are commonly used to do impact analysis. This type of analysis measures the change in output and employment of projected demand changes. These demand changes could be known for particular industries or could be planned as new production activities.

567 IO analysis gives us two important pieces of information that can help assist policy makers.

- They attempt to quantify the extent of the multiplier, at the industry level, due to a policy change and;
- They give some insight into the industry-wide effects of policies.

568 We can formally write the contents of Table 8.1 as follows:

$$X_1 = O_{11} + O_{12} + O_{13} + C_1 + I_1 + G_1 + E_1 \quad (6.1)$$

$$X_2 = O_{21} + O_{22} + O_{23} + C_2 + I_2 + G_2 + E_2 \quad (6.2)$$

$$X_3 = O_{31} + O_{32} + O_{33} + C_3 + I_3 + G_3 + E_3 \quad (6.3)$$

where the  $X$ 's are production levels,  $O_{ij}$  are intermediate requirements from industry  $i$  per unit of output of industry  $j$ , and where the  $C$ 's,  $I$ 's,  $G$ 's and  $E$ 's are the levels of final demand from the demand categories (consumers, investment, government and exports).

569 A simple input output model can be developed on the basis of the input output table, given the following assumptions:

- i. The intermediary inputs to a sector are a fixed share of total output of that sector, e.g.  $O_{11} = a_{11}X_1$  and  $O_{12} = a_{12}X_2$ , which gives  $a_{11} = O_{11}/X_1$  and  $a_{12} = O_{12}/X_2$ .
- ii. Imports to a sector are a fixed share of the output of the sector, e.g.
- iii.  $M_1 = a_{41}X_1$  or  $M_2 = a_{42}X_2$
- iv. The primary inputs are a fixed share of total output of a sector.
- v. The final demand categories per sector are a fixed share of the total of each final demand category (except for exports).
- vi. Exports per sector are given from outside the model (exogenous):  $E_1$ ,  $E_2$  and  $E_3$ .

570 Given these assumptions the input output model can be formulated as follows:

$$\begin{aligned}
 X_1 &= a_{11} X_1 + a_{12} X_2 + a_{13} X_3 + a_{14} C + a_{15} I + a_{16} G + E_1 \\
 X_2 &= a_{21} X_1 + a_{22} X_2 + a_{23} X_3 + a_{24} C + a_{25} I + a_{26} G + E_2 \\
 X_3 &= a_{31} X_1 + a_{32} X_2 + a_{33} X_3 + a_{34} C + a_{35} I + a_{36} G + E_3 \\
 M &= a_{41} X_1 + a_{42} X_2 + a_{43} X_3 \\
 \text{depr} &= a_{51} X_1 + a_{52} X_2 + a_{53} X_3 \\
 w &= a_{61} X_1 + a_{62} X_2 + a_{63} X_3 \\
 \text{prof} &= a_{71} X_1 + a_{72} X_2 + a_{73} X_3
 \end{aligned} \tag{6.4}$$

The coefficients  $a_{ij}$  above are called the input-output coefficients or technical coefficients.

571 Assume you had the IO table as given in table 8.1 but which has values and after diving through (i.e.  $O_{11} = a_{11} X_1$ , which gives  $a_{11} = O_{11}/X_1$ , etc), you get the following information;

$$\begin{aligned}
 X_1 &= 0.025 X_1 + 0.03 X_2 + 0.08 X_3 + 0.1 C + 0.00 I + 0.21 G + E_1 \\
 X_2 &= 0.050 X_1 + 0.06 X_2 + 0.10 X_3 + 0.3 C + 0.66 I + 0.16 G + E_2 \\
 X_3 &= 0.100 X_1 + 0.06 X_2 + 0.02 X_3 + 0.2 C + 0.00 I + 0.10 G + E_3 \\
 m &= 0.150 X_1 + 0.12 X_2 + 0.04 X_3 + 0.4 C + 0.33 I + 0.53 G \\
 \text{depr} &= 0.025 X_1 + 0.02 X_2 + 0.02 X_3 \\
 w &= 0.500 X_1 + 0.60 X_2 + 0.60 X_3 \\
 \text{prof} &= 0.150 X_1 + 0.11 X_2 + 0.14 X_3
 \end{aligned} \tag{6.5}$$

This model consists of 7 equations and 7 unknown variables, i.e.  $X_1$ ,  $X_2$ ,  $X_3$ ,  $M$ ,  $\text{depr}$ ,  $w$  and  $p$ . In this simple model it is assumed that  $C$ ,  $I$ ,  $G$ ,  $E_1$ ,  $E_2$  and  $E_3$  are given. The model makes it possible to solve for the unknown variables  $X_1$ ,  $X_2$ ,  $X_3$ ,  $M$ ,  $\text{depr}$ ,  $w$ , and  $\text{prof}$ .

572 We can use the model to answer questions such as:

- What will be the required level of total production of each industry (direct and indirect) to satisfy that final demand vector?
- What will happen if the exports of any of the sectors e.g. manufacturing are increased by 10%?

573 We can now solve the equation system and analyse what will be the new levels of  $X_1$ ,  $X_2$ ,  $X_3$ ,  $M$ ,  $w$  and profits. The model can be solved by means of substitution, or by using the computer in excel. In order to make our model easier to solve, we must develop it further by expressing it in matrix algebra.

574 Input-Output Model Expressed in Vector Notation

The input output model can also be expressed in vector notation, where  $X$  is a vector

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad (6.7)$$

$A$  is a matrix of technical coefficients

$$a = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$D$  is vector of final demand per sector:

$$D = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix}$$

with  $D_1 = C_1 + I_1 + G_1 + E_1$ ,  $D_2 = C_2 + I_2 + G_2 + E_2$  and  $D_3 = C_3 + I_3 + G_3 + E_3$

575 From the equations above, we can see that;

$$\begin{aligned} X &= AX + D \\ \Rightarrow D &= (I - A)X \end{aligned} \quad (6.8)$$

where  $I$  is an identity matrix.

$$\Rightarrow X = (I - A)^{-1} D \quad (6.9)$$

$(I - A)^{-1}$  is called the Leontief inverse. Thus

$$\Delta X = (I - A)^{-1} \Delta D \quad (6.10)$$

576 Equation (6.9) above is the basic material balance equation of the IO model. It says that total output of sector  $i$  ( $X_i$ ) is distributed between intermediates ( $a_{ij}X_j$ ) and final demand ( $D$ ).  $(I-A)$  gives the direct and indirect output requirement per unit of final demand.

577 The fact that the Leontief inverse is non-negative means that it is feasible to get a mathematical solution for any  $D$ . This may not be feasible economically. But we could use the estimate as a consistency check.

### 6.2.3 Solving IO Models

578 Equation (6.9) is the solution of the static IO model. Given exogenously specified final demand, the equation can be used to determine production requirements necessary to satisfy the demand. Given final demand targets, the Leontief inverse  $(I-A)^{-1}$  allows calculation of the implied targets for sectoral production, which can then be evaluated for 'reasonableness', perhaps by comparing with projected sectoral capacities.

579 One can use the inverse to calculate 'multipliers' that give the impact on endogenous variables of shifts in exogenous elements of final demand. Therefore, given  $D_1, D_2, D_3$  demands, we can replicate individual industries' output (based on the  $a_{ij}$ 's). Also, if we consider a change in demand, we can now calculate how industrial outputs change proportionately.

#### (a) Model Validation

580 We can determine the impact that changes in final demand,  $\Delta D_i$  have on industry output,  $X_i$ . Suppose a planner wishes to produce 10 units of manufacturing goods and 15 units of agricultural goods for final consumption from the table below, how much gross output needs to be produced?

	Manufacturing	Agriculture
Manufacturing	150	500
Agriculture	200	100
Labour	200	300
Capital	350	100
Total	900	1000

581 A key question is what are the direct input requirements to produce 1 unit of output? For instance, how much of inputs does manufacturing require to produce a unit of its output? This is given by the technical coefficient matrices. Using the table above, the direct input requirements are as in Table 6.3:

	Manuf	Per unit	Agric	Per unit
Manufacturing	150	150/900=0.17	500	500/1000=0.5
Agriculture	200	200/900=0.22	100	100/1000=0.1
Labour Income	200	200/900=0.22	300	300/1000=0.3
Capital Income	350	350/900=0.39	100	100/1000=0.1
Total	900	R1	1000	R1

582 Each coefficient can be interpreted as the proportion of industry  $j$ 's production input supplied by industry  $i$ . Thus, the value of  $a_{11}$  implies that manufacturing requires 0.17 units of manufactured inputs to produce a unit worth of its output.

583 To compute this in Excel, we need to create a 2x2 identity matrix, which we use to calculate  $(I-A)$ , which we then invert. For our example, the *Leontief inverse* is:

$$\begin{bmatrix} 1.41 & 0.78 \\ 0.35 & 1.30 \end{bmatrix}$$

#### 6.2.4 Impact Analysis

584 Suppose we wish to produce 10 units of manufactured goods and 15 units of agricultural goods for final consumption. How much gross output needs to be produced? The answer is:

$$\Delta X = (I - A)^{-1} \Delta D$$

$$\text{that is } \begin{bmatrix} 1.41 & 0.78 \\ 0.35 & 1.30 \end{bmatrix} \begin{bmatrix} 10 \\ 15 \end{bmatrix} = \begin{bmatrix} 25.8 \\ 23 \end{bmatrix}$$

585 To realise the additional consumption requirement of the economy, we need to produce additional 25.8 units of manufactured goods and 23 units of agricultural goods.

#### 6.2.5 Multiplier Analysis

##### a) Output Multipliers

586 The basic question to be asked is what is the impact of a 1 unit increase in the final demand of agriculture? The output multipliers are directly derivable from the Leontief inverse matrix. They are calculated by summing down columns of the inverse. The simple (Type 1) output multiplier is (direct output + indirect output)/direct output. The total output multiplier is (direct output + indirect output + induced output)/ direct

output. Denote the simple output multiplier for sector  $j$  as  $O_j$ , and the  $ij$ th element of the Leontief inverse as  $a_{ij}$ . Then formally, the output multiplier for sector  $j$  is:

$$O_j = \sum_{i=1}^n a_{ij} \quad (6.10)$$

587 Denoting the total output multiplier for sector  $j$  with a bar, then

$$\bar{O}_j = \sum_{i=1}^n \bar{a}_{ij}. \quad (6.11)$$

588 Using the Manufacturing-Agriculture example, the answer is given by:

$$\Delta X = (I - A)^{-1} \Delta D \quad (6.12)$$

that is 
$$\begin{bmatrix} 1.41 & 0.78 \\ 0.35 & 1.30 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.78 \\ 1.3 \end{bmatrix}$$

589 Agriculture output multiplier =  $0.78+1.3=2.08$ . This is also equal to the sum of the second column of  $(I-A)^{-1}$  illustrating that you can read off output multipliers from the Leontief inverse. They are calculated by summing down columns of the inverse.

## b) Income Multipliers

590 At each stage of the cycle, income is generated, i.e., change in final demand leads to change in  $X$ , which leads to change in  $L$ . Total  $L$  Income = Labour income coefficient\* change in  $X$

$$\Delta L = L_Y (I - A)^{-1} \Delta D \quad (6.13)$$

591 Income multiplier of manufacturing is given by:

$$\begin{bmatrix} 0.22 & 0.3 \end{bmatrix} \begin{bmatrix} 1.41 & 0.78 \\ 0.35 & 1.3 \end{bmatrix} \begin{bmatrix} 10 \\ 0 \end{bmatrix} = 0.42$$

592 This is the income multiplier for manufacturing. It is called a Type I multiplier. To get a type II multiplier, simply scale by factor coefficient, i.e., Type II=Type I/ $L_Y$

## c) More Multiplier Analysis

593 There are many varieties of multipliers, all appropriate in some circumstances. It is important to choose the correct multiplier for the multiplicand under analysis. Some commonly used multipliers are described below.

**Initial effect:** The exogenous change in final demand.

**Total effect:** The total consequences of the change in final demand (may be direct + indirect, or direct + indirect + induced).

**Direct effect:** The direct effect is the first round of response to the initial effect. It may be the same as the initial effect, if the initial change in final demand is entirely absorbed domestically. Otherwise the direct effect will be smaller than the initial effect. The direct effect may be in different units than initial effects. The direct effect may be in income or employment while the initial effect was expressed in output, for example. **Indirect effect:** The impact on the economy as the demand for intermediate inputs changes.

**Induced effect:** The change in the economy due to the re-spending of income earned as a result of the initial and subsequent changes.

**Simple effect:** The sum of direct and indirect effects.

**Simple multiplier:** The sum of direct plus indirect effects divided by direct effects also known as the Type I multiplier.

**Total multiplier:** The sum of direct plus indirect plus induced effects divided by direct effects. Also known as the Type II multiplier.

## 6.2.6 Strengths and Weaknesses of IO Models

594 Input-output analysis is particularly strong in describing the interactions between the various branches of industry in an economy. Changes in one sector will affect all the other sectors and changes in prices of a certain category of inputs will affect all other prices in the economy. Input output tables are also very useful to describe the quantitative relations in an economy even with no explicit modeling. If these tables are made every year, it can be observed how economic activity and emissions of pollutants are developing over time.

595 Input output analysis has some serious drawbacks. The assumption of fixed technical coefficients and the assumed Leontief production structure (which does not allow for substitution of production factors) are not realistic. If the year of analysis is close to the year in which the data is available, before much growth in labour capital, etc. occurs, then the results will be good approximations. However, it is possible to develop input-output models with endogenously changing technical coefficients, in order to reduce the rigidity of the models.

596 Input-output models are strongly demand driven. If additional demand for exports occurs, the model will calculate higher levels of production, regardless of the supply side of the economy. To include the supply side in a realistic way the input output analysis needs further extensions, by including the labour market, the capital market and the process of capital accumulation by means of saving. The basic structure of input

output analysis allows for these extensions, but the models become very large and difficult to empirically validate.

### 6.3 Introduction to Social Accounting Matrix

#### 6.3.1 What is a SAM?

597 Social Accounting Matrices (SAMs) document national income and product accounts in a way that highlights the interdependence among the various 'sectors' (producers, markets, households, and institutions) in economic systems (Stone, 1986). The SAM is a database/model in table format based on data from the National Accounts and statistics about Households and other institutions, showing:

- i. Production structure of each industry;
- ii. Supply of products from imports and domestic production;
- iii. Generation of income for each industry;
- iv. Redistribution of incomes among households, from government to households, etc; and
- v. Detailed expenditure patterns of households and other institutions.

598 SAMs have been prepared for many countries around the world. They are used for multiplier analyses (e.g., Defourney and Thorbecke, 1984) or as a basis for computable general equilibrium models.

599 In brief, social accounting matrix construction is quite similar to double-entry bookkeeping. The difference is that left and right columns representing receipts and expenditure in double-entry accounts are displayed as rows and columns in the square SAM matrix. An element of the matrix represents the transaction between two accounts, by convention, to *i* from *j*. The sum across row entries is total incomings or receipts, and the sum of column entries is total outgoings or expenditures. Thus, the SAM is a single-entry accounting representation of the flow of goods and services and payment between sectors, classes of economic actor and other accounts. The SAM is a representation of the basic old adage in economics that: "There is only one fundamental law of economics: for every income there must be a corresponding expenditure. No economic theory can be considered complete unless all incomes and outlays are accounted for."

600 The SAM contains information on the flow of goods and payments between institutions in the economy. It distinguishes between current (flow) transactions and capital account transactions, where the latter alter the stocks of assets and liabilities of agents. Most SAMs are "flow SAMs" which record the flow of capital account transactions (e.g. the flow of savings and investment) but do not track the balance sheet consequences of these net changes to stocks of real and financial assets.



601 In Table 6.4 below we present a simple hypothetical SAM for an economy. The table should be read following the principle that columns pay rows and where each column adds up to the same number as the corresponding row.

		PAYMENTS BY								TOT
		ACT	COM	FAC	FIRM	GOV	HH	CAP	ROW	
RECEIPTS BY	ACT		Gross Output (Make Table)							Total Sales
	COM	Interm. Demand (Use Table)				HH Cons.	Gov Cons.	Investment	Exports	Aggr. Dem
	FAC	Value Added							Factor Service Exports	Factor Inc.
	FIRM			Gross profits			Transfers			Firm Inc.
	HH			Wages	Distrib. profits		Transfers		Foreign Remitt.	HH Inc.
	GOV	Indirect tax	Tariffs	factor taxes	firm taxes	direct taxes				Govt Inc.
	CAP				Retain Earning	HH saving	Govt saving		Foreign savings	Total Saving
	ROW	Imports		factor service imports		transfer abroad	Transfer abroad	Transfer abroad		Forex payment
TOT	Total Costs	Aggr. Supply	Factor Expen.	Firm Exp.	HH Exp.	Gov Exp.	INV	Forex receipt		

602 Table 6.4 above shows 8 accounts typically found in a SAM. The first 4 accounts describe the supply side (Activities, Commodities and Factors) while the remaining 5 accounts describe institutions (firms, households, government, capital account and rest of the world).

**Activities (ACT)** account is responsible for production in the economy. It buys intermediate inputs and hires factor services to produce commodities, generating value added. It pays indirect taxes to the government. It delivers gross output to the commodities account. Goods sold by activities are valued at producer prices in the SAM.

**Commodities** (COM) combine domestic output with imports so that we have aggregate supply (taking into account tariffs). Total supply is used to satisfy intermediate demand, household demand, government demand, investment and exports. Commodities are valued at purchaser prices in the SAM.

**Factors** (FAC) are a set of accounts for the expenditure and receipts of the factors of production: labour, land and capital.

**Firms** (FIRM) are owned by households, or government, or the rest of the world. They consume factors and intermediate inputs to produce goods and services for the market. They may save, pay taxes and pay dividends

**Households** (HH) are usually distinguished by type (e.g., rich or poor, rural or urban, ethnicity, skill, etc). Households supply factors of production (capital, land and labour) to firms; consume goods and services in the market; pay taxes to and receive subsidies from government; consume public goods; make net current transfers to the rest of the world; save and invest.

**Government** (GOV) represents the activities of government. It levies taxes on households, firms, and commodities. It undertakes current consumption; makes transfers to households, firms, and the rest of the world. Finally, the government saves and invests (in sector specific capital and public infrastructure).

**Capital Account** (CAP) summarises savings and investment in the economy. The account receives payments in the form of domestic and foreign savings and these are used to finance investment and any transfers abroad.

**The rest of the world** (ROW) supplies goods to domestic markets (imports) and consumes domestic output (exports); makes net transfers (remittances, grants, net interest payments); provides saving (official concessional lending or private savings).

603 The SAM is organised as a square table, so that it can be used for mathematical calculations. Briefly, the rows and columns can be interpreted as follows:

Row sums = income, sales

Column sums = purchases, payments

604 At a given point in time, three fundamental economic balances are satisfied simultaneously (you need to always verify that this is, indeed, the case using the numbers in the SAM):

- i. Commodity Balance – For each commodity and factor: **supply = demand**;

- ii. Flow of Funds Balance – For each institution: **total income = total expenditure** (current and capital). This is also known as the Accounting Identity; and
- iii. Macroeconomic Balance – **Balance of Payments: Saving = Investment.**

605 It should be noted that there is a close relationship between macroeconomics and the SAM. For instance, there is a close relationship between a SAM table and Gross Domestic Product (GDP). GDP is defined as the value of the final goods and services produced in a given year. In order to extract information on GDP from a SAM table, we need to eliminate intermediate production (goods that are used to produce other goods). There are 2 approaches to measuring GDP: product approach and earnings approach. Product approach includes summation of private consumption, investment, government spending, and net exports. This approach is represented by a usual formula for GDP:

$$GDP = C + I + G + EX - IM$$

606 Earnings approach is based on counting earnings or costs, such as wages, indirect taxes, depreciation, profits, and other income. All the major macroeconomic identities are embedded within the SAM. To illustrate, refer to the SAM in Table 6.5. This is an aggregated SAM with only 1 activity, 1 household and 1 factor of production. The following identities suggest themselves (recalling that row sum must equal column sum):

		PAYMENTS BY						
		ACT	COM	HH	GOV	CAP	ROW	
RECEIPTS BY	ACT		D				E	D+E
	COM			C	G	I		C+G+I
	HH	Y						Y
	GOV	Tx		Ty				Tx+Ty
	CAP			Sh	Sg		Sf	Sh+Sg+Sf
	ROW		M					M
		Y+Tx	D+M	C+Ty+Sh	G+Sg	I	E+Sf	

- i. *Commodity Balance* – for each commodity, supply is equal demand, that is:  
 $D+M = C+I+G$
- ii. *Household balance*: Household income must equal expenditure, that is  
 $Y = C+Ty+Sh$
- iii. *Government Budget*: Revenue equals expenditure,  $Tx+Ty=G+Sg$
- iv. *Savings Investment Balance*:  $I = Sh+Sg+Sf$
- v. *Balance of Payments (Trade balance)*:  $E+Sf=M$

607 Finally, to illustrate the SAM, Table 6.6 below shows a SAM for a hypothetical economy with numbers. The economy is assumed to be closed and with no government and firms. For simplicity, the activities and commodities accounts are lumped together as Activities. There are 2 Activities accounts: transport and non-transport. There are 2 factor accounts: labour and capital. There are 2 household types: rich and poor. Finally there is 1 capital account.

		PAYMENTS BY								TOTAL
		ACTIVITIES		FACTORS		HOUSEHOLDS CAP				
RECEIPTS BY	ACT	Transport	Non Transport	Labour	Capital	Poor	Rich	Accum.		
		FACT	Non Transport					20	5	5
	Labour		20	25						45
	Capital		10	25						35
	HHL	Poor			30	10				40
		Rich			15	25				40
		CAP						5	10	15
	TOTAL		30	50	45	35	40	40	15	

### 6.3.2 SAM as a Modeling Devise

608 The SAM can also be used as a basis for modeling. Each cell in a SAM, which is a transaction, can be thought of as the outcome of an underlying optimisation problem of the relevant institution(s). We can represent the flow in the cell as

$$t_{ij} = t(p, q; V, \theta)$$

where  $\mathbf{p}$  and  $\mathbf{q}$  are, respectively, vectors of relative prices (for goods and factors) and quantities. The vector  $\mathbf{V}$  is a vector of exogenous factors and  $\theta$  is a vector of parameters defining the relevant functional form.

609 A major contribution of SAM analysis compared to the IO model is that it allows changes in income to feed back into additional demand in the model. This was not possible in the IO model where income changes could be computed but remained exogenous to the model. Now, we distinguish consumer demand from final demand. Shocks to final demand lead to changes in sectoral production. To satisfy the change in production, firms adjust their demands for factors of production, which in turn influence household income. The change in household income further influences production in its own right, which in turn adjusts its demand for factors of production, influencing household income, giving a new round of demand changes.

610 Thus, the income effect is now endogenised in the model. To do this, follow exactly same procedure as in IO model, but include HH in the matrix 'A'. This involves extending the IO table by adding a household income row and expenditure column. For instance, the SAM could be represented as:

Table 6.7: Hypothetical SAM with numbers, Example 2						
		PAYMENTS BY				
		ACTIVITIES		Consumer Demand	Other Demand	TOTAL
		Transport	Non Transport		FD	
RECEIPTS BY	Transport	100	200	400	300	1000
	Non Transport	500	150	200	50	900
	HH Income	300	300			600
	Other Payments	100	250			350
TOTAL		1000	900	600	350	

611 We used the SAM to endogenise this effect.

$$\begin{aligned} (1) \quad X_1 &= a_{11}X_1 + a_{12}X_2 + h_{13}Y_{hh} + D_1 \\ (2) \quad X_2 &= a_{21}X_1 + a_{22}X_2 + h_{23}Y_{hh} + D_2 \\ (3) \quad Y_{hh} &= h_{31}X_1 + h_{32}X_2 + D_3 \end{aligned}$$

612 In matrix terms, we can write:

$$X = AX + D$$

where  $x$  is the vector of levels of production and household income,  $d$  is the vector of other final demands (excluding household demand) and  $A$  is the augmented input-output coefficients matrix (also has household expenditure coefficients).

613 The SAM model can also be expressed in vector notation, where  $x$  is a vector as;

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$A$  is a matrix of technical coefficients

$$a = \begin{bmatrix} a_{11} & a_{12} & h_{13} \\ a_{21} & a_{22} & h_{23} \\ h_{31} & h_{32} & 0 \end{bmatrix}$$

D is vector of final demand per sector:

$$D = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix}$$

with  $D_1 = I_1 + G_1 + E_1$ , etc. The equations for  $X_1$ ,  $X_2$ , and  $Y_{hh}$  can then be expressed as

$$X = AX + D$$

614 This equation system can be solved, using matrix algebra simply as

$$X = (I - A)^{-1} D$$

615 We can solve this using the Neumann iteration method. Because of the properties of  $A$ , we can write

$$(I - A)^{-1} \cong I + A + A^2 + A^3 + \dots + A^n$$

$$X = (I + A + A^2 + A^3 + \dots + A^n) D = D + AD + A^2 D + A^3 D + \dots + A^n D$$

This illustrates the material balance issue. Starting with the vector of final demands, we can work out the successive rounds of gross outputs necessary to achieve it. As we include further rounds, this converges on 'equilibrium'. The fact that the Leontief inverse is non-negative means that it is feasible to get a mathematical solution for any D. This may not be feasible economically. But we could use the estimate as a consistency check.

616 Using our example above, the Augmented Leontief Inverse (L) is given as

$$L = \begin{bmatrix} 3.02 & 1.86 & 2.64 \\ 2.51 & 2.93 & 2.65 \\ 1.74 & 1.53 & 2.67 \end{bmatrix}$$

### 6.3.3 SAM Multiplier Analysis

617 An economy is characterised by a circular flow of income and expenditure. Thus, a change in the demand for the product of one industry will not only affect that industry. All other activities will be affected indirectly through intermediate demand, factor demand, household demand, and government demand changes. Since some of an economy's activities are exogenously determined, less than 100% of a shock remains local. For example, if 40% of activity is exogenous, each successive round of indirect effects is

only 60% of the previous round. Thus, the effects eventually die out. A multiplier is the cumulative sum of the endogenous effects. It is inversely related to the exogenous portion of local activity. As before, let  $X$  denote a vector of activity levels (in value terms) in an economy. A matrix  $A$  denotes the amount of each activity used by each other, at rates which are assumed to be independent of the levels of activity in  $X$  (constant returns to scale). A vector  $D$  denotes a column of exogenous uses of each activity. Total activity  $X$  satisfies endogenous ( $AX$ ) and exogenous ( $D$ ) uses:

$$X = AX + D$$

Assuming  $A$  is parametric; any change in  $D$  must be accommodated by a corresponding change in  $D$ .

618 Compared to input-output analysis, the assumption that  $A$  is parametric is the 'fixed coefficients' assumption. Also, the assumption that total activity ( $X$ ) can passively accommodate a change in  $D$  relies on the assumption that factor supplies are perfectly elastic. Both of these assumptions are based on the long-run assumption that all prices (and wages) ultimately remain the same. Otherwise, a change in relative prices could mean that coefficients change. Or, an increase in local factor demand may simply drive up local factor prices rather than expand local output.

619 Solving for  $X$ , the relationship between  $D$  and activity vector  $X$  is easily shown to be:

$$X = (I - A)^{-1} D$$

where the term  $(I - A)^{-1}$  is known as the multiplier matrix. The multiplier matrix shows the cumulative effect on all activities of a given change in exogenous accounts (under the assumptions noted). In particular, the elements of a particular column of the multiplier matrix show the dollar effects of a dollar change in the exogenous part, on each activity.

620 The next step in SAM multiplier analysis is to calculate the multiplier matrix. To find the endogenous portions of total activity, the ( $A$ ) coefficients, divide the cell entries in each column by the corresponding column sums. The [ $A$ ] matrix, as explained above, is parametric. This implies that we assume that each account will continue to allocate constant proportions of their totals across activities regardless of exogenous shocks. With respect to government accounts, this is equivalent to assuming that the underlying preferences over public expenditure alternatives are represented by Cobb-Douglas (fixed share) functions (Adelman and Robinson, 1986).

621 The last step in constructing the multiplier matrix is to subtract the coefficient matrix from the identity matrix,  $(I-A)$ . Finally, the infinite sum of effects of changes in  $D$  on  $X$  is given by the inverse of that net matrix, which gives us the multiplier matrix:

$$M = (I-A)^{-1}$$

### 6.3.4 Interpretation of SAM Multipliers

622 As explained above, the elements in a specific column account of the multiplier matrix show the effects on the row sectors and institutions of a  $R1$  (or \$1, or whatever relevant currency) change in exogenous activity (extra-regional demand or federal government expenditure) in the column account. Thus, the column coefficients show the backward linkages of a sector (purchases from sectors upstream). A specific account row shows how that account is affected by  $R1$  changes in the column accounts, or, the forward linkages of the sector (sales to sectors downstream). This is known as linkage analysis.

### 6.3.5 SAM Closure Rules

623 For multiplier analysis using a SAM, a key task is to distinguish endogenous from exogenous accounts. Once this choice is made, the basic SAM can be partitioned: the columns specified as exogenous ( $D$ ) (and the corresponding rows) are excluded from the SAM, leaving only the endogenous ( $X$ ) accounts. This critical choice is called the macro closure, and it should relate to the objective question (Adelman and Robinson, 1986; Thorbecke, 1994). The resulting multiplier matrix is very sensitive to the closure choice. As noted above, multipliers are inversely related to the portion of activity, which is exogenous.

### 6.3.6 Numerical Example of a SAM Model

624 Recall the example used to illustrate IO modeling given in Table 6.7. The basic question to be asked is what is the impact of a one unit increase in final demand of Transport? The output multipliers are directly derivable from the augmented Leontief inverse matrix. They are calculated by summing down columns of the inverse. They can also be computed by multiplying the augmented Leontief inverse by a new demand vector showing an increase in transport demand of one unit.

625 The Output and Income Multipliers for Transport are computed as follows:

$$\Delta X = (I - A)^{-1} \Delta D$$



$$\begin{bmatrix} \Delta X_1 \\ \Delta X_2 \\ \Delta Y_{hh} \end{bmatrix} = \begin{bmatrix} 3.02 & 1.86 & 2.64 \\ 2.51 & 2.93 & 2.65 \\ 1.74 & 1.53 & 2.67 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

626 Solving this gives us:

$$= \begin{bmatrix} 3.02 \\ 2.51 \\ 1.74 \end{bmatrix}$$

627 Hence, the Transport Output Multiplier is 5.53 (=3.02+2.51). The Transport Income Multiplier is 1.74. The Output and Income Multipliers for Non - Transport are computed as follows:

$$\Delta X = (I - \bar{A})^{-1} \Delta D^*$$

$$\begin{bmatrix} \Delta X_1 \\ \Delta X_2 \\ \Delta Y_{hh} \end{bmatrix} = \begin{bmatrix} 3.02 & 1.86 & 2.64 \\ 2.51 & 2.93 & 2.65 \\ 1.74 & 1.53 & 2.67 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

628 Solving this gives us:

$$= \begin{bmatrix} 1.86 \\ 2.93 \\ 1.53 \end{bmatrix}$$

629 Hence, the Non - Transport Output Multiplier is 4.79 (=1.86+2.93). The Non-Transport Income Multiplier is 1.53. The Output and Income Multipliers for Households are computed as follows:

$$\Delta X = (I - \bar{A})^{-1} \Delta D^*$$

$$\begin{bmatrix} \Delta X_1 \\ \Delta X_2 \\ \Delta Y_{hh} \end{bmatrix} = \begin{bmatrix} 3.02 & 1.86 & 2.64 \\ 2.51 & 2.93 & 2.65 \\ 1.74 & 1.53 & 2.67 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

630 Solving this gives us:

$$= \begin{bmatrix} 2.64 \\ 2.65 \\ 2.67 \end{bmatrix}$$

631 Hence, the Household Output Multiplier is 5.29 (=2.64+2.65). The Household Income Multiplier is 2.67. We can use these results to rank/prioritize policy. For instance, if we had only R1 to spend, the following prioritisation suggests itself:

632 In terms of impact on income:

1. Households
2. Transport
3. Non – Transport

633 In terms of impact on output:

1. Transport
2. Households
3. Non – Transport

### 6.3.7 Comparison of IO and SAM Modeling

634 The main difference between IO and the SAM is that the household income account is endogenised when using the SAM modeling technique. This enables the model to capture second order effects emanating from demand. Hence, the IO is likely to underestimate the full extent of the impact. Table 6.8 below illustrates this for our hypothetical example:

		Model Type	
		IO Model	SAM Model
<b>Output Multiplier</b>	Transport	2.08	5.53
	Non Transport	1.76	4.79
<b>Income Multiplier</b>	Transport	0.65	1.74
	Non Transport	0.57	1.53

### 6.3.8 Advantages of SAM

635 The SAM accounting framework offers many advantages for the analyst. One is that transactions are recorded only once. What is incoming to account  $i$  is outgoing from account  $j$ .

636 A second advantage is that all accounts can be reconciled simultaneously. The balance between income and expenditure in an account is the account's row sum minus its column sum. A SAM is "balanced" when all row-column sums equal zero simultaneously.

637 A third advantage is that by imposing SAM balance on the available data, we can solve for missing data about interactions between accounts. Even though agencies keep accounts in different ways according to their differing needs, using a SAM we can reconcile the potentially different reports.

638 The fourth advantage is that direct and indirect interactions are easily distinguished in SAMs. Direct interaction between  $i$  and  $j$  is shown by SAM element  $a_{ij}$  (or  $a_{ji}$ ). An indirect interaction is shown by using an intermediary account, e.g.,  $k$ , and two entries  $a_{ik}$  and  $a_{kj}$ . By the same token, an interregional flow can be directly articulated or indirectly articulated.

639 A fifth advantage of the SAM format is that it efficiently displays all interactions in a consistent manner, which greatly facilitates comparison across accounts or sets of accounts.

### 6.3.9 Disadvantages of SAM

640 Some of the disadvantages of SAM include:

- There are several major consumption patterns of the household.
- We still assume fixed coefficients.
- There is no supply constraint. This could be achieved in a similar way to the constrained input-output model – endogenise prices and exogenise outputs. Although prices would adjust, we would still not see the supply responses we might expect. Nor would we see household consumption patterns adjusting.

641 To include the supply side in a realistic way, SAM analysis needs further elaboration, by including the labour market, the capital market and the process of capital accumulation by means of saving. This leads to CGE models.

## 6.4 Computable General Equilibrium Models

### 6.4.1 What are CGE Models?

642 A Computable General Equilibrium (CGE) is a system of linear and nonlinear equations that are solved on the computer to simulate equilibrium of multi-markets. *Computable* reflects that the model is a computer representation of an economy while *general equilibrium* means that the model encompasses both the production and consumption side simultaneously.

643 The main form of analysis based on CGE models is that of *comparative statics*. In comparative statics, one starts from a base solution or scenario (also known as a counterfactual), then one of the exogenous variables or parameters is altered and the model calculates the new values for the endogenous variables. The comparison of the two sets of values of the endogenous variables suggests the estimated economic impact of the change in policy (intervention). The path taken by variables of interest to move from one equilibrium point to another is not of importance here, only the comparison of the two equilibria.

644 *Dynamic analysis*, on the other hand, is concerned with looking at the path from one equilibrium to another.

645 CGE models simulate a working economy by incorporating various institutional and structural characteristics that simple analysis fails to capture. CGE models are widely used by economists to analyse the magnitude and distributional effects of external shocks and policies, such as oil price shocks, and consequently to help formulating policies.

646 They are primarily based on neoclassical theory of general equilibrium, first formulated by Leon Walras in 1877 and later formalised by Arrow and Debreu (1954) and McKenzie (1954, 1959, 1981). Improvement in data collection and advances in computer technology and software has enhanced advanced methodology of applied policy work.

647 The principal advantage of using CGE models in policy analysis is that it permits taking into account interactions throughout the economy. If something is changed in only one part of the economy, such as the oil sector, due to oil prices and government policy response, then there will be effects on the other parts of the economy, and these are automatically taken into account when one computes direct and indirect effects using a general equilibrium model. Hence, this approach has both a sound theoretical structure as well as an exhaustive accounting strategy.

648 CGE models require a benchmark data presented in the form of a SAM. The latter is a set of accounts written in a condensed matrix form with an important property that the sum of the row elements is equal to the sum of the corresponding column elements. The SAM is therefore consistent in the sense that it describes a general equilibrium of the economy in question from which counterfactual analysis can be carried out.

649 The calibration of CGE models consists of finding parameters which permit the reproduction of the benchmark situation given by the SAM. The SAM should be as recent as possible although in situations where it is hard to find, it can be acceptable to work with SAMs even as old as five to ten years. Of course, the use to which the CGE model is put also dictates how recent a SAM should be. SAMs are ideally produced by statistics offices of countries, but it is common to find many other organisations producing SAMs. Such SAMs, not produced by the statistics office, have to be clearly documented and must use official statistics. Ideally, SAMs should be produced every two to three years, but because of their huge data requirements, this may not be possible in government statistics offices.

650 The CGE model is designed to mimic the working of a specific economy by incorporating various institutional and structural characteristics such as rigidities and constraints in different markets that simple theoretical analysis fails to capture. The model is later used to develop several counterfactual - "*what if*"- scenarios and provides interesting insights on the likely impact of macroeconomic shocks and policies on economic performance and income distribution. Generally, it is for this reason that even if a SAM is a number of years old, it can still be acceptable as discussed above.

651 In general, CGE models are used when proposed policy measures, or expected changes in exogenous conditions, are likely to have general equilibrium effects, that is, significant indirect effects with potentially strong effects on the allocation of scarce resources. Results generated by CGE models are often questioned on their validity and usefulness for policy decision making because of their underlying assumptions on optimising behaviour, competitive markets and flexible relative prices.

652 However, the strength of this class of economic models is its consistency with generally accepted microeconomic theory, significant structural detail, and the nature of general equilibrium, that is, changes in any one area of economic activity may have measurable impact in other areas. The lack of econometric estimation of key supply and demand parameters is also a weakness of CGE models. This call, for undertaking significant sensitivity analysis to see which parameters are key in altering the results and their interpretation (van der Mensbrugge, 1998).

653 As pointed out by Bergman and Henrekson (2003), "even if uncertainty about the numerical values of key parameters makes the magnitude of computed effects of policy

changes uncertain, the analyst may be able to safely conclude that the effects in question are small or big”.

#### 6.4.2 Why and When to Use CGE Models

654 One of the major reasons for using CGE models is that they provide a means for quantifying the costs and benefits of policy. They simulate (changing the values of exogenous variables and model parameters) the impact of a shock (non-marginal change of a policy variable) on the development of relative prices which leads to a new equilibrium with different levels of production, income and consumption. In most interesting policy settings, economic theory provides guidelines for judging if a policy will be beneficial or not for some households, but often cannot give a definite quantitative answer. CGE models allow parameter estimation evaluation so that a definite answer or number is provided.

655 Partial equilibrium analysis illustrates results for one market at a time. However, there often exist market interactions and thus, market feedback. Pricing outcomes in one market usually have effects in other markets, and these effects, in turn, create ripples throughout the economy, perhaps even to the extent of affecting the price-quantity equilibrium in the original market. To represent this complex set of economic relationships, it is necessary to go beyond partial equilibrium analysis and construct a model that permits viewing many markets simultaneously.

#### 6.4.3 Theoretical General Equilibrium Model

656 The basic theory behind Walrasian general equilibrium theory found in most microeconomics books is best illustrated using an example. Let us assume that there exists:

- i. One production sector producing output ( $q_s$ ).
- ii. Two factors of production, labour ( $L$ ) and capital ( $K$ ).
- iii. One representative consumer who owns all factors and consumes the goods produced.

657 Starting on the production side, we assume that the firm faces a Cobb Douglas, constant returns to scale production function where technical progress ( $A$ ) can affect productivity. The firm is assumed to maximise profit subject to the production function by choosing optimal combinations of labour and capital and taking all prices as given:

$$\text{Max } \{ \text{profit} = pX - wL - rK \quad \text{s.t.} \quad X = AL^a K^{1-a} \} \quad (6.14)$$

Where  $p$  is output price,  $w$  is wage rate and  $r$  is return on capital.  $A$  is a technical progress parameter and  $a$  is a share parameter.

658 This maximization yields factor demand functions of the form:

$$L = \left[ \frac{\alpha p}{w} \right] X \quad K = \left[ \frac{(1-\alpha)}{r} \right] X$$

659 We assume exogenous factor supplies so that:

$$L_S = \bar{L} \quad (4) \quad K_S = \bar{K}$$

660 The sole household owns all the labour and capital and supplies it to the firm in exchange for wage ( $w$ ) and profit ( $r$ ). The household spends all its earnings on the single goods ( $y_D$ ).

$$y = wL + rK \quad y_D = \frac{y}{P}$$

661 We have 3 markets: labour market, capital market and commodity market. We assume that all markets clear by imposing the equilibrium condition for each market:

$$L_S = L_D \quad K_D = K_S \quad y_D = y$$

662 This completes the specification of the model. Adding things up, we notice that we have 7 equations and 7 variables ( $y, q, p, k, l, w, r$ ). However, one of them is redundant given Walras' Law which establishes that for  $n$ -markets, we need only  $n-1$  equilibrium conditions. Thus, any of the market clearing conditions can be dropped! Also, the model is homogeneous of degree zero in prices. To assure that only one solution exists, we need to choose one price as the numeraire (means we fix one of the prices, for example  $p$ ). Choosing a numeraire and deleting the corresponding good market clearing equation leaves us with 6 equations and 6 variables, which is a well-defined model. Given the definition of the price normalisation equation, all simulated price changes can be directly interpreted as changes vis-à-vis the wage or CPI, etc.

#### 6.4.4 CGE Models

663 So far, we have described a very small model. However, applied economy-wide models tend to be large, thus making the use of computational techniques unavoidable. Not all of the parameters in the model can be determined independent of the restrictions imposed during specification of a model - yet all parameters must be assigned appropriate values. A number of ways suggest themselves: guesstimate, econometric guess or *calibration* (a.k.a. point estimation).

664 As Shoven and Whalley (1992) phrase it: "Numerical, empirically based general equilibrium models can be used to evaluate concrete policy options by specifying

production and demand parameters and incorporating data reflective of real economies” and “Calibration is most easily understood as the requirement that the entire model specification be capable of generating a base-year equilibrium observation as a model solution”. This implies that the model should be able to calculate the initial situation, using the data for the base year and the calibrated parameter values. If this replication check fails, there must be an error in the model specification. For the calibration of a comparative-static<sup>5</sup> CGE, three types of data are needed: (i) data describing the initial situation (the current allocation of resources); (ii) data describing the impulse; (iii) data describing the reaction of agents to changes in the circumstances.

665 The first type of data needed in CGEs is the data describing the *initial situation*. This data is usually taken from the National Accounts and describe a historical year. It includes expenditure by the production sectors and households on the various goods and the division of the production factors over the producers.

666 The data describing the *impulse* is the quantitative representation of some new policy, such as the change in a tax rate. Alternatively, the impulse can describe a change in the availability of one or more of the endowments, like an increase in total labour supply.

667 Finally, data is needed that describe the *reactions of the agents* to the impulse given. These reactions are usually specified in terms of elasticities. For the households, price and income elasticities govern the change in the demand for goods if a price changes or if income changes. Note that the demand for one type of goods may not only change due to changes in the price of the same goods, but also due to changes in the prices of other goods (so-called cross-price elasticities).

668 The *initial situation* described above is often described by the SAM that we have discussed already. Each cell in a SAM, which is a transaction, can be thought of as the outcome of an underlying optimisation problem of the relevant institution(s). We can represent the flow in the cell as follows;

$$t_{ij} = t(p, q; V, \theta)$$

where  $\mathbf{p}$  and  $\mathbf{q}$  are, respectively, vectors of relative prices (for goods and factors) and quantities. The vector  $V$  is a vector of exogenous factors and  $\theta$  is a vector of parameters defining the relevant functional form. A CGE model is simply a formalisation of this general representation of each combination together with the sectoral and macroeconomic balance constraints defined in the SAM. CGE models differ in the way in which we define the structure of  $t_{ij} = t(p, q; V, \theta)$ .

<sup>5</sup> The term ‘comparative’ is used to indicate that two distinct equilibria are compared.



### 6.4.5 Short Run CGE Model

669 In this section, we give illustrations of a short run model in which capital is immobile across sectors. The overall structure of the model can be read from Panel 1 below. On the production side, sectoral output is produced using a Cobb Douglas function [equation 1 in the pane A]. Maximising profit subject to this production function yields the labour demand function in [equation 2]. [Equation 3] is the return on capital. Because capital is fixed, the capital factor market equilibrating variable is its price, that is  $r_j$ . This rate of return is determined as a residual (what is left of firm revenue after deducting labour costs expressed as a function of capital stock). Next, we assume that all factor income accrue to the household in [equation 4]. Note that labour income is  $w \sum_j LD_j$  because the wage rate is the same across all activities because if there is any wage differentials, labour will migrate to compete those differentials. However, for capital, since it is sector specific and hence immobile, its price will differ by sector. This is taken into account in the income equation by summing the value of capital (note it is not factored out as was done with  $w$ ), i.e.,  $\sum_j r_j KD_j$ .

670 The household allocates its income so as to maximise utility taking prices as parameters. There are two demand functions, one for goods/sector as shown in [equations 5 and 6]. The functional form shown in these equations comes from assuming that the utility function is Constant Elasticity of Substitution (CES). The parameter in these functions is known as the elasticity of substitution.

671 There are two markets, a commodity market and a factor market. Each market is in equilibrium at the point where demand and supply coincide. This is represented by equations 7 and 8. Note that we do not specify the capital market in these equilibrium conditions - this is because capital is fixed and so capital supply and demand are the same. As soon as we make capital mobile, we should also expect a capital market equilibrium condition.

672 As modelers, we are always careful about the structure of our model. This helps us to be consistent with our modeling and also to ensure that our models are solvable. To ensure the latter, we track closely the number of equations and variables. According to our count performed in Panel A below, we have 12 equations and 12 variables. Everything seems okay but not quite. The problem is that Walras Law tells us that if there are  $n$  markets and  $n-1$  of them are in equilibrium, then the  $n$ th one must automatically be in equilibrium. Technically, this means one of the equilibrium conditions is redundant and can be dropped. But this upsets our equation equals variables condition. We have one less equation than variables. Also, we should note that this model is *homogeneous of degree zero* in prices i.e. a proportional change of all prices would not affect the solution of the real variables. Hence, the price level cannot be determined as there is an infinite

combination of prices which are consistent with a solution to the model. It is therefore necessary to impose a restriction which ensures a unique solution.

673 There are 2 ways which are common in modeling that are used to get around these technical problems. The first is that we define a numeraire and fix it, which is tantamount to dropping one variable - as a result equations and variables equality is restored. All experiments results should then be interpreted as relative price change vis-à-vis the exogenised price. Note that the choice of numeraire does not affect the solutions of the model as far as the real variables are concerned.

674 Another approach is to keep the redundant equation in but add a new variable called Walras or Leon. This variable is used to check Walras Law: if it is zero then the law is confirmed and your model is likely to be correctly specified.

#### 6.4.6 Extensions of CGE Models

675 Static CGE models described above have the disadvantage that they do not consider accumulation effects. As a result, they cannot say anything about the effects of transitory policies. One needs a dynamic CGE for this kind of analysis. Generally there are two types of dynamic CGE models, each premised on different assumptions on expectations and hence, giving different equilibrium concepts. The most common of the two is recursive dynamic CGE model. In this model the economy is a sequence of inter-periodic temporary equilibria. Essentially, these are "lurching equilibria," where *within* period (static) equilibrium is attained first, then the model, "lurches" forward to the next period, and another (static) equilibrium is attained. These are usually obtained by defining stock variation of both capital and population.

676 A second type of dynamics specification is via the inclusion of inter-temporal dynamics with agents that are forward-looking. In this kind of dynamic CGE model, consumers are assumed to maximise inter-temporal utility and producers are assumed to maximise profits subject to the usual technical constraints. Thus, consumers can take actions to prevent future adverse effects or to take advantage of the future, hence more realistic than recursive dynamic models. One can specify two different types of behavior on the supply side. One is to use the firm's optimisation function to get an inter-temporal investment equation - as in Tobin's  $q$  investment theory. These are thus, representations of the neoclassical CGE growth models with exogenous technical progress. The second specification is grounded in the new growth theory with endogenous technical progress. As these dynamic models are even more demanding in terms of data, they are also not as commonly uncounted as static models.

#### 6.4.7 Advantages and Disadvantages of CGE Models

677 One of the main advantages of CGEs is that they describe the entire economy (in contrast to partial equilibrium models, which describe one or a few markets). All the

economic agents are taken into account. Another major advantage of CGE models is that they have an elaborate microeconomic foundation (neo-classical theory). Using well-known theories on firm and household behaviour at the level of individual agents, a model for the whole economy can be derived. This is a property of CGEs that neo-Keynesian models and input-output models lack. Common applications of CGEs are tax reforms, trade liberalisation, economic development, and distributional and allocational impact of policy measures. A third advantage of CGEs is that, compared to input-output models, the variables that represent the behaviour of firms and households are endogenous: agents respond to price changes by changing their production or consumption patterns. Moreover, CGE models use resource constraints in their description of markets. The extension to sophisticated analysis is therefore straightforward.

678 Until recently, the main disadvantage of CGEs was that they could not be solved numerically (and the models are too complex to solve analytically). One needed good computers to calculate the model results, as the CGE models are large systems of highly non-linear interdependent equations. However, with the development of capable solvers and the increase in computational power of personal computers, numerically solving CGEs is nowadays not much of a problem. Another aspect of CGEs is that they require a lot of data: not only is data needed for the initial situation, but also for the reactions of the firms and households to the policy changes. This data is often hard (though not impossible) to find.

#### 6.4.8 Solving the Models Using Software

679 The analysis of economy-wide models is a particularly demanding topic in policy analysis, since it involves the study of interdependence. It implies a move to the realm of multiple heterogeneous agents, sectors and institutions interacting in complex ways. While there are some analytical methods and results available to help us in such endeavor, computational methods become necessary when we move from medium to large size models or when we have to deal with particularly complex ones. Our goal here is to provide a basic introduction to the art of economy-wide modeling. Having presented a sequence of small models, we now show how to implement them in Excel and GAMS and we perform and suggest some experiments. We start with an introduction to Excel and then show step by step how the IO model and SAM models can be solved using this software. We then move to introduce GAMS and show how a General Equilibrium model in which prices and quantities are determined simultaneously given technology, preferences and endowments can be solved.<sup>6</sup>

680 Finally, we refer the interested reader to real life examples of IO, SAM and CGE models of four (4) MEFMI member countries, namely Botswana, Namibia, Lesotho and

---

<sup>6</sup> We will present models in a sequence reflecting mainly their computational complexity in terms of degree of non-linearity and size. The order of the sequence does not mean historical or theoretical precedence of one type of model over the others, or a ranking of practical relevance.

Swaziland that were prepared by Rob Davies and commissioned by the World Bank, and is provided as an accompanying CD to this *manual* resource.

### 6.4.8.1 Introduction to Basic Excel Operations

681 We start with relatively simple introduction to Excel as a way of beginning with solving the IO and SAM models introduced in this chapter. Excel is a software paradigm that is well known and accessible to almost everybody. It is useful to solve small models that do not involve simultaneous nonlinear optimization solver which can handle constrained optimisation problems and is very handy to set up and solve interesting models. This software system is useful to deal with models or problems involving intensive use of vector and matrix operations. Three basic techniques in Excel are required to solve the IO and SAM models just discussed; namely (a) entering matrices and simple matrix operations (addition and subtraction), (b) matrix multiplication, and (c) matrix inversion. Below is a step-by-step illustration on each of these and how they can be used to solve the small hypothetical IO Model and SAM Models introduced earlier.

#### A. Matrices Entry and Basic Matrix Operations

Setting up a matrix in Excel is easy. You simply enter each entry in a cell. Let us illustrate using two hypothetical examples below, namely Matrix A and B:

A =

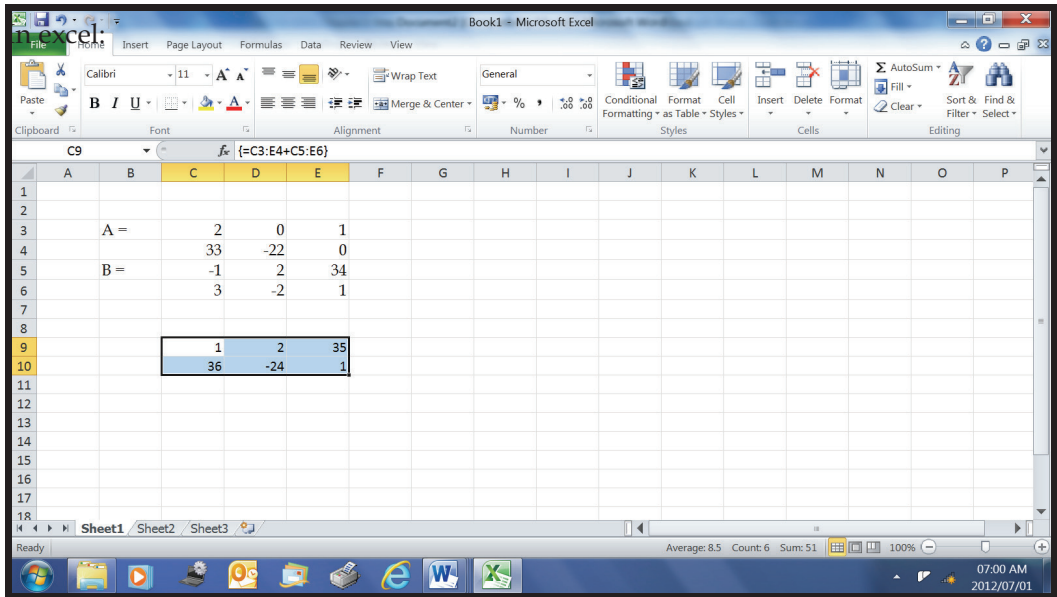
2	0	1
33	-22	0

B =

-1	2	34
3	-2	1

682 To Add or Subtract Matrices, it is instructive to use whole arrays as illustrated in the interface below. The procedures are as follows:

1. Highlight the whole answer array (highlighted in blue by cells C9:E10).
2. Type "=", highlight the whole of *A*, type "+", then highlight the whole of *B*.
3. Instead of hitting the "Enter" button, hit simultaneously "Ctrl Shift Enter".



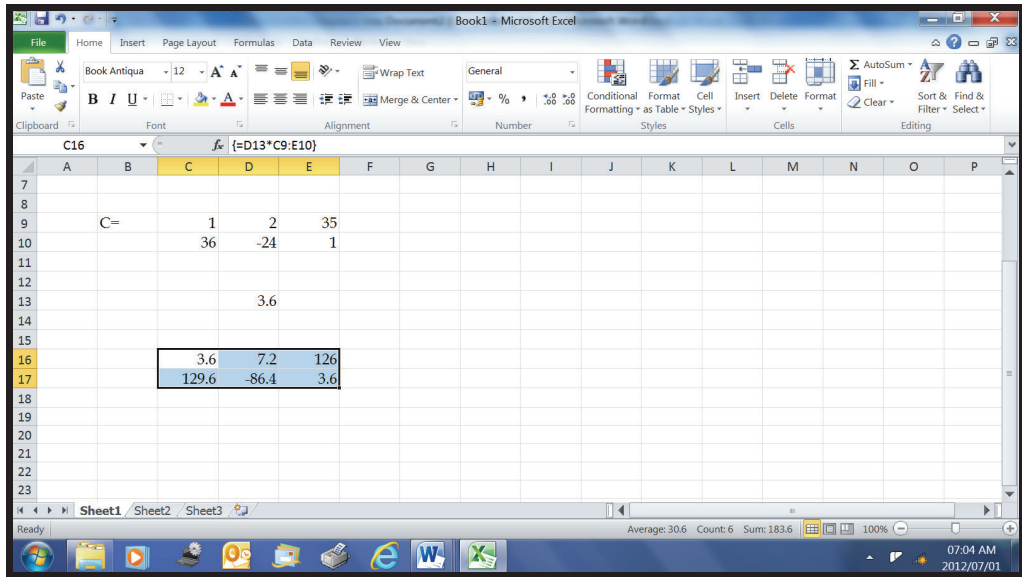
4. You should get the following answer in the area you highlighted for the answer:

$$A + B = C =$$

1	2	35
36	-24	1

683 For scalar multiplication, let's suppose we want to compute  $3.6 \cdot C$  where  $C$  is the matrix that we have just computed. Proceed as follows:

1. Highlight the whole answer array 2.  
Type " $=3.6 \cdot$ ", highlight the whole of array  $C$ , hit simultaneously "Ctrl Shift Enter"



3. You should get the following answer:

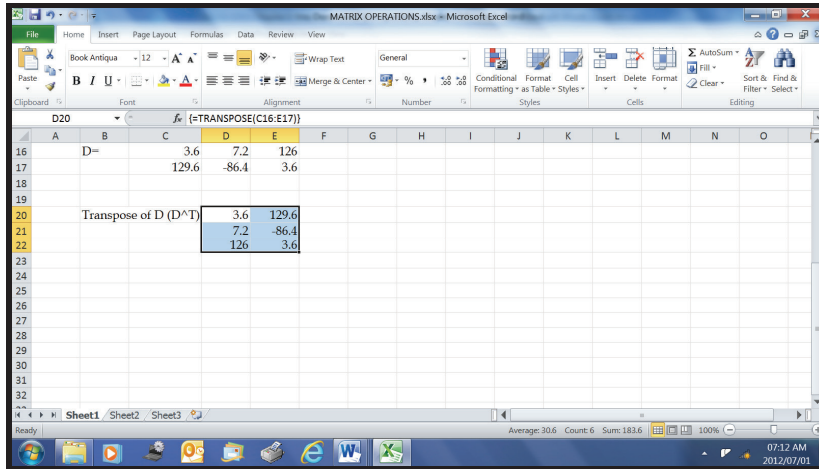
$$3.6 * C =$$

3.6	7.2	126
129.6	-86.4	3.6

684 Finally, for transposing a matrix  $A^T$ , let's suppose we want to transpose matrix D which we have just computed ( $3.6 * C$ ). Proceed as follows:

1. First select an appropriately sized block for the answer, e.g., the transpose of a  $2 \times 3$  matrix is a  $3 \times 2$  matrix.
2. Highlight the whole answer array.
3. Type " $=\text{TRANSPOSE}(",$  highlight the whole of D and type  $)"$ .
4. Simultaneously press "Ctrl Shift Enter".

685 In excel:



5. You should get the following answer:

$$D^T = \begin{bmatrix} 3.6 & 129.6 \\ 7.2 & -86.4 \\ 126 & 3.6 \end{bmatrix}$$

## B. Matrix Multiplication

686 To find the product AB of a row A and a column B in Excel,

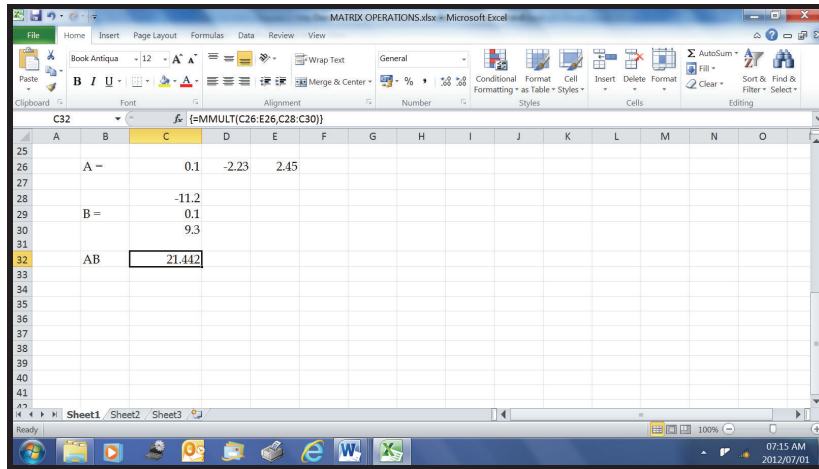
$$A = \begin{bmatrix} 0.1 & -2.23 & 2.45 \end{bmatrix}$$

$$B = \begin{bmatrix} -11.2 \\ 0.1 \\ 9.3 \end{bmatrix}$$

687 Proceed as follows:

1. Highlight a cell where you want the product.
2. Type “=MMULT(”, and highlight the whole of A, type “,”, then highlight the whole of B, type “)”.
3. Simultaneously press “Ctrl Shift Enter”.

688 In excel:



4. You should get the following answer:

$$AB = 21.442$$

689 For multiplying a matrix by a matrix, consider matrices A (2x4) and B (4x3) given below:

A =

2	0	-1	3
1	-1	2	-2

B =

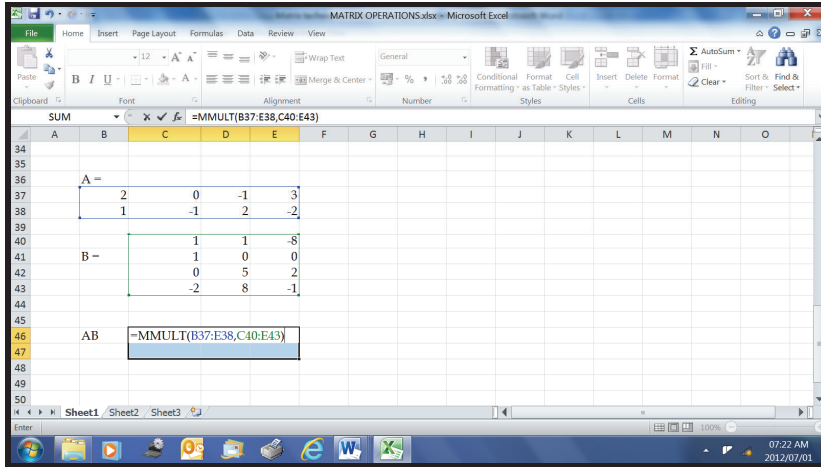
1	1	-8
1	0	0
0	5	2
-2	8	-1

690 The product AB will be a 2x3 matrix (as many rows as A and as many columns as B). Thus, select a 2x3 block for the product and proceed as follows:

1. Highlight the whole selected block.
2. Type "=MMULT(", and highlight the whole of A, type ",", then highlight the whole of B, type ")".
3. Hit simultaneously "Ctrl Shift Enter".



691 In excel:



4. You should get the following answer:

$$AB = \begin{bmatrix} -4 & 21 & -21 \\ 4 & -5 & -2 \end{bmatrix}$$

### C. Matrix Inversion

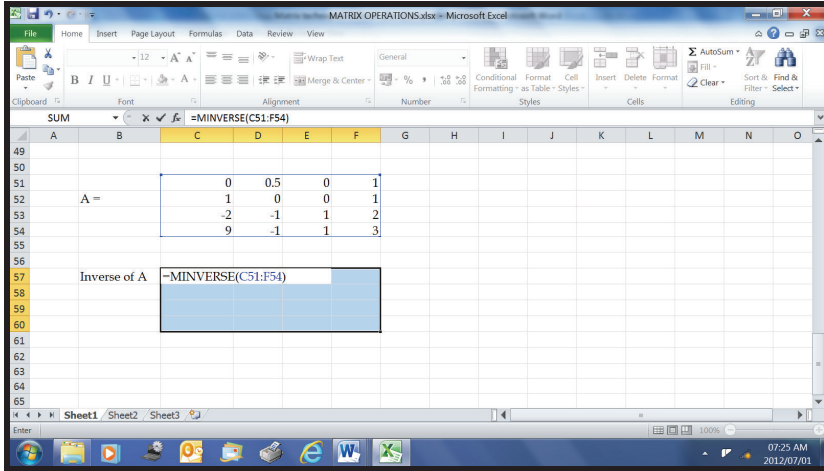
692 We now want to use Excel to invert a given matrix A given below:

$$A = \begin{bmatrix} 0 & 0.5 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ -2 & -1 & 1 & 2 \\ 9 & -1 & 1 & 3 \end{bmatrix}$$

A should be square and non-singular. Proceed as follows:

1. Highlight a block for the inverse (same dimension s the original matrix).
2. Type “=MINVERSE(”, and highlight the whole of A, type “)”
3. Hit simultaneously “Ctrl Shift Enter”.

693 In excel:



4. You should get the following answer:

$$A^{-1} =$$

0	-0.1	-0.1	0.1
2	-2.2	-0.2	0.2
2	-4.6	0.4	0.6
0	1.1	0.1	-0.1

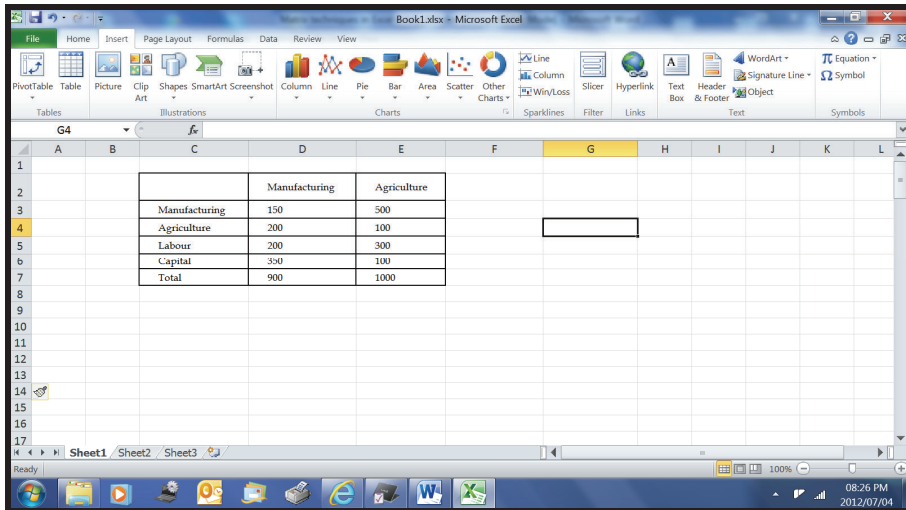
Solving a stylised IO Model in Excel

694 Refer to the example discussed earlier, which is now repeated here. The problem was to find out the production in gross output required to meet an exogenous increase in demand of R10 for manufacturing and R15 for agriculture.

	Manufacturing	Agriculture
Manufacturing	150	500
Agriculture	200	100
Labour	200	300
Capital	350	100
Total	900	1000

695 **Step 1:** The first step, using the techniques learnt above is simply to enter this matrix in Excel. As we saw, this is easy. You simply enter each entry in a cell. This is shown in the interface below:

696 In excel:

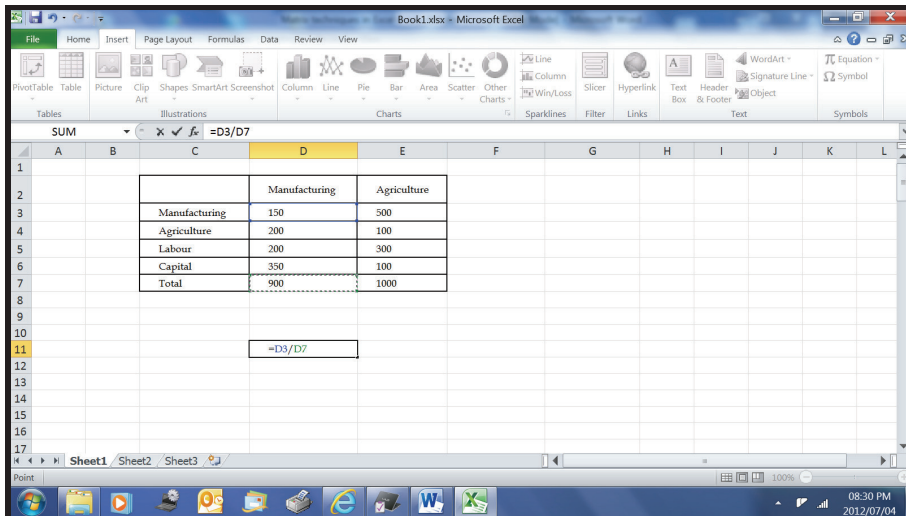


The screenshot shows a Microsoft Excel spreadsheet with a table containing the following data:

	Manufacturing	Agriculture
Manufacturing	150	500
Agriculture	200	100
Labour	200	300
Capital	350	100
Total	900	1000

697 **Step 2:** Next, we want to compute the direct input requirements matrix or the direct coefficient matrix. This requires Matrix division, one of the operations again shown earlier.

698 In excel:



The screenshot shows the same Microsoft Excel spreadsheet as above, but with a formula cell in row 11, column D containing the formula  $=D3/D7$ .

	Manufacturing	Agriculture
Manufacturing	150	500
Agriculture	200	100
Labour	200	300
Capital	350	100
Total	900	1000

Below the table, in cell D11, the formula  $=D3/D7$  is displayed.

699 You get:

The screenshot shows an Excel spreadsheet with the following data in cells C3 to E7:

	Manufacturing	Agriculture
Manufacturing	150	500
Agriculture	200	100
Labour	200	300
Capital	300	100
Total	900	1000

Below this table, in cells C11 to E12, are the following values:

0.166666667	0.5
0.222222222	0.1

700 **Step 3:** Next, we require the identity matrix. Identity matrices for these small examples can be generated by typing 'ones' along the diagonal and 'zeros' elsewhere.

701 In excel:

The screenshot shows the same Excel spreadsheet as above, but with an identity matrix added in cells C10 to E11:

	Manufacturing	Agriculture
Manufacturing	150	500
Agriculture	200	100
Labour	200	300
Capital	300	100
Total	900	1000

Below this table, in cells C10 to E11, is the identity matrix:

1	0
0	1

702 **Step 4:** Use the I matrix and Matrix SUBTRACTION learnt above to compute (I-A)

703 In excel:

The screenshot shows an Excel spreadsheet with the following data and formulas:

	A	B	C	D	E
4			Agriculture	200	100
5			Labour	200	300
6			Capital	350	100
7			Total	900	1000
10			Identity matrix	1	0
11				0	1
13			Coefficient matrix (A)	0.166666667	0.5
14				0.222222222	0.1
18			Identity minus A (I-A)	0.833333333	-0.5
19				-0.222222222	0.9

704 **Step 5:** Invert (I-A) by using Matrix Inversion in Excel introduced earlier. Highlight a block for the inverse, then type “=MINVERSE(”, and highlight the whole of I-A, type “)”. Hit simultaneously “Ctrl Shift Enter”.

705 Proceed as follows in Excel:

1. Highlight a block for the inverse.
2. Type “=MINVERSE(”, and highlight the whole of A, type “)”.
3. Hit simultaneously “Ctrl Shift Enter”.

The screenshot shows the Excel spreadsheet after the matrix inversion. The formula bar shows the formula: `=MINVERSE(D18:E19)`. The resulting inverse matrix is shown in cells D22 and E23:

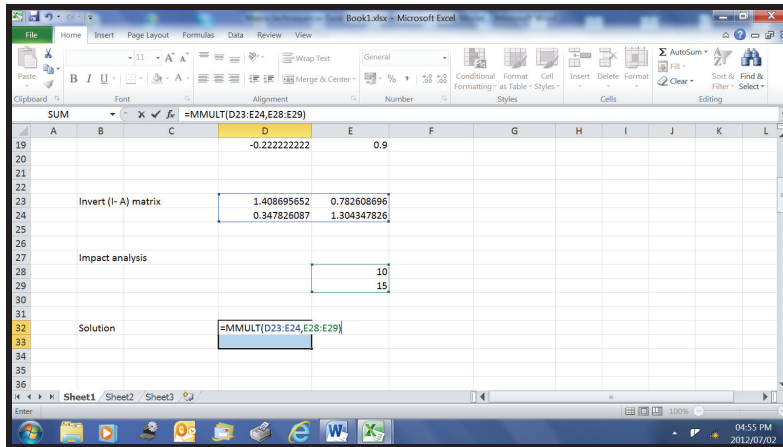
	A	B	C	D	E
13			Coefficient matrix (A)	0.166666667	0.5
14				0.222222222	0.1
18			Identity minus A (I-A)	0.833333333	-0.5
19				-0.222222222	0.9
22			Invert (I-A) matrix	1.408695652	0.782608696
23				0.347826087	1.304347826

706 **Step 6:** Impact Analysis requires matrix multiplication techniques just learnt. We now want to multiply the inverse of  $(I-A)$  just computed and the by exogenous demands of R15 and R10, that is:

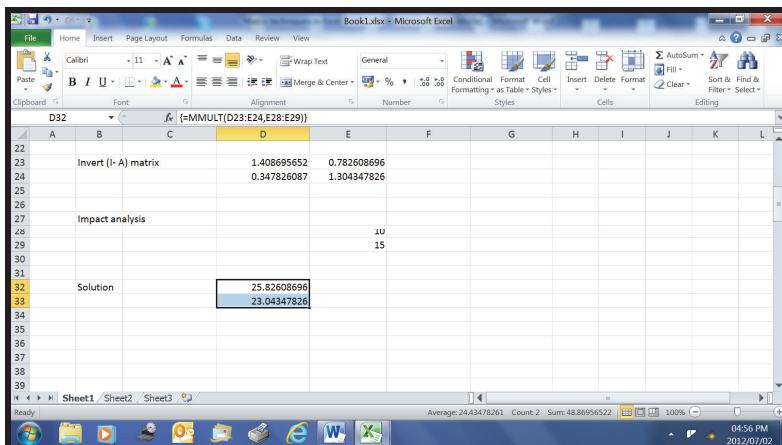
$$\Delta X = (I - A)^{-1} \Delta D$$

707 Proceed as follows in Excel:

1. Highlight a cell where you want the product.
2. Type “=MMULT(”, and highlight the whole of A, type “,”, then highlight the whole
3. of B, type “)”
4. Simultaneously press “Ctrl Shift Enter”.



708 You should see the following in excel:



709 Therefore, to realise the additional consumption requirement of the economy, we need to produce additional R25.8 of manufacturing and R23 of agriculture. This is the *answer to the problem we wanted to solve in Excel*.

### 6.4.8.2 Solving a stylised SAM Model in Excel

710 Using the SAM below let us illustrate how to calculate SAM multipliers in Excel interface.

		PAYMENTS BY				
		ACTIVITIES		Consumer Demand	Other Demand	TOTAL
		Transport	Non Transport		FD	
RECEIPTS BY	Transport	100	200	400	300	1000
	Non Transport	500	150	200	50	900
	HH Income	300	300			600
	Other Payment	100	250			350
TOTAL		1000	900	600	350	

711 **Step 1:** First we compute the augmented coefficient matrix.

712 In excel we use the DIVISION operation introduced earlier:

The screenshot shows an Excel spreadsheet with the following data:

	Transport	Non Transport	Consumer Dem	Other Demand	TOTAL
Transport	100	200	400	300	1000
Non Transport	500	150	200	50	900
HH Income	300	300	0	0	600
Other Payments	100	250	0	0	350
TOTAL	1000	900	600	350	

Augmented coefficient matrix (A)					
Transport	0.1	0.22222222	0.66666667		
Non Transport	0.5	0.16666667	0.33333333		
HH Income	0.3	0.33333333	0		

The Excel interface also shows the formula bar with  $=D47/$D$51$  and the status bar with Average: 0.291338025, Count: 9, Sum: 2.622222222, 100%, 08:28 PM 2012/07/01.

713 **Step 2:** Construct the I Matrix.

714 **Step 3:** Calculate the **(I-A)** Matrix as before in order to get the augmented.

### Leontief Matrix

715 In excel:

	A	B	C	D	E	F
54	Augmented coefficient matrix (A)					
55						
56		Transport		0.1	0.22222222	0.66666667
57		Non Transport		0.5	0.16666667	0.33333333
58		HH Income		0.3	0.33333333	0
59						
60						
61						
62		Identity Matrix (I)		1	0	0
63				0	1	0
64				0	0	1
65						
66		I-A matrix		0.9	-0.22222222	-0.66666667
67				-0.5	0.83333333	-0.33333333
68				-0.3	-0.33333333	1
69						
70						

716 **Step 4:** Calculate the inverse of the **(I-A)** Matrix, that is, the augmented Leontief Matrix using MINVERSE as before.

717 In excel:

	A	B	C	D	E	F
60						
61						
62		Identity Matrix (I)		1	0	0
63				0	1	0
64				0	0	1
65						
66		I-A matrix		0.9	-0.22222222	-0.66666667
67				-0.5	0.83333333	-0.33333333
68				-0.3	-0.33333333	1
69						
70						
71		Inverse I-A		3.023255814	1.860465116	2.635658913
72				2.511627907	2.930232558	2.651162791
73				1.744186047	1.534883721	2.674418605
74						
75						
76						





720 The Transport Income Multiplier is 1.74.

721 You can calculate the remaining output and income multipliers to get the results reported above.

#### 6.4.9 Getting Familiar with GAMS Architecture

722 GAMS was developed at the World Bank by Alexander Meeraus and his colleagues (see reference at the end). GAMS was designed as a "set driven" high level language which would facilitate the development of linear and non-linear programming models of industry and agriculture. Thus, it was not necessary to write out a separate equation for each commodity, time period, plant and market but rather only to create equations and variables indexed on these sets and defined over them. Thus, a model with thousands of equations could be represented in a GAMS statement with only a few set of specifications, variables and equations which might fit on a single page. This not only decreased the tedious labour intensive parts of model development but also substantially decreased the likelihood of errors in the model specification.

723 Later, the language became widely used for almost any computable model which could be expressed in algebra. In particular, there are now many computable general equilibrium models and financial models in GAMS, as well as a wide variety of other types of economic models. For a listing of more than a hundred GAMS models see the GAMS library at: <http://www.gams.com>

724 Rather than giving a comprehensive overview, we will motivate basic GAMS language components and GAMS syntax rules along the model examples from the Reader. In order not to overload the reader with "unnecessary" information we explain the GAMS language just as much as is needed for translating our algebraic statement of problems into (GAMS) computer-readable format. We will use the GAMS implementation of a simple optimisation exercise. We believe that this approach accommodates a straight computational access to learning economywide modeling and does not distract people with programming details.

725 Before we start, we want to recall why we consider GAMS as a very convenient modeling system for doing economy-wide modeling:

- The GAMS syntax is very close to the standard algebraic representation of economic models and largely self-evident. Therefore, the GAMS user will find it quite intuitive to translate algebraic problem into an equivalent GAMS program.

- The use of (algebraic) sets and detached-coefficient matrix notation makes GAMS a very efficient tool for representing large dimensional models using only a few lines of code.
- Separation of logic and data allows that a problem - once correctly specified in logical terms - can be easily scaled up to higher dimensions. This provides an efficient environment for developing large-scale models from little "toy" models.
- The problem (model) formulation under GAMS is independent of the numerical solution. GAMS passes on the problem definition to separate solver programs which return solution values back to the GAMS program. The solution process is a black-box to the model builder who is typically not interested in algorithmical details. The only thing a model builder has to care about is that the solver he selects must be consistent with the mathematical problem type underlying his problem formulation. The commercial solvers available under GAMS allow to tackle types of models belonging to different classes of mathematical programming such as (non) linear programming, (relaxed) mixed integer programming, non-linear systems of equations, mixed complementarity problems, etc.

726 Typically, a model programmed in GAMS can be decomposed into 3 parts corresponding to data entry (Part 1), model specification (Part 2) and solve procedure (Part 3).

1. **Stage 1: Data**
  - *SETS* declaration and definition
  - *PARAMETERS* declaration and definition
  - Data assignment
  - Intermediate displays
2. **Stage 2: Model**
  - *VARIABLES* declaration
  - *EQUATIONS* declaration
  - Equations definition
  - *MODEL* definition
3. **Stage 3: Resolution**
  - *SOLVE* statement
  - Results display

727

Some common rules, caveats and tips:

1. As a general rule, it is necessary to proceed to the declaration before using any element. Sets and parameters must be declared at the very beginning of the program.
2. Although not always necessary, a good habit is to end every statement by a semi-colon in order to avoid unexpected compilation errors.
3. GAMS allows for statements on several lines or several statements on the same line. This property can help to reduce the length of the code or facilitate printing.
4. Capital and small letters are not distinguished in GAMS.
5. Names for Set, Parameter, Variable, Equation or Model must start with a letter and cannot contain more than 9 characters. GAMS keywords cannot be used to declare an element.
6. Finally, the main mathematical functions are described below:

Multiplication	*	Equality in an operation	=	Logarithm	LOG (.)
Subtraction	-	Summation	SUM (set domain, element)	Maximum	MAX (.,.)
Addition	+	Product	PROD (set domain, element)	Minimum	MIN (.,.)
Division	/	Absolute value	ABS (.)		
Exponent	**	Exponential	EXP (.)		

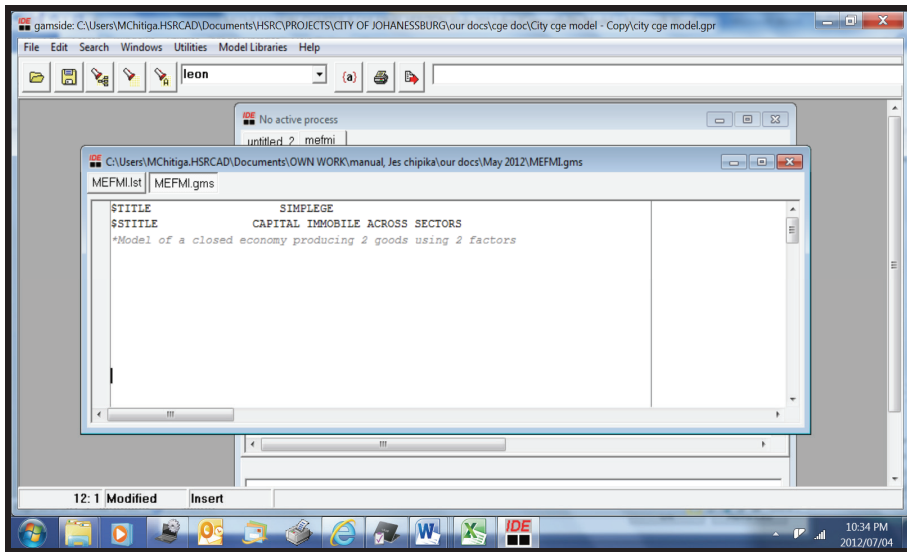
#### 6.4.9.1 Solving General Equilibrium Model in GAMS

728

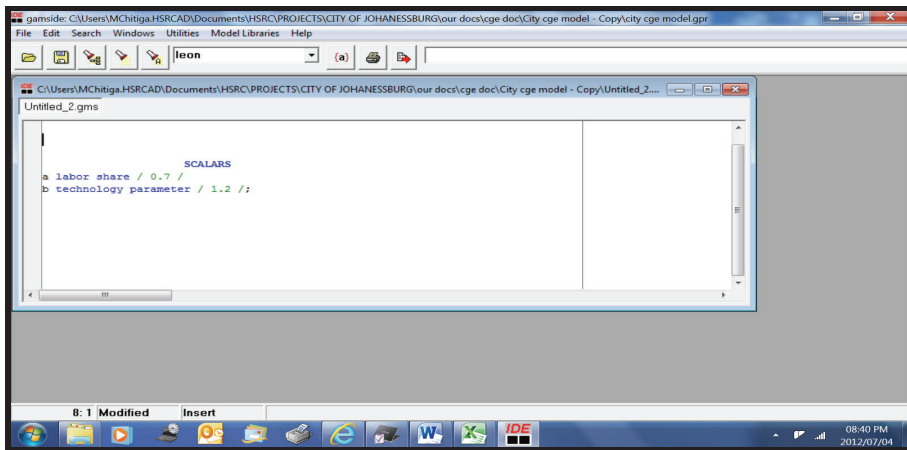
The GAMS representation of the theoretical model presented is shown below. Arbitrary numbers have been chosen for the parameters and for the labour and capital stocks. As illustrated earlier, this simple model has 7 variables and 7 equations. However, one of them is redundant, since “Walras law” establishes that for  $n$ -markets we need  $n-1$  equilibrium conditions only. Also, since this model determines relative prices ( $p$ ,  $w$  and  $r$ ), we need to fix one of them as the numeraire. Thus, by choosing one price as the numeraire (say we fix  $p = 1$ ) and deleting the corresponding good market clearing equation, we are left with a 6-variable 6-equation well defined model. The GAMS representation of the model using the architecture just discussed is shown below. Arbitrary numbers have been chosen for the parameters and for the labor and capital stocks.

729

**Step 1:** First you type the title (using \$TITLE command) and subtitle (\$STITLE command) as shown in the interface. Remember, you can comment out any additional information by starting every line with a star (we have chosen to describe the structure of the model).

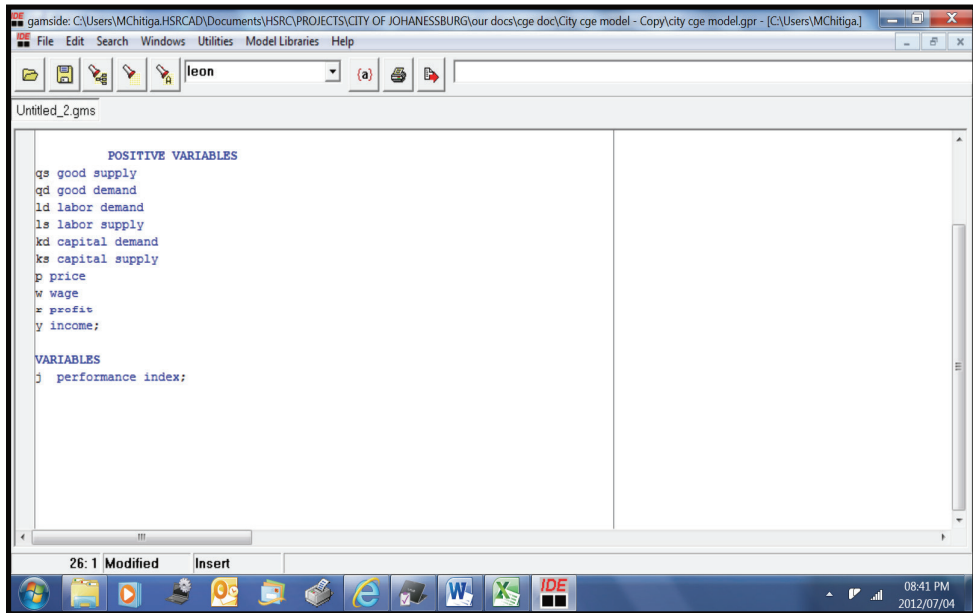


730 **Step 2:** Next you introduce the data into the model. For this small model, its easy to type in the data in various forms such as scalar form since its very simple. For bigger models, you can use GDX interface, or other ways of calling data directly from Excel, see GAMS manual, ([www.gams.com](http://www.gams.com)).



731 After these two steps you are now ready to type in your model, starting with the definitions of variables, then of equations, then typing the actual equations.

732 **Step 3:** Type in the variables in order to introduce these to GAMS and in this step you can also do the calibration of the unknown parameters. After the definition, remember to end with a semicolon. Notice that to avoid negative values which make no economic sense we had to define all variables but the performance index as positive variables.



The screenshot shows the GAMS IDE window titled 'Untitled\_2.gms'. The code defines positive variables and a variable:

```

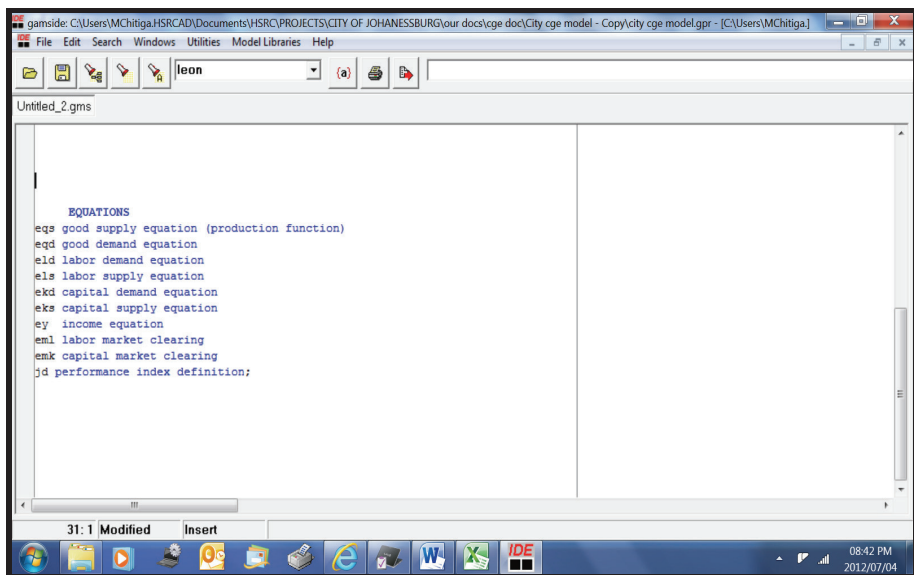
POSITIVE VARIABLES
qs good supply
qd good demand
ld labor demand
ls labor supply
kd capital demand
ks capital supply
p price
w wage
z profit
y income;

VARIABLES
j performance index;

```

The status bar at the bottom indicates '26: 1 Modified Insert' and the system clock shows '08:41 PM 2012/07/04'.

733 **Step 4:** Define the equations, remembering to end with a semicolon!



The screenshot shows the GAMS IDE window titled 'Untitled\_2.gms'. The code defines equations for the model:

```

EQUATIONS
eqs good supply equation (production function)
eqd good demand equation
eld labor demand equation
els labor supply equation
ekd capital demand equation
eks capital supply equation
ey income equation
eml labor market clearing
emk capital market clearing
jd performance index definition;

```

The status bar at the bottom indicates '31: 1 Modified Insert' and the system clock shows '08:42 PM 2012/07/04'.

734 **Step 5:** Finally, you type in the equations ending each equation with a semicolon.

```

jd..      j =E= 0;

eqs..    qs =E= b * ld**a * kd**(1-a);
eld..    ld =E= a * qs * p / w;
els..    ls =E= 2;
eml..    ld =E= ls;
ekd..    kd =E= (1-a)* qs * p / z;
eks..    ks =E= 1;
emk..    kd =E= ks;
sy..     y =E= v * ld + x * kd;
eqd..    qd =E= y / p;

*lower bounds to avoid division by zero
p.lo = 0.001; w.lo = 0.001; z.lo = 0.001;

```

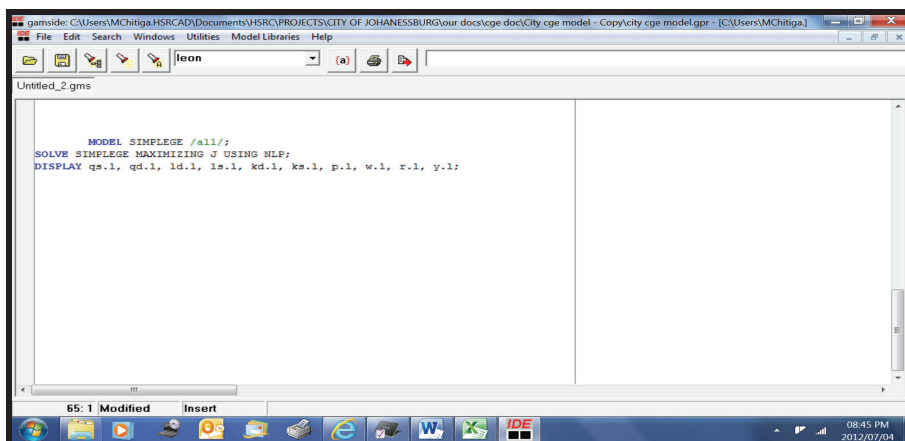
735 **Step 6:** You now need to set the numeraire. In this case we chose one price as the numeraire (i.e., we fix  $p = 1$ ) using the `.fx` command. It is here where you can also define closure rules for larger complex models.

```

*numeraire
p.fx = 1;

```

736 **Step 7:** You now declare to GAMS the name of the model, equations and which algorithm to use in solving the model. In this case, the name of the model is “SIMPLEGE”, and GAMS has been requested to solve it using Non-Linear Programming (NLP). Note, we have added a display command that you can use to view the results of the variables of interest, particularly the endogenous variables in the model. Again, for bigger models, it is advisable to send the results directly to Excel using for example GDX command.

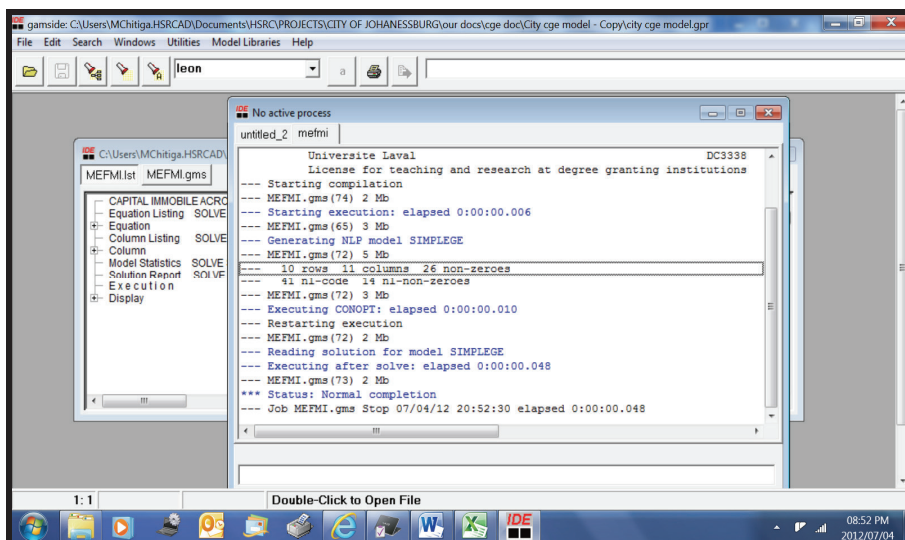


```

MODEL SIMPLEGE /all/;
SOLVE SIMPLEGE MAXIMIZING J USING NLP;
DISPLAY qs.1, qd.1, ld.1, ls.1, kd.1, ks.1, p.1, w.1, r.1, y.1;

```

737 **Step 8:** Finally, you are now ready to solve the model using GAMS. This is done by clicking the red arrow button or pressing F9 on your keyboard. You should get the following report if your model has solved correctly:



```

Universite Laval DC3338
License for teaching and research at degree granting institutions
--- Starting compilation
--- MEFMI.gms (74) 2 Mb
--- Starting execution: elapsed 0:00:00.006
--- MEFMI.gms (65) 3 Mb
--- Generating NLP model SIMPLEGE
--- MEFMI.gms (72) 5 Mb
--- 10 rows 11 columns 26 non-zeros
--- 41 nl-code 14 nl-non-zeros
--- MEFMI.gms (72) 3 Mb
--- Executing CONOPT: elapsed 0:00:00.010
--- Restarting execution
--- MEFMI.gms (72) 2 Mb
--- Reading solution for model SIMPLEGE
--- Executing after solve: elapsed 0:00:00.048
--- MEFMI.gms (73) 2 Mb
*** Status: Normal completion
*** Job MEFMI.gms Stop 07/04/12 20:52:30 elapsed 0:00:00.048

```



738 The results are as follows:

	LOWER	LEVEL	UPPER	MARGINAL
--- VAR qs	.	1.949	+INF	.
--- VAR qd	.	1.949	+INF	.
--- VAR ld	.	2.000	+INF	.
--- VAR ls	.	2.000	+INF	.
--- VAR kd	.	1.000	+INF	.
--- VAR ks	.	1.000	+INF	.
--- VAR p	1.000	1.000	1.000	EFO
--- VAR w	0.001	0.682	+INF	.
--- VAR r	0.001	0.585	+INF	.
--- VAR y	.	1.949	+INF	.
--- VAR j	-INF	.	+INF	.

qs good supply  
qd good demand  
ld labor demand  
ls labor supply  
kd capital demand

739 The solution values for the general equilibrium model that we have been looking for are therefore:

qs.L = 1.949 good supply  
qd.L = 1.949 good demand  
ld.L = 2.000 labor demand  
ls.L = 2.000 labor supply  
kd.L = 1.000 capital demand  
ks.L = 1.000 capital supply  
p.L = 1.000 price  
w.L = 0.682 wage  
r.L = 0.585 profit  
y.L = 1.949 income

740 It is important to perform some basic checks on the workings of the model. For instance, since we assumed market clearing, we have to verify that supply equal demand in each market. Also, when increasing the value of the numeraire, all quantity variables should remain the same, while nominal variables (prices and income) should increase proportionally.

741 Some interesting experiments can be performed with this model. For example, the economy-wide effects of technological progress can be simulated by increasing the value of the  $b$  parameter. Also, you could change the supply of labor or the supply of capital and see how wage and profits are affected. If you do so, you will observe that, in

the simple model that quantities do not change, only the wage and the profit rate do. Quantities would change if you specify elastic labor and capital supply functions, instead of the fixed supplies that have been assumed. As an exercise, you could add a few lines to run a simulation increasing labour supply by 10%, say due to increased immigration or education policy.

---

---

**SELECTED BIBLIOGRAPHY**

- Adams, F. G., & Vial, J. (1991). Comparisons of macroeconomic models of developing countries. In L. R. Klein (Ed.), *Comparative performance of US econometric models* (pp. 260–285). Oxford: Oxford University Press.
- Adelman, I. and Robinson, S. "The Application of General Equilibrium Models to Analyze U.S. Agriculture". Division of Agriculture and Natural Resources, Working Paper No. 423, 1986.
- Anderson, L. C., & Carlson, K. M. (1970). A monetarist model for economic stabilisation. *Federal Reserve Bank of St. Louis Review*, 52(4), 7–25.
- Backus, D.K., Kehoe, P.J. and Kydland, F.E. (1994), "Dynamics of the Trade Balance and the Terms of Trade: The J-Curve ?," *American Economic Review*, 84, 84-103.
- Baillie, R.T. (1996), "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5-59.
- Bautista, R. M. (1988). Macroeconomic models for east Asian developing countries. *Asian-Pacific Economic Literature*, 2(2), 1–25.
- Behrman, J. R., & Hanson, J. A. (1979). *Short-term macroeconomic policy in Latin America*. Cambridge, MA: Ballinger.
- Bergstrom, A. R. (1990). *Continuous time econometric modelling*. Oxford: Oxford University Press.
- Blanchard, O. J., & Quah, D. (1989). The dynamic effects disturbances. *The American Economic Review*, 79(4), 655–673.
- Bodkin, R. G. (1988). A survey of non-Dutch European macroeconomic models: Some international perspective. In W. Driehuis, M. M. G. Fase, & H. D. Hartog (Eds.), *Challenges for macroeconomic modelling*. Amsterdam: Elsevier Science Publishers.
- Bodkin, R. G., & Marwah, K. (1988). Trends in macroeconomic modelling: The past quarter century. *Journal of Policy Modeling*, 10(2), 299–315.
- Bodkin, R. G., Klein, L. R., & Marwah, K. (1986a). Keynes and the origins of macroeconomic modelling. *Eastern Economic Journal*, XII(4), 442–450.
- Bodkin, R. G., Klein, L. R., & Marwah, K. (1986b). Macroeconomic modelling as a background to development planning. *International Journal of Development Planning Literature*, 1(1), 39–56.
- Bodkin, R. G., Klein, L. R., & Marwah, K. (1991). *A history of macroeconomic model-building*. Aldershot: Edward Elgar.
- Bollerslev, T., Chou, R.Y., Kroner, K.F. (1992), "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5-59.
- Bollerslev, T., Engle, R.F. and Nelson, D.B. (1994), "ARCH Models," in R. Engle and D. McFadden (eds.), *Handbook of Econometrics*, Vol. 4. Amsterdam: North-Holland.
- Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (1994), *Time Series Analysis, Forecasting and Control*, Third Edition. Englewood Cliffs, N.J.: Prentice Hall.

Brayton, F., Levin, A. Tryon, R. and Williams, J.C. (1997), "The Evolution of Macro Modeling at the Federal Reserve Board," *Finance and Economics Discussion Series no. 1997-29*, Federal Reserve Board, Washington, DC.

Burns, A.F. and Mitchell, W.C. (1946), *Measuring Business Cycles*. New York: *National Bureau of Economic Research*.

Campbell, J.Y. and Mankiw, N.G. (1987), "Are Output Fluctuations Transitory?," *Quarterly Journal of Economics*, 102, 857-880.

Canova, F., Finn, M., and Pagan, A.R. (1994), "Evaluating a Real Business Cycle Model," in C.P. Hargreaves (ed.), *Nonstationary Time Series and Cointegration*. Oxford: Oxford University Press.

Capros, P., Karadeloglou, P., & Mentzas, G. (1990). An empirical assessment of macroeconomic and CGE approaches in policy modelling. *Journal of Policy Modeling*, 12(3), 557-585.

Challen, D. W., & Hagger, A. J. (1983). *Macroeconometric systems: Construction, validation and applications*. London: Macmillan.

Charemza, W. W., & Deadman, D. F. (1992). *New directions in econometric practice*. Aldershot: Edward Elgar.

Chiang, A. (1984), *Fundamental Methods of Mathematical Economics*, Third Edition, McGraw-Hill.

Chow, G.C. and Kwan, Y.K. (1997), "How the Basic RBC Model Fails to Explain U.S. Time Series," *Journal of Monetary Economics*.

Christiano, L.J. and Eichenbaum, M. (1992), "Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations," *American Economic Review*, 82, 430-450.

Clements, M P and Hendry, D F (2002), 'Pooling of forecasts, *Econometrics Journal*, Vol. 5, pages 1.26.

Cooley, T.F. (ed.) (1995), *Frontiers of Business Cycle Research*. Princeton: Princeton University Press.

Corden, M. (1985). *The relevance for developing countries of recent developments in macroeconomic theory*. Cambridge, MA: Harvard University Press.

Darnell, A. C., & Evans, J. L. (1990). *The limits of econometrics*. Aldershot: Edward Elgar.

Defourney, J. and E. Thorbecke. "Structural Path Analysis and Multiplier Decomposition within a Social Accounting Matrix Framework." *The Economic Journal* vol. 94, pp. 111-136, March 1984.

Diebold, F. (1998). The past, present and future of macroeconomic forecasting. *Journal of Economic Perspectives*, 12(2), 175-192.

Diebold, F.X. (1998), *Elements of Forecasting in Business, Economics, Government and Finance*. Cincinnati, Ohio: South-Western College Publishing.

Diebold, F.X. and Kilian, L. (1997), "Measuring Predictability: Theory and Macroeconomic Applications," Manuscript, Department of Economics, University of Pennsylvania.

Diebold, F.X. and Rudebusch, G.D. (1989), "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics*, 24, 189-209.

Diebold, F.X. and Rudebusch, G.D. (1996), "Measuring Business Cycles: A Modern Perspective," *Review of Economics and Statistics*, 78, 67-77.

- Diebold, F.X. and Rudebusch, G.D. (1998), *Business Cycles: Durations, Dynamics, and Forecasting*. Princeton: Princeton University Press.
- Diebold, F.X. and Senhadji, A. (1996), "The Uncertain Unit Root in U.S. GNP: Comment," *American Economic Review*, 86, 1291-1298.
- Diebold, F.X., Ohanian, L. and Berkowitz, J. (1997), "Dynamic Equilibrium Economies: A Framework for Comparing Models and Data," *National Bureau of Economic Research Technical Working Paper No. 174*, 1995, revised April 1997.
- Dixon, P. B., & Malakellis, M. (1995). Investment behaviour in the MONASH model of the Australian economy. In T. Vlacic, T. Nguyen, & D. Cecez-Kecmanovic (Eds.), *Modelling and control of national and regional economies*. Postprint volume from an IFAC Symposium, Gold Coast, Australia, 2-5 July 1995.
- Doan, T., Litterman, R. and Sims, C.A. (1984), "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1-144.
- Duesenberry (1947), *Income, Savings and the Theory of Consumer Behaviour*. Harvard University Press, Cambridge, Mass.
- Eckstein, O. (1983). *The DRI model of the U.S. economy*. New York: McGraw-Hill.
- Eckstein, Z. and Wolpin, K.I. (1989), "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models," *Journal of Human Resources*, 24, 562-598.
- Economic Commission for Asia and the Far East [ECAFE]. (1968). *Feasible growth and trade gap projections in the ECAFE region*. Bangkok: ECAFE.
- Efron, B. and Tibshirani, R.J. (1993), *An Introduction to the Bootstrap*. New York: Chapman and Hall.
- Engle, R. F., & Hendry, D. F. (1993). Testing superexogeneity and invariance in regression models. *Journal of Econometrics*, 56(1), 119-139.
- Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.
- Engle, R.F. and Granger, C.W.J. (1987), "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251-276.
- Ericson, R. and Pakes, A. (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62, 53-82.
- Fair, R. C. (1971). *A short-run forecasting model of the United States economy*. Lexington: Heath.
- Fair, R. C. (1974). An evaluation of a short-run forecasting model. *International Economic Review*, 15(2), 285-304.
- Fair, R. C. (1987). Macroeconometric models. In J. Eatwall, et al. (Eds.), *The new Palgrave: A dictionary of economics* (Vol. 3, pp. 269-273). London: Macmillan.
- Fair, R. C. (2004). *Estimating how the macroeconomy works*. Boston: Harvard University Press.
- Fair, R.C. (1984), *Specification, Estimation, and Analysis of Macroeconometric Models*. Cambridge, Mass.: Harvard University Press.
- Fair, R.C. (1994), *Testing Macroeconometric Models*. Cambridge, Mass.: Harvard University Press.

- Favero, C. A. (2001). *Applied macroeconometrics*. New York: Oxford University Press.
- Fisher, D. (1983). *Macroeconomic theory: A survey*. London: Macmillan.
- Forni, M. and Reichlin, L. (1997), "Let's Get Real: A Dynamic Factor Analytical Approach to Disaggregated Business Cycles," Manuscript, University of Modena and University of Bruxelles.
- Friend, I., & Taubman, P. (1964). A short run forecasting model. *Review of Economics and Statistics*, 46(3), 229-236.
- Frisch, R. (1933), "Propagation Problems and Impulse Problems in Dynamic Economics," in *Economic Essays in Honor of Gustav Cassel*. London: Allen and Unwin.
- Fuller, W.A. (1976), *Introduction to Statistical Time Series*. New York: John Wiley and Sons.
- Gali, J. (1992). How well does the IS-LM model fit postwar U.S. data. *Quarterly Journal of Economics*, 107(2), 709-738.
- Garratt, A., et al. (2003). A long run structural macroeconomic model of the UK. *The Economic Journal*, 113(487), 412-455.
- Geweke, J. (1977), "The Dynamic Factor Analysis of Economic Time-Series Models," in D.J. Aigner and A.S. Goldberger (eds.), *Latent Variables in Socioeconomic Models*. Amsterdam: North-Holland, 365-383.
- Gourieroux, C. and Monfort, A. (1996), *Simulation-Based Econometric Methods*. Oxford: Oxford University Press.
- Granger, C.W.J. (1969), Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, *Econometrica*, 37, 424-438.
- Granger, C.W.J. (1981), "Some Properties of Time Series Data and their Use in Econometric Model Specification, *Journal of Econometrics*, 16, 121-130.
- Granger, C.W.J. (1989) *Forecasting in Business and Economics*. Academic Press.
- Granger, C.W.J. and Joyeux, R. (1980), "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, 1, 15-39.
- Granger, C.W.J. and Teräsvirta, Y. (1993), *Modeling Nonlinear Economic Relationships*. Oxford: Oxford University Press.
- Griliches, Z. (1986). Economic data issues. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of econometrics*. Amsterdam: North-Holland.
- Hall, S. (1995). Macroeconomics and a bit more reality. *The Economic Journal*, 105(431), 974-988.
- Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357-384.
- Hansen, L.P. and Heckman, J.J. (1996), "The Empirical Foundations of Calibration," *Journal of Economic Perspectives*, 10, 87-104.
- Hansen, L.P. and Sargent, T.J. (1980), "Formulating and Estimating Dynamic Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 2, 7-46.

- Hansen, L.P. and Sargent, T.J. (1998), *Recursive Models of Dynamic Linear Economies*. Princeton: Princeton University Press.
- Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Hendry, D. F. (1980). Econometrics - Alchemy or science? *Economica*, 47(188), 387–406.
- Hendry, D. F., & Richard, J. F. (1983). The econometric analysis of economic time series. *International Statistical Review*, 51, 111–163.
- Hendry, D.F. (1995), *Dynamic Econometrics*. Oxford: Oxford University Press.
- Hildreth, C. (1986), *The Cowles Commission in Chicago, 1939-1955*. New York: Springer-Verlag.
- Ingram, B. and Whiteman, C. (1994), “Supplanting the ‘Minnesota’ Prior: Forecasting Macroeconomic Time Series Using Real Business Cycle Model Priors,” *Journal of Monetary Economics*, 34, 497-510.
- Houthakker, H. S. (1956). The Pareto distribution and the Cobb–Douglas production function in activity analysis. *Review of Economic Studies*, 23, 27–31.
- Hunt, B., Rose, D., & Scott, A. (2000). The core model of the Reserve Bank of New Zealand’s forecasting and policy system. *Economic Modelling*, 17(2), 247–274.
- Ichimura, S., & Matsumoto, Y. (Eds.). (1994). *Econometric models of Asian-Pacific countries*. New York: Springer-Verlag.
- Intriligator, M. D., Bodkin, R. G., & Hsiao, C. (1996). *Econometric models, techniques, and applications*. New Jersey: Prentice-Hall.
- Jansen, E. S. (2000). Statistical issues in macroeconomic modelling. In *Arbeidsnotat*. Norges Bank Paper No. 2000/12. Presented to the Fall Meeting of the World Project LINK, Oslo.
- Judd, K. (1998), *Numerical Methods in Economics*. Cambridge, Mass.: MIT Press.
- Kapetanios G., V. Labhard and S. Price (2007), “Forecast combination and the Bank of England’s suite of statistical forecasting models”, Price. Bank of England Working Paper No. 323, May 2007.
- Karingi and Ndungu 2000, ‘Macro Models of the Kenyan Economy’, KIPPRA Discussion paper Series, No. 2
- Khayum, M. F. (1991). *Macroeconomic modelling and policy analysis for less developed countries*. Boulder: Westview Press. 280 A. Valadkhani / *Journal of Policy Modeling* 26 (2004) 265–281.
- Kim, C.-J. and Nelson, C.R. (1998a), “Business Cycle Turning Points, A New Coincident Index, and Tests of Duration Dependence Based on A Dynamic Factor Model with Regime-Switching,” *Review of Economics and Statistics*.
- Kim, C.-J. and Nelson, C.R. (1998b), *Dynamic Time Series Models and Markov Switching: Classical and Gibbs Sampling Approaches with Applications*. Manuscript, Department of Economics, University of Washington
- King, R. G., et al. (1991). Stochastic trends and economic fluctuations. *American Economic Review*, 81(4), 819–840.
- King, R.G., Plosser, C.I., Stock, J.H., and Watson, M.W. (1991), “Stochastic Trends and Economic Fluctuations,” *American Economic Review*, 81, 819-840.

- Klein, L. R. (1965). What kind of macroeconomic model for developing economies? *The Econometric Annual of the Indian Economic Journal*, XIII(3), 313–324.
- Klein, L. R. (1983). *Lectures in econometrics*. Amsterdam: North-Holland.
- Klein, L. R. (1989a). Development and prospects in macroeconomic modelling. *Eastern Economic Journal*, 15(4), 287–304.
- Klein, L. R. (1989b). Econometric aspects of input–output analysis. In R. E. Miller, et al. (Eds.), *Frontiers of input-output analysis* (pp. 3–21). Oxford: Oxford University Press.
- Klein, L. R. (1999). Economic stabilisation policy: Pitfalls of parsimonious modelling. *Journal of Quantitative Economics*, 15(2) (special issue on Macroeconomics Policy Modelling).
- Klein, L.R. (1946), *The Keynesian Revolution*. New York: MacMillan.
- Klein, L.R. and Goldberger, A.S. (1955), *An Econometric Model of the United States: 1929-1952*. Amsterdam: North-Holland.
- Kloek, T. (1988). Macroeconomic models and econometrics. In W. Driehuis, et al. (Eds.), *Challenges for macroeconomic modelling* (pp. 343–388). Amsterdam: North-Holland.
- Kmenta, J., & Ramsey, J. B. (Eds.). (1981). *Large scale macro-econometric models*. Amsterdam: North-Holland.
- Kydland, F.E. and Prescott, E.C. (1982), “Time to Build and Aggregate Fluctuations,” *Econometrica*, 50, 1345-1371.
- Kydland, F.E. and Prescott, E.C. (1996), “The Computational Experiment: An Econometric Tool,” *Journal of Economic Perspectives*, 10, 69-86.
- Leamer, E. E. (1983). Let’s take the con out of econometrics. *The American Economic Review*, 73(1), 31–44. Reprinted in *Modelling Econometric Series: Readings in Econometric Methodology*.
- Leeper, E.M. and Sims, C.A. (1994), “Toward a Modern Macroeconomic Model Useable for Policy Analysis,” in O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual*. Cambridge, Mass.: MIT Press, 81-117.
- Leeper, E.M., Sims, C.A. and Zha, T. (1996), “What Does Monetary Policy Do?,” *Brookings Papers on Economic Activity*, 2, 1-78.
- Lucas, R. E., Jr. (1976). Econometric policy evaluation: A critique. In K. Brunner & A. H. Meltzer (Eds.), *The Phillips curve and labor markets*. Amsterdam: North-Holland.
- Lucas, R.E. (1972), “Expectations and the Neutrality of Money,” *Journal of Economic Theory*, 4, 103-124.
- Lucas, R.E. (1976), “Econometric Policy Evaluation: A Critique,” in K. Brunner and A. Meltzer (eds.), *The Phillips Curve and the Labor Market* (Carnegie-Rochester Conference Series, Volume 1). Amsterdam: North-Holland.
- Lucas, R.E. and Prescott, E.C. (1971), “Investment Under Uncertainty,” *Econometrica*, 39, 659-681.
- Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, I., Parzen, P., & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: results of a forecasting competition. *Journal of Forecasting* 1, 111-153.
- Mankiw, N. G. (1988). Recent developments in macroeconomics: A very quick refresher course. *Journal of Money Credit and Banking*, 20(3), 436–449.



- Mankiw, N.G. and D. Romer, eds. (1991), *New Keynesian Economics*. Cambridge, Mass.: MIT Press.
- McCallum, B. T. (1994). A Reconsideration of the Uncovered Interest Rate Parity Relationship," *Journal of Monetary Economics*, 33, 105(132)
- McCullough, B. D. (2006). The unreliability of Excel's statistical procedures. *Foresight: The International Journal of Applied Forecasting*, 3 (February), 44-45.
- McGrattan, E.R., Rogerson, R. and Wright, R. (1997), "An Equilibrium Model of the Business Cycle with Household Production and Fiscal Policy," *International Economic Review*, 38, 267-290.
- McNees, S. K. (1979). The forecasting record for the 1970s. *New England Economic Review*, (September/ October), 33-53.
- McNees, S. K. (1982). The role of macroeconomic models in forecasting and policy analysis in the US. *Journal of Forecasting*, 1(1), 37-48.
- Moore, G.H. (1983), *Business Cycles, Inflation, and Forecasting*. Cambridge, Mass.: Harper and Row.
- Murphy, C., & Brooker, R. (1994). Murphy model and microeconomic reform. In C. Hargreaves (Ed.), *A comparison of economy-wide models of Australia*. EPAC Commission Paper No. 2, AGPS, Canberra.
- Muth, J.F. (1960), "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 55, 299-305.
- Muth, J.F. (1961), "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29, 315-335.
- Narasimham, N. V. A. (1956). *A short-term planning model for India*. Amsterdam: North-Holland.
- Nelson, C.R. (1972), "The Prediction Performance of the F.R.B.-M.I.T.-Penn Model of the U.S. Economy," *American Economic Review*, 62, 902-917.
- Nelson, C.R. and Plosser, C.I. (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139-162.
- Newbold, P., & Granger, C. W. I. (1974). Experience with forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistical Society (A)* 137; 131-165.
- Nugent, J. B. (1975). Policy-oriented macroeconomic models for development and planning. *Annals of Economic and Social Measurement*, 4(4), 509-529.
- Oduor, J. (2008) *Impacts of Real exchange rate misalignment on Trade Creation and Diversion within Regional Trading Blocs: The Case of COMESA*. Shaker Verlag, Germany.
- Pagan, A. (1987). Three econometric methodologies: A critical appraisal. *Journal of Economic Surveys*, 1(1), 3-24.
- Pagan, A. (1994), "Calibration and Econometric Research: An Overview," in A. Pagan (ed.), *Calibration Techniques and Econometrics*, special issue of *Journal of Applied Econometrics*, 9, S1-S10.
- Pesaran, M. H. (1995). New direction in applied dynamic macroeconomic modelling. In *The international conference on economic policy evaluation models in theory and practice*. Tunis: The Arab Planning Institute.
- Pesaran, M. H., & Smith, R. P. (1985). Evaluation of macroeconomic models. *Economic Modelling*, 2(2), 125-134.

- Pindyck, R.S. and Rubinfeld, D.L. (1991). *Econometric Models and Economic Forecasts*. 3rd edition, McGraw-Hill.
- Phelps, E.S., et. al. (1970), *Microeconomic Foundations of Employment and Inflation Theory*. New York: W.W. Norton and Company.
- Phillips, A. W. (1954). Stabilisation policy in closed economy. *Economic Journal*, 64, 290–323.
- Phillips, A.W. (1957). Stabilisation policy and the time forms of lagged responses. *Economic Journal*, 67, 265–277.
- Prescott, E.C. (1986), “Theory Ahead of Business Cycle Measurement,” *Quarterly Review*, Federal Reserve Bank of Minneapolis, 9-33.
- Quah, D. and Sargent, T.J. (1993), “A Dynamic Index Model for Large Cross Sections,” in J.H. Stock and M.W. Watson (eds.), *Business Cycles, Indicators and Forecasting*. Chicago: University of Chicago Press for NBER, 285-310.
- Quah, D.T. (1995), “Introduction,” in D. Quah (Ed.), *Business Cycle Empirics – Calibration and Estimation*, “Controversy” section of *Economic Journal*, 105, 1594-1596.
- Rios-Rull, J.-V. (1995), “Models with Heterogeneous Agents,” in T.F. Cooley (ed.), *Frontiers of Business Cycle Research*. Princeton: Princeton University Press, 98-125.
- Rosen, S., Murphy, K.M. and Scheinkman, J.A. (1994), “Cattle Cycles,” *Journal of Political Economy*, 102, 468-492.
- Rosenzweig, M.R. and Wolpin, K.I. (1993), “Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-income Countries: Investments in Bullocks in India,” *Journal of Political Economy*, 101, 223-244.
- Rotemberg, J.J. and Woodford, M. (1996), “Real Business-Cycle Models and the Forecastable Movements in Output, Hours, and Consumption,” *American Economic Review*, 86, 71-89.
- Rotemberg, J.J. and Woodford, M. (1997), “An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy,” in O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual, 1997*. Cambridge, Mass.: MIT Press.
- Rothwell, G. and Rust, J. (1995), “A Dynamic Programming Model of U.S. Nuclear Power Plant Operations, Manuscript, Department of Economics, University of Wisconsin.
- Rust, J. (1994), “Structural Estimation of Markov Decision Processes,” in R. Engle and D. McFadden (eds.), *Handbook of Econometrics*, Vol. 4. Amsterdam: North-Holland.
- Rust, J. (1996), “Numerical Dynamic Programming in Economics” in H. Amman, D. Kendrick and J. Rust (eds.), *Handbook of Computational Economics*. Amsterdam: North-Holland.
- Sargan, J.D. (1964), “Wages and Prices in the United Kingdom: A Study in Econometric Methodology,” in P.E. Hart, G. Mills and J.N. Whitaker (eds.), *Econometric Analysis for National Economic Planning*. London: Butterworths.
- Sargent, T. (1976). A classical macroeconometric model for the United States. *Journal of Political Economy*, 84(2), 207–232.

- Sargent, T.J. and Sims, C. (1977), "Business Cycle Modeling Without Pretending to Have Too Much a Priori Theory," in C. Sims (ed.), *New Methods of Business Cycle Research*. Minneapolis: Federal Reserve Bank of Minneapolis.
- Sargent, T.J. and Wallace, N. (1975), "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," *Journal of Political Economy*, 83, 241-254.
- Sastry, V.K. (1975). The use of macroeconomic regression models of developing countries: A comment. *Oxford Economic Papers*, 27(1), 156-165.
- Schlicht, E. (1985). *Isolation and aggregation in economics*. Berlin: Springer-Verlag.
- Schnaars, S. P. (1986). A comparison of extrapolation procedures on yearly sales forecasts. *International Journal of Forecasting* 2, 71-85.
- Seers, D. (1963). The limitations of the special case. *Bulletin of Oxford Institute of Economics and Statistics*, 25(2), 77-98.
- Shourie, A. (1972). The use of macro-economic models of developing countries for forecasts and policy prescription: Some reflections on current practice. *Oxford Economic Papers*, 24(1), 1-35.
- Sims, C. A. (1982). Policy analysis with econometric models. *Brookings Papers on Economic Activity*, 0(1), 107-152.
- Sims, C.A. (1972), "Money, Income and Causality," *American Economic Review*, 62, 540-552.
- Sims, C.A. (1980), "Macroeconomics and Reality," *Econometrica*, 48, 1-48.
- Sims, C.A. (1996), "Macroeconomics and Methodology," *Journal of Economic Perspectives*, 10, 105-120.
- Sims, C.A. and Zha, T. (1996), "Does Monetary Policy Cause Recessions?," Manuscript, Department of Economics, Yale University.
- Sims, C.A. and Zha, T. (1997), "System Methods for Bayesian Forecasting Models," Manuscript, Yale University and Federal Reserve Bank of Atlanta.
- Soludo C. C., (2002), "Macroeconomic Modelling and Economic Policy Making: A Survey of Experience in Africa", *Department of Economics University of Nigeria*.
- Slutsky, E. (1927), The Summation of Random Causes as the Source of Cyclic Processes," *Econometrica*, 5, 105-146.
- Smith, R. P. (1984). The evaluation and comparisons of large macroeconomic models. In P. Malgrange & P. A. Muet (Eds.), *Contemporary macroeconomic modelling*. London: Blackwell.
- Sowey, E., & Hargreaves, C. (1991). Dissemination of macroeconomic models: Is a new age dawning? *Journal of Policy Modelling*, 13(4), 599-621.
- Stiglitz, J. (1969). Distribution of income and wealth among individuals. *Econometrica*, 37(3), 382-397.
- Stock, J.H. and Watson, M.W. (1988a), "Variable Trends in Economic Time Series," *Journal of Economic Perspectives*, 2, 147-174.
- Stock, J.H. and Watson, M.W. (1988b), Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097-1107.

Stock, J.H. and Watson, M.W. (1997), "Adaptive Diffusion Indexes," Manuscript, Kennedy School, Harvard University, and Woodrow Wilson School, Princeton University.

Stock, J.H. and Wise, D. (1990), "Pensions, the Option Value of Work, and Retirement," *Econometrica*, 58, 1151-1180.

Stock, J.H., and Watson, M.W. (1989), "New Indexes of Coincident and Leading Economic Indicators," in O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual*. Cambridge, Mass.: MIT Press, 351-394.

Stone, R. "The Accounts of Society" *Journal of Applied Econometrics*, 1(1), pp 5-28, 1986.

Tashman, L. J., & Leach, M. L. (1991). Automatic forecasting software: a survey and evaluation. *International Journal of Forecasting* 7, 209-230.

Taylor, L. (1990), *Socially Relevant Policy Analysis*, MIT Press, Cambridge Mass.

Taylor, J. B. (1993). "Discretion versus Policy Rules in Practice," Carnegie-Rochester Conference Series on Public Policy, 39, December, 195-214

Thorbecke, E. *Intersectoral Linkages and Their Impact on Rural Poverty Alleviation*. United Nations Industrial Development Organization (UNIDO), 1994.

Tong, H. (1990), *Non-linear Time Series*. Oxford: Clarendon Press.

Uebe, G. (Ed.). (1995). *World of economic models: A catalogue of typical specifications of economic models*. Aldershot: Avebury.

United Nations Conference on Trade and Development [UNCTAD]. (1968). *Trade prospects and capital needs of developing countries*. New York: United Nations.

Valadkhani A. (2004), "History of Macroeconometric Modelling: Lessons from Past Experience". *Journal of Policy Modeling*.

Vokurka, R. J., Flores, B. E., & Pearce, S. (1996). Automatic feature identification and graphical support in rule based forecasting: a comparison. *International Journal of Forecasting* 12, 495-512.

Walras, L. (1954). *Elements of pure economics* (W. Jaffee, Trans.). London: Allen and Unwin.

Walras, L. (1969), *Elements of Pure Economics*, Augustus M. Kelley Publishers.

Weiss, A. A., & Anderson, A. P. (1984). Estimating time series models using relevant forecast evaluation criteria. *Journal of the Royal Statistical Society*

Whittle, P. (1983), *Prediction and Regulation by Linear Least-Square Methods*, Second Edition. Minneapolis: University of Minnesota Press.

Willman, A., et al. (2000). The BOF5 macroeconomic model of Finland, structure and dynamic microfoundations. *Economic Modelling*, 17(2), 275-303.

Yule, G.U. (1927), "On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers," *Philosophical Transactions*, 226A.

Zarnowitz, V. (1992), *Business Cycles: Theory, History, Indicators, and Forecasting*. Chicago: University of Chicago Press.

Zarnowitz, V. (1978). On the accuracy and properties of recent macroeconomic forecasts. *The American Economic Review*, 68(2), 313-319.

Zellner, A. (1992), "Statistics, Science and Public Policy," *Journal of the American Statistical Association*, 87, 1-6.

## APPENDIX A: Tables of Critical Values

### Table A1: Critical Values of Cointegration Test

n	Model	Point	$\phi_\infty$	SE	$\phi_1$	$\phi_2$
1	No constant, no trend	1	-2.5658	0.0023	-1.96	-10.04
		5	-1.9393	0.0008	-0.398	0
		10	-1.6156	0.0007	-0.181	0
1	Constant, no trend	1	-3.4336	0.0024	-5.999	-29.25
		5	-2.8621	0.0011	-2.738	-8.36
		10	-2.5671	0.0009	-1.438	-4.48
1	Constant+ trend	1	-3.9638	0.0019	-8.353	-48.44
		5	-3.4126	0.0012	-4.039	-17.83
		10	-3.1279	0.0009	-2.418	-7.58
2	No constant, no trend	1	-3.9001	0.0022	-10.534	-30.03
		5	-3.3377	0.0012	-5.967	-8.98
		10	-3.0462	0.0009	-4.069	-5.73
2	Constant+ trend	1	-4.3266	0.0022	-15.531	-34.03
		5	-3.7809	0.0013	-9.421	-15.06
		10	-3.4959	0.0009	-7.203	-4.01
3	No constant, no trend	1	-4.2981	0.0023	-13.79	-46.37
		5	-3.7429	0.0012	-8.352	-13.41
		10	-3.4518	0.001	-6.241	-2.79
3	Constant+ trend	1	-4.6676	0.0022	-18.492	-49.35
		5	-4.1193	0.0011	-12.024	-13.13
		10	-3.8344	0.0009	-9.188	-4.85
4	No constant, no trend	1	-4.6493	0.0023	-17.188	-59.2
		5	-4.1	0.0012	-10.745	-21.57
		10	-3.811	0.0009	-8.317	-5.19
4	Constant+ trend	1	-4.9695	0.0021	-22.504	-50.22
		5	-4.4294	0.0012	-14.501	-19.54
		10	-4.1474	0.001	-11.165	-9.88
5	No constant, no trend	1	-4.9587	0.0026	-22.14	-37.29
		5	-4.4185	0.0013	-13.641	-21.16
		10	-4.1327	0.0009	-10.638	-5.48
5	Constant+ trend	1	-5.2497	0.0024	-26.606	-49.56
		5	-4.7154	0.0013	-17.432	-16.5
		10	-4.4345	0.001	-13.654	-5.77
6	No constant, no trend	1	-5.24	0.0029	-26.278	-41.65
		5	-4.47048	0.0018	-17.12	-11.17
		10	-4.4242	0.001	-13.348	0
6	Constant+ trend	1	-5.5127	0.0033	-30.735	-52.5
		5	-4.9767	0.0017	-20.883	-9.05
		10	-4.6999	0.0011	-17.445	0

Source: Mackinnon (1991)

## APPENDIX B: Participants during the Manual Review Workshop in Harare – August 2011

Country	Name of Participant	Institution Represented
Botswana	Batane Walter Matekane (Mr)	Ministry of Finance and Development
	Moemedi Phetwe (Mr)	Bank of Botswana
Kenya	Maureen Were (Dr/Ms)	Central Bank of Kenya
	Esman Nyamongo (Dr)	Central Bank of Kenya
	Johnstone Poipoi (Mr)	Kenya National Bureau of Statistics
	Peter Chacha Wankuru (Mr)	Ministry of Finance
Lesotho	Maselone Thite (Mr)	Ministry of Finance and Development Planning
	Molapo Senei (Mr)	Ministry of Finance and Development Planning
Mozambique	Amorim Pery (Mr)	Ministry of Planning and Development
Rwanda	Regina Nyirakanani (Ms)	National Bank of Rwanda
	Valence Kimenyi (Mr)	National Bank of Rwanda
	Jean C. Rwililiza (Mr)	National Bank of Rwanda
Swaziland	Patrick Ndzinisa (Mr)	Central Bank of Swaziland
Tanzania	Wilfred Mbowe (Dr.)	Bank of Tanzania
	Salome Kingdom (Mr)	President's Office – Planning Commission
Uganda	Elizabeth Kasekende (Ms)	Bank of Uganda
	Mustapha Achidri (Mr)	Ministry of Finance, Planning and Economic Development
Zambia	Edward Kapwepwe (Mr)	Ministry of Finance and National Planning
	Chungu Kapembwa (Mr)	Bank of Zambia
Zimbabwe	Kennedy M. Kupeta (Mr)	Reserve Bank of Zimbabwe
	Prudence Stephen Moyo (Mr)	Reserve Bank of Zimbabwe
	Nebson Mupunga (Mr)	Reserve Bank of Zimbabwe
	Paul Mukoki (Mr)	Reserve Bank of Zimbabwe
	Cresncia Gapare (Ms)	Ministry of Finance
	Everjoice Chinoda (Ms)	Ministry of Economic Development and Investment Promotion
South Africa (Reviewer)	Dr. Christopher Malikane	Macro-Financial Analysis Group, School of Economic and Business Sciences, University of The Witwatersrand
Kenya (MEFMI Fellow)	Kethi Ngoka (Mrs)	Central Bank Of Kenya
Zambia (MEFMI Fellow)	Peter Zgambo (Mr)	Bank of Zambia

## APPENDIX C: Developing a VECM

The reduced form VAR can easily be formulated from a sytem of equations. Assuming we have three variables  $x_1, x_2$  and  $x_3$  in our model and they are related to each other as:

$$x_1 = A_{11}x_{1t-1} + A_{12}x_{2t-1} + A_{13}x_{3t-1} + \varepsilon_{1t}$$

$$x_2 = A_{21}x_{1t-1} + A_{22}x_{2t-1} + A_{23}x_{3t-1} + \varepsilon_{2t}$$

$$x_3 = A_{31}x_{1t-1} + A_{32}x_{2t-1} + A_{33}x_{3t-1} + \varepsilon_{3t}$$

We can write the above system of equations in a VAR form as;

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

This is same as  $x_t = Ax_{t-1} + \varepsilon_t$ , where  $A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$

The modeler is interested in obtaining the coefficients in the matrix  $A$ .

Subtracting  $x_{t-1}$  on both sides we have:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

We can then factor out  $x_{t-1}$  on the right hand side as;

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

This can then be written as;

$$\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} = \begin{bmatrix} A_{11} - 1 & A_{12} - 0 & A_{13} - 0 \\ A_{21} - 0 & A_{22} - 1 & A_{23} - 0 \\ A_{31} - 0 & A_{32} - 0 & A_{33} - 1 \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

Simplifying further it becomes;

$$\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} = \begin{bmatrix} A_{11} - 1 & A_{12} & A_{13} \\ A_{21} & A_{22} - 1 & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

We let  $\pi = A - I = \begin{bmatrix} A_{11} - 1 & A_{12} & A_{13} \\ A_{21} & A_{22} - 1 & A_{23} \\ A_{31} & A_{32} & A_{33} - 1 \end{bmatrix}$



However, due to the Granger Representation theorem,  $\pi = A - I$  is normally written

$$\text{as } \pi = -(1 - A) = - \begin{bmatrix} 1 - A_{11} & -A_{12} & -A_{13} \\ -A_{21} & 1 - A_{22} & -A_{23} \\ -A_{31} & -A_{32} & 1 - A_{33} \end{bmatrix}$$

The VECM can therefore be written as:

$$\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} = - \begin{bmatrix} 1 - A_{11} & -A_{12} & -A_{13} \\ -A_{21} & 1 - A_{22} & -A_{23} \\ -A_{31} & -A_{32} & 1 - A_{33} \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

In compact form, the matrix above can be written as:

$$\Delta x_t = \pi x_{t-1} + \varepsilon_t$$

The expression  $\pi x_{t-1}$  is the error correction term. This equation is similar to the ADF model. It is also similar to the standard ECM model. Our interest here would be the rank of the matrix  $\pi$ . If the  $\text{rank}(\pi) = 0$ , all the variables in the model are unit root processes and are not cointegrated. The model therefore reduces to a normal VAR in first differences, just like in the EG two step method. If  $\text{rank}(\pi) = n$ , then all the variables are stationary with  $n$  cointegrating vectors. When  $1 < \text{rank}(\pi) < n$ , then there are more than one but less than  $n$  cointegrating vectors. The number of cointegrating vectors equals the rank of  $\pi$ , which is basically the number of the characteristic roots of the  $\pi$  matrix that are different from zero. In the case where  $\text{rank}(\pi) = 0$ , all the characteristic roots will equal to zero. The multivariate generalization of the model above can be given as:

$$x_t = \pi x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + \varepsilon_t$$

$$\text{where } \pi = - \left( I - \sum_{i=1}^p A_i \right) \text{ and } \pi_i = - \sum_{j=i+1}^p A_j$$

## ECO-AUDIT

## Environmental Benefits Statement

MEFMI is committed to preserving endangered forests and natural resources. The Publications & Networking office has chosen to print this manual on recycled paper.

# Notes

---

---

A series of horizontal dotted lines for writing notes, spanning the width of the page.



# Notes

---

---

A series of horizontal dotted lines for writing notes.

## Notes

---

---

A series of horizontal dotted lines for writing notes, spanning the width of the page.



# Notes

---

---

A series of horizontal dotted lines for writing notes, spaced evenly down the page.