2006

Modeling of Development and Carbon Dioxide Emission in Bangladesh: A Bootstrap Approach

Sharker, Md. Abu Yushuf

University of Rajshahi

Modeling of Development and Carbon Dioxide Emission in Bangladesh: A Bootstrap Approach

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
MASTER OF PHILOSOPHY
in
THE FACULTY OF SCIENCE

by

Md. Abu Yushuf Sharker
M. Phil Researcher
Roll No. 80, Session: July-2000

Department of Statistics, University of Rajshahi
Rajshahi-6205, Bangladesh

May 2006

Summary

In the thesis we study the relationship between CO_2 emission per capita and GDP per capita in context of Bangladesh. First exploratory data analysis (EDA) is used to uncover the hidden information carried by the observed data. An attempt is made to fit Environmental Kuznets Curve model by means of classical techniques as well as bootstrap techniques. EDA shows that both CO_2 Emission per capita and GDP per capita are trended. There is no steady state of the variables within their sample period. These variables does not follow EKC. To test the presence of stochastic trend of the variables we use unit root tests. To mitigate small sample limitations of our data we first design a simulation based study to compare the performance of classical tests with bootstrap tests. Our simulation based result provides that CADF (Covariate Augmented Dickey Fuller Test) test has higher power and BCADF (Bootstrapped CADF) test has less size distortion for testing unit roots for small sample of size 30. Unit root tests suggest that both the series are nonstationary, i.e., CO_2 is accumulating in the atmosphere and GDP is also experiencing accumulation. In the succession of time series modeling ARIMA model is fitted to both of the series. Both CO_2 emission per capita and GDP per capita of Bangladesh follow ARIMA(0,1,1) model. Forecasting by bootstrap produces better results sometimes. In quest of dynamic regression model cointegration is checked. Classical, bootstrap and double bootstrap techniques are used to test whether there exists any cointegrating relationship among CO_2 emission per capita, GDP per capita and it's square. Result shows that these variables are nonstationary but not cointegrated. There exists no long run equilibrium relationship between CO_2 emission per capita and GDP per capita.

Contents

List of Tables

List of Figures

Declaration

I hereby declare that the thesis entitled "**Modeling of Development and Carbon Dioxide Emission in Bangladesh: A Bootstrap Approach**" is the result of my own research work undertaken by the Department of Statistics, University of Rajshahi, Bangladesh. I further declare that this thesis has not concurrently been submitted in part or in full previously for any degree or diploma in either this university or any other university.

Md. Abu Yusuf Sharker
M. Phil Researcher
Roll No. 80, Session: July 2000
Department of Statistics
University of Rajshahi
Bangladesh.

Certificate

This is to certify that this thesis entitled "**Modeling of Development and Carbon Dioxide Emission in Bangladesh: A Bootstrap Approach**" is a research work carried out by **Md. Abu Yushuf Sharker** for the degree of Master of Philosophy undertaken by the Department of Statistics, University of Rajshahi, Bangladesh. Any part or whole of this thesis has not been submitted elsewhere for any degree or award.

Md. Nasser

M. Nasser
Principal Supervisor
Associate Professor
Department of Statistics
University of Rajshahi
Bangladesh.

Certificate

This is to certify that this thesis entitled "**Modeling of Development and Carbon Dioxide Emission in Bangladesh: A Bootstrap Approach**" is a research work carried out by **Md. Abu Yushuf Sharker** for the degree of Master of Philosophy undertaken by the Department of Statistics, University of Rajshahi, Bangladesh. Any part or whole of this thesis has not been submitted elsewhere for any degree or award.

M. A. Razzaque 08/05/06

Professor M. A. Razzaque
Co-Supervisor
Pro-Vice Chancellor
Bangladesh Open University
Guzipur
Bangladesh.

Acknowledgments

First I wish to pay my thanks and gratitude to my supervisors, M. Nasser and Professor M. A. Razzaque for their guidance and help throughout my research.

I acknowledge the valuable suggestions of Professor R. Imon, Assistant professor M. Jahanur Rahman, Assistant professor Mesbahul Alam of this department. I am also grateful to my friend M. A. Khalek, lecturer, and all teachers, staffs and youngers of the department of statistics for their moral support.

I am very much grateful to my parents and parents-in-law. Without their help and encouragement, the research work was not possible.

Finally I would like to pay my thanks to my wife for her understanding and mental support during the work.

DEDICATED

To

Dr. Kazi Motahar Hussain

Who is my inspiration

and

Late Professor M. G. Mostafa

Who taught me Econometrics

Chapter 1

Introduction

1.1 Introduction

Sustainable development is now a global concern. The term 'sustainable development' was appeared in 1970 and focused by International Union for the Conservation of Nature and Natural Resources in 1980. The term gained wide popularity after the publication of Brundtland report in 1987. They defined sustainable development as "Development that meets the needs for the present without compromising the ability of future generations to meet their own needs." This definition has drawn attention to the concern with the long-term implication of the present-day-development. For a state-of-the-art discussion of sustainable development in an economic or ecological point of view see Arrow et al. (1996). The Earth Council completed this concept with more meaningful refinement. They indicated that sustainable development is related to three sustainability. 1) Economic sustainability 2) Social sustainability and 3) Environmental sustainability.

The natural environment provides some fundamental life support services and maintain the suitable atmospheric and climatic conditions. The climate regulation provided by natural ecosystem is the key life support service both on a local and on a global scale. But increasing emission of greenhouse gases (CO_2, SO_x, NO_x) pose

a serious threat to the global climate. Certain levels of these gases exist naturally, trapping the sun's heat and making the planet inhabitable. However, increasing emissions of these gases also have the potential to trap too much heat, altering the climate in major and unpredictable ways which includes increasing strom severity, sea level rise, spread of infectious diseases etc. The last two decades have witnessed an unprecedented global warming. This has brought about great concerns over its causes and consequences. Scientists claimed that the increasing carbon dioxide emission (CO_2) produced a massive build-up of greenhouse gas, which gave rise to recent warm temperature (IPCC 1995; Watson et al. 1996).

The main source of CO_2 is fossil fuel. Unlike other automotive and industrial air pollutants, that are transformed by chemical reactions or washed out of the atmosphere by rainfall or other natural procedures, CO_2 is extremely stable gas that accumulates in the atmosphere once emitted. Of total fossil carbon used probably less than 1-2% is recycled, whereas the remainder is transformed into wastes, including pollutants. There is some compensation for future generations through the planting of additional forest and the development of solar and wind energy, but overall there is no conformity to a steady state operationalization.

Modeling CO_2 emission and its effects on ecosystem is much tough job. Different scientific paper have suggested different techniques and variables to assess CO_2 emission locally and globally. UNEP/UNSTAT (1993) consultative expert group selected the indicators for reporting the state of environment. They include energy state as total per capita primary energy use, total per capita domestic energy use, energy intensity of industrial production and atmosphere condition judged from CO_2 emission per person, production and import of ozone-depleting chemicals, frequency of days of unacceptable air quality etc.

After the historic agreement of Kyoto Protocol in 1997, Gupta and Bhandari (1999) thought that it is reasonable to accept an individual as the unit of account of CO_2 emission, i.e., those which are generated by humans rather than by countries.

According to their view, global emissions have to be restricted to a pre-determined level with the widest participation (Annex I and Non-annex I countries). He aimed to strike a balance between the concerns relating to the choice of the numeraire, viz., population and GDP in determining emissions targets.

The size of a nation's economy is traditionally denoted by its yearly gross domestic product (GDP), the value of all final goods and services produced within its borders and by gross national product (GNP), GDP plus net income from abroad. Per capita GDP and GNP are often used, as a measure of an average resident's economic well being. These indicators are frequently used as yardsticks of economic progress and as a basis for international comparisons. (Repetto, 1989).

Several recent and often-cited papers on the relationship between pollution and economic growth have shown that the higher the income level, the greater would be the environmental degradation (Daly, 1977, Georgescu and Roegen, 1971, Hall et al., 1986). Many forms of air and water pollution "initially worsen but then improve as incomes rise" (World Bank, 1992). "At higher levels of development, structural changes towards information-intensive industries and services, coupled with increased environmental awareness, enforcement of environmental regulations, better technology and higher environmental expenditures, result in leveling off and gradual decline of environmental degradation" (Panayotou, 1993). Grossman and Krueger (1995), in particular, reported that for most pollutants, the turning point in environmental quality typically occurs at income levels below $8000 per capita. Because of its similarity to the pattern of income inequality documented by Kuznet (1955), this inverse-U-shaped pollution-income pattern is sometimes called an "Environmental Kuznets Curve." In response to these empirical findings, Selden and Song (1994), Shafik (1994), Holtz, Eakin and Selden (1995), Jones and Manuelli (1995), Hilton and Levinson (1998) have sought further evidence for inverse-U-shaped pollution-income relationships. Jones and Manuelli (1995), Selden and Song (1995), Chaudhuri and Pfaff (1997), Stokey (1998), Jaeger (1998), Andreoni and Levinson (2001) have

proposed theoretical explanations for the relationship between pollution and economic growth. Based on some existing researches, some policy analysts for example, Beckerman (1992), Bartlett (1994) have concluded that developing countries will automatically become cleaner as their economies growing. Some argued that it is natural for the poorest countries to become more polluted as they develop. These types of conclusions depend on the apparently growing conventional wisdom that pollution follows a deterministic inverse-U-shaped environmental Kuznets curve. Birdsall and Wheeler (1991) studied air pollution in the manufacturing sector of Latin America using developed and developing country data. They described pollution intensity in terms of GDP. Day and Grafton (2002) showed that only CO appears to be declined relation in the long run with increase of real GDP per capita for Canada.

Some researchers has ben given the emphasis on population growth as a factor of increasing carbon emission. According to Smil (1990), Bongarrts (1992), Dietz and Rosa (1994), Engelman (1994, 1998) and O'Neill et al. (2001), population growth is one of the major factors in causing carbon emissions in both developed and developing countries. Anquing Shi (2001) concluded that (1) One percent of population growth is associated with a 1.28 percentage increase in emissions on average. (2) The impact of population pressure on emissions has been more pronounced in developing countries than developed countries. (3) It is estimated that global emissions will reach as high as 13.72 gigatons in 2025 under the business-as-usual assumptions this magnitude more than doubles the emissions level of 1990 and half of the gains will be attributed to the future population growth alone. (4) Rising income levels have been associated with a monotonically upward shift in emissions.

1.1.1 Drivers of Environmental Change

The first major factor affecting environmental change is *population*. The environmental impacts of human activities depend strongly on the number of people that inhabit in a particular part of the planet. So projection of future population trend is

an important for environmental forecasts.

The second important factor is the *standard of living.* This is usually measured in economic terms such as the gross domestic product (GDP) per capita. The more affluent the population is the more goods and services they demand, and consequently the greater the resulting environmental impacts per person. Thus change in economic conditions and standard of living over time are another dimension of environmental forecasting.

A third critical factor is *technology.* Technology includes the methods used to provide food, shelter comfort, transportation etc. The design and deployment of modern technology gives rise to environmental issue. So, expectations of technological change over time are thus the third element needed to describe an environmental future. Among many measurements of technological change are the change of environmental emissions, energy requirements, and natural resource requirements of particular technology, as well as changes in the nature of types of technologies in use at the future time.

Of course, these three factors alone do not tell the whole story of environmental futures. Rather, they are the principal drivers that determine future land use patterns, natural resource requirements, and pollutant emissions to air, water and land.

1.1.2 Modeling Environmental Process

To assess the resulting environmental impacts (such as pollutant concentrations), additional mathematical models and data are required. Many of these scientific models of environmental process are dynamic. They can predict how factors like pollutant concentrations will change over time in response to a specified input. To develop such type of model in Bangladesh context, I have considered population, GDP and CO_2 emission. This type of model so far, does not exist for Bangladesh. For our study we considered only the first two factors (Population, and GDP per capita as a Standard of living) to explain the CO_2 emission of Bangladesh.

1.1.3 Why CO_2 Emission in Study

Over the past century, the quantities of green house gases emitted to the atmosphere from human activity have increased dramatically. The principal culprit is CO_2. CO_2 is emitted from the combustion of the oil, gas and coal that supply roughly 75% of the total energy of the world. Current projections anticipate a doubling of atmospheric CO_2 concentration by the year 2100, bringing the global average temperature increase of about 2.4°C (IPCC, 1996). Though seemingly small, a temperature change of this magnitude in just one century would be unprecedented in human history and would profoundly affect the earth's climate and inhabitants. The chief concerns include sea level rise and flooding of low-lying regions; increased precipitation and severity of Strom events; increased drought and some ecological effects as plants and animals attempt to cope with rapid change.

Although, there are many other green house gases (CH_4, N_2O etc.), CO_2 is the major concern. Among all GHG CO_2 is the most stable and non-reacting chemical. The time scale for removal of natural process is very long. Again methane (CH_4)and carbon monoxide (CO) are oxidized to CO_2 by chemical reaction. The Global Warming Potential Index is sometime used to estimate the CO_2 equivalence of different gases. So, to forecast the future climatic change due to GHG, it is necessary to study the CO_2 emission.

1.2 Bangladesh in Short

Bangladesh is a South Asian developing country covering an area of 147,570 sq. km. Its population is 130 million and enjoys a very low per capita gross domestic product, i.e., $350 dollars (US). The current estimated population growth rate is around 1.6% and at this rate the population will be nearly 160 million in the next 10 years. The populations with a very low per capita gross domestic product, i.e., $350 dollars (US). Vulnerability of Bangladesh to climate change and sea level rise is well recognized.

Agriculture, manufacturing industry and various service sectors are the economic backbone of the country. Flat terrain, low economic growth, high population density, intensive dependence on agriculture and agricultural production, lack of institutional infrastructure etc. combine to make the country vulnerable to any nuance in climate change and sea level rise. Agriculture, manufacturing industry, and various services (such as transport, trade services, and housing) are the major economic sectors in Bangladesh. There is a falling trend in agriculture, and yet, despite this, agriculture remains of paramount importance because of the dependence of most other sectors or activities either on processing its products or on servicing it. According to the Bangladesh Bureau of Statistics, agriculture's share of GDP fell from 25.6% in 1991-92 to 21.3% in 1995, adjusted to 1989-90 constant market prices.

1.2.1 Energy Production and Consumption in Bangladesh

Bangladesh has one of the lowest per capita commercial energy consumption rates in the world with about 75 kilograms oil equivalent (kg/oe) per year (1994-95). Therefore, Bangladesh's energy sector contributes a relatively small amount of greenhouse gases (GHG) to the global atmosphere. On an average, more than 60% of total energy comes from renewable energy sources either in the form of biomass or hydropower. Similarly more than 55% of commercial energy comes from natural gas which is known as cleaner fuel.

Bangladesh depends heavily upon traditional biomass fuel. But the proportion of commercial fuels is greatly increasing due to increased use of indigenous natural gas and the limited increase in consumption of biomass fuels. Commercial energy consumption by different sector has gained pace after 1984-85 and increasing trend has been observed for all sectors except commercial and service sectors. Significant increase has been observed in transport and residential sectors which exceed the consumption level of industrial sector in 1992-93. In 1994-95, among the consumers of different commercial fuels, 35% of total commercial fuels was used for non-energy

purposes (e.g., fertilizer production). The residential sector consumed 19%, industry used 17%, and transport accounted for 19%. Commerce and services sectors accounted for 2% and agriculture sectors used the remaining portion. Commercial energy for the country comes from both indigenous and imported sources. Natural gas, hydropower, and a small amount of petroleum products are indigenous and the country imports a significant amount of petroleum. Over time, the share of oil in total energy has declined, and the use of natural gas has increased. In 1980-81, the share of oil in the total energy mix was about 56%, but by 1994-95 it had dropped to 36%. On the other hand, share of natural gas has increased from 32% to 54% during the same time.

1.2.2 Greenhouse Gas Emissions

A number of studies have estimated greenhouse gas emissions, but the Asian Least-cost greenhouse Gas Abatement Strategy (ALGAS) project conducted a comprehensive inventory of emissions in 1990 for energy, agriculture and livestock, and forestry and land-use changes using IPCC methodology. ALGAS studies revealed that per capita emission of greenhouse gas for Bangladesh in terms of CO_2 equivalent is less than a ton, about 670 kg per year. The detailed estimate revealed that 21,186 kt (kilo ton) of CO_2 equivalent greenhouse gas is released from the energy sector, of which 4,392 kt (kilo ton) is emitted by energy and transformation, 3,050 kt (kilo ton) from the industry sector, 1,875 kt (kilo ton) from the transport sector, and the remaining amount from small combustion and fugitive emissions. Combustion of commercial energy and use of forestry products are the main sources of CO_2 that accumulates in the atmosphere. The emission inventory of the forestry sector for 1990 revealed net emission of carbon. The existing forests can not absorb more carbon than the rate of removal from the forests. The present net emission from the forestry and land-use change sector is 19,738 kt (kilo ton) in CO_2 equivalent. Total methane emissions from livestock due to enteric fermentation is estimated at about 519 kt (kilo ton)

of which contribution by cattle is approximately 374 kt, nearly 72% of total emissions from enteric fermentation. Within the cattle population, dairy and non-dairy cattle accounted for 73 kt (kilo ton) and 301 kt (kilo ton), respectively. The goat population contributed about 123 kt (kilo ton) (24%) of total emissions. Buffaloes contribute about 3%, and the remaining 1% comes from sheep. The contribution of cattle population is nearly 83% of which dairy cattle account for about 31% and non-dairy cattle emit 52%. The contribution from poultry is small (4%). Methane gas emissions due to rice cultivation has been estimated at 767 kt (kilo ton). Much of it is due to non-irrigated rain-fed and deep water rice cultivation contributing about 518 kt (kilo ton) (68%) and 120 kt (kilo ton) (17%) of total emissions, respectively. The remaining methane gas is emitted from irrigated rice fields.

For emission of non-CO_2 gases was done in terms of field burning of paddy-straw only. The amount of CO_2 released was not estimated assuming that it would be balanced by growing plants during the next cropping season. The estimated amounts of CO, CH_4, N_2O, and NO_X released due to field burning of biomass reveal that about 695.4 kt (kilo ton) carbon and about 9.7 kt (kilo ton) nitrogen are being released annually from field burning in the form of 4.63 kt (kilo ton) of CH_4, 97.33 kt (kilo ton) of CO, 0.11 kt (kilo ton) of N_2O and 3.87 kt (kilo ton) of NO_x. Much of the CH_4 emissions (82%) in Bangladesh come from rice cultivation and livestock management. Of this, 43% is due to fermentation and manure management of livestock and the rest comes from the rice field.

1.3 Review of Literature

The environmental Kuznets curve (EKC) hypothesis proposes that there is an inverted U-shape relation between environmental degradation and income per capita. This has been taken to imply that economic growth will eventually redress the environmental

impacts of the early stages of economic development. The EKC is named for Kuznets (1955, 1963) who hypothesized that the relationship between a measure of inequality in the distribution of income and the level of income is an inverted U shape curve.

Grossman and Krueger (1991)[1] estimated EKCs for SO_2, dark matter (fine smoke), and suspended particles (SPM) using the GEMS data set. Each regression involves a cubic function of PPP per capita GDP and various site-related variables, a time trend, and a trade intensity variable. The estimated turning points for SO_2 and dark matter was at income level \$4000-5000. Income variables in each of their regressions were highly significant. The concentration of suspended particles appeared to decline even at low income levels. Both the time trend and the trade intensity variables had a significant negative coefficient in the SO_2 regression. Neither the time trend nor the trade variable was significant in the equation explaining the concentration of dark matter. The time trend was significant in the suspended particles regression but again the trade variable was insignificant. At income levels over \$10000-15000. Grossman and Krueger's estimates show increasing levels of all three pollutants. Though economic growth at middle income levels would improve environmental quality, growth at high income levels would be detrimental.

Beckerman (1992) explained "although economic growth usually leads to environmental degradation in the early stages of the process, in the end the best - and probably the only - way to attain a decent environment in most countries is to become rich." That is the core theme of EKC was forward by the World Bank's World Development Report 1992 (IBRD 1992). The authors noted that "The view that greater economic activity inevitably hurts the environment is based on static assumptions about technology, tastes and environmental investments" (p 38) and that "As incomes rise, the demand for improvements in environmental quality will increase, as will the resources available for investment" (p 39).

[1]This was the first EKC study. The paper was later published as Grossman and Krueger (1994). See also Grossman and Krueger (1995).

These views have been countered by critics of the EKC concept and literature (e.g. Arrow et al., 1995; Stern et al., 1996). The main arguments against the EKC are: much of the empirical evidence is weak and statistical techniques inappropriate, the static relationship between rich and poor countries does not necessarily tell us about dynamics as countries experience economic growth. Again EKC relationships have been found for only a subset of indicators and growth does not always improve the leves of environmental indicators (Stern, 1998).

Shafik and Bandyopadhyay (1992) estimated EKCs for ten different indicators: lack of clean water, lack of urban sanitation, ambient levels of suspended particulate matter, ambient sulfur oxides, change in forest area, the annual rate of deforestation between 1961 and 1986 (i.e., observations for each year), dissolved oxygen in rivers, faecal coliform in rivers, municipal waste per capita, and carbon emissions per capita. The results were used in the 1992 World Development Report (IBRD 1992). Data coverage and sources varied between the different indicators. They used three different functional forms: log-linear, log-quadratic and, a logarithmic cubic polynomial in untransformed PPP (Purchasing Power Parity), GDP per capita, site related variables and a time trend. They also carried out a number of additional regressions adding various policy variables such as trade orientation, electricity prices, etc.

Lack of clean water and lack of urban sanitation were found to decline uniformly with increasing income over time. Both measures of deforestation were found to be insignificantly related to the income but these data were notoriously poor (Stern et al., 1996). River quality tends to worsen with increasing income. The two air pollutants, however, conformed the EKC hypothesis. The turning points for both pollutants were found for income levels of between $3000 and $4000. The time trend was significantly positive for faecal coliform and significantly negative for air quality. Finally, both municipal waste and carbon emissions per capita increase unambiguously with rising income. Stern et al. (1994) criticized the model as the broader range of indicators examined by Shafik and Bandyopadhyay clearly shows a much more ambiguous

picture of the relationship between environment and development than indicated by Grossman and Krueger's more limited study Panayotou (1993) estimated EKCs for SO_2 , NO_x , SPM, and deforestation. This study employed only cross sectional data and GDP was in nominal at 1985 US dollars. The three pollutants were measured in terms of emissions per capita on a national basis. Data for developing countries were estimated from fuel use and fuel mix data. Deforestation was measured as the mean annual rate of deforestation in the mid 1980's plus unity. Since estimated, there were some problems with those data (see Stern et al., 1996). There were 68 countries in the deforestation sample and 54 countries in the pollution sample. The fitted equations for the three pollutants were logarithmic quadratics in income per capita. All the estimated curves were inverted U's. For the sample mean population density, the turning point for deforestation was $823 per capita. For SO_2 emissions the turning point was around $3000 per capita, for NO_x around $5500 per capita, and for SPM around $4500. The official exchange rates used by Panayotou tend to lower the income levels of developing countries and raise those of the developed countries relative to the PPP values. Despite this the turning points for the pollutants were in a similar range to those, reported by Grossman and Krueger and Shafik and Bandyopadhyay. This may be because Panayotou uses emissions per capita rather than ambient concentrations.

Shukla and Parikh's (1992) paper was primarily aimed to find the relationship between city size and ambient pollution levels for SO_2, particulates and smoke. Using cross-sectional data from WRI (1989) they found that pollution rose with city size. However, when they added GDP per capita and its square to the regression an inverted U was found with respect to city size. The EKC relationship holding city size constant was, however, U shaped, though except for in the particulates regression few coefficients were significant at conventional levels.

Selden and Song (1994) estimated EKCs for four airborne emissions: SO_2, NO_x, SPM, and CO on longitudinal data from World Resources Institute (WRI 1991). They

estimated a variety of specifications and presented their results for a fixed effect model including a population density variable. Authors suggested that in countries with low population densities, there will be less pressure to adopt stringent environmental standards and emissions due to transportation will be higher. The estimated turning points were all very high compared to other studies: SO_2, \$8709; NO_x, \$11217; SPM, \$10289 and CO, \$5963. Selden and Song reported that this was because ambient pollution levels were likely to decline before aggregate emissions. There was some support for this interpretation from Panayotou's (1993) results.

Cropper and Griffiths (1994) estimated three regional (Africa, Latin America, and Asia) EKCs for deforestation using panel data for 64 countries over a thirty year period. The dependent variable was the negative of the percentage change in forest area between two years. The independent variables in each regression were: rural population density, percentage change in population, timber price, per capita GDP and percentage change in per capita PPP, square of per capita PPP, a dummy variable for each country, and a time trend. Neither the population growth rate nor the time trend was significant in either Africa or Latin America, and the price of tropical logs was insignificant in Africa. The coefficients in these regressions were significantly different from zero. None of the coefficients in the Asian regression were significant. For Africa the turning point was \$4760, and for Latin America \$5420. These levels are very much higher than those from either Panayotou's or Shafik and Bandyopadhyay's results. Cropper and Griffiths concluded that economic growth will clearly not solve the problem of deforestation.

Lopez (1994) did a theoretical analysis of environment-growth relationships at a fairly high level of generality. The model used to analyze the relationship between pollution emissions and income has two production sectors, weak separability between pollution and the conventional factors of production, constant returns to scale, quasi-fixed inputs of capital and labor, exogenous technical change, and exogenous output prices. Preferences were a function of revenue, pollution, and the output price vector.

If producers pay a zero or fixed pollution price, then increases in output unambiguously result in increases in pollution in this system, irrespective of the features of the technology or preferences. However, when producers pay the social marginal cost of pollution then the relation between emissions and income depends on the properties of the technology and preferences. If preferences are homothetic, increasing output again results in increasing pollution. However, when preferences are nonhomothetic, as is likely in reality (Pollak and Wales, 1992), the response of pollution to growth depends on the elasticity of substitution in production between pollution and the conventional inputs, and the degree of relative risk aversion, i.e., the rate at which marginal utility declines with rising consumption of produced goods. The faster marginal utility declines and the more substitution is possible in production the less pollution will tend to increase with production. For empirically reasonable values of these two parameters pollution may increase at low levels of income and fall at high levels - the inverted U. This result is interesting, but its relevance is limited if the price of pollution is not socially optimal. Command and control measures on pollutants that show inverted us may result in effective prices that are close to being socially optimal while the effective price of pollutants such as carbon dioxide is generally close to zero. Also in the latter case the elasticity of substitution is probably lower and the apparent damage less evident to consumers, both implying a higher turning point. Lopez also constructs a model for deforestation, where as might be expected, if the stock effects of the forest on agricultural production are internalized then growth results in less deforestation and vice-versa.

Westbrook (1995) estimated a number of regressions explaining CO_2 emissions on a panel for 56 developing countries between 1971 and 1991. Of most interest was a regression of log emissions per capita on GNP per capita and its square and the shares of GNP in agriculture and services. All the coefficients in the regression were significantly different from zero. The emission-income relationship was an inverted U and as would be expected, the signs of the industrial structure coefficients are nega-

tive reflecting the lower emissions of agriculture and services relative to the industrial sector. The implication was that though industrial structure is a significant explanatory factor, other factors also contribute towards the inverted U relationship. In this paper generality was restricted by the omission of developed countries - though many other EKCs omit developing or low income countries instead.

Holtz, Eakin and Selden (1995) estimated quadratic EKCs for carbon dioxide emissions on panel data confirming the very high ($35000 in a level regression - $8 million in a logarithmic regression) turning points for this pollutant. They utilized a wide range of diagnostic tests and statistics. They also projected emissions over the next century under a number of different assumptions and a convergence based economic growth model. Antle and Heidebrink (1995) estimated EKCs for afforestation and national parks on cross-sectional data. They find a U shape curve for both indicators with turning points of 2000 and 1200 at 1985 US dollars respectively.

Becker (1996) examined some of the theoretical determinants of the EKC and their influence on possible growth and environmental quality trajectories. Compared to Lopez (1994), the framework used by Becker was not very general specific functional forms are introduced at the outset. The model consists of a transformation frontier between environmental quality (E) and other goods (G). Welfare is defined in the same arguments. Technical change can increase the potential to produce G but E cannot be improved beyond its pristine state. The paper examined the influence of different patterns of technical and preferential change in influencing the paths the economy takes. Becker argued for an inverted U shape path if preferences change over time in favor of environmental quality.

Rock (1996) estimated regressions where the dependent variables are two indicators of toxic intensity of GDP also used by Lucas et al. (1992). These equations included the quadratic function in income per capita, the share of manufacturing in GDP, and four different indicators of trade orientation. A separate regression was estimated. The inverted U is present and pollution is rising with the share of man-

ufacturing in GDP. A dummy variable for closed vs. open economies showed that closed economies had lower toxic intensities of GDP, while the growth rate of exports and the growth rate of the share of exports in GDP are both positively related to the pollution indices. The Dollar (1992) index of trade orientation hhas an insignificant coefficient.

Komen et al. (1996) estimated an EKC for public R and D expenditures on environmental protection in a group of OECD countries. The main result of the paper was that the elasticity of expenditures with respect to income was less than one but positive. The authors recognized that public expenditures are only a small part of total environmental R and D expenditures. Also R and D is only a small part of total expenditure on environmental protection and may or may not actually result in improved environmental quality. Nonetheless this was one of the links in an investigation of the empirical determinants of the EKC.

De Bruyn et al. (1996) estimated the following regression individually for West Germany, the Netherlands, the UK, and the USA for groups of 17 to 29 observations over the period 1960 to 1993;

$$ln(E_{jt}/E_{jt-1}) = \beta_{0j} ln(Y_{jt}/Y_{jt-1}) + \beta_{1j} + \beta_{2j} ln(Y_{jt-1}) + \beta_{3j} ln(P_{jt}/P_{jt-1}) + \varepsilon_{jt} \quad (1.1)$$

for CO_2, NO_x and SO_2. E is emissions, Y is income, and P is energy prices. Their result was that β_{0j} was significant and it was positive. In most of the case β_{1j} was zero or negative. β_{2j} was negative or zero more significantly for all three pollutants in Germany. β_{3j} was insignificant except for CO_2 in the USA. They argued that the declination in pollution seen in developed countries since the early 1970s was due to the slow rate of economic growth during this period which had not overcome the ongoing effects of the level of income in these countries in reducing pollution. They also calculate the economic growth rates that are compatible with zero emissions growth. For CO_2 these are 1.8% in the UK and Netherlands, 2.9% in Germany, and 0.3% in the US where the price effect has had most influence in reducing emissions. Zero emissions growth rates are much higher for the other two pollutants but lowest

for the UK, which has the least effective legislation. They also determined the long run EKC effect by $\beta_{1j} + \beta_{2j} ln(Y_{jt-1})$. They restricted one of those parameters to zero for most of the regressions.

Unruh and Moomaw (1996) also look at the effect of prices on CO_2 emission. They used exploratory data analysis of CO_2 emissions per capita in various countries rather than econometric analysis. They found that the transition to lower per capita emissions levels can happen at varying income levels and tends to happen fairly quickly.

Torras and Boyce (1996) used variety of impacts including the GEMS data analyzed by Grossman and Krueger (1994) and access to safe water and access to sanitation at the national level from the Human Development Report as dependent and income per capita, square of income per capita, cubic term if income per capita, the Gini coefficient of income distribution, literacy, an index of civil liberties, and the and some control variables similar to those used by Grossman and Krueger. They found for evidence of N shaped EKC. They used panel data. The parameter was estimated in the model in level and does not used any regression diagnostic. So the result demands more analysis.

Kaufmann et al. (1996) examined the impact of the spatial intensity of economic activity (GDP/Area) on ambient SO_2 concentrations for a panel of mostly developed and middle income countries between 1974 and 1989. The estimated equation is:

$$C_{ij} = \alpha + \beta_1 (Y/P)_{jt} + \beta_2 (Y/P)^2_{jt} + \beta_3 (Y/A)_{jt} + \beta_4 (Y/A)^2_{jt} + \beta_5 (S/GDP)_{jt} + \beta_6 t + \varepsilon_{jt} \tag{1.2}$$

Where Y/P is GDP per capita, Y/A is GDP per area, and S/GDP is steel exports as a percentage of GDP, which was intended to capture the effects of trade. The authors argued that this approach was superior to include population density as an RHS variable because the impact of population density would be expected to vary with the level of income per capita. The model was estimated by OLS, fixed effects, and random effects, and for both national average levels of Y/P and city specific levels

of Y/A. The authors found that concentrations were a U shape function of income per capita and an inverted U function of income per area. The former result is obviously diametrically opposed to the standard EKC results for sulfur dioxide concentrations. These findings were similar to Shukla and Parikh's (1992) results.

Liddle (1996) examined changes in the consumption/production ratio of various metals and paper in the OECD countries. In most cases there is no clear trend. This result would be evidence against a major role for trade in determining the EKC. He also predicted unabated levels of NO_x emissions for the same group of countries using the Dignon and Hameed (1989) method and compares these results with the reported estimates of emissions in those countries. The percent of emissions abated is constant or slightly increasing in most countries. Given that the Dignon and Hameed model predicts emissions as a constant fraction of fossil fuel consumption this implies a role for abatement technology in producing the downward sloping part of the EKC.

In contrast to Liddle (1996), Suri and Chapman (1996) argued for an important role for trade in generating the EKC relationship. They showed that the ratio of manufacturing exports to manufacturing imports has increased in many developing countries while it has decreased in many developed countries. They estimate an EKC for energy use per capita for the period 1970 to 1991 for 34 countries ranging from Bangladesh to the US in income levels. They estimated a logarithmic quadratic EKC with the addition of the following variables: M/MFG - imports of all manufactures as a share of domestic manufacturing production; X/MFG - exports of all manufactures as a share of domestic manufacturing production; and MFG/GDP - share of manufacturing in GDP. The estimates indicate an inverted U (turning point more than $50000) with the expected signs for the auxiliary variables but only the coefficient for X/MFG among the auxiliary variables are significant. Exclusion of the quadratic term results in all the coefficients being significant showing that these variables covary strongly with squared GDP per capita. A number of variant models are also estimated. A model with an interaction term M, GDP/MFG investigates whether

the effect of imports on lowering energy use is greater as income rises. This seems to be the case and the M/MFG coefficient falls to zero. A third version with a dummy variable for high income countries in the interaction term provides a better fit. This is the strongest evidence so far that trade effects are one of the causes of the flattening or downward slope of the EKC.

Horvath (1996) quantified a relationship for the growth rates of E/GDP with respect to income per capita. This relation was monotonically declining but the multiple correlation coefficient was almost zero. The intercept of the regression curve was very close to zero suggesting that even if an inverted U does exist for E/GDP, it is largely an artefact and individual countries all see declining or constant levels of E/GDP. This is an interesting result and should be examined for more conventional EKC relationships - de Bruyn et al. (1996) only look at rates of change in a few developed countries individually. It is normally argued that E/GDP in individual countries follows a pronounced inverted U curve (e.g. Reddy and Goldemberg, 1990).

Rothman (1996) contributed critique and estimation of some novel EKCs, which are intended to empirically, examine the determinants of the EKC. He updated some of the earlier critiques of the EKC arguing that much of the analysis to date has not really dealt with the main or interesting issues. He argued that one way to look at the trade issue is to look at the environmental impacts generated by consumption rather than by production activities in a country. To this end he estimated some EKCs for expenditure on resource intensive consumption goods. Only expenditure on food, beverages, and tobacco showed an inverted U within the sample range of income. The other categories of expenditure are monotonically increasing. The data were calculated on the basis of international PPP prices but there was no guarantee that the prices of individual commodities within groups reflect their environmental impact or that the impacts from goods consumed in the developed countries are equal to those in the developing countries.

Alberti and Layton (1996) were generally more positive about the utility of EKCs

than some of the other discussions. They point out that there are data on many environmental indicators, which have not yet been investigated using the EKC methodology. They argued that the EKC literature may be useful in choosing sustainability indicators. For example, an indicator such as sulfur dioxide concentrations would be a poor indicator of overall environmental impact because it is likely to decline with increasing development at high income levels while other impacts are increasing.

Perman and Stern (1999) have done an empirical study on EKC relationship using time series approach. They used panel data of 74 (OECD and Non-OECD) countries over 30 years time period. They choose to examine sulphur emissions.

To identify the time series properties they first use ADF test for testing the unit root of individual countries after subtracting a common time trend from data. Lag length for individual countries selected using Hall (1991) procedure. Their test infers that both the variables emission and income per capita are I(1) process for most of the countries. Only 18 countries out of 74 shown that income per capita is not integrated. The full search result is that only 6 definitely rejects the null hypothesis of unit root and 4 borderline case.

Again they use panel unit root test because conventional unit root test on a single time series, such as augmented Dickey Fuller test (ADF) procedure some time suffer from unacceptably low power when applied to a series of moderate length and are susceptible to large size distortion (especially in the presence of moving average error). The test result suggests that all three series for all country in the panel containing a single unit root. The statistics reinforce the findings of the individual country ADF test statistics. The only statistic suggesting a stationary process is the group statistic from a model including both heterogeneous trends and time dummies for $ln(M/P)$. Next they tried to determine the cointegrating relationship in modeling EKC. Test statistics did not provide strong evidence for cointegration in individual countries in the panel. In particular, in just under one half of all cases (35 out of 74 countries), none of the models exhibited cointegration between emissions per capita and the first

and second powers of income per capita. Cointegration was found most frequently where the regressions included a deterministic trend. The support for cointegration was substantially weaker in this case where the relationship allows for non-linearity. They also used panel cointegration test because, panel cointegration tests test a particular form of hypothesis, power comparisons are somewhat questionable (Maddala, 1998). Their empirical result was that there were some support for the hypothesis that there was a cointegrating quadratic relationship between emissions per capita and first and second powers of income per capita over the panel as a whole. This contention may be surprising given the country-by-country findings that gave little or no support for cointegration. Next they tried to fit a dynamic EKC in search of common long run parameter under the assumption that there exist a cointegrating relationship. They estimated an unrestricted dynamic EKC model for each country, and test whether the individual countries' emissions/income relationships converge to a common cointegrating vector. The estimated equation is an autoregressive distributed lag (ADL) model, parameterised in error correction form:

$$\begin{aligned}\Delta Y_{it} &= \{\alpha_i Y_{it} - \beta_{1,i} X_{1,it} - \beta_{1,i} X_{1,it}\} + \sum_{j=1}^{p-1} \chi_{ij} \Delta Y_{i,t-j} + \sum_{j=0}^{q-1} \delta_{1,ij} \Delta X_{1,i,t-j} \\ &\quad + \sum_{j=0}^{r-1} \delta_{2,ij} \Delta X_{2,i,t-j} + \mu_i + \eta_t + \varepsilon_{it} \qquad (1.3)\end{aligned}$$

Where Y, X_1 and X_2 stands for ln(Emission per capita), ln(Income per capita) and ln[(income per capita)^2] respectively. μ_i and η_t are country and time specific intercepts. They estimated the model using maximum likelihood and inference procedures and the statistical adequacy of this regression model assessed using conventional diagnostic statistics. This is in contrast to static regressions (including fixed effects) where, they are inapplicable. The hypothesis of a common cointegrating vector is decisively rejected when even when the sample is restricted to either OECD or non-OECD countries alone. Their Hausman (1978) test also supports rejection of homogeneity of the sub group of OECD and Non-OECD countries. Rejection of the homogeneity restrictions implies that each country has a unique β_1, β_2 parameter pair; strictly

speaking, each country has, therefore, a unique turning point in its EKC. At last they estimated a pooled mean group estimates under the null that the long-run parameters are constant over the panel but all other parameters including the speed of adjustment parameter and error variance varying over countries to find a common turning point. For the whole panel, the static fixed effect model yields a turning point at $82,746. For the non-OECD sample, the pooled mean group turning point estimate is $28 792 vs. $116 619 for the fixed effect estimator. Both estimates are out of sample and imply a monotonic emissions-income relation. The turning point implied by the average of the unrestricted estimates is a minimum at $403. This, too, implies an essentially monotonic relation. Only Tanzania and Myanmar had income lower than this and even then only for a few years in the 1960.

Finally they emphasized on the following findings on their concluding remark: (a) It is important to take difference to estimate and inference techniques if the data are non-stationary rather than stationary. (b) Their data used in their study stochastically non-stationary. They also suspect that many other indices of environmental pressure are also integrated variables. (c) There is weak support for the contention that there exist cointegrating relations between sulphur emissions per capita and income per capita in individual countries. Such cointegrating relations do exist in the panel as a whole. (d) A large minority of countries has basic shapes of emission/income relationships that do not have the EKC form. (e) They have no strong evidence that all individual members of the panel converge to a common cointegraing vector. (f) Static regressions such as simple fixed or random effects are badly misspecified, and are inappropriate for statistical inference (because of (a) and (b) above).

There is no doubt that Perman and Stern's (1999) work is progressive. They were much aware about the sample size. They used panel unit root and panel cointegration test to gain the efficiency over conventional ADF and cointegration tests that are much affected by small size sample. They might check how much efficiency achieved using panel unit root and cointegration test. Again they did not report any model

diagnostic result. Also they could use cross-validation technique so that their might be more reliable.

Bousquet and Favardy (2000) tried to investigate the link between pollution and income theoretically. Their study shows how income inequality affects environmental policies and therefore pollution. The Environmental Kuznets Curve (EKC) hypothesis proposes that there is an inverted U-shape relation between environmental degradation and income per capita. This paper invalidates this common result. Indeed they find for a set of parameters of two-hump curve.

Their result for a given inequality index is as follows;

First both income thresholds are decreasing functions of the inequality index. This means that a society crosses the line between the development phase and the environment phase for a lower average income when inequality is large. This is due to the redistribution impact of environmental policies, which are more stringent in highly unequal societies.

Second, they draw on the same figure the income inequality relationship according to the Kuznet's hypothesis. It is not possible to determine the sign of the variation for pollution, when average income increases and income inequality decreases. For particular values of the parameters it is then possible to have an increasing part for the pollution curve in this phase. In order to obtain a two hump curve for pollution it is necessary (not sufficient) that the maximum inequality occurs in this phase. Finally that in the two others phases pollution is always independent of income inequality. If income inequality is maximum in one of these two regions then pollution follows a simple inverted U curve.

Harbaugh et al.(2001) reexamined the empirical evidence for an environmental kuznets curve using the air pollution data studied by World Bank(1992) and Grossman and Kruenger(1995) with the benefits of retrospective data cleaning and additional ten years of data. They used robust approach to decide about the sensitivity

of the model. In their study they used data on ambient pollution level collected by the GEMS (Global Environmental Monitoring System) sponsored by WHO and UN. The EPA maintains these data in its aerometric information retrieval system (AIRS). Grossman and Kruenger(1995) used the same data but the new data contains substantially more usable observation than were originally available. Since the earlier release of GEMS data contained missing descriptive statistic, Grossman and Kruenger (1995) used indicator variables when covariates were unavailable. The latest data has no such gap. So Harbaugh et al. (2001) dropped the corresponding indicator variables. In addition missing observations for existing cities have been filled by the World Health Organization. They analyze three common air pollutants: SO_2, smoke and total suspended particulate (TSP). They also included trade intensity and the democracy index as additional covariates.

Using the same observations and econometric specification as was Grossman and Kruenger(1995), the changes in the data yield large change in the regression result and the shape of the predicted pollution-income relationship. Rather than increasing and then peaking at $4000, declines initially, then starts to increase at about $7000, at nearly the same point where the second regression line was actually decreasing at its highest rate. The line then starts to decrease again at about $14,000, about where the second regression line starts to increase. Using the most recent AIRS data and all available observations from 1971 to 1992 they showed that the individual GDP coefficients are generally highly significantly different from zero, which was not true of all of the preceding regressions. Again, the estimated pollution-income equations change significantly from those fitted using the original data, though the changes examined are minor. TSP and smoke, there were fewer changes to the data and therefore the regression results are less sensitive to those changes. In both the original and new data TSP decreases monotonically with GDP, although the slopes at $10,000 and $12,000 are smaller in the new data. For smoke, in both the original and new data the pollution concentrations exhibit an inverted-U, with a peak at about $6000.

Finally, they used chi-squared statistics from a Hausman test of whether the random-effects error terms are uncorrelated across the monitoring stations. In three of the four samples, this hypothesis rejected, suggesting that fixed monitoring-station effects are more appropriate. Changes such as logarithm of variables substantially alters the pattern of GDP coefficients; however, they do alter the predicted pollution-income paths. The logarithmic specification yielded an S-shaped curve with a peak at $3000 and a trough at about $13,000. Constructing confidence bands, they found that they were wide enough to incorporate a variety of GDP-pollution paths over the relevant range of GDP. So, monotonically rising or falling pollution-income paths, U-shaped or inverted U-shaped paths, or more complicated relationships all can easily fit within the 95 percent confidence bands, further demonstrating the extent of the uncertainty about the relationship between economic growth and pollution. These results suggest an alternative shapes of the pollution-income paths at intermediate incomes are not spurious results of multi-collinearity, the cubic functional form, or clustering of the data alone. However, the varying predictions at income levels outside this middle range, particularly at high incomes, may well be driven by the cubic function and by coefficient estimates that are imprecise due to multi-collinearity. In sum, the results here suggest that neither outliers in the data nor collinearity have caused the fragility of the early results depicting environmental Kuznets curves. Rather, the largest variations in the predicted pollution-income path have been the result of revisions and additions to the underlying data.

Their alternative specifications yield such drastically different patterns demonstrates the fragility of those earlier results. The key insight of this literature so far has been that pollution does not necessarily increase *deterministically* with economic growth. In fact, if it is true that pollution does not inevitably increase with economic growth but rather declines at some point, then the pollution-income path must be inverse-U-shaped. The estimated relationship between pollution and GDP is sensitive to both sample selection and empirical specification. Their conclusion is simply that,

for these pollutants, the available empirical evidence cannot be used to support either the proposition that economic growth helps the environment, or the proposition that it harms the environment. The next important empirical step for this line of research to take will be to categorize those pollutants and countries for which pollution has already begun to decline, simultaneously with economic growth. Rather than trying to fit a universal, reduced-form, pollution-income relationship, Such detail will better allow the various theoretical explanations to be tested.

Day and Grafton (2001) have done a work on EKC using the time series data of Canada. The purpose of their study is to assess the nature of the long-term relationship between per capita income and Canadian measures of environmental degradation and to determine whether increases in per capita income are associated with reductions in environmental degradation in Canada. They examined the relationship for four indicators of environmental degradation in Canada to income per capita, as measured by real GDP per capita. The four environmental indicators were emissions of CO_2 and concentrations of CO, SO_2 and total suspended particulate matter (TSP). For all four measures, an increase in the indicator implies an increase in environmental degradation. They first estimated a standard reduced form model of the relationship between environmental degradation and per capita income, and then they evaluate the model using various econometric tests. Their model includes per capita income term in levels and squared and cubed with a time trend.

Their model was of the form;

$$LED_{it} = \alpha_i + \alpha_2 LY_{it} + \alpha_3 LY_{it}^2 + \alpha_4 LY_{it}^3 + \alpha_5 t + \varepsilon_{it} \tag{1.4}$$

Where LED was the natural logarithm of the measure of environmental degradation and LY was the natural logarithm of real GDP per capita. They used the method of Ordinary Least Squares (OLS) to estimate the coefficients of equation 1.4. The sample sizes range from 38 observations for carbon dioxide emissions to 24 for concentrations of CO, SO_2 and TSP. Despite the fact that all the equations were estimated using time series data.

The reduced form results suggest that a comprehensive long-term beneficial relationship between per capita income and measures of environmental degradation does not appear to exist for Canada. They further investigated whether their data were non-stationary or not. To check stationarity of the data they used the Augmented Dickey-Fuller test and the Phillips-Perron test, carried out for various lag lengths. In all cases, the test equation includes both a constant and a linear deterministic trend. They fail to reject the null hypothesis of nonstationarity at either the 5% or 10% levels of significance for all cases. The log of per capita GDP was also found to be nonstationary. The Engle-Granger test and the maximum eigenvalue test suggest that per capita income and the measures of environmental degradation are not cointegrated, or that a long-term relationship between the variables does not exist. Causality tests also indicate a bi-directional causality, rather a uni-directional causality, from income to the environment. The results suggest that Canada does not have the luxury of being able to grow out of its environmental problems. In concluding remark they suggests that such form of standard reduced form model is spurious.

Day and Grafton (2001)used time series sequentially. The drawback of their article is that All test procedure they used e.g. unit root tests, tests for cointegration, causality and feedback check etc. are affected by sample of small size as they used in their study. No model diagnostic were available.

For snap views of some empirical paper along with their criticism are summarized in table 1.3. From the table 1.3 we saw that, most of the time the sample size, time series properties, multicollinearity problem an diagnostics were over looked for model estimation and test. In the thesis we tried to overcome all of those problems. Small sample problem is handled by most recently developed resampling techniques named bootstrap (discussed in chapter 2) approach first introduced by Efron(1979). It is a computer intensive approach of statistical inference. Bootstrap technique is very much helpful when appropriate sampling distribution is unavailable.

Table 1.1: Some empirical work summarization

Authors	Variables	Time Period	Countries/ Cities	Estimation Techniques	Comment
Grossman and Krueger (1991)	SO_2, TSP, Water Quality, GDP per capita and it's lag	'77, '82, '88	Various Countries from GEMS data set	Panel regression based on GLS	Most of the variables distribution are time dependent. So the study did not focused the long-run relationship
Shafik and Bandopadhyay (1992)	SPM, SO_2, Fecal Coliform in rivers, Sanitation, Municipal waste, Carbon emissions, Deforestation	1960-1990	149 countries	Panel regression based on OLS, log linear, quadratic and cubic	The data was not analyzed in time series point of view. Diagnostic was not available
Selden and Song (1994)	Panel of NO_x, SPM CO, SO_2	1973-75, 1979-81, 1982-84	22 OECD countries and 8 developing countries	Pooled cross section, fixed effect and random effect	Heteroscedasticity was absent and time dependency was not considered. So the long run forecast is not possible
Rock (1996)	Toxic intensities of GDP, Income per capita, 4 trade orientation indicators and a dummy variable of closed and open economy		cross sectional data of different countries	OLS	No diagnostics
De Bruyn et al. (1996)	CO_2, NO_x, SO_2, Income, Energy Price	1960-1993	West Germany, Netherlands, the UK and the USA	OLS	Only deterministic time trend considered, So for stochastic trend the regression may be spurious. Small sample time series problem was not considered

(Continued from previous page)

Authors	Variables	Time Period	Countries/ Cities	Estimation Techniques	Comment
Kaufman et al. (1996)	SO_2 concentration, population density, Income per capita	1974-1989	A panel of nearly developed, developing and middle income countries	OLS, fixed effect and random effect	Time series properties was not considered
Perman and Stern (1999)	SO_2 emission	30 years	Panel of 74 OECD and Non-OECD Countries	Used latest time series analyzing techniques. Model were estimated in classical approach.	Small sample time series problem faced by classical approach. Multicollinearity was not checked. Diagnostic results were not reported.
Harbaugh et al.(2001)	SO_2, Smoke, SPM		Data used by Grossman and Krueger (1991) with additional 10 years of data from GEMS, AIRS	Fixed effects, panel with polynomial in GDP and lagged GDP. Robust approach is used to decide about the sensitivity of the model	Have diagnostics. It's an extended work of Grossman and Krueger (1991)
Day and Grafton (2001)	CO_2 emission, CO concentration, SO_2 concentration and TSP	38 years for CO_2 and 24 years for the rest	Canada	Analyzed data using time series approach. OLS used for model estimation	Faced the problem of applying time series an lysing techniques for small sample. Multicollinearity not checked. Diagnostics unavailable.

1.4 Objectives

A well-established relationship between some measures of emission and economic development indicators like GDP per capita begs the question, Is there any relationship between economic growth and CO_2 emission for Bangladesh? If yes, what is the nature of the relationships and what is the future of emission in Bangladesh depending on economic development? Identifying such relation and providing the relationship on a solid foundation of statistics is the main objective of the thesis.

Objectives of the thesis are as follows;

- To seek the separate behavior of both CO_2 emission and GDP
- To study whether any relationship exists between CO_2 emission per capita and GDP per capita.
- To ascertain the nature of relationship.
- To apply tools of modern exploratory data analysis.
- To apply both classical and bootstrap methodologies for study of relationship.
- To provide model adequacy.
- To put forward policy suggestions.

1.5 Layout of the Thesis

We present a review of past works in chapter 1 togather with some criticism of growth environment relationships. Here we show that the relationship is based on unsound foundation considering the time series point of view. In chapter 2 we describe the analysis techniques used in the thesis to model the relationship between CO_2 emission per capita and GDP per capita. Both classical and bootstrap methods are described in an elaborated form. In chapter 3 we try to recognize something necessary information about data by exploratory data analysis. The conventional EKC model is fitted for Bangladesh data in chapter 4. To check the stationarity of data used in the study in chapter 5we first check which tests perform better. This is done by simulation.

Next we apply the tests for testing unit roots for our data series. In chapter 6 we fit ARIMA models for the data. In chapter 7 we test the cointegrating relations among the study variables and finally in chapter 8 we conclude the thesis.

Chapter 2

Methodology

2.1 Introduction

Generally statistical procedures are designed to be used with data originated from a series of independent experiment. The collected data or samples are taken to be the representative of some population. The *order* in which the sample is prsented to the statistician is irrevelent to the statistician. These classical techniques remains no longer relevent if the data appear orderly by time. The methods devised to deal with such data has developed from the last three decade and gained the ground of a special subject of research.

In classical statistics one has a vital concept of population and sample. The equivalent concepts with time ordered data are the stochastic process and its realization or observed series. The initial objective of time series analysis is to make inference about the properties of the stochastic process from the information contained in the observed series. The eventual aim is to construct a model and it is hoped that the model has similar properties to those of the generating mechanism of the stochastic process. Once a model has been obtained, the model then can be used either to test some hypothesis or theory about the generating mechanism of the process. It can also be used to forecast future values of the series and/or may be used to decide on

a system to control future values.

In the thesis, the recent techniques of time series data analysis has been used. The most recent developed computer based resampling techniques has also been used to overcome small sample time series problem. To introduce those techniques used for the thesis, the chapter is organized as follows;

The next section 2.2 describes the data used for analysis. Section 2.3 introduces some commonly used terms used in time series analysis, section 2.4 illustrates the unusual data events, section 2.5 describes the Box-Jenkins modeling strategy. In section 2.6 an overview is given on ARIMA modeling. Section 2.7 introduces the nonstationarity problem of time series, section 2.8 illustrates the idea of some dynamic models, section 2.9 pointes up the cointegration analysis, section 2.10 demonstrates the model adequacy checking, section 2.11, section 2.12 and section 2.13 describe the bootstrap approach and its application in time series estimation and hypothesis testing and finally in section 2.14 the necessary softwares are introduced to execute the whole analysis conveniently.

2.2 Data Source

Bangladesh is one of the newly born countries. The infrustracture of collecting the historical data is not well established. Ministry of government of Bangladesh has a department of environment. But their documentation is not well enough to use for such analysis. So data are collected from www.unep.org Geo data set. GDP per capita and CO_2 emission per capita are used mainly for analysis.

GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Money value has evaluated at constant 1995 US$ per person available from 1960-2001. Data has been collected

from World Development Indicators 2003 provided by The World Bank.

The second variable is Carbon Dioxide Emissions per Capita. It's unit is Metric Tons of CO_2 per Person. The data source is United Nations Framework Convention on Climate Change - United Nations Department of Economic and Social Affairs (UNFCCC-UNDESA/Statistics Division) provided by United Nations Statistics Division. The data is available for years 1972-2001 for Bangladesh.

2.3 Basic Definitions in Time Series

2.3.1 Time Series Data

A time series is a sequence of values or readings ordered by a time parameter (Granger and Newbold, 1977). In time series analysis the order is an important factor. So the classical statistical techniques are not sometimes relevant. If observations are taken every moment of time called continuous time series denoted by Y(t). Most of the time series data are collected after a certain time interval is called discrete time series denoted by Y_t.

Time series is thought of as a realization of a stochastic process. generally for every values of t a time series Y_t is generated by random input. So all values of t, Y_t is a random variable and a time series $Y_{t_1}, Y_{t_2}, \ldots, Y_{t_N}$ is a group of random variable. Apparently it is much tough to characterize such series because only one realization of each variable is obtained. Theoretically the distribution will exist.

Again since it is sometime impossible to get more realization of the variables under the stochastic process, time series analysts adopte the restriction *stationarity*.

2.3.2 White Noise Process

A white noise process is a sequence $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ whose elements have zero mean, constant variance and uncorrelated over time.

2.3.3 Lag Operator

In time series analysis the lag operator is a highly useful operator. Let us consider we have two sequence $\{x_t\}_{t=-\infty}^{\infty}$ and $\{y_t\}_{t=-\infty}^{\infty}$ where, the values of y_t at date t is equal to the values of x_t at date $t-1$.

$$y_t = x_{t-1} \tag{2.1}$$

The equation 2.1 is described by using lag operator to $\{x_t\}_{t=-\infty}^{\infty}$ as

$$L(x_t) \equiv x_{t-1} \tag{2.2}$$

where L is used as lag operator. Similarly we can apply lag operator twice will produce $L(L(x_t)) \equiv L^2(x_t) = x_{t-2}$ two lag of x_t. Double operation is indicated by "L^2" similarly k-th lag operation is indicated as "L^k" and $L^k(x_t) = x_{t-k}$

The lag operator follows some algebraic law. The lag operator and multiplicative operator are commutative. The lag operator is distributive over additive operator. We can freely use the commutative, associative and distributive law of algebra for multiplication and addition to express the compound operator. Sometimes polynomial in lag operator appeares in the time series. Those are algebraically similar to the polynomial defined on a scalar. The difference is that the simple polynomial refers to a particular number whereas a polynomial in lag operator refers to an operator that produce a new time series. For more see Hamilton(1994).

2.3.4 Stationarity

A time series process y_t is said to be covariance stationary or weakly stationary if its mean and variance is independent of time and covariance between y_t and y_{t-j} depends only on lag length j not on the time. Again a time series is said to strictly stationary if the joint distribution of y_t is independent of time. Next, the term stationary is taken to be covariance stationary. If the mean is constant over time, we may use all

sample observation to forecast it. If the process mean were different at each time period we cannot get useful estimates of its value since, in practice, we get only one observation for each time period. Non-stationarity occur due to deterministic or stochastic trend in the data. Process with stochastic time trend is termed as unit root process discussed in section 2.7. If a series is not stationary, we have to modify the series to be stationary. Since we know the modification, we can reverse them later to the original metric.

2.3.5 Box-Cox Transformation for Variance Stabilization

Many other transformations of data are possible. The log and square root transformations are members of a family of power transformations called the Box-Cox transformation (Box and Cox, 1964). With this transformation we define a new transformed series y'_t as

$$y'_t = \frac{y_t^{\lambda} - 1}{\lambda} \tag{2.3}$$

where λ is a real number and y_t is a time series. Note that y_t must not be negative. If some values of y_t are negative, we add some constant to y_t so all values are positive. After modifying the data, we may than return forecasts of this series to the correct overall level by subtracting some constant from the forecasts. For the case $\lambda = \frac{1}{2}$, for example, (2.3) gives the square root transformation and in the limiting case $\lambda \to 0$ Box-Cox transformation is continuous and leads to the natural log transformation.

Inspection of the time series plot may suggest an appropriate transformation. If the variance tends to rise as the level of the series rises, setting $\lambda < 1$ is called for. If the variance tends to fall as the level of the series rises, setting $\lambda > 1$ is called for. Further inspection of the plots of the data after various transformations may confirm which transformation seems best.

2.3.6 Stationary Mean

A series with a stationary mean returns fairly quickly to a constant mean. When a series is non stationary, usually we can create a new series with a constant mean by differencing the data for all t, as follows:

$$w_t = y_t - y_{t-1} \tag{2.4}$$

Performing this calculation once, for all t, is called first differencing. If the resulting series does not yet have a constant overall mean, we then compute the first differences of the first differences for all t. That is, the first difference of w_t series are

$$z_t = w_t - w_{t-1} \tag{2.5}$$

$$= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \tag{2.6}$$

The resulting series is called the second differences of y_t. Let d denote the degree of differencing. For first differencing $d = 1$. For second differencing $d = 2$. If the original data lack a constant mean, usually setting $d = 1$ will create a new series with a constant mean; setting $d > 2$ is all most never needed. When differencing is needed to achieve the stationarity, the series is said to be integrated, detail discussion is in section 2.7.

Seasonal Differencing

Usually, seasonal differencing induces a constant mean in a series that shift in a seasonal fashion. To performing seasonal differencing, we compute the successive changes between observations separated by s time periods, s is the number of seasons. For quarterly data, $s = 4$, for monthly data $s = 12$; and so forth. A series may be differenced non-seasonally only, seasonally only or, in both ways. Let D denote the degree of seasonal differencing. If $d = 0$, a seasonal differencing series ($D = 1$) is computed for all t as

$$w_t = z_t - z_{t-s} \tag{2.7}$$

Almost always, setting $D = 1$ remove any large seasonal shifts in the level of the series. If both non-seasonal and seasonal differencing are used, either one may be done first; the result is always the same.

2.3.7 Autocorrelation

Autocorrelation measures the direction (positive or negative) and strength of the relationship among observations within a single time series y_t when the observations are separated by k time periods, for $k = 1, 2, 3, \ldots, k$. For each k we offset the column of y_t observations by k time periods to create the column y_{t+k}; Thus we can have many autocorrelation coefficients for a single data series y_t, one for each k. Since we lose another observation on y_{t-k} each time k increases by one, the maximum useful value of k is somewhat less than n; a rough rule is to choose $k \leq n/4$, where n is the sample size A study of the autocorrelation patterns in a data series often can lead us to identify an ARIMA model for that series. We use sample data to obtain the information of the population autocorrelation coefficient at various lags $k = 1, 2, 3, \ldots, k$. This theoretical coefficient is defined as

$$\rho_k = cov(y_t, y_{t+k})/\sigma_y^2 \tag{2.8}$$

where $\sigma_y^2 = E(y_t - \mu_y)^2$, $\mu_y = E(y_t)$ and $cov(y_t) = E\left[(y_t - \mu_y)(y_{t+k} - \mu_y)\right]$.
For a stationary series $cov(y_t, y_{t+k})$, and therefore ρ_k , are independent of t. they depend only on k, the number of time periods separating y_t and y_{t+k}.

The sample autocorrelation coefficient, which provide an estimate of ρ_k, is usually computed as

$$\hat{\rho}_k = \frac{\sum_{i=1}^{n-k}(y_t - \overline{y})(y_{t+k} - \overline{y})}{\sum_{i=1}^{n}(y_t - \overline{y})^2} \tag{2.9}$$

The resulting set of values is the sample autocorrelation function, abbreviated SACF. This and other formulas for are discussed by Jenkins and Watts (1968). Any $\hat{\rho}_k$ is only a sample value that might differ from zero just because of sampling variation.

We can get some idea of the importance of the sample statistic by comparing it with its standard error. An approximation standard error for $\hat{\rho}_k$, due to Bartlett (1946), is

$$s\left(\hat{\rho}_k\right) = \left(1 + 2\sum_{j=1}^{k-1} \hat{\rho}_j^2\right)^{\frac{1}{2}} n^{-\frac{1}{2}} \tag{2.10}$$

To test for a linear association in the population between y_t and y_{t+k}, we test the null hypothesis

$$H_0; \rho_k = 0$$

against the alternative

$$H_1; \rho_k \neq 0$$

We then compute the approximate t statistics,

$$\hat{t} = (\hat{\rho}_k - \rho_k)/se(\hat{\rho}_k) \tag{2.11}$$

If t is significant at $\alpha\%$ (generally 5% or less), we reject null hypothesis.

2.3.8 Partial Autocorrelation Coefficient

Another useful measure of autocorrelation for stationary series is the partial autocorrelation coefficient. One way to think this coefficient is to consider the set of K regression equations:

$$\begin{aligned} y_t &= C_1 + \phi_{11} y_{t-1} + e_{1t} \\ y_t &= C_2 + \phi_{21} y_{t-1} + \phi_{22} y_{t-2} + e_{2t} \\ &\vdots \\ y_t &= C_k + \phi_{k1} y_{t-1} + \phi_{22} y_{t-2} + \cdots + \phi_{kk} y_{t-k} + e_{kt} \end{aligned} \tag{2.12}$$

The population partial autocorrelation coefficient at lag $k = 1, 2, \cdots, K$ is the last coefficient (ϕ_{kk}) in each equation. Each population coefficient is estimated for a given data set by its sample counterpart $\hat{\phi}_{kk}$. The resulting set of values is the sample partial

autocorrelation function abbreviated SPACF. In computing ρ_k we considered only two random variables y_t and y_{t+k}, and we ignored the intervening random variables $y_{t+k-1}, y_{t+k-2}, \cdots, y_{t+1}$. But in computing ϕ_{kk} we simultaneously take into account the role of theses intervening random variables, We can gauge the significance of each by comparing it with the standard error,

$$s(\hat{\phi}_{kk}) = n^{-\frac{1}{2}} \tag{2.13}$$

2.3.9 Correlogram

It is convenient to present the SACF and SPACF in graphical form. For $k = 1, 2, \cdots, K$ for example, the values of $\hat{\rho}_k$ and $\hat{\phi}_{kk}$ are represented on the graph by the length of the bars at lag k. If $\hat{\rho}_k$ is plotted against k, the graph is known as the sample correlogram and its population counterpart is known as the population correlogram. If the mean of a series is stationary, then the SACF and SPACF will tend to decay quickly toward zero. In practice, a quick decay means that the autocorrelation coefficients and partial autocorrelation coefficients are well below their two standard error limits by about lag 5 or 6. The ratio of the coefficients to their standard errors (the approximate t values) should fall to about 1.6 or less by about lag 5 or 6 available

2.4 Outliers in Time Series

2.4.1 Outlier

Outliers are unusual observation, which are different from the remaining part in the scence that they stands apart from the pattern shown by bulk of the observation. Outliers affects the validity of the subsequent analysis of the variable. There are three type of outliers occur in time series.

1. Additive Outlier (AO)
2. Innovative outlier (IO)
3. Level Shift (LS)

All the outliers may be described in a generalized form as follows;

Let us consider a process y_t contained the original series z_t and the contaminated part $f(t)$. The process can be written in the form:

$$y_t = f(t) + z_t \tag{2.14}$$

$f(t)$ can be of various forms. Here we will consider $f(t)$ of the type

$$f(t) = \frac{\delta(L)}{\phi(L)} I_t \tag{2.15}$$

where I_t is a binary variable. The outlier will occur in z_t at time $t = i$, $I_t = 1$ at $t = i$,$I_t = 0$ otherwise. $\delta(L)$ and $\phi(L)$ are the lag polynomial. For expositional brevity let us consider a simplest form of $f(t)$

$$f(t) = \frac{\delta}{(1 - \phi(L))} I_t \tag{2.16}$$

From equation 2.14 and equation 2.16, the process becomes

$$y_t = \frac{\delta}{(1 - \phi L)} I_t + z_t \tag{2.17}$$

- **Additive Outlier**

  If 2.17 occures with $\phi = 0$ y_t is shifted at time $t = i$. The shift being upward or downward depends on $\delta > 0$ or $\delta < 0$. Such an external event is called additive outlier. A common cause of AO is data recording error.

- **Innovative Outlier**

  Consider again equation 2.17 for $0 < \phi < 1$, like additive outlier y_t shifted at time $t = i$ and the shifting decays gradually with a Koyck type decay response. The decay rate is ϕ. This type of outlier is called innovative outlier.

- **Level Shift**

  Again in the contamination function in equation 2.17 for $\phi = 1$, and where $I_t = 0$ for all t and $I_t = 1$ for $t = i$, The series is shifted up or down permanently.

2.5 Box-Jenkins Modeling Strategy

To build a dynamic regression model the most popular strategy is associated with the name Gorge E. P. Box and Gwilym M. Jenkins, so called Box-Jenkins strategy. It has the following steps:

I. Identification

From the theory and practice a useful class of model is selected at the stage. To understand the time structured pattern exploratory data analysis "Letting data talk to us" is helpful to identify subclass of these too extensive model. The identification process can be used to yield rough preliminary estimates of the parameter in the model.

II. Estimation

At the stage we estimate the parameter of the tentative model. These parameters may be 1) Regression type coefficients or 2) ARIMA type coefficients. The regression type coefficients help to identify the input-output relationship. The ARIMA type coefficients help to understand the disturbance series autocorrelation pattern.

III. Diagnostic Checking

Diagnostic checks are applied with the object of uncovering possible lack of fit. If no lack of fit is indicated, the model is ready to use. If any inadequacy is found, the repetition cycle of identification, estimation, and diagnostic checking is undertaken until a suitable representation is found.

IV. Principle of Parsimony

Box and Jenkins emphasized the principle. According to the principle, a model should include the *smallest possible* number of parameters for adequate representation. The central role played by the principle of *parsimony* in the use of parameters will become clearer. The main objective of B-J modeling strategy is to obtain adequate and parsimonious models. Forecasting and control procedures could be seriously deficient if models are either inadequate or unnecessarily prodigal in the use of parameters. So, care and effort is needed for model selection. The process of selection is necessarily iterative, that is to say, it is a process of evolution, adaptation, or trial and error. The schematic diagram of the ittarative procedure is illustrated in figure 2.1

2.5.1 Dynamic Regression Model

Dynamic regression model (DRM) is a special family of statistical models. Such model states how an output (Y_t) is linearly related to current and past values of one or more input variables ($X_{1.t}, X_{2.t}, \ldots$) under the assumption that observation of the various series occur equally spaced time intervals. There is another crucial assumption is that *the inputs are not affected by the output.* This means that we are limited to single equation models.

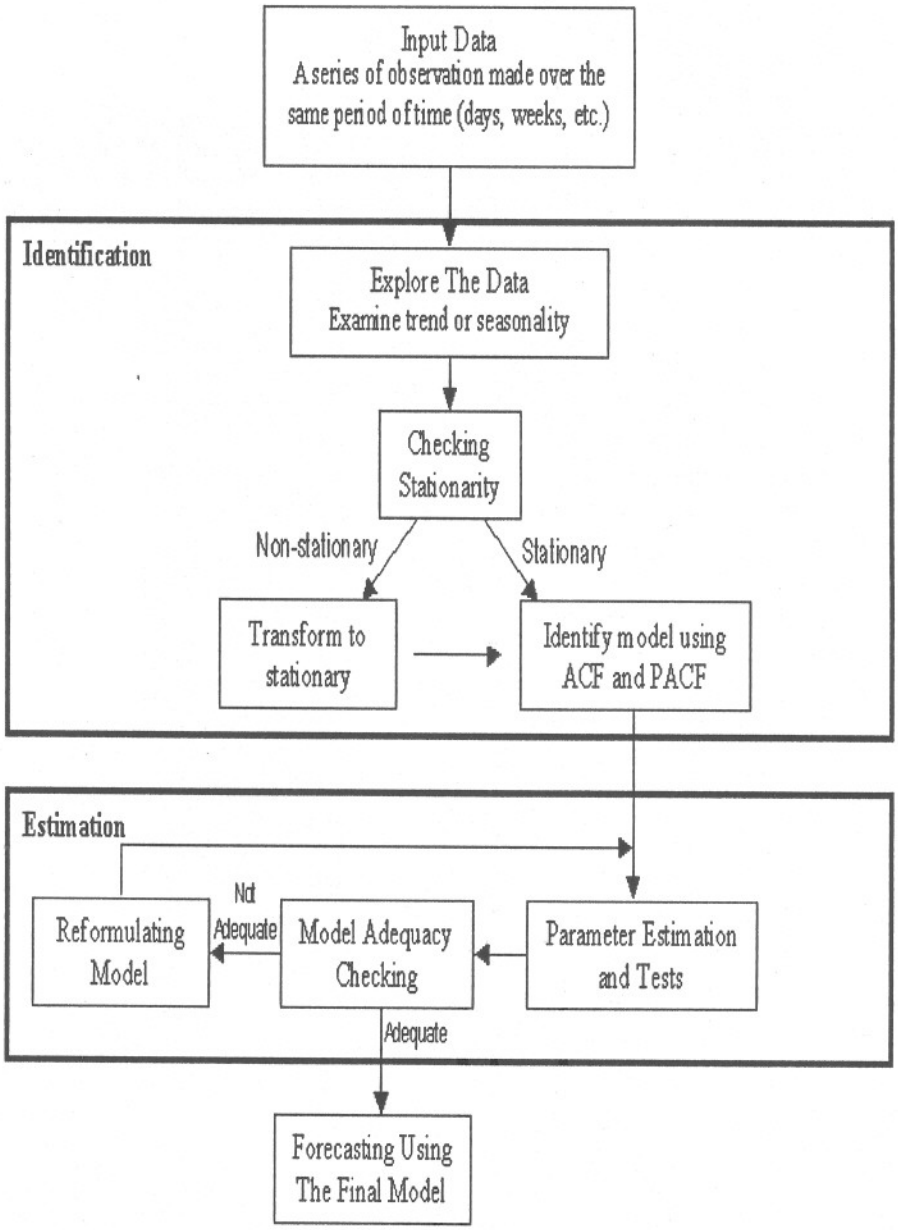


Figure 2.1: Functional diagram of the Box-Jenkins modeling strategy

2.6 ARIMA Model

Let us suppose that we have a time series observation of a variable. We want to forecast it. We may simply use the arithmetic mean of the data to forecast. But there may exits a dependency structure to its past values. It is appealing to give more weight to the more recent observation than the past. But the arithmetic does not consider it. However, The Box-Jenkins modeling strategy produces the best-weighted average for a single time series. These forecasts produced by the single equation Autoregressive Integrated Moving Average model. A single equation ARIMA model states how any value in a single time series is linearly related to its past values. It helps to choose an appropriate model. In theory, if the correct model is chosen, the ARIMA weighted average forecasts are "best" (The minimum of mean squared forecasts error).

There are some reasons to study ARIMA in connection with dynamic regression.

1. If the ARIMA forecasts better than DR then we may not go with DR due to extra trouble of working.

2. In a DR model the disturbance may be autocorrelated. This can be treated using ARIMA, which may improve the model and its forecasts.

3. To forecast the output variable with DR, we may need to forecast the input. Often ARIMA can produce forecasts conveniently.

4. Often ARIMA can announce some interesting feature of the input and output variable. This may help to build a DR.

5. Finally to perform diagnostics checking for DR model adequacy ARIMA models for stochastic inputs are needed.

2.6.1 Assumptions of ARIMA Model

Standard ARIMA process stands on a simple assumption that the process generated the single time series is stationary. There are two forms of stationarity a)Strong form b) Weak form. In practice we will go with weak form. If the random shocks are Gaussian then the two forms are identical.

Modification to induce stationary variance should be applied, if needed, before any further modifications or analysis of data in ARIMA model. We may be able to induce a constant variance by transforming the data. For example,

1. If the standard deviation of a series is proportional to its level, natural logarithms yield a new series with constant variance.

2. If the variance of the original series is proportional to its level, square root transformation induces a constant variance.

These two (especially the log transformation is both common and interpretable) are often useful in practice.
Usually we assume that the data are normally distributed. The usual inferential procedures depend on the normality assumption. Often a transformation to induces a constant variance brings the data closer to normality, but this is not always happen. The most popular transformation is the Box-Cox transformation.

To stabilize the mean over time (if necessary), differencing or seasonal differencing or both is used after stabilizing the variance.

2.6.2 Some Stationary ARIMA Process

The combined multiplicative non-seasonal $ARIMA(p, d, q)$ is as follows;

$$\phi(L)\Delta^d z_t = C + \theta(L)e_t \tag{2.18}$$

$\Delta^d = (1 - L)^d$ (The d order differencing operator)
$\phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \ldots - \phi_p L^p)$ (The p order AR operators)

$\theta(L) = (1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3 - \ldots - \theta_q L^q)$ (The q order MA operators)

e_t = Random Shocks, C is the constant and z_t is any time series.

When difference is not necessary to achieve stationarity, $d = 0$ and the model reduced to ARMA.

2.6.3 Theoritical Correlogram for ARIMA Model Identification

An ARIMA model is based on the available data. its theoretical counterpart is an ARIMA process. Each ARIMA model has a theoretical correlogram for associated autocorrelation function (ACF) and partial autocorrelation function (PACF). To identify an ARIMA model in practice we first contract the correlogram for Sample Autocorrelation Function (SACF) and Sample Partial Autocorrelation Function (SPACF) for a given data series. Then we compare the correlogram of SACF and SPACF with some common theoretical correlograms of ACF and PACF. In the following we present some common stationary ARIMA process and their associated theoretical correlograms of ACF and PACF.

The theoretical correlogram associated with AR(p) process will show for any value of p ACF decays exponentially to zero, while PACF will show significant spike at lag $1, 2, \ldots, p$. When $\phi_i > 0$ the ACF decays all on the positive side, and the PACF has p number of significant spike is positive. When any $\phi_i < 0$ ACF decays with alternating signs. stationary AR process of order p have these characteristics:

1. For the theoretical correlogram ACF decays, either exponentially or with a dumped sine wave pattern or with both of these patterns.

2. For the theoretical correlogram PACF has spikes through lag p, then all zeros. (Some values before lag p could be zero; the main point is that the last nonzero value occurs at lag p.)

The theoretical correlogram of MA(q) process has the form;

1. For an MA theoretical correlogram ACF has spike through lag q, then all zeros. (Some values before lag q could be zero; the main point is that the last nonzero values occurs at lag q.)

2. For an MA theoretical correlogram PACF decays.

3. For a mixed theoretical correlogram ACF and PACF both decay.

The SACF for integrated series decays slowly and thus obscures the rest of the patterns that may be present. Differencing permits us to uncover those other patterns.

In practice if we compare a sample correlogram that look much like one of the theoretical correlogram. The model we choose corresponds to the process whose theoretical correlogram match the sample correlogram. Then we estimate and check the tentative model if it is adequate.

2.6.4 ARIMA Modeling

The SACF and SPACF can tell much about the pattern of the data series. But they are based on sample information. So we must make allowance on sampling variation. In practice we should pay special attention to the nonseasonal $\hat{\rho}_k$ values that are about 1.6 times or more their standard errors in absolute value and seasonal $\hat{\rho}_k$ values that are about 1.25 or more their standard errors in absolute value. In the SPACF we should pay special attention to $\hat{\phi}_{kk}$ values that are two or more times their standard error. These are not rigid rules, they are guidelines.

Along with these practical rules, we should consider the overall pattern of $\hat{\rho}_k$ and $\hat{\phi}_{kk}$ coefficients. Not every moderately large spike or wave in an SACF and SPACF is worth our attention, since SACF coefficients are often correlated with each other. Thus an SACF can show moderately large waves that reflect only sampling variation.

When the model identification is complete, we used OLS or MLE to estimate the parameter. If the estimated residuals are gaussian white noise, we can use the

conventional t or F statistic to check the goodness of fit. Model adequacy checking and forecasting effeciency of the model helps to select the final model.

2.6.5 Model Checking

An ARIMA represents the autocorrelation pattern available in the data. So there should be no further significant autocorrelation pattern left in the residual series. Then we construct the SACF for the residual. This ensures whether the model adequately represents the autocorrelation pattern in the data. Again a standard assumption is that the random shocks are normally distributed. This permits us to perform approximate t tests for the coefficient significance at the estimation stage. One way to check the normality is to examine a histogram of the residuals. Another is to plot the model residuals in a normal probability plot. Both these procedure give a helpful representation of the data. To create a normal probability plot, the residuals are standardized and ordered from lowest to highest value. These values are plotted against a set of theoretical ordered normal values. If the residuals are normally distributed, the plots form a straight line. It is to be noted that Gaussian behavior of the estimated noise ultimately guarantees the adequacy of a model. So it is important for a researcher to seek for a model until he/she gets gaussian noise.

2.6.6 Other Model Selection Criteria

Another approach is to use the *Akike Information Criterion* (AIC) (Sakamoto et al., 1986) for checking model adequacy as well as models lag order selection. AIC is defined as;

$$AIC = log\left(\frac{\sum \hat{\varepsilon}_i^2}{N}\right) + \frac{2k}{N} \tag{2.19}$$

where $\sum \hat{\varepsilon}_i^2$ is the sum of the squared residuals. In principle, one could select a lag structure by increasing the number of lags up to the point where the AIC reaches a minimum value. Also there are some other statistic for model selection like *Schwartz Criterion* (SC) which is closely related to AIC and *Baysian Information Criterion*

(BIC) (Schwarz, 1978). Also there are some contradictions among econometricians about which criterion performs better. However, in the thesis, we use AIC only as our model selection criterion.

2.6.7 Forecasting

A correct ARIMA model gives minimum mean squared error forecasts among all linear univariate models with fixed coefficients. For each time period we can produce point forecast. We can also construct a confidence interval around each point forecast to give us an interval forecast. Interval forecast are sometime useful because it provides the possible degree of error associated with the point forecast. The 95% intervals for each forecast is $f \pm 2s$. where f denotes a forecast and s is its standard error. For a stationary model the forecasts converges to the mean of the series. How quickly or how slowly depends on the nature of the model and on how close the most recent observations are to the mean. For a nonstationary model, the forecasts do not converges to the mean.

Interpreting ARIMA Model Forecast

An ARIMA forecast is the best weighted average of the past observations. So it is easy to interprete. There are several other interesting and useful interpretations of ARIMA model forecasts that arise in special cases. one of such case is *exponentially weighted moving average* (ARIMA(0,1,1)) that arises often in practice. Their forecast have a special interpretation. For exposition, let us consider an ARIMA(0,1,1) model of the form taking the constant zero;

$$(1 - L)y_t = (1 - \hat{\theta}L)\hat{\varepsilon}_t \tag{2.20}$$

Multiplying both side of (2.20) by $(1 - \hat{\theta}L)^{-1}$ to get

$$(1 - \hat{\theta}L)^{-1}(1 - L)y_t = \hat{\varepsilon}_t \tag{2.21}$$

If $|\hat{\theta}| < 1$, using Taylor series expansion, we may write 2.21 as $(1 + \hat{\theta}_1 L + \hat{\theta}_1^2 L + \hat{\theta}_1^3 L + \ldots)(1 - L)y_t = \hat{\varepsilon}_t$. The first factor is equivalent to an AR operator of infilitely high order. The coefficients have special pattern $\hat{\phi}_i = -\hat{\theta}_1^i$. The model 2.20 is a parsimonious form of 2.21.

2.7 Unit Root and It's Consequences

2.7.1 Unit Root

Time series can be characterized in many ways. First, we want to focus on the presence of trends in the time series. There are two types of trends: (i) deterministic trends and (ii) stochastic trends. A stochastic trend is a random walk, which may or may not contain deterministic or stochastic drift. A time series that contain a random walk process is termed as a unit root process.

Now let us consider a univariate time series Y_t follows ARMA (p, q) process. The process can be written in the form;

$$(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p)y_t = c + (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q)\varepsilon_t \qquad (2.22)$$

provided that the roots of $(1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p)$ lies outside the unit circle. Dividing both sides of 2.22 with $(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p)$ we obtain

$$Y_t = \mu + \psi(L)\varepsilon_t \qquad (2.23)$$

where

$$\psi(L) = \frac{(1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q)}{(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p)}$$

and

$$\mu = c/(1 - \phi_1 - \phi_2 - \ldots - \phi_p)$$

$\sum_{j=1}^{\infty} \mid \psi_j \mid < \infty$. Roots of $\psi(z) = 0$ are outside the unit circle, $\{\varepsilon_t\}$ is a white noise sequence with $E(\varepsilon_t) = 0$ and $V(\varepsilon_t) = \sigma^2$. Two important features of such processes

are i) $E(Y_t) = \mu$ and ii) $\lim_{s\to\infty} \hat{Y}_{t+s/t} = \mu$, i.e., unconditional mean is independent of time and the forecast values converges to an unconditional mean. Such process is said to be a stationary process and stationary of an ARMA process depends only on autoregressive parameters ϕ_i not on the moving average parameters θ_j. This assumption is quite unappealing for many econometric time series. Most of the time series trends over time. To describe such trend a deterministic time trend is added to equation 2.23

$$Y_t = \alpha + \beta t + \psi(L)\varepsilon_t \tag{2.24}$$

The mean μ of the stationary process is replaced by $\alpha + \delta t$. Such a process is called trend stationary because the deduction of the term $\alpha + \delta t$ produces a stationary time series.

Another process of describing the trend is unit root process, i.e., first difference is stationary written in the form

$$(1 - L)Y_t = \delta + \psi(L)\varepsilon_t \tag{2.25}$$

where the assumption $\psi(1) \neq 0$ assures that the series remains non-stationary. The mean of $(1 - L)Y_t$ is denoted by δ rather than μ. Letting $\psi(L) = 1$ equation 2.25 can be written as

$$Y_t = Y_{t-1} + \delta + \varepsilon_t \tag{2.26}$$

The process is known as random walk with drift δ. The process 2.26 is also termed as an integrated process of order one shortly I(1).

2.7.2 Effects of Unit Root in Time Series Analysis

It is much important to check the stationarity of the time series. If the innovation ε_t is Gaussian and the process y_t generated by equation 2.24 contain a simple time trend then the OLS estimate of coefficients are Gaussian. Their test statistic t and F will follow exactly small sample distribution t and F (Hamilton 1994). But if there

exists a unit root in the process, the consequences is listed by Bierens H. J. (1999) as follows;

1. Regression involving unit root process may provide spurious regression because time series data often tend to move in the same direction. Consequently this may show a higher R^2 and lower Durbin-Watson (DW) statistic, which may not indicate the true degree of association between the study variables.

2. For two or more unit root process there may exist a linear combination, which may be stationary, and this can be interpreted as a long-run relationship. This circumstance will be called cointegration.

3. For a non-stationary time series y_t, if one would fit the model $y_t = \rho y_{t-1} + \varepsilon_t$ and test the null hypothesis H_0; $\rho = 1$ in the above AR(1) model, the null distribution is non-normal and it follows the Dickey-Fuller distribution.
   In short if a time series is generated by a unit root process, the conventional test procedures remains no longer valid.

So, it is important to check whether a time series is stationary or not. The detailed discussion are in chapter 5.

2.8 Some Dynamic Regression Models

Some other dynamic regressions models are in vogue along with univariate time series model like ARIMA. The models are much efficient to study the relationship among time dependent variables.

2.8.1 Vector Autoregressive Model

A vector autoregression system is a system in which each variable is regressed on a constant and p of its own lag as well as p lags of each of the other variables in the

VAR. The *pth-order vector autoregression* denoted by VAR(p) is as follows;

$$y_t = \pi + \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \ldots + \Pi_p y_{t-p} + \varepsilon_t \tag{2.27}$$

where $t = 1, 2, \ldots, T$, π denotes and $(n \times 1)$ vector of constants and Π_j is an $(n \times n)$ matrix of coefficients for $j = 1, 2, \ldots, p$ and

$$\varepsilon_t \sim \mathrm{IN}_p[0, \Omega_\varepsilon]$$

with Ω_ε an $(n \times n)$ positive definite matrix. The equation 2.27 can be written in the form

$$[I_p - \Pi_1(L) - \Pi_2(L^2) - \ldots - \Pi_p(L^p)]y_t = \pi + \varepsilon_t \tag{2.28}$$

or

$$\Pi(L)y_t = \pi + \varepsilon_t \tag{2.29}$$

where, $\Pi(L)$ indicates an $(n \times n)$ matrix polynomial in the lag operator L. The dynamic stability of the of the process 2.28 can be investigated by calculating the roots of the characteristic polynomial of the VAR

$$\Pi(z) = (I_p - \Pi_1 z - \Pi_2 z^2 - \ldots - \Pi_p z^p)$$

The roots of $|\ \Pi(z)\ | = 0$ contains all necessary information about the stability of the process. In econometrics it is more convenient to discuss the stability in terms of companion matrix of the process which can be obtained rewriting equation 2.28 as follows;

$$\begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p+1} \end{pmatrix} = \begin{pmatrix} \Pi_1 & \Pi_2 & \Pi_3 & \ldots & \Pi_p \\ I_p & 0 & 0 & \ldots & 0 \\ 0 & I_p & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & I_p \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ \vdots \\ y_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2.30}$$

Here the first block is the VAR process and the other are identity for y_{t-1}, y_{t-2} and so on. Stability depends on the eigen values of the coefficient matrix of 2.30 which is called the companion matrix of the process. If all eigen values of the companion matrix are inside the unit circle, then $\{y_t\}$ is stationary. If all eigen values are inside or on the unit the unit circle, then $\{y_t\}$ is nonstationary, else $\{y_t\}$ is explosive.

When there are unit roots in the model, we reformulate the VAR into an equilibrium correction model. There are different ways to reformulate the VAR described in the following;

For first order integrated process, The VAR expressed by equation 2.28 can be written in the form;

$$y_t = \Phi_1 \Delta y_{t-1} + \Phi_2 \Delta y_{t-2} + \ldots + \Phi_{p-1} \Delta y_{t-p+1} + \Pi y_{t-1} + \pi + \varepsilon_t \qquad (2.31)$$

where

$$\Pi = \Pi_1 + \Pi_2 + \ldots + \Pi_p$$

and

$$\Phi_s = -[\Pi_{s+1} + \Pi_{s+2} + \ldots + \Pi_p] \qquad \text{for} \qquad s = 1, 2, \ldots, p-1$$

Subtracting y_{t-1} from equation 2.31 produces

$$\Delta y_t = \Phi_1 \Delta y_{t-1} + \Phi_2 \Delta y_{t-2} + \ldots + \Phi_{p-1} \Delta y_{t-p+1} - \Pi y_{t-1} + \pi + \varepsilon_t \qquad (2.32)$$

where Π is redefined as $(\Pi = I_n - \Pi_1 - \Pi_2 - \ldots - \Pi_p)$. Equation 2.32 is the VAR in first order difference.

Another reformulation of VAR model using the second order difference (Acceleration rates) in level is of the form;

$$\Delta^2 y_{t-1} = \Phi \Delta y_{t-1} + \Phi_2 \Delta y_{t-2} + \ldots + \Phi_{p-1} \Delta y_{t-p+1} - \Pi y_{t-1} + \pi + \varepsilon_t \qquad (2.33)$$

where

$$\Phi = \Phi_s - I_n = -I_n - [\Pi_{s+1} + \Pi_{s+2} + \ldots + \Pi_p]$$

This presentation is useful when y_t contains time series variables of I(2). The general condition for $y_t \sim I(0)$ is that Π is non-singular. Stationary variables cannot grow

systematically over time. So for stationary y_t we can take expectation on both sides of 2.32 that yields;

$$0 = -\Pi E(y_{t-1}) + \pi \tag{2.34}$$

or

$$E(y_{t-1}) = \Pi^{-1}\pi \tag{2.35}$$

So when Π has full rank, the levels of stationary variables has a unique equilibrium mean. This is preciously why stationary is so unresonable for economic variable. When Π is not full rank (y_t exhibits I(1) behavior), 2.34 leaves some of the levels indeterminate. At the other extreme $\Pi = 0$, The VAR becomes one in the differences Δy_t, and these are stationary if $\Phi = \Phi_s - I_n$ has full rank, in which case $y_t \sim I(1)$. Notice that if $\Phi_s = I_n$ when $\Pi = 0$ makes Δy_t a vector of random walk, so $y_t \sim I(2)$.

2.8.2 Environmental Kuznets Curve (EKC) Model

A dispute issue of the environmental statisticians is the model named Environmental Kuznets Curve model. The basic EKC model is given by:

$$ln\left(\frac{M}{P}\right)_{it} = \alpha_i + \chi_t + \delta_i t + \beta_{1j} ln\left(\frac{Y}{P}\right)_{it} + \beta_{2j}\left[ln\left(\frac{Y}{P}\right)_{it}\right]^2 + \varepsilon_{it} \tag{2.36}$$

Where M is some measure of pollutant emissions or concentration (or some other index of environmental pressure), Y is national income, P denotes the population size of a country, and t is a deterministic time trend. The variables are observed over a panel of countries ($i = 1, \ldots, N$) and time periods ($t = 1, \ldots, T$). The random disturbances are assumed to be independent across countries. Non zero α_i terms allow for country-specific effects. The χ_t terms are time-specific dummy variables, usually interpreted as disturbances affecting all countries in the panel at some point in time in a common way and δ_{it} allows heterogeneous linear time trends over the sample of countries. Some or all of these country-specific or time-specific effects, or time trends, may be restricted either on the basis of prior information or after some specification search process.

A weaker version of the EKC hypothesis is that the EKC has a common form, equivalent to the restriction that $\beta_{1i} > 0$ and $\beta_{2i} < 0$ for all i, But these parameters have different values in different countries. A stronger version of the hypothesis is that the weak condition is satisfied and $\beta_{1i} = \beta_1$ and $\beta_{2i} = \beta_2$ for all i.

2.9 Cointegration

This section is concerned with the statistical analysis of multivariate systems of a special class of non-stationary time series using the concept of cointegration and common trend in the framework of the autoregressive model with Gaussian error. The methodology is to formulate economic hypotheses as parametric restrictions on the Gaussian VAR model and analyze these submodels. The section is written with the help of Hamilton (1994) and seminal papers of Hendry and Juselius (2000, 2001) part I and part II.

2.9.1 Description of Cointegration

An $(n \times 1)$ vector time series y_t is said to be cointegrated if each of the series taken individually is $I(1)$, that is nonstationary with a unit root, while some other linear combination $a'y_t$ is stationary, for some nonzero $(n \times 1)$ vector a. a is called the cointegrating vector. Cointegration means that although many developments can cause permanent change in the individual elements of y_t, there are some long run equilibrium relation tying the individual components together that is represented by their linear combination $a'y_t$. Existence of cointegrating relation between variables is not identifiable from the exploratory data analysis. The only way to find out such relation is through careful statistical analysis.

General Characteristic of Cointegrating Vector

The cointegration vector a is not unique. For if $a'y_t$ is stationary, then so is $ba'y_t$ for any nonzero scalar b. If a is a cointegrating vector then so ba. So, an arbitrary normalization is made such that the first element of a is unity.

If there are more than two variable in y_t, then there may be two nonzero $(n \times 1)$ vector a_1 and a_2 such that $a_1'y_t$ and $a_2'y_t$ are both stationary. Also a_1 and a_2 are linearly independent, then there may be $h < n$ linearly independent $(n \times 1)$ vectors $(a_1, a_2, \ldots, a_h)$ such that $A'y_t$ is stationary $(h \times 1)$ vector. where A' is the following $(h \times n)$ matrix

$$A' = \begin{bmatrix} a_1' \\ a_2' \\ \vdots \\ a_h' \end{bmatrix} \tag{2.37}$$

If $A'y_t$ is stationary, then for any nonzero $(1 \times h)$ vector b' the scalar $b'A'y_t$ is also stationary. Then the $(n \times 1)$ vector π given by $\pi' = b'A'$ could also be described as a cointegrating vector.

2.9.2 Cointegration in VAR

There are two general conditions necessary for $y_t \sim I(1)$ expressed as being cointegrated. The first condition is needed to ensure the data are not I(0), that is Π of equation 2.32 has reduced rank $r < p$. so can be written as;

$$\Pi = -\alpha\beta' \tag{2.38}$$

where α and β are $p \times r$ matrices, both of rank r, substituting 2.38 into 2.33 that will deliver the cointegrated VAR model of the form;

$$\Delta^2 y_{t-1} = \Phi\Delta y_{t-1} + \Phi_2\Delta y_{t-2} + \ldots + \Phi_{p-1}\Delta y_{t-p+1} + \alpha(\beta' y_{t-1}) + \pi + \varepsilon_t \tag{2.39}$$

An important feature of the matrices α and β are that they have orthogonal complements.

The second condition, which is needed to ensure that the data are not I(2) which requires that a transformation of Φ in (2.39) must be of full rank. That means If $r = p$, then y_t is stationary and standard inference procedure is applicable. If $r = 0$ then Δy_t is stationary, but it is not possible to obtain stationary relations between the levels of the variables by linear combinations. Such variables do not have any cointegration relations and hence, cannot move together in the long run. In this case each of 2.32 and 2.33 becomes a VAR model in difference and standard inference is still applicable as $\Delta y_t \sim I(0)$. If $p > r > 0$, then $y_t \sim I(1)$ and there exist r directions in which the process can be made stationary by linear combinations $\beta' y_t$ and 2.39 is said to be cointegrated VAR. Some time it is called vector error corrections model. The model allow us to investigate both short-run and long-run effects in the data.

2.9.3 Role of Deterministic Components in Cointegrated VAR

It is noticeable that the constant term as well as the other deterministic terms like dummy variables play a dual role in the dynamic regression model. This is also true for the cointegrated model like equation 2.39. When two or more variables share the same stochastic or deterministic trends, it is possible to find a linear combination that cancels both the trends. The resulting cointegrating relation is not trending, even if the variables by themselves are. In the cointegrated VAR model this situation can be accounted for by including a trend in the cointegration space. In other cases, a linear combination of variables removes the stochastic trend(s), but not the deterministic trend, so we again need to allow for a linear trend in the cointegration space. Similar arguments can be used for an intervention dummy. So it is important to understand the role of stochastic and deterministic components in the model. One can obtain biased parameter estimate if the deterministic components are incorrectly formulated. The asymptotic distribution of the cointegration tests are not invariant to

the specification of the components. Furthermore, the properties of the resulting formulation may prove undesirable for forecasting by inadvertently retaining unwanted components.

2.9.4 Restrictions on Constant Terms

To illustrate the idea let us consider a p dimensional cointegrated VAR with a constant and a linear trend of the form;

$$\Delta y_t = \Phi_1 \Delta y_{t-1} + \Phi_2 \Delta y_{t-2} + \ldots + \Phi_{p-1} \Delta y_{t-p+1} + \alpha(\beta' y_{t-1}) + \pi + \delta t + \varepsilon_t \quad (2.40)$$

All variables appeared in equation 2.40 are stationary. Without loss of generality the two $(p \times 1)$ vectors π and δ can be decomposed into two new vectors, of which one is related to the mean value of the cointegration relation $\beta' y_{t-1}$ and the other to growth rate in Δy_t. Later we described five baseline cases frequently occured for VAR specification.

Case 1. Trends and intercepts are unrestricted in the VAR model. This implies linear trend in the differenced series and thus, quadratic trend in the levels of the variable.

Case 2. Trend is restricted to lie in the cointegration space but the constant is unrestricted in the model. This restriction allows a linear trend in the model but precludes the quadratic trend in the data. $E(\Delta y_t) \neq 0$ implies linear deterministic trend in the level of y_t. In addition if $\mu \neq 0$, these linear trends in the variable do not cancel in the cointegrating relations, so the model contain 'trend-stationary' relations which can describe either a trend stationary variable or an equilibrium relation contains the trend. Therefor the hypothesis that a variable is trend stationary can be tested with the form of model.

Case 3. $\delta = 0$ and the constant term is unrestricted, In the situation, there are still linear trend in the data but, there in no deterministic trend in the cointegrating

relations. Also $E(\Delta y_t) \neq 0$, is consistent with the linear deterministic trend in the variable but, since $\mu = 0$ these trends cancel in the cointegrating relations.

Case 4. The constant is restricted to lie in the cointegration space. In this case there is no linear deterministic trend in the data, consistent with $E(\Delta y_t) = 0$. The only deterministic components in the model are the intercepts in any cointegrating relations, implying that some equilibrium means are different from zero.

Case 5. The model excludes all deterministic components in the data with both $E(\Delta y_t)$=0 and $E(\beta' y_t) = 0$, implying no growth and zero intercepts in every cointegrating relations. Since an intercept is generally needed to account for the initial level of measurements, y_0, only the exceptional case when the measurement starts from zero, or when the measurements cancel in the cointegration relations, in those situation the restriction is justified.

The two components δ and π play an important role in the cointegrated VAR and it is necessary to ascertain whether they are significantly different from zero or not at the outset of empirical analysis. Further topic of cointegration like tests for conintegration in VAR in both classical and bootstrap approachs are described in chapter 7.

2.10 Diagnostic Checking

Diagnostic checking is an important step of modeling. It is necessary to check the model whether the model is adequate or not. Any inadequacies revealed may suggest and alternative model specification. Residual analysis is usually the best way for diagnostic checking.

2.10.1 Checking White Noise Error

A residual analysis is usually based on the fact that the residuals of an adequate model should be approximately white noise. Autocorrelations of the estimated residuals from the model are checked. For a white noise process autocorrelations are zero. Therefore

the significance of the residual autocorrelations is often checked by comparing with approximate two standard error bounds $\pm 2/\sqrt{T}$, where T is the sample size used in computing the estimates. Plots of autocorrelation along with the approximate standard error bounds is helpful to take the decision. Ljung and Box (1978) statistic is also used to check the overall acceptability of the residuals autocorrelations.

2.10.2 Checking Normality

A standard assumption is that the random shocks are normally distributed. This permits us to perform approximate t tests on coefficient significance at the estimation stage. One way to check the normality is to examine the histogram of the residuals. Another is to plot the residuals in a normal probability plot (See Weisberg (1980) and Liu and Hudak (1986)). Both of the procedure provide a helpful graphical presentation of the data; they do not, however, provide any formal test for normality. There are various formal test for normality. The most recent and efficient test for normality of the fitted residuals is described below.

2.10.3 RM Test for Normality Check

There are various tests e.g. Jarque-Bera (1987), Bowman and Shenton (1975) for testing normality . It is well known that these tests suffers from possessing a lower size than the nominal level when the sample size is small. Again, these methods were proposed for testing the normality of the sample obtained randomly. But in regression, since the true innovations are unobserved, it is a common practice to use estimated residuals from the fitted model which may be correlated even, if the real innovations are not. In such situation, Imon (2003) proposed a new technique using the rescaled moment. The rescaled moments are used to estimate the coefficients of skewness and kurtosis. The test statistic of the test is as follows;

$$\upsilon = \frac{b_1^{*2}}{\sigma_1^2} + \frac{(b_2^* - 3)^2}{\sigma_2^2} \quad \sim \quad \chi^2_{(2)} \tag{2.41}$$

where $\sigma_1^2 = \frac{6}{n}$ and $\sigma_2^2 = \frac{24}{n}$. n is the size of the sample. b_1^* and b_2^* are the measure of skewness and kurtosis estimated from the rescaled moments defined as

$$b_1^{*2} = c^3 b_1$$

and

$$b_2^* = c^2(b_2 - 3) + 3$$

b_1 and b_2 are the Pearsons measure of skewness and kurtosis and $c = \frac{n}{n-p}$ where p is the number of estimated coefficient.

Imon (2003) showed that the test procedure is much more effective in non-normal cases when n is relatively small. The test have a slightly lower size than nominal but, they are much more accurate in size than the other techniques.

2.10.4 Checking Heteroskedasticity

A Important question often the applied worker facing is whether the variance of the residuals is independent of covariates or time. If heteroskedasticity does not arise, OLS produces BLUE of the coefficients as well as unbiased variance estimates. However if it goes undetected, our least square estimator will not be the best and is likely to lead misleading inference. On the other hand corrections for heteroskedastic error disturbance can lead to more efficient parameter esitmates. There are various kinds of test to check the heteroskedasticity of the residuals. Goldfeld-Quandt (1972) test and Breusch-Pagan (1979) test are popular.

There is another kind of heteroskedasticity present in which the variance of the regression error depends on the volatility of the errors in the recent past. A widely used model of such heteroskedasticity was developed by Robert Engle (1982) named Autoregressive conditional heteroskedasticity (ARCH) model. ARCH model would lead to increased efficiency. Tim Bollerslev (1986) proposed a modification of ARCH named generalized autoregressiove conditional heteroskedasticity moldel (GARCH). Both of the models have a wide application in the field of econometric research.

2.10.5 Checking Outliers

Outliers are unusual observations which are different from the majority. Those could really be a part of the data or could possibly be due to gross error such as incorrect key punching. Of course the latter kind are easily fixed. Real ourliers, however, are more difficult to handle. (References include Fox (1972), Kleiner et al. (1979), Martin (1980), Martin et al. (1983)) Undetected outliers may affect every subsequent analysis of the data series.

A careful inspection of time series plot is helpful for evaluating whether there are outliers in time series. Standardized residual plot is also preferable to detect the presence of outliers of the data. Standardized residuals greater than some number such as 3.0, 3.5 or 4.0 may be considered important. Of course, some residuals may be large just by chance. Formal outlier detection procedures are also available. Chang et al. (1988) proposed an iterative detection procedure using the estimated residuals from any DR. The method is not free of problem. The DR model identification may affected by outliers. Chen et al. (1990) discussed another iterative procedure designed to deal with this problem.

2.10.6 Checking Multicollinearity

In regression analysis, if the independent variables are correlated, is called multicollinearity problem. Multicollinearity violates no regression asumptions. Unbiased, consistent estimates will occur, and their standard erros will be estimated correctly. The only effects is to make hard to get estimates with small standard error. Multicollinearity is a question of degree and not of kind. The meaningful distinction is not between the presence and the absence of multicollinearity, but between its various degrees. It is essentially a sample phenomenon and we do not have any unique method of detecting multicollinearity of measuring its strength. We use observed high R^2 and few significant t ratios as a rule of thumb of detecting multicolinearity. Also eigen values and conditional index is used to measure the strength of multicollinearity.

2.11 Bootstrap

2.11.1 Basic Philosophy of Bootstrap

According to the view of Politis et al. (1999), Bootstrap is the statistic of 21^{st} century. The basic philosophy of bootstrap is quoted by the inventors own word in the preface of their book (Efron and Tibshirani 1993) as follows;

"The traditional road to statistical knowledge is blocked, for most, by a formidable wall of mathematics. Our approach here avoids that wall. The bootstrap is a computer based method of statistical inference that can answer many real statistical questions without formulas"

Bootstrap was first introduced by Efron (1979) for assigning the measure of accuracy of the estimates using the idea of resampling from a sample. It is indissolubly linked with computer to obtain a reliable estimate. The key idea is to resample from original data to create replicate datasets from which variability of the quantities of interest can be assessed without long winded analytical calculation. It established a new atmosphere for simulation a computer based technique. (See more Efron, B., and R. J, Tibshirani 1993). During the past four decades technological progress has made massive increase possible in the speed of digital computers. For scientific computing, a typical PC of today is several hundred times as fast as a typical PC of just a decade ago, although it costs less than half as much. Now a day a simple PC performs a huge amount of calculation in a single second. Again computer is also very cheap now. Development of computer technology has made econometric analysis an easy task to the applied econometricians.

The reason for using bootstrap inference is that hypothesis tests and confidence intervals based on asymptotic theory can be seriously misleading when the sample size is not large. For example, see Davidson and MacKinnon (2002a), Davidson and MacKinnon (1992). Of course, asymptotic tests are not always misleading. In many cases, a bootstrap test will yield essentially the same inferences as an asymptotic test

based on the same test statistic. Even in some situation when asymptotic theory fails or inapplicable, bootstrap produce better result. Although this does not necessarily imply that the asymptotic tests are unreliable, the investigator may reasonably feel greater confidence in the results of asymptotic tests that have been confirmed in this way. The main advantage of bootstrapped method is that any inference is possible for both which probability distribution is completely known or not. Again in classical approach, inference depends on normality of the residuals. But in real world situation, it is sometime impossible.

2.11.2 Simulation vs Bootstrap

The process of interest is usually called system. In order to study a system scientifically, we often made some assumptions about how it works. These assumptions which usually take the form of mathematical or logical relationship constitute a model. Generally the model is used to try to gain some understanding of how the corresponding system behaves. In simulation we use a computer to evaluate a model numarically, and data are gathered in order to estimate the desired true characteristic of the model.

In real world situation, the assumptions that constitutes a model remains unknown. Then we try to estimate the model from the observed sample. When the simulation is done using the estimated model, the simulation is termed as bootstrapping. A schematic difference of simulation and bootstrap is presented in figure 2.2

2.11.3 Bootstrapping IID Data

The technique of bootstrap is simple. Let us consider, we have a sample x which is independently and identically distributed. We just know the empirical distribution $\hat{F}_x$ of x. Let $\hat{\theta}(x)$ be the statistic of interest. We want to know the distribution of $\hat{\theta}$ so that we can draw the inference on θ. The bootstrap algorithm to draw the distribution is as follows.

1. Select B independent random samples depending on $\hat{F}_x$ $x_1^*, x_2^*, \ldots, x_B^*$ from given

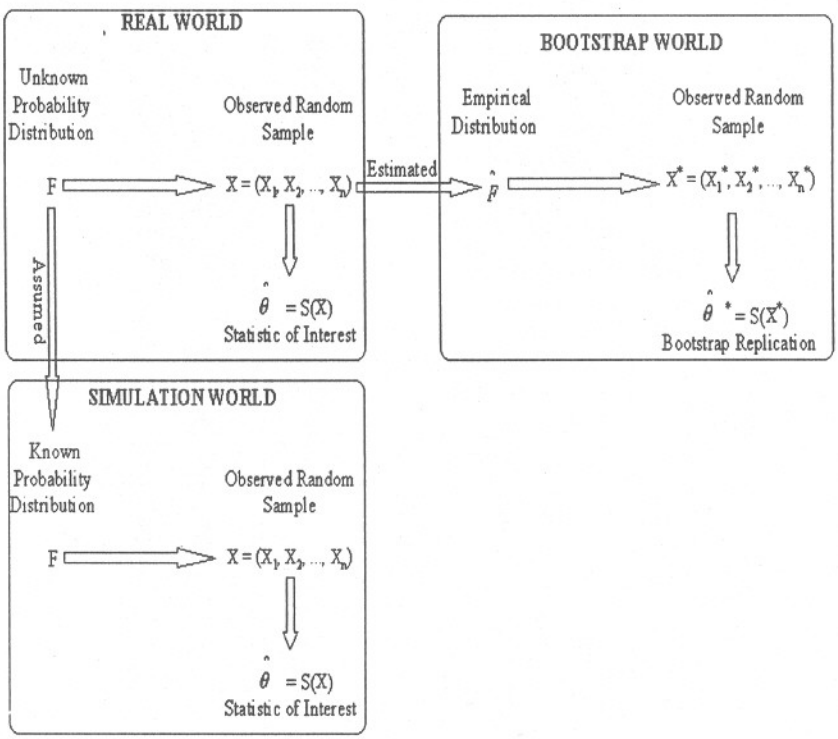


Figure 2.2: A schematic diagram of bootstrap and simulation as it applied to one sample problem.

sample x

2. Evaluate bootstrap replication corresponding to each bootstrap samples
3. Estimate the statistic of interest and realize the distribution of the statistic

The Illustration of bootstrap algorithm for estimating the standard error of an estimate is presented in the figure 2.3

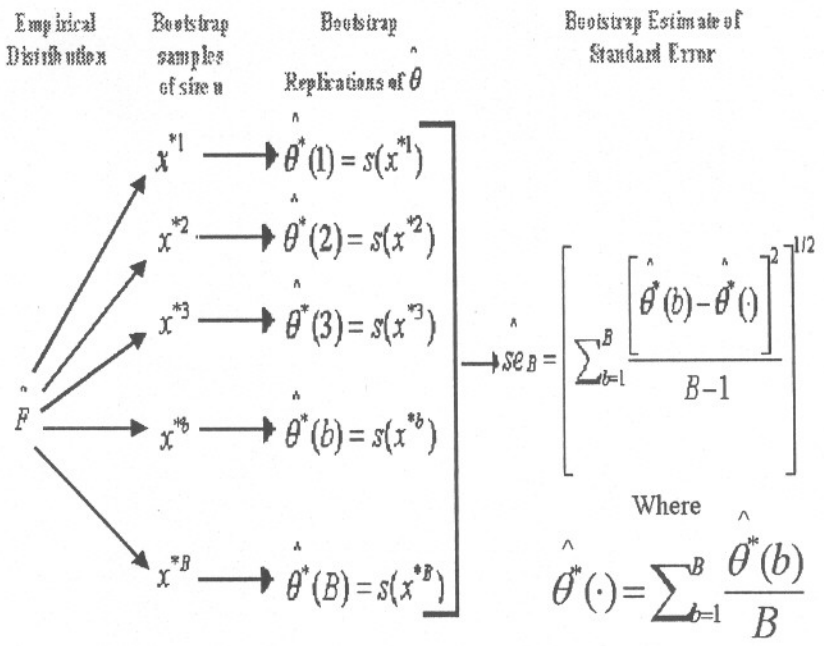


Figure 2.3: Bootstrap algorithm for estimating the standard error.

The procedure is termed as non-parametric bootstrap. Here it is notable that, the sample is independently and identically distributed. Each estimate from bootstrapped sample is called replication.

2.11.4 Parametric Bootstrap

Bootstrap sample can be carried out parametrically. In this procedure $\hat{F} = F_{\hat{\theta}}$ is an estimate of $F = F_{\theta}$ derived from parametric model to the data.

2.11.5 Bootstrapping Regression

In regression analysis, the mean of random response Y observed at value $x = (x_1, x_2, \ldots, x_p)$ of the explanatory vector variable is

$$E(Y/x) = \mu(x) = x'\beta$$

The model is complicated by specifying the nature of random variation. Linear regression with normal random error having constant variance, the least squares theory of regression estimation and inference provides exact method for analysis. But for generalizations of non-normal error, and non-constant variance, exact method rarely exists and we are faced with approximate methods based on linear approximations to estimators and central limit theorems.

Model Based or Fixed X Approach

The linear regression data set contains n data points x_1, x_2,...,x_n where x_i is itself a pair.

$$x_i = (c_i, y_i) \tag{2.42}$$

Here c_i is a $1 \times p$ vector of covariates, and y_i is the response. The probability structure of the linear model can be expressed as

$$y_i = c_i\beta + e_i \quad \text{for} \quad i = 1, 2, \ldots, n \tag{2.43}$$

The error terms e_i in equation 2.43 are assumed to be a random sample from an unknown distribution F having expectation 0. Thus the probability model $P \rightarrow x$ for linear regression has two components

$$P = (\beta, F)$$

where β is the parameter vector of regression coifficients and F is the probability distribution of the error terms. The general bootstrap algorithm requires to estimate P. The bootstrap algorithm for regression by estimating $\hat{P}$ is as follows;

1. Estimate the model coefficients $\hat{\beta}_i$ and $\hat{e}_i$. The obvious estimate of F is the empirical distribution on the $\hat{e}_i$.

2. With $\hat{P} = (\hat{\beta}, \hat{F})$ Select a random sample of bootstrap error terms

$$\hat{F} \rightarrow (e_1^*, e_2^*, \ldots, e_n^*) = e_i^*$$

3. The bootstrap responses y_i^* are generated according to 2.43 as

$$y_i^* = c_i\hat{\beta} + e_i^*$$

4. Fit the linear regression for $x^* = (c_i, y_i^*)$ and obtain $\hat{\beta}^*$.

5. Repeat 1-4 a number of times to obtain the bootstrapped distribution of $\hat{\beta}$.

Random X Approach

In this approach we can imagine the data as a sample from some $(P+1)$ dimensional distribution F. In such situation the regression coefficients are viewed as satatistical functions of F with no assumptions on e_t. For bootstrapping, it is appropriate to take $\hat{F}$ to be the empirical distribution of the data pair $(c_1, y_1), (c_2, y_2), \ldots, (c_n, y_n)$. The resampling simulation therefore involves sampling pairs with replacement from (c_i, y_i) so called pair wise bootstrap. Bootstrapping algorithm is as follows;

1. Sample $i_1^*, i_2^*, \ldots, i_n^*$ randomly with replacement from $\{1, 2, \ldots, n\}$;

2. For $j = 1, 2, \ldots, n$ set $c_j^* = c_{i_j^*}$, $y_j^* = y_{i_j^*}$;

3. Fit the regression to $(c_1^*, y_1^*), (c_2^*, y_2^*), \ldots, (c_n^*, y_n^*)$ giving the necessary estimates of interest.

4. Repeat 1-3 a number of times to obtain the distribution of estimates

2.12 Bootatrapping Time Series

Bootstrapping time series is not possible in the process because, time series has a complex time dependency structure. If one applies the IID bootstrap to the time series and if the data contain heteroscedasticity or autocorrelation, the randomly resampled data will not preserved the properties. So the statistic estimated from the sample will produce inconsistent result (Shing (1981), Babu and Shing (1983)). Thus, the IID bootstrap fails for time-dependent data. In such situation all the dependency information is lost. So the estimate σ_n^2 does not converges to the correct limit given by $\sigma_\infty^2 = \sum_{-\infty}^{\infty} \gamma(s)$ rather it converges to $\gamma(0)$ where $\gamma(s)$ is the autocovariance at lag s and $\gamma(0)$ is the variance of the series (Politis 2003). So algorithm of bootstrapping time series is a liitle bit complicated than IID data bootstrap and has a wide variey described latter.

2.12.1 Parametric or Model Based Bootstrap

Bootstrap in ARMA

Almost immidiately after Efron's (1979) paper Freedman (1981, 1984) introduced the residual based bootstrap for linear regression and bootstrap for autoregression was provided by Efron and Tibshirani (1986, 1993). They reduced the complex dependency of the data to an IID structure by fitting a parametric model. The residual based bootstrap is same as IID bootstrap of the estimated residuals from the fitted model. So it is sometimes called model based bootstrap or residual based bootstrap.

Let us consider a time series y_t follows AR(p) of the form;

$$\phi(L)y_t = \varepsilon_t \tag{2.44}$$

where $\phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p)$ is an invertible polynomial. The assumptions for the residual based bootstrap is that $\varepsilon_t \sim$ IID, $\mathrm{E}(\varepsilon_t) = 0$ and $\mathrm{E}(\varepsilon^2) < \infty$ The

series generated by the process 2.44 may be bootstrapped as follows;

Step-1 Determine the order p of the AR(p) process

Step-2 Estimate the parameter $\hat{\phi}(L)$ of the model

Step-3 Generate the bootstrap innovations ε_t^* from the residuals $\hat{\varepsilon}_t = \hat{\phi}(L)y_t$

Step-4 In such way the initial p observations is unobtainable because $Y_0^* = (y_1^*, y_2^*, \ldots, y_p^*)$ is unknown. It can be obtained by a random draw from y_t or can be used any other method of initialization described below at the subsection "Initialization Problem".

Step-5 Generate pseudo-data $\hat{\phi}(L)y_t^* = \varepsilon_t^*$ conditional on Y_0^*

Step-6 Then we estimate the bootstrap estimate of the parameter $\hat{\phi}^*(L)$ from the bootstrapped data series y_t^*.

Step-7 The procedure 1 to 6 repeated many times to build the empirical distribution of $\hat{\phi}(L)$.

Generally, AR(p) models are estimated by least square, Yule-Walker method may be used for small sample. If no intercept included in the regression model, the residuals must be centered by $\varepsilon_t^* - \bar{\varepsilon}$ to ensure the zero population mean. $\bar{\varepsilon}$ is the mean of $\hat{\varepsilon}_t$. Sometime residuals are rescaled by a factor $[T - p/(T - p - d)]^{\frac{1}{2}}$ to give the desired variance of the residuals. d denotes the number of estimated coefficients. (Stine(1987), Peter and Freedman(1984)). If the assumed model holds true, then the above procedure of bootstrap works well for sample mean or more complicated statistic even it enjoys a higher order accuracy property as was in Efron's IID. bootstrap. See Bose(1988).

An alternative class of parametric models are stationary MA(q) models of the forms;

$$y_t = \theta(L)\varepsilon_t \tag{2.45}$$

$\theta(L)$ is the polynomial of the lag operator. MA(q) models are rarely bootstrapped in econometric practice. Simulation results for the finite order stationary MA(1) model can be found in De Wet and van Wyk (1986) and Bose (1990). Bose (1990) proved that

the bootstrap approximation of the parameter estimates in moving average models are accurate to the order $o(T^{-\frac{1}{2}})$. In contrust, the asymptotic normal approximation is accurate only to the order $O(T^{-\frac{1}{2}})$.

Chatterjee (1986) applied the bootstrap algorithm to general ARMA(p,q) models of the form;

$$\phi(L)y_t = \theta(L)\varepsilon_t \tag{2.46}$$

The series generated by the process 2.46 may be bootstrapped as follows

Step-1 Determine the order p and q of the ARMA(p,q) process

Step-2 Estimate the parameter $\hat{\phi}(L)$ and $\hat{\theta}(L)$ of the model

Step-3 Generate the bootstrap innovations ε_t^* from the residuals $\hat{\varepsilon}_t = \hat{\theta}^{-1}(L)\hat{\phi}(L)y_t$

Step-4 Choose a large positive integer τ, set $y_t^* = 0$ for $t < \tau$ and generate IID draws for ε_t^* for $t = -\tau, \ldots, T$.

Step-5 Generate pseudo-data $y_t^* = \hat{\phi}^{-1}(L)\hat{\theta}(L)\varepsilon_t^*$ for ε_t^* for $t = -\tau, \ldots, T$ and retain the last T values of y_t^*

Step-6 Calculate the bootstrap estimate of the parameter $\hat{\phi}^*(L)$ and $\hat{\theta}^*(L)$ from the bootstrapped data series y_t^*.

Step-7 The procedure 1 to 6 repeated many times to build the empirical distribution of $\hat{\phi}^*(L)$ and $\hat{\theta}^*(L)$.

Under regularity conditions, Kreiss and Franke (1989) proved the asymptotic validity of the bootstrap approximation for ML estimators in the finite-order stationary ARMA models.

Bootstrap in VAR

Superficially, the bootstrap algorithm for VAR models is similar to the familiar algorithm for the regression model with fixed regressors. However, in autoregressive models the OLS estimates of the slope coefficients are systematically biased away from their population values. As a result, the standard bootstrap algorithm used by

Runkle (1987) may be misleading in small samples. The size of the bias depends on the sample size, the persistence of the data generating process and whether a deterministic time trend is included in the regression.

2.12.2 Initialization Problem

The effect of p initial conditions $Y_0 = (y_1, y_2, \ldots, y_p)'$ like all observations for Y_t are stochastic. Though the effect of conditioning on a particular set of initial conditions is asymptotically negligible, it is not appropriate to condition on Y_0 in order to generate bootstrap replicates. This type of initialization may generate a non stationary data series. To avoid such problem it is better to start the series in equilibrium or to generate a longer series of innovations and start Y_j^* for $j = 1, \ldots, t$ at $j = -k$. The 'burn-in' period $-k, \ldots, 0$ is chosen large enough to ensure that the observations $y_1^*, \ldots, y_t^*$ are essentially stationary. We can discarded the first $y_{-k}, \ldots, y_0$ observations.

Again bootstrap replicates conditional on Y_0 disapproves the randomness of Y_0^*. (Lütkepohl (1991), p. 496) suggested a way to randomize Y_0^* is to set

$$Y_0^* = \hat{\Gamma}^{-\frac{1}{2}} \left(\hat{\Sigma}_\varepsilon^{-\frac{1}{2}} \varepsilon_t^* \right)$$

where Γ is the estimate of $E(y_t y_t')$ defined by $\hat{\phi}(L)$

This procedure preserves the second moment structure in the data. The main drawback with this method is that it requires the estimated process to be stationary. For nonstationary coefficient estimates, the procedure breaks down because Γ is noninvertible. Even for borderline stationary processes. Rayner (1990) suggest an alternative method from the estimated moving average representation as $Y_0^* = \hat{\phi}^{-1}(L)\varepsilon_t^*$. But this requires a truncation of the infinite sum. Stine (1987) used another approach for initialization. He divided the observed data into $T - p + 1$ overlapping block of length p and randomly select one block for Y_0^* with replacement.

2.12.3 Lag Order Problem in Parametric Bootsrtap

The bootstrap can only be expected to perform well when the parametric model provides a good approximation to the true model. Determining the correct orders of an ARMA or AR model is thus a crucial issue. For example, Chatterjee (1986) reported simulation results for ARMA(1 ,1), ARMA(2,0) and ARMA(0,2) models. He compared bootstrap and asymptotic estimates of standard errors. Chatterjee's (1986) bootstrap results was quite satisfactory, but observed that much of the attraction of this method depends on selecting the right order. He noted that the bootstrap performs poorly if the selected order is not correct. Recent work by Kilian (1996a) offers some guidance on selecting the lag order. For the AR(p) model, the lag order selection criterion need not be consistent for the lag order for the bootstrap algorithm to be asymptotically valid. However, it is necessary that the probability of underestimating the true lag order is asymptotically zero. Provided that the range of lag orders considered includes the true lag order, this suggests that a wide range of information-based lag order selection criteria including the Akaike Information Criterion (AIC) are potentially valid criteria.

Kilian (1996a) also points out that the consequences of bootstrapping an over parameterized VAR model may be very different from those of bootstrapping an under-parameterized model. This suggests that lag order selection criteria such as the Schwarz information Criterion (SIC), which are known to be biased downward in small samples, will result in poor bootstrap estimates. Kilian's simulation results confirm that in small and moderate samples the coverage accuracy of bootstrap confidence intervals for VAR impulse response estimates is much closer to nominal coverage for the AIC than for more parsimonious criteria such as the SIC or the Hannan-Quinn Criterion.

2.12.4 Non-parametreic Bootstrap for Time Series

Sieve Bootstrap

If the true model is infinite ordered, the asymptotic justification of the bootstrap approximation proposed by Bose (1988) and Kreiss and Franke (1992) is no longer valid. Bootstrap algorithms designed for this class of processes: the sieve bootstrap proposed by Bühlmann (1996a). He considered a class of linear, infinite dimensional process which can be approximated by a sequence of finite-dimensional autoregressive approximations. This so-called sieve bootstrap is model free within the class of linear MA(∞) processes with polynomial decay. Bühlmann (1996a) proved that the sieve bootstrap gives correct approximations to the distribution of smooth functions of linear statistics of the data. Bühlmann (1996b) studied the sieve bootstrap for autoregressive models including a deterministic time trend.

Cholesky Factor Bootstrap

Diebold, Ohanian and Berkowitz (1995) formulated a bootstrap algorithm which does not require conditioning on any particular parametric model of the VARMA type. The context is the vector covariance stationary MA(∞). Any finite realization of length T thus has representation:

$$Y = P\varepsilon \tag{2.47}$$

where Y is $rt \times 1$ vector of time series variable, P is $rT \times rt$ and ε is $rT \times 1$. The bootstrap procedure is as follows;

1. Consistently estimate $COV(Y) = \Sigma$
2. Take the Cholesky decomposition $\hat{P}\hat{P}' = \hat{\Sigma}$
3. Resample from the normal distribution $\varepsilon^* \sim N(0, \hat{\Sigma})$
4. Generate pseudo-data: $y^* = \hat{P}\varepsilon^*$

5. Calculate bootstrap statistics of interest

6. Repeat steps 3-5 many times and build up the empirical distribution of interest.

This "Cholesky factor" algorithm is a model-free method for generating pseudo data focusing on the second moment properties of the observed data. Note that in place of the lag order selection problem in parametric models, the nonparametric Cholesky factor bootstrap requires a bandwidth choice. Data-based bandwidth selection procedures for consistent covariance matrix estimation may be found, for example, in Andrews (1991), Andrews and Monahan (1992) or Newey and West (1994).

Block Bootstrap

A different strategy has focused on resampling blocks of contiguous time series observations. In the method $b = Tk + 1$ blocks of data is defined. The blocks may be overlapping or not. The method resamples the blocks of data (moving blocks). For more see Künsch (1989) and Liu and Singh (1992). Resampling overlapping blocks may provide somewhat higher bootstrap estimation efficiency than non-overlapping blocks, although the available evidence indicates that the efficiency gain is small (e.g., Hall, Horowitz and Jing (1995)).

2.12.5 Forecasting with Bootstrap

An AR(p) process forecasts is a conditional distribution of all past observation, is the same as the conditional distribution on the last p observation of the sample realization. Bootstrapping such conditional distribution requires the last p observation identical to the original series. So the model based bootstrap algorithms are not appropriate for forecasting. Thombs and Schucany (1990) proposed a method to construct forecast confidence interval with bootstrap approach using backdraw representation of a time series (Box and Jenkins, 1976, pp. 197-200). The correlation structure of forward representation and backward representation remains the same. So, y_t can be expressed

as;

$$\phi(B)y_t = a_t \tag{2.48}$$

where B indicates the forward difference operator, $B = L^{-1}$ and $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$, i.e., y_t expressed as a linear combination of future values plus error term. Resampling (IID) can be used for backdraw residuals a_t and hence reproduce the time series fixing the last p observations as initial observation. Thombs and Schucany (1990) algorithm for bootstrap prediction is summarized in the following;

1. Determine the order of AR(p) process.

2. Estimate the parameter $\hat{\phi}(L)$ from the observed data. Obtain the forward residuals $\hat{\varepsilon}_t$ and backward residuals $\hat{a}_t$. Let $\hat{F}_\varepsilon$ and $\hat{F}_a$ be the empirical cdf of the centered and rescaled forward and backward residuals respectively.

3. Then we generate the bootstrap innovations a_t^* from the backward residuals $\hat{a}_t$

4. Generate bootstrap realization $\{y_t^*\}$ for $\{y_t\}$ keeping fixed the last p observation of the samples via

$$y_t^* = y_t \quad \text{for} \quad t = T, T-1, \ldots, T-p+1 \tag{2.49}$$

$$y_t^* = \hat{\phi}(B)y_t^* + a_t^* \quad \text{for} \quad t = T-p, T-p-1, \ldots, 1 \tag{2.50}$$

5. Calculate $\hat{\phi}^*(L)$ for y_t^* and compute the bootstrap future values using the model

$$y_{t+l}^* = \hat{\phi}^*(L)y_{t+l-j}^* + \hat{\varepsilon}_{t+l}^* \tag{2.51}$$

6. Repeat steps 3-5 many times and buildup the conditional empirical distribution of the l-step ahead forecasts $\hat{y}_{T+h}$.

Breidt, Davis, and Dunsmuir (1992, 1995) also proposed a method for conditional bootstrap prediction which is little bit different from Thombs and Schucany (1990). McCullough (1994) finds that the conditional forecast distributions implied by the Thombs and Schucany (1990) procedure are very different from those implied by

the Breidt et al. procedure. Sharker and Nasser (2004) extend the algorithm to ARMA(p,q) empirically for small sample. The efficiency of their technique is not well justified.

2.13 Hypothesis Testing with Bootstrap

Let us consider a sample $\{x_1, \ldots, x_n\}$ be an IID sample n random variables with distribution function F and the parameter $\theta(F)$ which is a real valued functional statistic to be tested for $H_0 : \quad \theta(F) = \theta_0$. To test this hypothesis, Beran (1984, 1986, 1988) proposed two alternative approaches described in the following;

2.13.1 The Test Statistic Approach

This approach approximates the exact critical region of the test by the percentile of its bootstrap distribution. The unknown distribution of the sample is replaced by its empirical distribution.

Let $T_n = T_n(X_1, \ldots, X_n)$ be the test statistic for the null hypothesis $H_0 : \quad \theta = \theta_0$ vs. $H_1 : \quad \theta \neq \theta_0$ and $F_T(\xi; \theta_0) = Prob(T_n < \xi)$ is the distribution of T_n under H_0. The critical region of the test is given by

$$d_L(\frac{\alpha}{2}; \theta_0, F) = F_T^{-1}(\frac{\alpha}{2}; \theta_0) \tag{2.52}$$

$$d_U(1 - \frac{\alpha}{2}; \theta_0, F) = F_T^{-1}(1 - \frac{\alpha}{2}; \theta_0) \tag{2.53}$$

where α is the significance level of the test. In most application $F_T(\cdot)$ is unknown and therefore have to be replaced by estimates

$$d_L^*(\frac{\alpha}{2}; \theta_0, \hat{F}) = F_T^{*-1}(\frac{\alpha}{2}; \theta_0) \tag{2.54}$$

$$d_U^*(1 - \frac{\alpha}{2}; \theta_0, \hat{F}) = F_T^{*-1}(1 - \frac{\alpha}{2}; \theta_0) \tag{2.55}$$

where $F_T^*(\cdot)$ is the bootstrap distribution of T_n based on the empirical distribution of X given H_0 is true.

2.13.2 The Confidence Region Approach

In this approach a confidence region for the estimated parameter is constructed and the null hypothesis is rejected if θ_0 lies outside the confidence interval. Whenever the parameters of interest can be transformed to a normal pivot, an appropriate critical region can be derived without knowledge of the exact functional form of the transformation.

Let $R_n(X,\theta)$ be a pivot for θ where $X=(X_1,\ldots,X_n)$ and $F_R(\xi;\theta)$ is the corresponding c.d.f. of R_n. The critical values of the confidence region approach are

$$c_L(\frac{\alpha}{2};\hat{\theta},F) = F_R^{-1}(\frac{\alpha}{2};\hat{\theta}) \tag{2.56}$$

$$c_u(1-\frac{\alpha}{2};\hat{\theta},F) = F_R^{-1}(1-\frac{\alpha}{2};\hat{\theta}) \tag{2.57}$$

Replacing F_R by its bootstrap approximation yields

$$c_L^*(\frac{\alpha}{2};\hat{\theta},\hat{F}) = F_R^{*-1}(\frac{\alpha}{2};\hat{\theta}) \tag{2.58}$$

$$c_u^*(1-\frac{\alpha}{2};\hat{\theta},\hat{F}) = F_R^{*-1}(1-\frac{\alpha}{2};\hat{\theta}) \tag{2.59}$$

It is sometimes difficult to find the pivot $R(X,\theta)$ and for some parametric families no such pivot exists. Furthermore, in non-parmetric cases pivots should be constructed which hold for all possible distributions. Efron (1979) considered a class of estimators $\hat{\theta}$ for which a monotonic increasing function $g(\cdot)$ exists, such that the transformed quantities $\phi=g(\theta)$ and $\hat{\phi}=g(\hat{\theta})$ satisfy

$$(\hat{\phi}-\phi)/\tau \sim N(0,1) \tag{2.60}$$

where τ is the standard error of $\hat{\phi}$. Since the percentile method for constructing bootstrap confidence interval is transformation invariant, the normalization to the pivot (2.60) is automatically incorporated. Efron (1982, 1987) extends the results to

$$BC: \quad (\hat{\phi}-\phi)/\tau \sim N(-z_0,1) \tag{2.61}$$

and

$$BC_a: \quad (\hat{\phi}-\phi)/\tau \sim N(-z_0(1+a\phi),(1+a\phi)^2) \tag{2.62}$$

where z_0 and a are suitable chosen constants.

2.13.3 Double Bootstrap

As with most statistical methods, the bootstrap does not provide exact answers. Beran(1988) showed that the bootstrap inference is refined when the quantity bootstrapped is asymptotically pivotal. A statistic is said to be asymptotically pivotal if its limiting distribution does not depend on DGP $\mu \in M$ where M is a model, a set of DGP. Most statistic commonly used in econometric practice is asymptotically pivotal. Many statistics of interest based on AR(p) and ARMA(p,q) models are asymptotically normal and can be studentized to make them asymptotically pivotal.

If an asymptotic pivot $t(x)$ is not and exact pivot, its distribution depends on which particular DGP $\mu \in M$ generates the data used to compute it. In this case bootstrap inference is no longer exact in general. The bootstrap samples used to estimate the finite-sample distribution of $t(x)$ are generated by *bootstrap DGP* which although it usually belongs to model M. Davidson and Mackinnon (2000) proposed a method of calculating bootstrapped p-values in such situation named *Fast Double Bootstrap*. It is a modification of bootstrapped p-value. The procedure is closely related to double bootstrap developed by Beran (1988). (Davidson and Mackinnon (2000) used the term because, Beran (1988) proposed the genuine double bootstrap which is computationally expensive. They used the technique with more efficiently and renamed. Fast is for "Fast version") Their technique is as follows;

1. Obtain B_1 first level bootstrap from the DGP $\hat{\mu}$ in the usual way and use them to compute bootstrap statistics $t(x^{*b})$ for $b = 1, 2, \ldots, B_1$ and the bootstrap p-value $\hat{p}^*$

2. For each first level bootstrap sample b, compute the second level bootstrap DGP μ_b^* and use it to compute B_2 second level bootstrap samples. These samples are used to compute a test statistic $t(x^{**bl})$ for $l = 1, 2, \ldots, B_2$. These are used in the Fast double bootstrap procedures (FDB)

3. For first level bootstrap sample b, compute the second level bootstrap p-value

$$\hat{p}_b^{**} = \frac{1}{B_2}\sum_{l=1}^{B_2} I(t(x^{**bl}) \leq t(x)) \tag{2.63}$$

I is the indicator variable. if $t(x^{**bl}) \leq t(x)$ holds true $I = 1$ else $I = 0$

4. Finally compute the double-bootstrap p-value

$$\hat{p}^{**} = \frac{1}{B_1}\sum_{l=1}^{B_1} I(\hat{p}_b^{**} \leq \hat{p}^{*}) \tag{2.64}$$

Thus the double bootstrap p-value is equal to the proportion of the second level bootstrap p-values that are more extreme than the first level bootstrap p-value. The procedure does not requires any sort of independence between the bootstrap DGP and the test statistic.

2.14 Softwares Used

SPlus 2000 professional is most frequently used for simulation and bootstrap programming. Sometimes Eviews, EasyReg, SPSS, Statistica, Excel etc. are used for more convenience. For word processing LaTex is used. MS Paint and MS word are also sometimes used for picture processing used in LaTex.

Chapter 3

Exploratory Data Analysis

Abstract

In this chapter data has been explored and analysed using visual inspection. Time series plots shows that CO_2 emission per capita, GDP per capita and Total energy consumption of bangladesh have trended over time. There may exist a long-run equilibrium relationship between CO_2 emission per capita and GDP per capita. First difference of the variables are moving around some central value. Phase diagram illustrates that none of the variables create an attractor. Both the variables are increasing at some constant rate. Decadal box plot study shows that both median of CO_2 emission per capita and variation of emission have increased in consecutive decades. Scatter plot study demonstrates that there may exist any causal relationship from GDP per capita to CO_2 emission per capita.

3.1 Introduction

Exploratory data analysis (EDA) utilizes minimum prior assumptions and thus allow the data to guide the choice of appropriate models (Tukey, 1977, Hoaglin, Mosteller, Tukey, 1985). Exploratory Data Analysis (EDA) reveals the behavior of the data and the structure of the analysis. It is a robust technique. Using the visual inspection

of the data series one can decide about data and structure of analysis. It helps to ensure that a few extraordinary data values do not influence the result unduly.

The chapter includes some exploratory presentation of the study variables. The next section 3.2 data explored with time series plot. Section 3.3 describes the phase diagram, section 3.4 uses box plot to see the decadal variation and finally in section 3.5 scatter plot is used to explore the causal relationship between the two variables.

3.2 Time Series Plot

Collecting data applied researchers and econometrician are interested to see the visual structure of the data. In time series plot, the data are plotted against their occurrence time. Vertical axis is the scale of the variables and horizontal axis represents the observation time. If the data have a strong up or down trend we may suspect the nonstationarity of the data series. In such situation we should use the transformed series taking difference a lag from current time values for analysis. Sometime data series may be nonstationary without showing any upward or downward trend. The graphical Illustration 3.1 represents the untransformed plot of the study variables CO_2 emission per capita in metric tons, GDP per capita in constant 1995 $US, squared GDP per capita and Total energy consumption in unit 1000 tons of oil equivalent. All the variables have strong trend. These trends may be stochastic or deterministic. Figure 3.2 illustrates the simultaneous dynamics of log transformed CO_2 emission per capita and log transformed GDP per capita. The plotting has been done under different scale. The left vertical axis represents the scale log(CO_2 emission per capita) and the right vertical axis is for log(GDP per capita). The figure shows that there may exists a common trend in both of the transformed series. The figure 3.3 is the time series plots of differenced variables in different scale. The figure shows that time trend is removed after differencing. But their means likely nonzero. The figure 3.4 is the plot of all differenced variables in a common scale. Figure shows that all the log transformed differenced series have nonzero similar means.

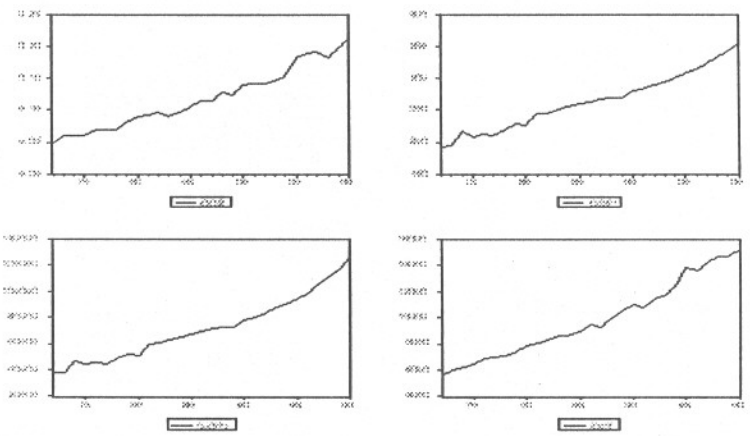


Figure 3.1: Graphical presentation of untransformed CO_2 emission per capita (unit: metric tons of CO_2 per person), GDP per capita (unit: constant at 1995 US$ per person), squared GDP per capita and total energy consumption (unit: thousand tons oil equivalent).

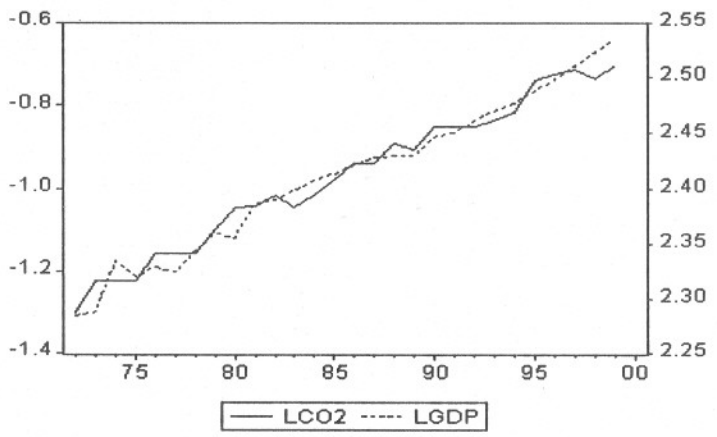


Figure 3.2: Graphical presentation of log transformed CO_2 per capita and GDP per capita.

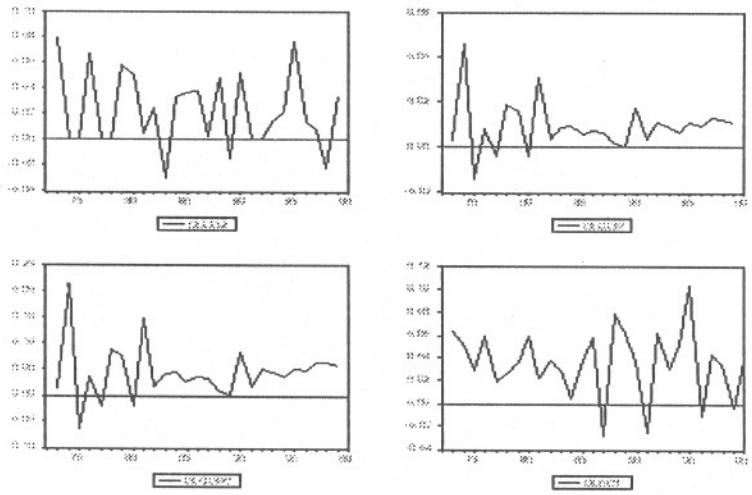


Figure 3.3: Graphical presentation of differenced log transformed variables in different scale.

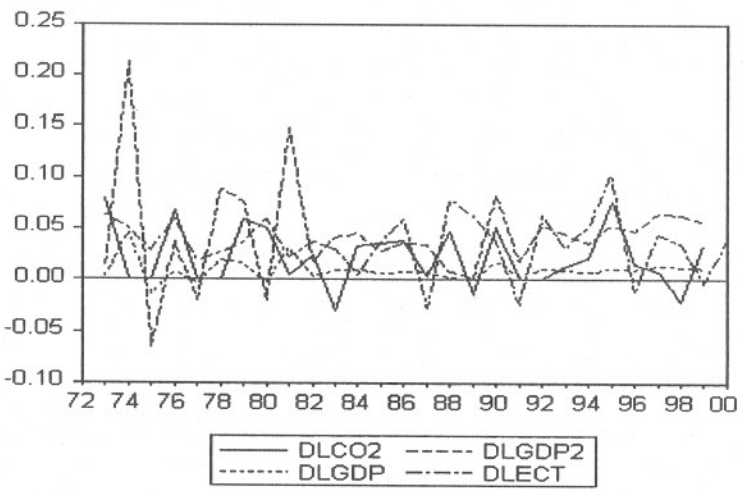


Figure 3.4: Graphical presentation of differenced log transformed variables in the same scale.

3.3 Phase Diagram

Dynamical systems are nonlinear feedback systems that sometime produce complex behavior from relatively simple functions. These systems are generally characterized by no single solution rather by multiple or even an infinite number of solutions indicating that a multitude of states are possible. Researchers have therefore relied on phase space diagrams to identify possible limits to the range of potential solutions (Cambel, 1993). A useful approach for this analysis is a time-based space comparing variables in the previous year (y-axis) with those in the current year (x-axis). The dynamics of a system, in this case emissions of an economy, then traces out a trajectory phase space which can reveal whether the measure is changing in a systematic or irregular fashion. Systems will often be "attracted" to a region of the phase space indicating that the variable is fluctuating around an average value. A description of attractor can provide a classification for a given dynamical system (Peters, 1991). If variables level simply increased at a constant rate, the path would trace a straight line. Phase diagram reveals the behavior of an individual variable through time. For this reason considerable size of time series are most useful. Generally the concept of phase diagram was originated in particle physics. However, the econometric data sets which include time series with measures of variables before and after the transition to stable or declining levels are seldom found. Carbon dioxide emission per capita under an economy may be one of the more complete and is therefore suited to phase diagram.

3.3.1 Phase Diagram of CO_2 Emission Per Capita

Figure 3.5 displays the phase diagram of CO_2 emission per capita in Bangladesh. In the figure, the ordinate represents emission levels in the present year and the abscissa indicates emissions in the previous year. Individual points are labeled with the previous year, and are connected together in time sequence. The figure illustrates the

dynamics of CO_2 emission per capita. The figure shows that no significant attractor has developed within the studied time period. The measure is increasing in a systematic fashion and the path is nearly a straight line.

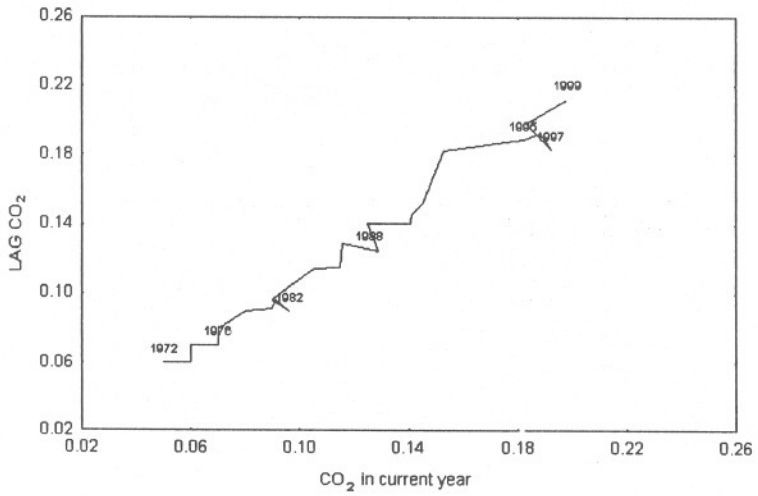


Figure 3.5: Phase diagram of CO_2 emission per capita.

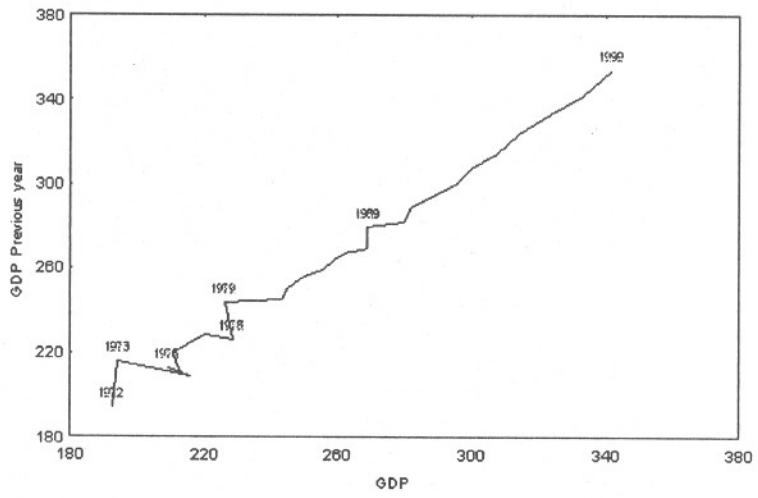


Figure 3.6: Phase diagram of GDP per capita.

3.3.2 Phase Diagram of GDP Per Capita

Figure 3.6 displays the phase diagram of GDP per capita in Bangladesh. The figure has been drawn in the same fashion described in previous subsection. Figure illustrates that except some previous years the line is straight and no attractor has been developed. So we may conclude that GDP per capita in Bangladesh has simply increased at constant rate. This may also happen due to non-stationarity liable to its non-constant mean over time.

3.4 Boxplot Study

The decadal change and variation is studied in the boxplot. For this purpose we have divided the data according to three successive decades. Boxlot of the three decades are shown in figure 3.7 in same scale. The figure illustrates that, both median emission and variation of emission have increased in consecutive decades. This is alarming as emission increases with instable fashion. This plots also showed the decadal trend. The plot reveal that there were no value abruptly fluctuated from the other.

3.5 Scatterplot Study

To find the relationship of rate of change of GDP per capita with changing rate of CO_2 emission per capita we have constructed a scatter plot. Figure 3.8 shows that there may exist a relationship that rate of CO_2 emission per capita is directed by the rate of change of GDP per capita.

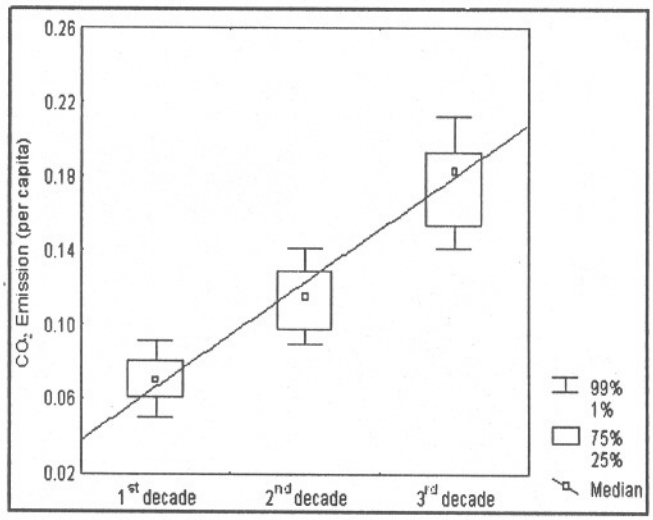


Figure 3.7: Decadal boxplot of CO_2 emission per capita.

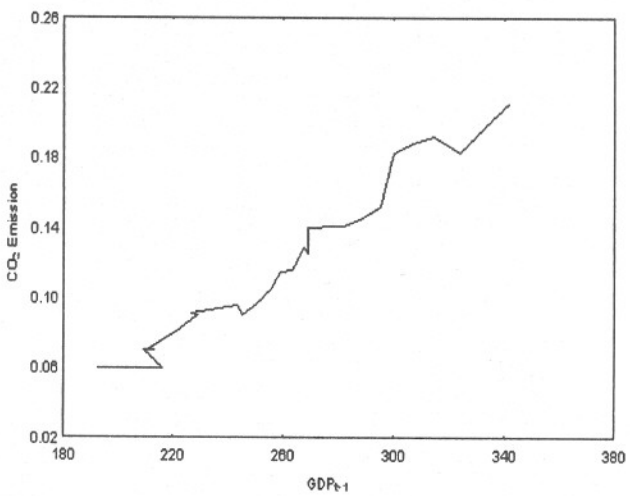


Figure 3.8: Scatter plot of $\Delta \text{GDP}_{(t-1)}$ per capita and ΔCO_2 emission per capita.

Chapter 4

Environmental Kuznets Curve

Abstract

In this chapter, we study the most familiar environment development relationship model Environmental Kuznets Curve (EKC) for CO_2 emission of Bangladesh. Classical method is used to estimate the parameters. To provide the validity of inference on parameter, bootstrap technique is used. Both classical and bootstrap result shows that per capita CO_2 emission of Bangladesh is better explained by time trend rather than GDP per capita and it's square. The CO_2 emission of Bangladesh does not follow EKC in static sense. The data has severe multicollinearity problem.

4.1 Introduction

Economic development affects the environment no doubt, but a fundamental question of economic development is to what degree increases in economic activity affects the natural environment. The common way to evaluate such growth environment relationship is to regress the environmental measurement against the measure of economic development. We have mentioned in chapter 1 that such models of growth and environment have been estimated for a large number of measures of environmental degradation in both panel and cross-country studies for a variety of different coun-

tries. Some of the studies, often with panel data for a group of countries, provide some evidence to support a so-called Environmental Kuznets curve (EKC) where environmental degradation initially increases with the level of per capita income, reaches a turning point, and then declines with further increases in per capita income. In the review of literature we have seen that EKC often does not exist for a number of important measures of environmental degradation, particularly water quality, and for many consumption-based measures of environmental degradation (Rothman 1998). The potential problems with estimating reduced form models of income-environment relationships are detailed by Stern, Common, and Barbier (1996). However, if the estimated EKC results are correct, the fact that many of the reported turning points are at a level greater than the current income of most countries then increasing per capita income implies declining environmental quality for poor and middle-income countries for the foreseeable future (Ekins, 2000). Even if an EKC exists for wealthier countries it may arise from the "export" of pollution-intensive industries and thus may represent the ability of rich countries to separate themselves from their own consumption by engendering environmental degradation in poor countries (Rothman 1998). The possible existence of an EKC for some measures of environmental degradation in a panel of countries begs the question, "what is the nature of the growth-environment relationships for Bangladesh?" I have not found any article on economic growth-environment relationship so far for Bangladesh.

The organization of the chapter is as follows; section 4.2 describes the model estimation and testing using classical approach. Section 4.3 describes some hypothesis testing for EKC in bootstrap approach and finally conclusion is in section 4.4

4.2 EKC for Bangladesh in Classical Approach

We examine the relationship of CO_2 emission and GDP per capita of Bangladesh. To study the relationship the following model is used;

$$(CO_2)_t = \alpha + \delta t + \beta_1 GDP_t + \beta_2 GDP_t^2 + \varepsilon_t \tag{4.1}$$

where CO_2 is the log of CO_2 emission per capita, GDP is the log GDP per capita. The sample size is 28 for both CO_2 emission per capita and GDP per capita. To estimate the parameter we use OLS technique under as usual assumptions. Then we evaluate the model using various classical econometric tests. The OLS results together with diagnostics are presented in table 4.1.

Table 4.1: Results of parameter estimation by OLS for EKC model

	Coefficients	SE	t-value	Pr(> $\|t\|$)
(Intercept)	-9.1393	4.4872	-2.0367	0.0524
Trend	0.0221	0.0041	5.3347	0.0000
β_1	6.6297	3.6850	1.7991	0.0841
β_2	-1.3946	0.7670	-1.8183	0.0810
Test Statistics	F: 604.4	RM: 0.6995	G-Q: 0.4204	J-B: 0.5221
p-value	0.00	0.352	0.433	0.770
RSE: 0.0224	R^2: 0.9864	D-W Stat: 1.805	k : 1996.87	CI: 3987472

4.2.1 Results in Classical Approach

Table 4.1 shows that about 98.64% variability of CO_2 emission is explained by the model. The time trend coefficient is highly significant. The other coefficients are significant at 10% level. The F-statistic implies that the overall regression is highly significant. Thus the model as a whole have an explanatory power. RM test (Imon,

2003) and Jarque-Bera test do not reject the null hypothesis that the residuals are normally distributed. The D-W statistics implies that the null hypothesis of no serial correlation can not be rejected at 5% level of significance. Plot of SACF of the residuals (figure 4.1) and their corresponding L-B statistics suggests that there is no unexplained serial correlation left by the model. Figure 4.2 is a scatter plot of residuals vs. GDP per capita. Figure suggests that the residual has constant variance. The Goldfeld-Quandt (G-Q) test also suggests that the null hypothesis that the error variance are homoscedastic may be accepted. Presence of outlier was checked by standardized residual plot. Figure 4.3 illustrates the standardized residual plot from the fitted EKC. Plot shows that there are no values abrupt fluctuation as no values is more than three standard deviation from zero.

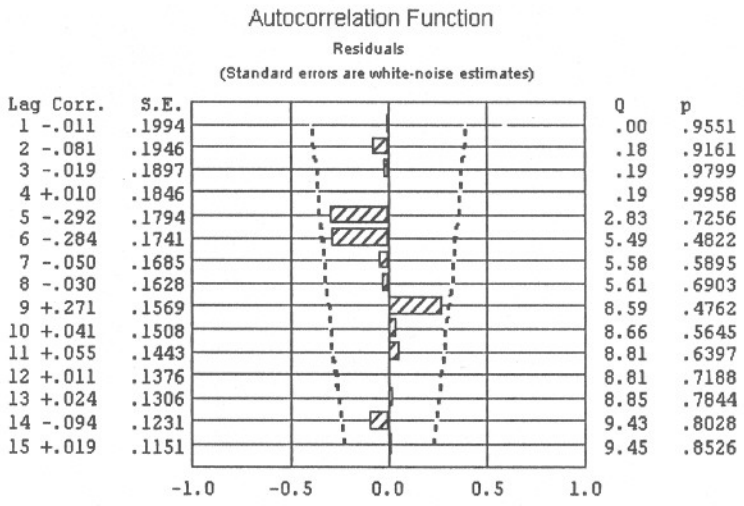


Figure 4.1: SACF plot of residuals.

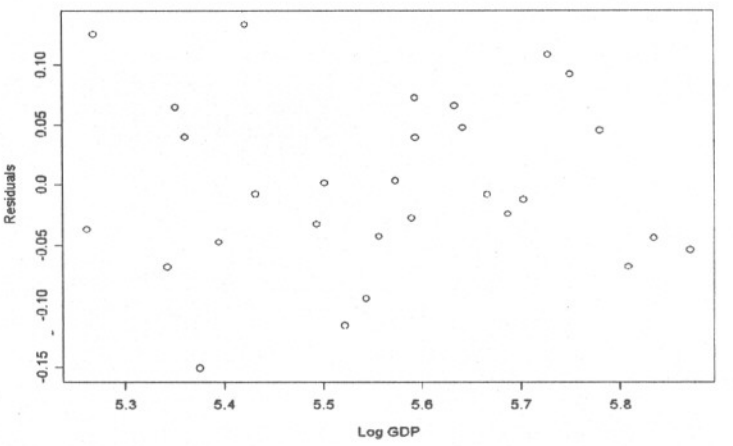


Figure 4.2: Scatterplot of residual vs. GDP per capita.

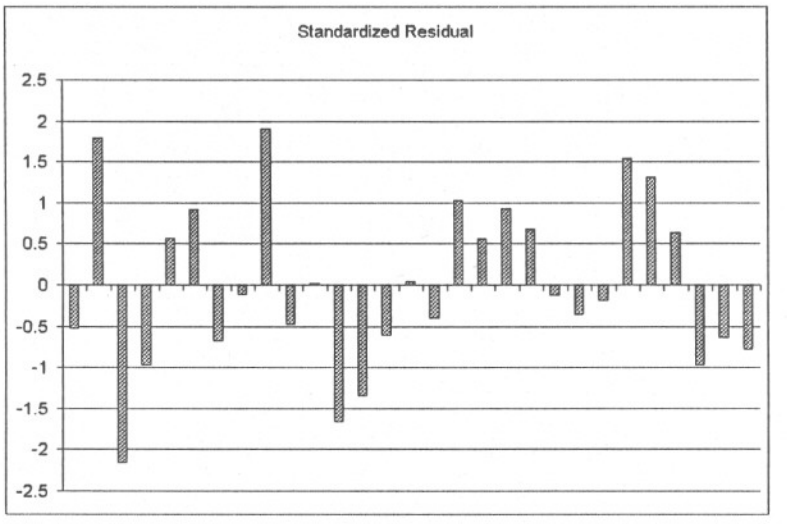


Figure 4.3: Standardized residual plot of the residuals obtained from the EKC model.

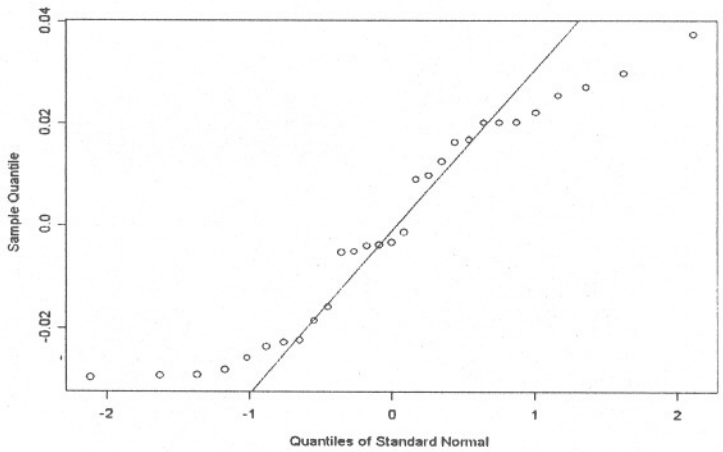


Figure 4.4: QQ plot of residuals.

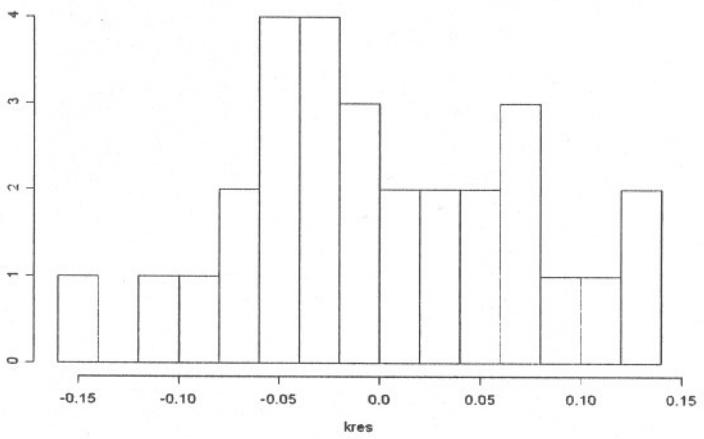


Figure 4.5: Histogram of residuals.

4.2.2 Drawbacks

The histogram (figure 4.5) and QQ plot (figure 4.4) of the residuals incites question on the validity of the inference of the RM test, JB test and DW test. RM test and JB test are asymptotic tests designed for large sample assuming that the residuals are independent. DW test depends on assumption of nonstochastic regressor. Our regressors are stochastic. So to draw a valid inference on the coefficients we use the resampling technique.

It is remarkable that $R^2 = 0.986$ is very high, F statistic has rejected the hypothesis that the partial slope coefficients are simultaneously equal to zero but the individual t statistic has shown that only one slope coefficient is significant. It is a symptom of presence of multicollinearity. So we calculate the eigen values of $X'X$ so as to find the *condition number* k and the *condition index* (CI). The value of $k = 1996.87$ and CI= 3987472 is alarming. Our k value exceeds 1000 and the CI value exceeds 30 therefore suggest the presence of severe multicollinearity.

4.3 EKC in Bootstrap Approach

It is also possible to use the bootstrap technique to construct empirical sampling distribution of the coefficients. Here we use the model based bootstrap with fixed regressor. We want to test the hypothesis H_0; $\beta_1 = 0$. So in our data generating process we set $\hat{\beta}_1 = 0$ so that we get the distribution of β_1 under the null hypothesis. 9999 replicates taken to draw the empirical distribution of the coefficient $\hat{\beta}_1$. The other hypothesis $\alpha = 0$, $\delta = 0$ and $\beta_2 = 0$ are tested using the same manner. The distributions of the coefficients under the null hypothesis are illustrated in figure 4.6 and figure 4.7. The histogram and QQ plot showed that the replicates of the coefficients under the null is distributed normally. The bootstrap test results has been summarized in table 4.2. Result shows that $\hat{\alpha}$, $\hat{\beta}_1$ and $\hat{\beta}_2$ is not a rare value under the null hypothesis hence the null hypotheses $\alpha = 0$, $\beta_1 = 0$ and $\beta_2 = 0$ are accepted at

5% level of significance. The trend coefficient is highly significant and the histogram of the trend coefficient under H_0 does not contain the value of $\hat{\delta}$. Thus, bootstrap test procedure supports the classical test result.

4.4 Conclusion

From the analysis it is clear that CO_2 per capita, GDP per capita and squared GDP per capita are trended over time. So there exists severe multicollinearity in the data. It is also clear that these variables does not follow the global environmental kuznets curve. Nevertheless, analysis suggests that the constant and trend coefficients are highly significant. The final conclusion is that CO_2 emission is better explained by time trend model rather than EKC. That is why, in the next chapter we try to model them separately by Box-Jenkins modeling philosophy.

Table 4.2: Test results of bootstrapped OLS for EKC model

Coeff	Estimates	Hypothesis	Rejection Point			
			1%	5%	95%	99%
Const.	-9.13931	$\alpha = 0$	-9.618145	-6.825136	6.794673	9.594075
Trend	0.02209	$\delta = 0$	-0.008851	-0.006376	0.006495	0.008847
β_1	6.62971	$\beta_1 = 0$	-8.016327	-5.721828	5.734302	7.883458
β_2	-1.39460	$\beta_2 = 0$	-1.606433	-1.168716	1.177770	1.660179

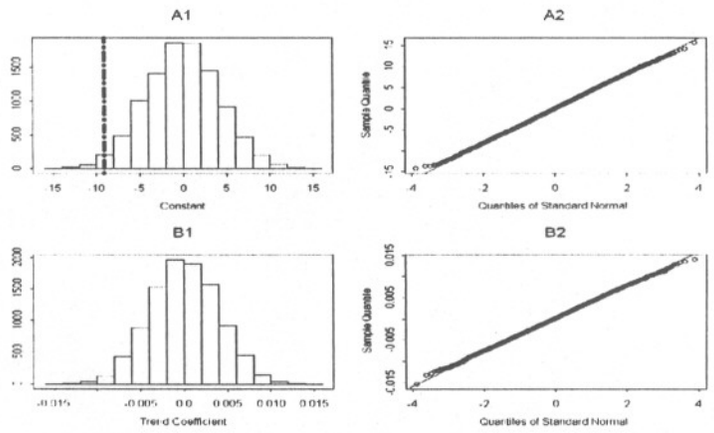


Figure 4.6: Histogram and QQ plot of constant (A1, A2) and trend coefficient (B1, B2) under H_0: $\alpha = 0$ and $\delta = 0$ respectively. The dashed varticle line points the observed value of the coefficients in A1, at B1 the estimated coefficient lies outside the range displayed in the histogram.

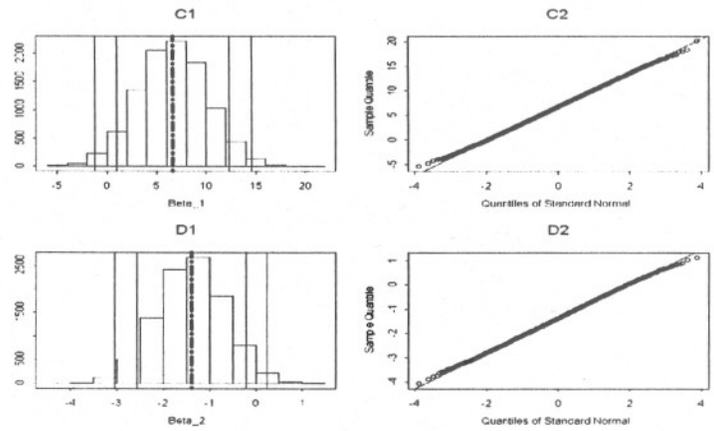


Figure 4.7: Histogram and QQ plot of $\hat{\beta}_1$ (C1, C2) and $\hat{\beta}_2$ (D1, D2) under H_0: $\beta_1 = 0$ and $\beta_2 = 0$ respectively. The dashed varticle line pointed the estimated coefficients and the solid varticle lines are the critical points of two side tests.

Chapter 5

Tests for Time Series Properties

Abstract

It is well known that the power of conventional unit root tests such as ADF, PP, CADF etc. are very low and suffer severe size distortion problem in small sample time series. A simulation based study is done to extract the performance of bootstrap on ADF and CADF test. It is found that for small sample ($n = 30$) BCADF performs relatively better than all other tests. Therefore, variables like GDP per capita and CO_2 emission per capita of Bangladesh where data are available from 1972-2000, i.e., only for 29 years it is wise to use BCADF test. The test result shows that CO_2 emission per capita and GDP per capita in Bangladesh are unit root process.

5.1 Introduction

Time series with unit root have great importance in the analysis of time series. Tests of hypothesis of unit root in a time series process is yet challenging because the question whether a time series has a unit root is "inherently unanswerable on the basis of a finite sample observation." (Hamilton, 1994, page-444) due to almost equal finite sample behavior of the processes with unit root and near unit root. The research on the subject has been extremely active both theoretically and empirically. Dif-

ferent type of unit root tests have been proposed and applied. The methods have some disagreements over how strong and extensive is the empirical evidence for the unit root in the process. The most frequently used method for the test of a unit root in a parametric framework is the Dickey-Fuller test developed by Dickey and Fuller (1979, 1981). This test is based on autoregression of finite order that is assumed to be known. In general, however, it is undesirable to test for the unit root within a specific parametric family, since misspecification could lead us to incorrect inference. Said and Dickey (1984) proposed to include the lagged difference term for correction of serial correlation in the residuals proved that such ADF (Augmented Dickey-Fuller) tests are valid for all finite ARMA processes of unknown order, if we increase the number of included lagged differences appropriately as the sample size gets large. However, several authors have argued that the tests may have considerable size distortions in finite samples, especially when the model has moving average components. See Leybourne and Newbold (1999). Another test for testing unit root is Covariate Augmented Dickey-Fuller (CADF) test developed by Hansen (1995). He showed that inclusion of related stationary covariates in the regression of ADF equations might lead to a more precise estimate of the autoregressive coefficient. He proposed to use CADF unit root test rather than conventional univariate unit root tests for small samples. He also analyzed the asymptotic local power functions for the CADF t-statistic and discovered that enormous power gains could be achieved by the inclusion of appropriate covariates. The major drawback of CADF test is that the limiting distribution of the CADF test statistic is dependent on the nuisance parameter that characterizes the measure of relative contribution of covariates to residuals of the equation without including covariates. Hansen (1995) suggested using critical values based on an estimated nuisance parameter to draw a valid statistical inference. Though relatively better from ADF test, size distortion problem is also present in CADF test for small sample realization. Chang and Park (2001) used sieve bootstrap method on ADF test for testing the unit root for infinite order autoregressive model.

They showed that the procedure improves the size distortion of ADF test for finite sample. In their simulation study they took the initial value $y_0^* = 0$ which violates the stochastic properties of y_0^*. They used sample size 50 as small sample. Chang et al. (2001) used bootstrap method for CADF test that is independent of nuisance parameter. They showed that the bootstrap CADF test is consistent and the critical values based on the empirical distribution of the test statistic obtained by bootstrap is asymptotically valid. They showed their test performance by simulation. Results provided that CADF test offers drastic power gains over the conventional ADF test specially when the covariates are highly correlated with the error. Bootstrap CADF test significantly improves the finite sample size performances of the CADF test. In simulation their sample size was n = 50, 100 and 250. They claimed that for n = 50 and 250 the results are "qualitatively similar."

In this chapter we compare four different types of test. ADF, Bootstrapped ADF (BADF), CADF and bootstrapped CADF (BCADF). The aim of the paper is to verify the result for sample size $n = 30$. In real world situation we have sometime not enough data to generate a longer series of innovation to discard the 'burn-in' period observations. So, Chang et al. (2001) initialization method used for their bootstrapping technique is difficult for small sample non-stationary time series. Again Bootstrapping ADF test as was done by Chang and Park (2001) is also problematic for sample size of less than 50 since a number of lag difference term included in regression to whiten the residual. Sufficient inclusion of differenced lag term reduces the effective size of residuals. The study methodology of the article is similar to Chang et al. (2001) except the initialization method and data generating process (DGP). The intuition is that, a random block from the main series will carry the properties of the series. So, a bootstrap replica of the series using such initial values, the generated series will preserve the properties of the original series. However, here Stine's (1987) approach for initialization is used. Stationary covariates are used for augmentation in the ADF model. It supports us to describe the unexplained serial correlation in

the white noise of ADF model. For Bootstrapped ADF test sieve bootstrap is used. Finally we decide to use the test procedure to test the unit root hypothesis for GDP per capita and CO_2 emission per capita of Bangladesh. The chapter organized as follows;

Section 5.2 introduces the unit root problem in time series and CADF test in presence of unit root along with conventional ADF test. Section 5.3 demonstrates the bootstrapping methods of time series, specially bootstrapping ADF and CADF test while section 5.4 launches the simulation algorithm and results, section 5.5 explicated an application to the data series GDP per capita and CO_2 emission per capita of Bangladesh and finally section 5.6 concluds.

5.2 ADF and CADF Test

The ADF test is proposed by Dickey and Fuller (1981). In ADF test the hypothesis that the data is a unit root process, is tested in the context of three different maintained hypotheses or models concerning the alternative hypothesis that the series is stationary. These differ according to the presence of an deterministic and/or intercept trend.

For the Augmented Dickey-Fuller test the following three models were considered:

$$\text{Model I} \quad \Delta y_t = \alpha + \beta t + \rho y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + u_t \tag{5.1}$$

$$\text{Model II} \quad \Delta y_t = \alpha + \rho y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + u_t \tag{5.2}$$

$$\text{Model III} \quad \Delta y_t = \rho y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + u_t \tag{5.3}$$

Δ is the difference operator. u_t is assumed to be Gaussian white noise. Tests may be carried out on the parameters α and β. The DF test does not include the lag-differenced term. The purpose of including the lagged variables is to provide a correction for serial correlation that may be present in residual. When lagged dependent variables are included in the regression, the test is known as the Augmented

Dickey-Fuller Test (ADF). In Model I the null is that y_t is a random walk process with drift while the alternative hypothesis is that y_t is a trend stationary process. In Model II the null is that y_t is a drift less random walk vs. a stationary process with mean α. Model III is a more powerful test for a unit root when the null is that y_t is a random walk with no drift against the alternative that y_t is a stationary process with mean zero. The test for a unit root is given by H_0 ; $\rho = 0$. If the series is I(0), against the hypothesis H_1; $\rho < 0$. To perform ADF test of a sample realization we first fit any one of the models (I, II or III) using OLS and estimate the test statistic τ. under H_0; $\rho = 0$, the test statistic is $\hat{\tau} = \hat{\rho}/\text{se}(\hat{\rho})$. Although the test statistic is calculated in the usual way, it's limiting distribution is Dickey-Fuller distribution rather than Gaussian. Statistical tables for Dickey-Fullar distribution for various sample size are reported in Hamilton (1994).

Hansen (1995) developed the test named Covariate Augmented Dickey-Fuller test. The CADF test is a unit root test procedure in multivariate context. Here more stationary variables are added to whiten the residuals from unexplained serial correlation in the models of ADF test.

Let us assume that the regression residual u_t in model (I, II, III) in equation (5.1, 5.2, 5.3) is serially correlated and also allowed them to be related to other stationary covariates. Let w_t be the m-dimensional stationary covariates specified as;

$$\delta(L)u_t = \gamma(L)'w_t + \varepsilon_t \tag{5.4}$$

Where L is the lag operator, $\delta(L) = 1 + \delta_1 L + \delta_2 L^2 + \ldots + \delta_p L^p$ and $\gamma(z) = \sum_{k=-r}^{q} \gamma_k z^k$
Using equation 5.4 in equation 5.1, 5.2, 5.3 the ADF models became

$$\text{Model I} \quad \Delta y_t = \alpha + \beta t + \rho y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \sum_{k=-r}^{q} \gamma_k' w_{t-k} + \varepsilon_t \tag{5.5}$$

$$\text{Model II} \quad \Delta y_t = \alpha + \rho y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \sum_{k=-r}^{q} \gamma_k' w_{t-k} + \varepsilon_t \tag{5.6}$$

$$\text{Model III} \quad \Delta y_t = \rho y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \sum_{k=-r}^{q} \gamma_k' w_{t-k} + \varepsilon_t \tag{5.7}$$

We wish to test the hypothesis H_0 ; $\rho = 0$ against the alternative hypothesis H_1; $\rho < 0$. The long run covariance matrix

$$\Omega = \sum_{k=-\infty}^{\infty} E\left(\begin{bmatrix} u_t \\ \varepsilon_t \end{bmatrix} [\, u_{t-k} \quad \varepsilon_{t-k} \,]\right) = \begin{pmatrix} \sigma_u^2 & \sigma_{u\varepsilon} \\ \sigma_{\varepsilon u} & \sigma_\varepsilon^2 \end{pmatrix} \tag{5.8}$$

$\lambda^2 = \sigma_{u\varepsilon}^2/\sigma_u^2\sigma_\varepsilon^2$ is a measures of relative contribution of w_t to u_t at the zero frequency. Under the null hypothesis the test statistic estimated by equation (5.7), $\tau(\hat{\rho}) = \hat{\rho}/\text{se}(\hat{\rho})$ has the distribution.

$$\tau(\hat{\rho}) \Rightarrow \lambda \frac{\int_0^1 W(r)dW(r)}{\left(\int_0^1 W^2(r)\right)^{\frac{1}{2}}} + (1-\lambda^2)^{\frac{1}{2}} N(0,1) \tag{5.9}$$

The asymptotic distribution depends on nuisance parameter λ^2. The critical values for the three models provided in the article of Hansen (1995) by Monte Carlo simulation for different values of λ^2. An enormous power achieved compared to Augmented Dickey-Fuller test is reported.

5.3 Bootstrapping Time Series

In iid. data structure one can create a bootstrap sample by random sampling with replacement. But it fails if data is time dependent due to heteroscedasticity or autocorrelation. In such situation if we use iid. bootstrap, all the dependency information lost (Politis, 2003). To preserve the dependency information, there are two approaches in resampling time series. 1) Parametric bootstrap and 2) Non-parametric bootstrap. Reviews of Barkoitz and Killian (1996) reviewed both parametric and non-parametric approach in bootstrapping time series while Bühlmann (2001) did review non-parametric approach only. In our article we use parametric bootstrap.

5.3.1 Bootstrapping ADF Test

To construct bootstrap ADF we first let, under the null hypothesis of unit root, $\Delta y_t = u_t$. Then we fit the autoregression model

$$\delta(L)u_t = \varepsilon_t \tag{5.10}$$

where ε_t is assumed to be Gaussian white noise. Usually, parameters of the model are estimated by OLS method. Yule-Walker method can also be used since, it always yields an autoregression that is invertible and it is asymptotically equivalent to the OLS method. Lag order is chosen by AIC. Basawa et al. (1991) showed that the samples generated without the unit root restriction do not behave like unit root processes. That makes the subsequent bootstrap techniques inconsistent. So it is important to base the bootstrap sampling on regression (5.10). The equation is as same as equation (5.3) with the restriction $\rho = 0$.

Next we obtain a bootstrapped sample ε_t^* from the centered residual and the boostsrapped realization of u_t^* and y_t^* is obtained by the equation

$$\hat{\delta}(L)u_t^* = \varepsilon_t^* \tag{5.11}$$

For initial vector u_0^* we used Stine's (1997) approach of a random block selection from the series u_t of size p. and

$$y_t^* = y_0^* + \sum_{k=1}^{t} u_k^* \tag{5.12}$$

which also requires the initial observation y_0^*. A random draw from the main sample realization is taken as the initial value of the bootstrapped sample y_t^*. Then ADF test applied to y_t^* and we obtained The bootstrapped τ statistic $\hat{\tau}^* = \hat{\rho}^*/\text{se}(\hat{\rho}^*)$.

This technique repeated a number of times to obtain the empirical distribution of $\hat{\tau}$. At the nominal size 1%, 5% and 10% the critical value of bootstrapped test is the 1%, 5% and 10% empirical percentile of the empirical distribution. If the test statistic $\hat{\tau} \leq c^*(\alpha)$ the bootstrapped critical value of size α we reject the null hypothesis of a unit root.

5.3.2 Bootstrapping CADF Test

We have described the test statistic of CADF test $\tau(\hat{\rho})$ in section 5.2. We want to draw the bootstrapped critical value for $\tau(\hat{\rho})$ bootstrapping the sampling distribution of $\tau(\hat{\rho})$. To construct the bootstrap CADF test, letting $u_t = \Delta y_t$ equation (5.7) can be written in the form:

$$\hat{u}_t = \sum_{k=1}^{p} \hat{\delta}_k u_{t-k} + \sum_{k=-r}^{q} \hat{\gamma}_k' w_{t-k} + \hat{\varepsilon}_t \tag{5.13}$$

Equation (5.13) is estimated under the null hypothesis by OLS method and $\hat{\varepsilon}_t$ obtained and $\hat{\eta}_t$ from the l^{th} order autoregression for m dimensional covariate of the form

$$w_{t+r+1} = \hat{\Phi}_1 w_{t+r} + \ldots + \hat{\Phi}_l w_{t+r-l+1} + \hat{\eta}_t \tag{5.14}$$

The combined innovations obtained from the model (5.13) and model (5.14) are defined as $\hat{\xi}_t = (\hat{\varepsilon}_t, \hat{\eta}_t')'$. (1 + m) dimensional bootstrap samples $\xi_t^* = (\varepsilon_t^*, \eta_t^{*\prime})'$ generated by resampling from the centered fitted residual vector $\left(\hat{\xi}_t - 1/n \sum_{t=1}^{n} \hat{\xi}_t\right)_{t=1}^{n}$. The bootstrapped sample of w_t^* obtained recursively from the fitted model

$$w_{t+r+1}^* = \hat{\Phi}_1 w_{t+r}^* + \ldots + \hat{\Phi}_l w_{t+r-l+1}^* + \eta_t^* \tag{5.15}$$

A new innovation is generated by

$$v_t^* = \sum_{k=-r}^{q} \hat{\gamma}_k w_{t-k}^* + \varepsilon_t^* \tag{5.16}$$

Using the innovation v_t^* we generate u_t^* by

$$u_t^* = \hat{\delta}_1 u_{t-1}^* + \hat{\delta}_2 u_{t-2}^* + \ldots + \hat{\delta}_p u_{t-p}^* + v_t^* \tag{5.17}$$

Finally y_t^* is generated by

$$y_t^* = y_0^* + \sum_{k=1}^{t} u_k^* \tag{5.18}$$

At all steps, initial values are taken as described in section 5.3.1. Using y_t^* and w_t^* we applied the classical CADF test and obtained the test statistic $\hat{\tau}^*$, the bootstrap replicate of $\hat{\tau}$ by

$$\tau^*(\hat{\rho}^*) = \frac{\hat{\rho}^*}{se(\hat{\rho}^*)} \tag{5.19}$$

$\hat{\rho}^*$ is the OLS estimate of ρ from the models of CADF test applied to bootstrapped samples. To implement the bootstrap CADF tests, we repeat the bootstrap sampling for the given original sample a large number of times and obtain critical value $b_n^*(\alpha)$ such that α is the prescribed size. The bootstrap critical values is determined in the same manner as described in the previous subsection 5.3.1. The bootstrap CADF test rejects the null hypothesis of a unit root if $\hat{\tau}(\rho) \leq b_n^*(\alpha)$

5.4 Simulation and Results

5.4.1 Simulation Algorithm

We wish to perform a set of simulation to investigate the relative performance of different (ADF, BADF, CADF, BCADF) unit root test. For simulation, we considered (y_t) as a unit root process of the form.

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

The error term $u_t = y_t - y_{t-1}$ for $\rho = 1$ generated by

$$u_t = \delta_1 u_{t-1} + v_t$$

where v_t is related to the covariate w_t by

$$v_t = \gamma w_t + \varepsilon_t$$

In our simulation the covariate w_t is assumed to follow AR(1) process of the form;

$$w_{t+1} = \varphi w_t + \eta_t$$

The innovations $\xi_t = (\varepsilon_t, \eta_t)'$ are iid. N(0,Σ) where

$$\Sigma = \begin{pmatrix} 1 & \sigma_{\varepsilon\eta} \\ \sigma_{\eta\varepsilon} & 1 \end{pmatrix}$$

Under this setup, we have the following covariate augmented Dickey-Fuller regression

$$\Delta y_t = \rho y_{t-1} + \alpha_1 \Delta y_{t-1} + \gamma w_t + \varepsilon_t$$

The correlation between v_t and w_t depends on two parameter values γ and φ. In the data generating process we used $\gamma = 0.8$ and $\varphi = 0.5$. The lagged difference term set to $\delta_1 = 0$ in our simulations. We want to test the hypothesis H_0; The process contains a unit root against the hypothesis H_1; The series is a stationary process. To investigate the power and size of the test we consider $\rho = 0.5$ and $\rho = 0.95$. The sample size is $n = 30$. In our total simulation we take $\sigma_{\eta\varepsilon} = \sigma_{\varepsilon\eta} = 0.5$. After generating the data we perform four test procedures on the artificial data. To perform ADF test we use the most powerful model $y_t = \rho y_{t-1} + u_t$. Reason behind choosing such model is that in our DGP we generate y_t in the same fashion. We want to use ADF procedure taking the exact model. Next we used bootstrapped ADF test on the artificially generated data. Our model for bootstrapped ADF test is as follows;

$$\Delta y_t = \rho y_{t-1} + \sum_{k=1}^{p} \delta_k \Delta y_{t-k} + \varepsilon_t \tag{5.20}$$

The test statistic $\hat{\tau}$ is estimated from the model and it is compared to the tabulated value provided in Hamilton (1994).

For CADF test we use the model as follows;

$$\Delta y_t = \rho y_{t-1} + \delta_1 \Delta y_{t-1} + \gamma w_t + \varepsilon_t \tag{5.21}$$

The test statistic $\hat{\tau}$ and λ^2 are estimated from the model and they are compared to the tabulated value developed by Hansen (1995).

In bootstrapped CADF test we use model (5.21) for $\hat{\varepsilon}$ and the model

$$w_t = \sum_{i=1}^{l} \varphi_i w_{t-i} + \eta_t \tag{5.22}$$

for $\hat{\eta}_t$. Equation (5.22) is sometimes different from our DGP because after data generation we fit AR model again to the artificial w_t using the Yule-Walker method for estimating the parameter. The maximum lag length is to be $10log_{10}(n)$. The

default value of SPlus is 14.77 for sample size 30. We used AIC for lag order selection. 10,000 simulations is done to draw the conclusion. The power and size of the tests ADF, CADF, BADF and BCADF are illustrated in figure-5.1 and figure-5.2.

5.4.2 Results

From Figure-5.1 we see that application of bootstrap technique did not improve the power of the test. In case of CADF test, the power of bootstrapped CADF was lower than classical CADF test at each prescribed level of significance. As level of significance increases, power of bootstrapped test also increases. At 10% level of significance the power of both classical and bootstrapped tests are literally the same. In case of ADF test, we see that, the power of bootstrapped ADF test and classical ADF test are the same and like BCADF test, as the level of significance increases, the power of bootstrapped tests also increases. There is another findings that if the residuals of ADF model is correlated with any other covariate or, if the model failed to explain the serial correlation completely, the power of ADF tests drastically falls. At that time The power of CADF and BCADF is comparatively very high. These results supports the result of Chang et al. (2001) for sample size $n = 50, 100$ and 250 under the same simulation setup.

Figure 5.2 shows that bootstrap techniques improves the size distortion problem of unit root test for finite size of sample. Specially bootstrapped CADF test reduces the size distortion which is amazing. The size distortion of both ADF and bootstrapped ADF tests are the same. Comparing ADF and CADF tests, the result shows that classical ADF and bootstrapped ADF tests size distorte less than both classical CADF and bootstrapped CADF test and bootstrap able to reduce size distortion. But the size distortion problem remains present in all of the tests considered in the article for sample of size $n = 30$. Chang et al. (2001) also found such type over rejection but less than our results. Their over rejection rate was less, may be, due to their large sample size.

5.5 Application to CO_2 Emission Per Capita and GDP Per Capita of Bangladesh

We apply all four tests to CO_2 emission per capita and GDP per capita of Bangladesh. Data series are graphically presented in figure 3.1. The length of both the series is 29. Data are collected from www.unep.org, Geo-Dataset. At first classical tests ADF and CADF tests are applied. Models for these tests consist of constant and time trend. The null hypothesis H_0; The series is a unit root process with drift is tested against the alternative hypothesis H_1; the process is trend stationary. To determine the bootstrapped critical values of ADF and CADF test, Models are estimated under H_0 and these models are used as DGP (Data Generating Process) for bootstrap replicates of $\hat{\tau}$. We use CADF test to test the same hypothesis. ΔGDP and ΔCO_2 used as covariates for testing CO_2 emission per capita and GDP per capita respectively. The model for testing CO_2 emission consists of ΔGDP, one period leads of ΔGDP, one lag of ΔGDP, and two lags of ΔCO_2, Constant and time trend. The model for testing GDP data series consists of ΔCO_2, one lag of ΔCO_2, two lags of ΔGDP, constant and time trend. The model estimates are presented in table (5.1, 5.2, 5.3 and 5.4). The critical values based on DF distribution, Bootstrapped Critical values, Critical values of CADF test based on estimated λ^2 and bootstrapped critical value of CADF test statistics are reported in table-(5.5) in appendix.

Result of Dickey-Fuller test shows that CO_2 emission in Bangladesh is a random walk process. Gaussianity assumption free ADF test based on bootstrapped critical value provides the result of ADF test. The CADF test and bootstrapped CADF test fails to reject the null hypothesis at 10% level of significance. Similarly ADF test and CADF test of GDP per capita does not reject the null hypothesis. The bootstrapped ADF and bootstrapped CADF test also provides the same result.

5.6 Conclusion

In this paper we try to know about the performance of Bootstrapped ADF and Bootstrapped CADF Compared to ADF and CADF developed by Hansen (1995) for sample size 30. 10,000 simulations are done. Analysis shows that bootstrap technique relatively reduces size distortion problem for small sample compared to their classical counterparts. Bootstrapped ADF and bootstrapped CADF test do not achieve radical power gain over their classical ones but at higher level of significance, bootstrapped and classical tests power are almost the same. Tests are applied on CO_2 emission per capita and GDP per capita of Bangladesh. All tests supports that CO_2 emission per capita and GDP per capita are explained better by random walk process.

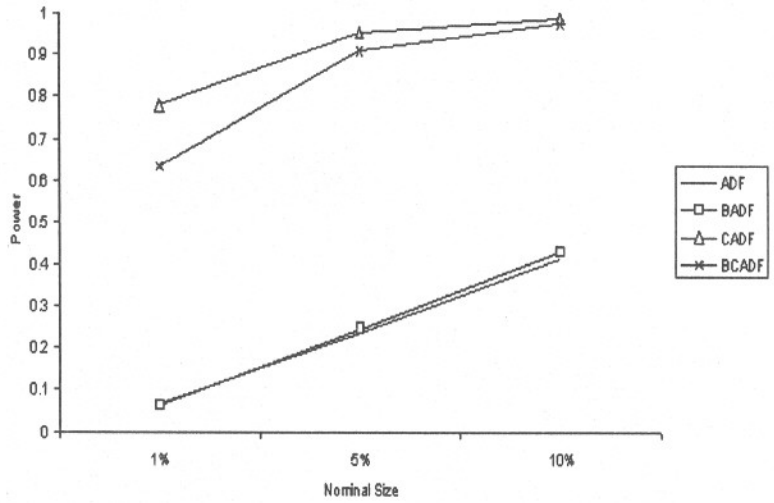


Figure 5.1: Graphical illustration of power comparison of tests.

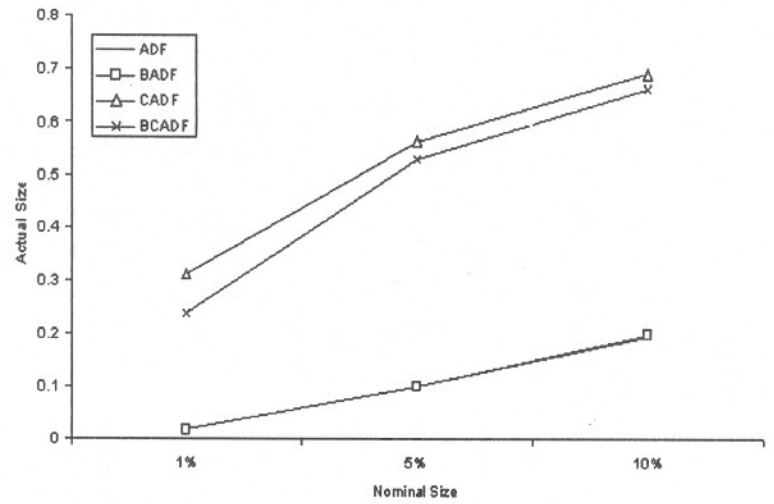


Figure 5.2: Graphical illustration of size distortion of tests corresponding to their nominal size.

Table 5.1: Model estimates for ADF test and DGP for BADF test for CO_2 emission per capita

	Estimates	SE	Test Statistic	Model Under H_0
ρ	-0.444938655	0.203785644	-2.1833660	
C	0.022655397	0.009410052	2.4075739	0.0032173227
Trend	0.002705245	0.001136209	2.3809382	0 .0002626649
δ_1	0.034362297	0.214102542	0.1604946	0 .1727955335

Table 5.2: Model estimates for ADF test and DGP for BADF test for GDP per capita

	Estimates	SE	Test Statistic	Model Under H_0
ρ	0.017340682	0.011102524	1.5618684	
C	-0.004535135	0.004208009	-1.0777390	3.31870845
Trend	0.255757718	0.153279584	1.6685700	0.34174135
δ_1	-0.437542807	0.172207222	-2.5407924	-0.42314426
δ_2	-0.041339108	0.169458140	-0.2439488	-0.03294488

Table 5.3: Model estimates for CADF test and DGP for BCADF test for CO_2 emission per capita

	Estimates	SE	Test Statistic	Model Under H_0
ρ	-0.5012830060	0.2247228011	-2.2306726	
C	0.0241944909	0.0097949570	2.4700967	0.00386278296
Trend	0.0029006816	0.0012098312	2.3975920	0.00025927046
δ_1	0.0761635767	0.2315364612	0.3289485	-0.18422440661
ΔGDP_t	0.0002655712	0.0004120174	0.6445631	-0.00000843938
ΔGDP_{t-1}	0.0001546688	0.0003252782	0.4754968	-0.00010027194

GDP per capita is used as covariate

Table 5.4: Model estimates for CADF test and DGP for BCADF test for GDP per capita

	Estimates	SE	Test Statistic	Model Under H_0
ρ	0.023142420	0.012829138	1.8038952	
C	0.005087908	0.004330911	-1.1747892	4.496682529
Trend	0.279509013	0.158488334	1.7635936	0.395094263
δ_1	0.474159612	0.179984758	-2.6344431	-0.457168450
δ_2	0.050960644	0.178401918	-0.2856508	-0.041346592
$\Delta(CO_2)_t$	0.002646021	0.002830244	-0.9349092	-0.002721990
$\Delta(CO_2)_{t-1}$	0.002055790	0.002829123	-0.7266527	-0.002132991

using total CO_2 emission per capita as covariate

Table 5.5: Table of critical values for both classical and bootstrapped tests

	Co2			GDP		
λ^2	0.7864965			0.6577151		
Tests	1%	5%	10%	1%	5%	10%
ADF	-4.3382	-3.5867	-3.2279	-4.3382	-3.5867	-3.2279
BADF	-4.3319	-3.5934	-3.2078	-4.4785	-3.5102	-3.2235
CADF	-3.8241	-3.2554	-2.9599	-3.7242	-3.1433	-2.8342
BCADF	-4.6336	-3.8911	-3.4598	-0.8831	-0.3904	-0.0986

Chapter 6

ARIMA Modeling

Abstract

In the chapter ARIMA models are fitted for CO_2 emission per capita and GDP per capita. Both classical and bootstrapped approach are used for parameter estimation, hypothesis testing and forecasting. We see that both CO_2 emission per capita and GDP per capita of Bangladesh are better explained by ARIMA(0,1,1) process. The further results can be summarize as follows: 1) The bootstrap method produce the same inference on parameter for both variables as the classical does. 2) The bootstrap method sometimes produce better forecasts and 3) Bootstrap forecast points histogram has a regularity among each other which demands further study on bootstrap forecasts.

6.1 Introduction

It is a good practice to build an ARIMA model for each stochastic series in a dynamic regression model. Such models give us a baseline model for the output series. We can compare the fit and forecast accuracy of the DR models with those of the ARIMA models. In addition, we may learn something useful about the series going through the ARIMA modeling. Finally, ARIMA models are needed to perform diagnostics

checks of the DR model's adequacy. Since our aim is to find a DR model of CO_2 emission per capita with the development related variables of Bangladesh, we fit suitable ARIMA models for those series. To stabilize the variance of the series we use log transformation. As in the previous chapter we have decided that CO_2 emission and GDP is better explained by a random walk process, we took first difference for those variables to stabilize the mean. These transformed series are used to fit ARIMA models. To identify the suitable ARIMA, we use SACF and SPACF. From those graphical presentation different models are chosen and fitted. This is done so because, SACF and SPACF illustrates sometime ambiguous, specially for small sample. Also efficiency of forecast of the models has been checked. Certainly we like to choose the model which have minimum root mean squared forecast error (RMSFE) amongst the set of models for an individual series. We also observe the residuals variance, autocorrelation plot, histogram, normal probability plot of residuals carefully to check a model adequacy. Residuals normality is checked also by RM test (Imon, 2003) and Jarke-Bera test. To check the presence of outliers in the series only standardized residual plot is used.

6.2 ARIMA Modeling (Classical)

6.2.1 ARIMA Model for CO_2 Emission Per Capita

The best suited ARIMA model for CO_2 emission per capita is ARIMA(0,1,1). The estimators are obtained by statistical software package SPlus-2000. Software gives the following estimates by conditional log-likelihood approximation method. The estimated model is as follows;

$$y_t' = 0.050357 + (1 - 0.6962L)\hat{\varepsilon}_t \tag{6.1}$$

6.2.2 ARIMA Model for GDP Per Capita

Within a set of possible model, the best fitted model to explain transformed GDP per capita (z_t') is as follows;

$$z_t' = 0.02004965 + (1 - 0.71347)\hat{\varepsilon}_t \tag{6.2}$$

The standard error t-statistic, p-values and 95% confidence intervals for both of the variables are reported in table 6.1. We used the t-values to test the null hypothesis H_0; $\theta = 0$. The p-value shows that all coefficients are highly significant. For both models $|\hat{\theta}| < 1$ that ensures the invertibility of our models. The residual standard

Table 6.1: Model summary of ARIMA for CO_2 emission per capita and GDP per capita

Variable	Coefficients	$\hat{\theta}$	SE($\hat{\theta}$).	t	p-value	Lower Conf.	Upper Conf.
CO_2	q(1)	0.6962	0.1531	4.54	0.0002	0.396	0.996
GDP	q(1)	0.7135	0.1494	4.78	0.0001	0.421	1.01

Before estimating the parameter, we center the transformed data to zero, so related statistics to the constant terms are not presented in the table though they are present in the model.

error of the shock term series is unknown: sample standard error is used to estimate it. The residual standard error for model 6.1 is $\hat{\sigma}_{y_t'}(\varepsilon) = 0.059$ and for model 6.2 is $\hat{\sigma}_{z_t'}(\varepsilon) = 0.02444292$

6.2.3 Model Checking

Outlier Checking

Outlier is checked by standardized residual plot. The standardized residual $(\hat{\varepsilon}_t/\hat{\sigma}_{\varepsilon_t})$ plot (figure 6.1) do not show any unusual observation that falls three standard deviations from zero for the CO_2 emission per capita model. Plot of standardized residual from the model 6.2 in figure 6.2 has only one value that falls a little more than three standard deviations from zero. Though in a roughly normal distribution, this is not unusual, RM test shows that the residuals distribution is non normal. So it might be worth while to do some research to find out if something unusual happened at time $t = 1973$.

Normality Checking

A standard assumption that the random shocks are normally distributed is checked. Histogram and normal probability plot (discussed in Weisberg, 1980) of the residuals are used to check the normality of the residuals. The graphical presentations are helpful but, they do no provide any formal test. The histogram (figure 6.3) and normal probability plot (figure 6.5) of residual for CO_2 emission model (equation 6.1) shows that the residuals may be normally distributed though they are little bit irrigular. For GDP per capita, the histogram (figure 6.4) of residual shows that the distribution is irrigular and the normal probability plot (figure 6.6) of residual is fragile and not a straight line. So we may conclude that the residual may be far from normal.
We also used formal test procedure to check the normality of the estimated residuals, The recently developed rescaled moment (RM) test (Imon, 2003) is used to check the normality. The value of RM test statistic $S = 8.545964$ with p-value 0.007. That means that the null hypothesis of normality of the residual is rejected at 5% level of significance. So we may conclude that the residuals obtained from model 6.1 may be non-normal. The same technique is also applied to the residuals estimated from the

model 6.2. The RM test statistic is $S = 32.901$ with p-value $= .0000$. Test implies that the probability of accepting the null hypothesis of normality of the residual from model 6.2 is zero.

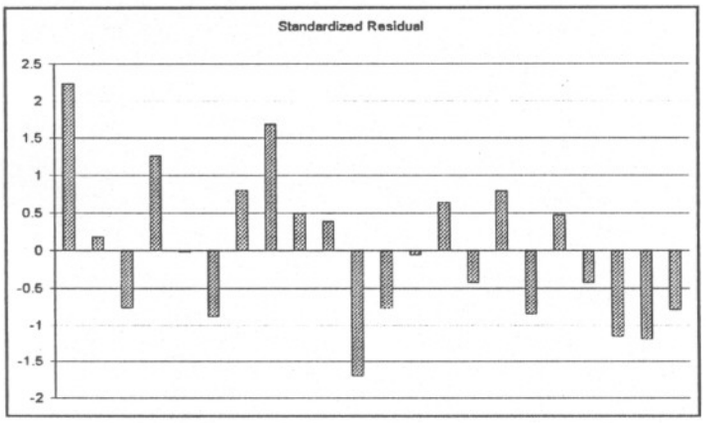


Figure 6.1: Standardized residual plot obtained from ARIMA(0,1,1) for log transformed CO_2 emission per capita.

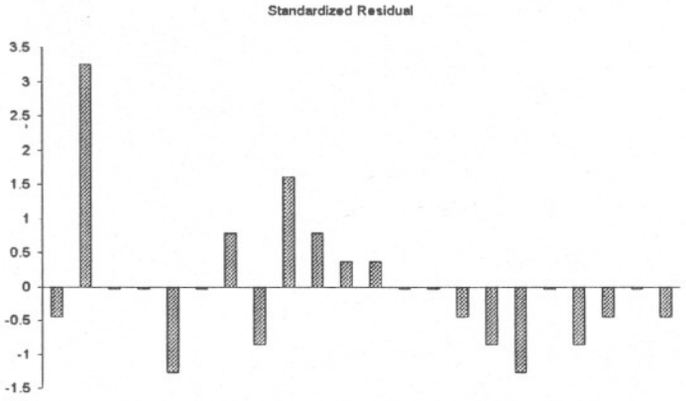


Figure 6.2: Standardized residual plot obtained from ARIMA(0,1,1) model from log transformed GDP per capita.

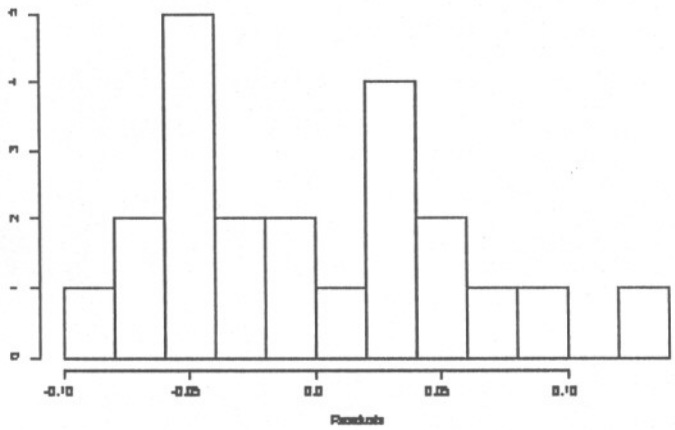


Figure 6.3: Histogram of residuals obtained from ARIMA(0,1,1) for log transformed CO_2 emission per capita.

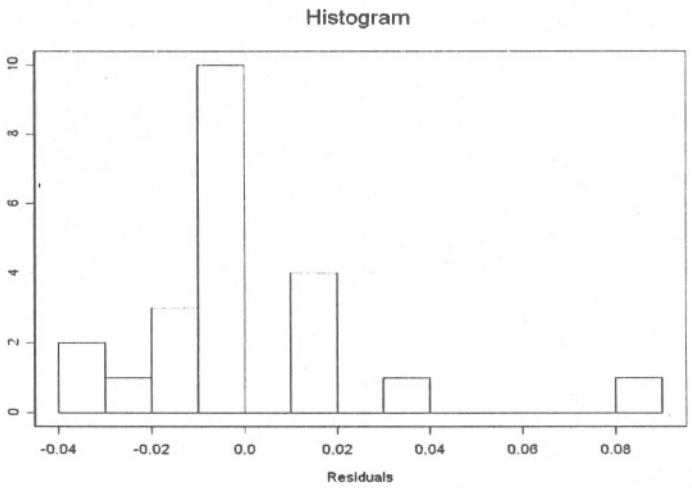


Figure 6.4: Histogram of residuals obtained from ARIMA(0,1,1) for log transformed GDP per capita.

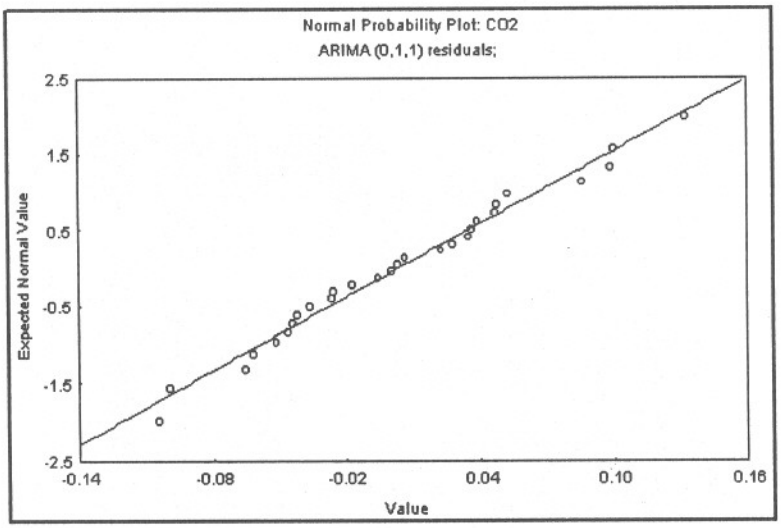


Figure 6.5: Normal probability plot of residuals obtained from ARIMA(0,1,1) for log transformed CO_2 emission per capita.

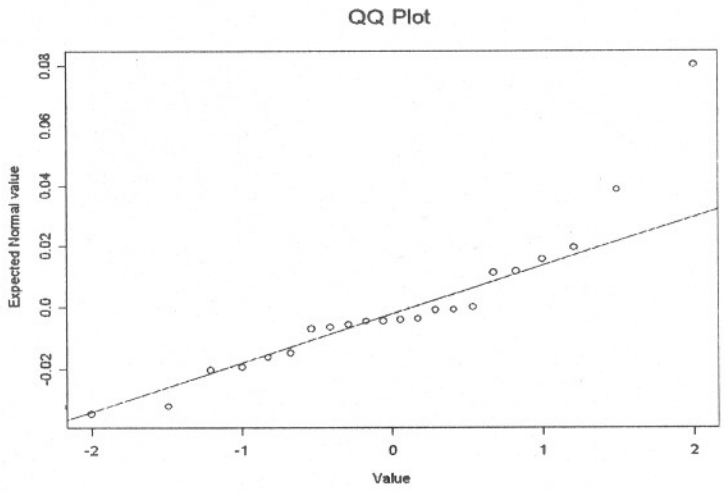


Figure 6.6: Normal probability plot of residuals obtained from ARIMA(0,1,1) for log transformed GDP per capita.

Checking the Presence of Serial Correlation

We construct the SACF of the residuals obtained from model 6.1 and model 6.2. Figure 6.7 is the SACF of the residuals achieved from the model 6.1. Figure illustrates that each residual autocorrelation is very small relative to its standard error. Each spike falls well short of two standard error limit. This suggests that the model adequately represents the autocorrelation pattern of the data. Ljung-Box test (Ljung-Box, 1987) statistic (Q statistic) has also studied to test the joint null hypothesis H_0; All autocorrelation coefficient are zero. The values of test statistic for different lag and their p-values are reported at the right side of the figure 6.7. p-values indicates that all Q statistics are insignificant. Thus the joint test suggest that the model 6.1 has adequately captured the autocorrelation pattern in the data.

In the same way the SACF of residuals from the model 6.2 is also checked and was presented in figure 6.8. The graphical illustration and the Ljung-Box Q statistic suggests that there is no serial correlation unexplained by the model.

6.2.4 Model Interpretation

Both model 6.1 and model 6.2 have an interesting interpretation. The models are exponentially weighted moving average of the available data. If we consider those model as autoregressive form, the models are the infinite order autoregressive models. The current and future output depends on infinite past lag. More weight to the nearest lag and less weight to the farthest lag.

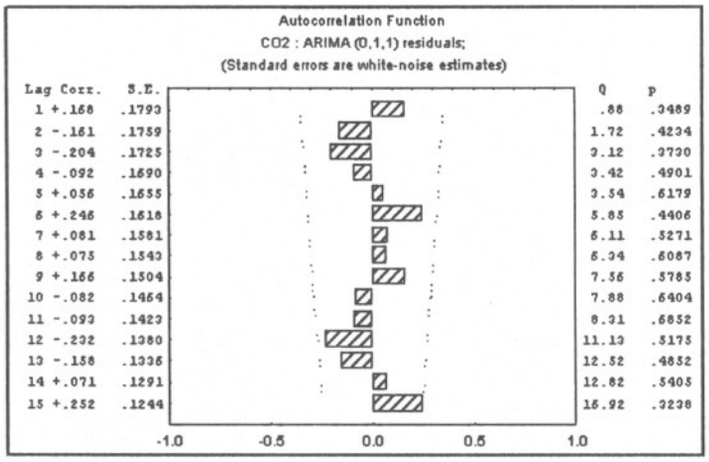


Figure 6.7: ACF of residuals from ARIMA(0,1,1) for log transformed CO_2 emission per capita.

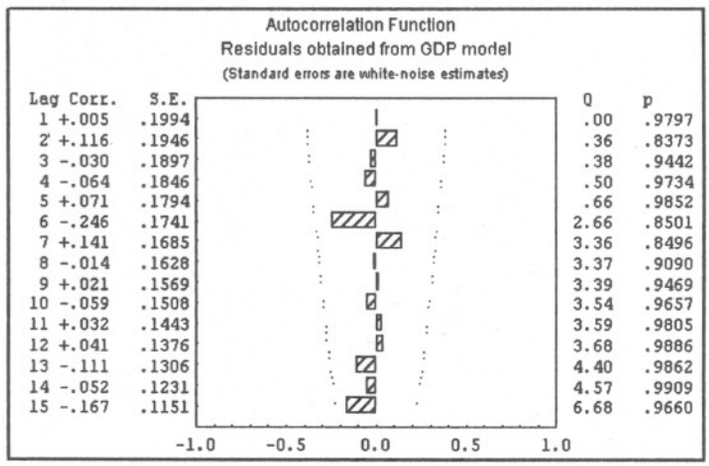


Figure 6.8: ACF of residuals from ARIMA(0,1,1) for log transformed GDP per capita.

At the left of figures, autocorrelation at different lag and their corresponding standard error was displayed. The right side of the figure shows the corresponding Ljung-Box's Q statistic and their respective p-values.

6.3 ARIMA Modeling (Bootstrap Approach)

We are not sure whether the estimated residuals are normally distributed. If the distribution is really non normal, the classical technique fails to infer properly. To check the validity of classical technique it is good practice to use simulation or bootstrap. In quest of better and efficient estimates of parameter and forecast, we use bootstrap.

6.3.1 Bootstrap Estimation

Parametric method is used for bootstrapping ARIMA(0,1,1) process for estimating the bootstrapped estimator of the moving average estimator for CO_2 emission per capita and GDP per capita. First of all, the residuals from the classically fitted models are centered to zero. Then the centered residuals are bootstrapped, The bootstrapped residuals and the classical model is used to get the replica of study variables (the transformed CO_2 emission per capita) $y_t'*$ and (GDP per capita) $z_t'*$. Series were initialized by generating an initial random state vector according to a state space form of the model.

By the process, *one thousand* sample is generated and used to estimate the bootstrap replicates of $\hat{\theta}$. The histogram of bootstrap replicates of MA estimator is shown in figure 6.9 for CO_2 emission per capita and figure 6.10 for GDP per capita. Generally the mean of the bootstrap replication is the bootstrapped estimate. In this thesis median of the bootstrap replicates are also used for bootstrapped estimator from the robust statistical point of view. Using the bootstrapped estimators of θ our model becomes

$$y_t' = 0.050357 + (1 - 0.74496L)\hat{\varepsilon}_t \tag{6.3}$$

$$y_t' = 0.050357 + (1 - 0.7436424L)\hat{\varepsilon}_t \tag{6.4}$$

Models of differenced log GDP per capita

$$z_t' = 0.02004965 + (1 - 0.7319L)\hat{\varepsilon}_t \tag{6.5}$$

$$z_t' = 0.050357 + (1 - 0.7424L)\hat{\varepsilon}_t \tag{6.6}$$

The bootstrapped estimates, bias, and standard error of the estimator are reported in the table 6.2. From the table we see that bias from mean and bias from median of the classical estimator are the same. The SE of the moving average estimator for CO_2 emission per capita is 0.2158 and for GDP per capita is 0.2201.

Table 6.2: Classical and bootstrapped estimates of MA parameter, bias and standard error for CO_2 per capita and GDP per capita

	Coef	Classical	Mean	Bias	Median	Bias from Median	SE
CO_2	MA	0.6962	0.745	-0.04877	0.7436424	-0.04747	0.2158
GDP	MA	0.7135	0.732	-0.0315	0.7424143	-0.0289	0.2201

From the equation 6.3 and equation 6.4 we obtain the residuals. The residual standard error from equation 6.3 and equation 6.4 are $\hat{\sigma}(\varepsilon)_{mean} = 0.05077902$ and $\hat{\sigma}(\varepsilon)_{median} = 0.06043072$ respectively. $\hat{\sigma}(\varepsilon)_{mean} = 0.05077902$ is much less than that of from classically estimated residual $\hat{\sigma}(\varepsilon) = 0.059$. RM statistic is tested for those residuals are $S = 1.638841$ for model 6.3 and $S = 1.636994$ for model 6.4 with p-values 0.2203434 and 0.220547 respectively. Test does not reject the null hypothesis that the residuals obtained from equation 6.3 and equation 6.4 may be normally distributed.

We do the same task for model 6.5 and model 6.6. We observe that the standard errors of residuals from those models are $\hat{\sigma}(\varepsilon)_{mean} = 0.02377872$ and $\hat{\sigma}(\varepsilon)_{median} =$

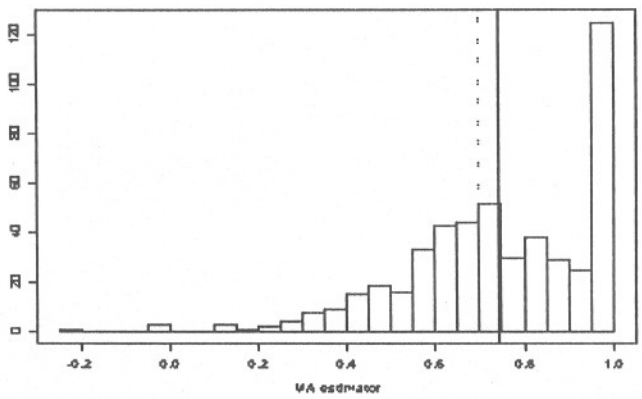


Figure 6.9: Histogram of 1000 bootstrap replication of MA parameter for CO_2 emission per capita. The dashed varticle line indicates the median bootstrapped estimate and the solid varticle line point indicates the mean bootstrapped estimate.

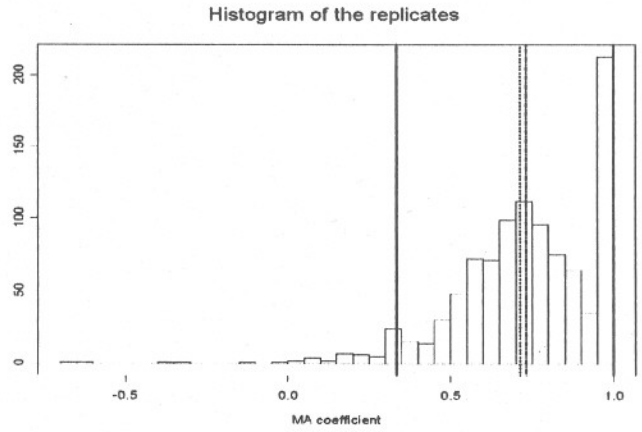


Figure 6.10: Histogram of 1000 bootstrap replication of MA parameter for GPD per capita. The dotted vertical line indicates the classical estimate and the dashed vertical line indicates the bootstrapped estimate.

0.02379658 respectively. Both are less than that of from the classically estimated residual standard error (0.02444292). Again using the median bootstrapped estimate do not improve the residual standard error compared to its mean bootstrapped estimate. So we discard the median bootstrapped estimate from our subsequent analysis. The RM test statistics $S = 24.24795$ with p-value= .000 and $S = 23.1911$ with p-value= 0.000 still rejects the null hypothesis of normality.

6.3.2 Hypothesis Testing with Bootstrap

To test the null hypothesis that the MA parameter $\theta = 0$, we use bootstrap technique in same fashion described above with a little change. Here we generate bootstrapped samples under the null hypothesis. So for data generation we set MA coefficient to zero. Finally we take one thousand bootstrap replicates of MA estimator and hence we get the distribution structure. The 2.5% and 97.5% empirical percentail is considered as the critical points of 5% size test.

For the model of the series y'_t, under the null hypothesis $\hat{\theta} = 0$, The 5% bootstrapped critical values are -0.6153793 and 0.5625293. Since our observed coefficient $\hat{\theta} = 0.6962$ (in classical method) stands outside the interval, we reject the null hypothesis and conclude that the coefficient is significantly different from zero. The test result coincides the classical inference described in table 6.1. The test result is illustrated graphically in figure 6.11.

For the series z'_t, the 5% bootstrapped critical values are -0.6153793 and 0.5625293. Both the classical and bootstrapped estimate lies in the critical region. So we may conclude that the estimated coefficients $\hat{\theta}$ and $\hat{\theta}^*$ are significantly different from zero.

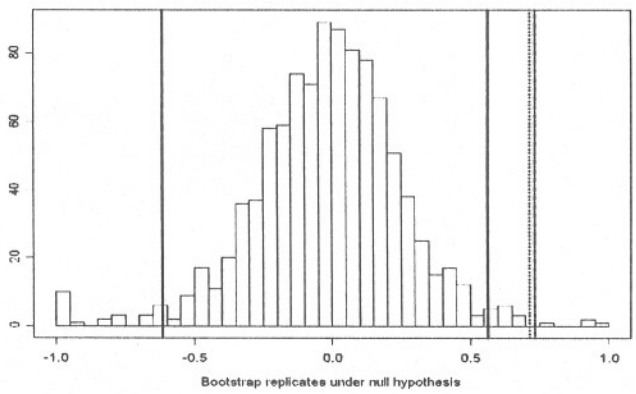


Figure 6.11: 1000 bootstrap replication of MA coefficients under the null hypothesis $\theta = 0$. The solid varticle line indicates the 5% critical points, the dashed varticle line point the estimated coefficient and the dashed varticle line indicates the bootstrapped estimate.

6.4 Forecasting with Classical and Bootstrap Techniques

There are two types of forecasting. one, we can produce a single value forecast for each point of time and two, we can also construct a confidence interval around each point forecast which gives the interval forecast. Interval forecasts are especially useful because they convey the possible degree of error associated with the point forecast. Here we estimated both point and interval forecast for the last six years of the observed sample. To compare the efficiency we compared our forecast to the observed data for those years. Our produced forecasts from the models are in differenced log metric form. The standard error of these forecasts are reported. But when we like to compare the forecasts to the observed data, the forecasts are transformed to the level.

6.4.1 Forecast Using Classical Method

In classical forecast we used the classical models (equation 6.1 for CO_2 emission per capita and equation 6.2 for GDP per capita). Forecast values , 95% upper and lower bound of each point, deviation from observed and residual mean squared forecast error is reported in table 6.3 for CO_2 emission per capita and table 6.4 for GDP per capita. The future values for these periods are assumed to be unknown at the time the forecasts are made. We see from both table that all forecasts lie within the 90% forecast interval. The root mean squared forecast error for CO_2 emission per capita model is 0.0106163 and for GDP per capita is 576.3413.

Table 6.3: Forecast of CO_2 emission using ARIMA model

Time	Lower 95%	Forecast	Upper 95%	Observed	Residual
24	0.148872	0.1662017	0.185548	0.1823	0.0002496
25	0.136969	0.1748694	0.223257	0.1886	0.0001997
26	0.126018	0.1839890	0.268628	0.1922	0.0000642
27	0.115942	0.1935843	0.323220	0.1830	-0.0001120
28	0.106672	0.2036800	0.388907	0.1982	-0.0000323
29	0.098143	0.2143022	0.467943	0.2122	-0.0000185
			RMSFE	0.0106163	

Table 6.4: Forecast of GDP per capita using ARIMA(0,1,1) model

Time	Lower 95%	Forecast	Upper 95%	Observed	Residual
24	293.7552	307.5010	321.8901	307	-0.501045
25	283.3318	313.7286	347.3864	314	0.271445
26	273.2782	320.0822	374.9023	324	3.917817
27	263.5813	326.5645	404.5976	333	6.435514
28	254.2285	333.1781	436.6451	342	8.821932
29	245.2076	339.9256	471.2310	355	15.074412
			RMSFE	7.7691	

6.4.2 Forecast with Bootstrap Method

In bootstrap technique to forecast, we bootstrapped the backward residual. The variance of forward residual and backward residuals for the model of y_t' are $\sigma_{\hat{\varepsilon}} = 0.058$ and $\sigma_{\hat{a}} = 0.057$ respectively which are literally same. Also we check it for z_t'. These are $\sigma_{\hat{\varepsilon}} = 0.00056$ and $\sigma_{\hat{a}} = 0.00059$. So our covariance remains almost same taking the froward residuals. Then we generated the bootstrap sample fixing the last observation for both the series. In our data generating process we used three estimates (classical, bootstrapped mean coefficient and bootstrapped median coefficient) of coefficients. Next we used bootstrap samples for forecasting the six year. One thousand forecast value for each time period was generated and the mean and median of the forecast replicates were used as bootstrap forecast. The 95% confidence intervals are achieved from the 2.5% and 97.5% empirical percentile of the forecast replicates. The graphical presentation of ten bootstrapped samples are shown in figure 6.12 generated from forward residual of y_t' model. At length we compared the forecast efficiency.

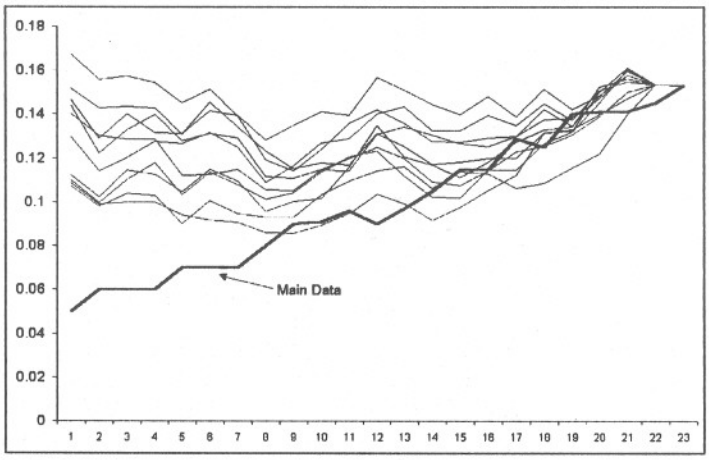


Figure 6.12: 10 bootstrapped sample. The solid thick line is the original series. The lines before the concave point are bootstrapped sample and lines right to the concave point are forecasts from the bootstrapped sample.

6.4.3 Comparative Results between Bootstrap approach and Classical Approach

For y_t' forecasting results and their standard error are presented in table 6.5. Table showed that the standard error of forecast at each point of time is minimum for the cases where mean of θ^* is used for data generating process when bootstrapped. For easy comparison the results of CO_2 forecasts combinedly illustrated in figure 6.13 in the original metric. Figure showed that the bootstrap forecast is almost nearer to the classical forecast. But the confidence interval of classical forecast is much wider than the bootstrap forecast interval. All combination of bootstrapping forecast(i.e., mean forecast replicate, median forecast replicate combined with classical estimator of θ, bootstrapped mean estimator of θ and bootstrapped median estimator of θ) lies approximately over the same line. On the whole, classical forecast is nearer to the original at first three points and far from the last three. But in the case of bootstrap forecast, the last three years forecast is nearer to the original. Among all the bootstrapped forecast, bootstrap with classical estimator of θ and the means of the forecast replicates produces the best bootstrap forecast. Also we mentioned earlier, the standard error is also minimum of the bootstrapped forecasts. The forecasts with different bootstrapped method with their 95% empirical percentile residuals and RMSFE are reported in table 6.7, table 6.8 and table 6.9. From those table we see that the bootstrap forecast where data is generated using classical estimate of θ and the mean points of the replicates are used as forecast produces the minimum $RMSFE = 0.01205$ (shown in table 6.7) among all other bootstrap forecast. If the classical technique is compared with bootstrap techniques, then classical technique produced the minimum $RMSFE = 0.0106163$ (shown in table 6.3)

Table 6.5: Table of bootstrap estimates of forecast mean and forecast median and forecast standard error with classical forecast for transformed series of CO_2 emission per capita

Classical forecast		Classical estimate of θ			Bootstrapped mean replicate of θ			Bootstrapped median replicate of θ		
Mean	SE	Mean	Median	SE	Mean	Median	SE	Mean	Median	SE
0.03193	0.0551	0.0122649	0.007976	0.06388	0.0002174	-0.0008352	0.05841	0.00007347	-0.0010595	0.05866
0.0	0.0671	-0.0021970	-0.0029150	0.05659	0.0021970	-0.0029150	0.05659	-0.00219703	-0.0029150	0.05659
0.0	0.0671	0.0002032	-0.0029150	0.05842	0.0002032	-0.0029150	0.05842	0.00020318	-0.0029150	0.05842
0.0	0.0671	0.0002055	-0.0005782	0.05614	0.0002055	-0.0005782	0.05614	0.00020550	-0.0005782	0.05614
0.0	0.0671	-0.0033792	-0.0029150	0.05745	0.0033792	-0.0029150	0.05745	-0.00337925	-0.0029150	0.05745
0.0	0.0671	-0.0026763	-0.0029150	0.05708	0.0026763	-0.0029150	0.05708	-0.00267628	-0.0029150	0.05708

For z_t', the forecast for transformed metric along with their standard error is reported in table 6.6. From table we see that bootstrap forecasts are very much nearer. The forecasts standard error are of bootstrapped cases are literally same. Comparatively standard error of the bootstrapped forecasts are smaller than that of classical . To compare the bootstrap forecasts to the classicals we drawn figure 6.14 in the original metric. figure shows that all forecast points (both classical and bootstrapped) approximately lie on the same line. In the figure we also illustrated the 95% lower and upper confidence limits. We see that the both the upper and lower point of bootstrapped confidence interval has and increasing tendency compared to the classical confidence interval. Also, the bootstrapped confidence interval is shorter

The forecasts with different bootstrapped method with their 95% empirical percentile residuals and RMSFE are reported in table 6.10, table 6.11 and table 6.12. Tables showed that the forecasts using the median of the forecast replicated did not necessarily improved the forecast error. Using the median of the bootstrapped replicates in data generating process is not so effective. The root mean squared forecast error is minimum in classical technique ($RMSFE = 7.769$). The mean of the forecast replicates also provides a good estimates of forecasts. The RMSFE for bootstrapped forecasts is $RMSFE = 9.078$

Next we construct the histogram of the replicates for each bootstrap techniques to know about the distributional structure of forecasts. Figure 6.15 represents the histograms of the forecast replicates at each point of time forecasted. The classical estimate of θ is used for data generating process. Figure 6.16 viewing the same type of histogram. Here bootstrapped mean estimate of θ is used and the figure 6.17 represents the histograms of the forecast replicates. Bootstrapped median estimate of θ is used for data generation. All histograms have some regularity pattern except the 24th prediction. The 24th prediction of figure 6.15 and figure 6.16 showed that they may be normally distributed.

Table 6.6: Table of bootstrap estimates of forecast mean and forecast median and forecast standard error with classical forecast for transformed series of GDP per capita

Classical forecast		Classical estimate of θ			Bootstrapped mean replicate of θ			Bootstrapped median replicate of θ		
Mean	SE	Mean	Median	SE	Mean	Median	SE	Mean	Median	SE
0.004646	0.022866	-0.000170	-0.004053	0.023759	-0.000072	-0.003932	0.023868	-0.000169	-0.004098	0.023824
0.000000	0.028089	0.001221	-0.003845	0.025394	0.001264	-0.003724	0.025150	0.001155	-0.003797	0.025611
0.000000	0.028089	-0.000818	-0.003993	0.023662	-0.000837	-0.003930	0.024057	-0.001045	-0.003955	0.024097
0.000000	0.028089	0.000306	-0.004099	0.025053	0.000265	-0.003973	0.025051	0.000507	-0.003911	0.024803
0.000000	0.028089	-0.001136	-0.004154	0.022909	-0.001276	-0.004249	0.022736	-0.001372	-0.004379	0.022896
0.000000	0.028089	-0.000628	-0.003887	0.022173	-0.000645	-0.003979	0.022230	-0.000281	-0.003894	0.022205

6.5 Conclusion

In this chapter we fitted ARIMA model for both CO_2 emission per capita and GDP per capita. At first we used classical approach for estimating the model parameters. Following the usual methods we select the models. We see that both CO_2 emission per capita and GDP per capita are better explained by exponentially weighted moving average model (ARIMA(0,1,1)). We also drawn inferences on parameters under the usual assumptions of Gaussianity of error. After diagnostic checking we used the models for forecast the next six year.

Since the sample realization is very small in time series analysis point of view, we used bootstrap technique for further justification. We used bootstrap technique for estimating and testing the models. Using the model estimated in bootstrap approach, we forecasted the last six years of the observed sample which was excluded from the analysis before. To forecast we used bootstrapped forecast technique. We also have tried to find the distributional structure of each forecast point using bootstrap. Finally we compared the classical forecasts and bootstrapped forecasts to the original observations.

Findings of the chapter are as follows;

1. Se(residual) from the model using bootstrap estimate is less than that of using classical estimate, i.e., models estimated by bootstrap approach can explain variables better than models estimated by classical approach.

2. Bootstrap forecast is nearly the same as classical forecast even, sometimes it produce better than classical forecast.

3. In bootstrap approach most of the time the forecast standard error was minimum.

4. Bootstrap forecast interval is shorter than classical forecast.

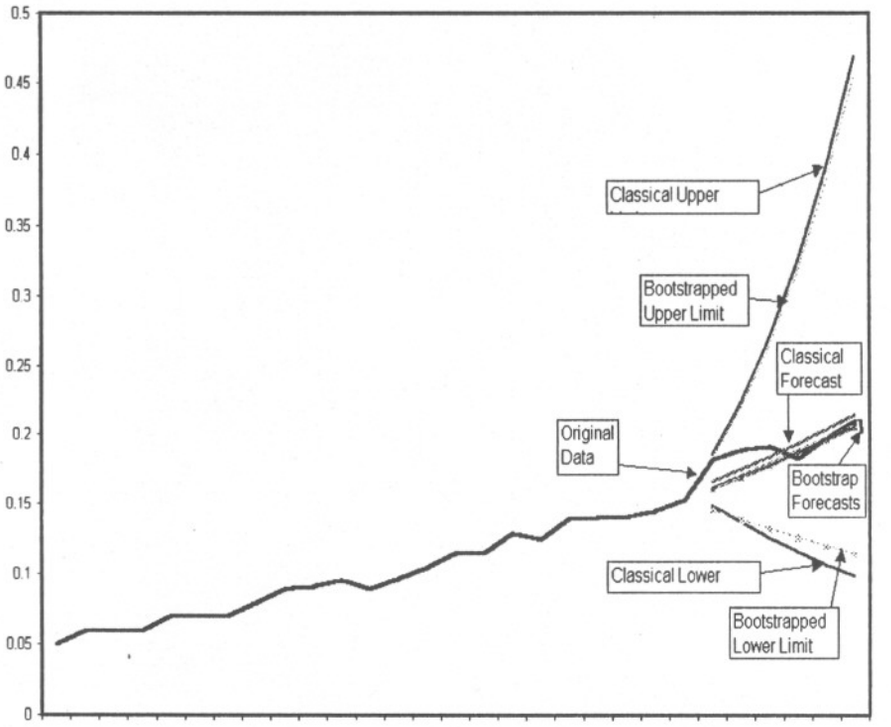


Figure 6.13: Forecasts with bootstrap and classical technique along with their 95% confidence intervals for CO_2 emission per capita.

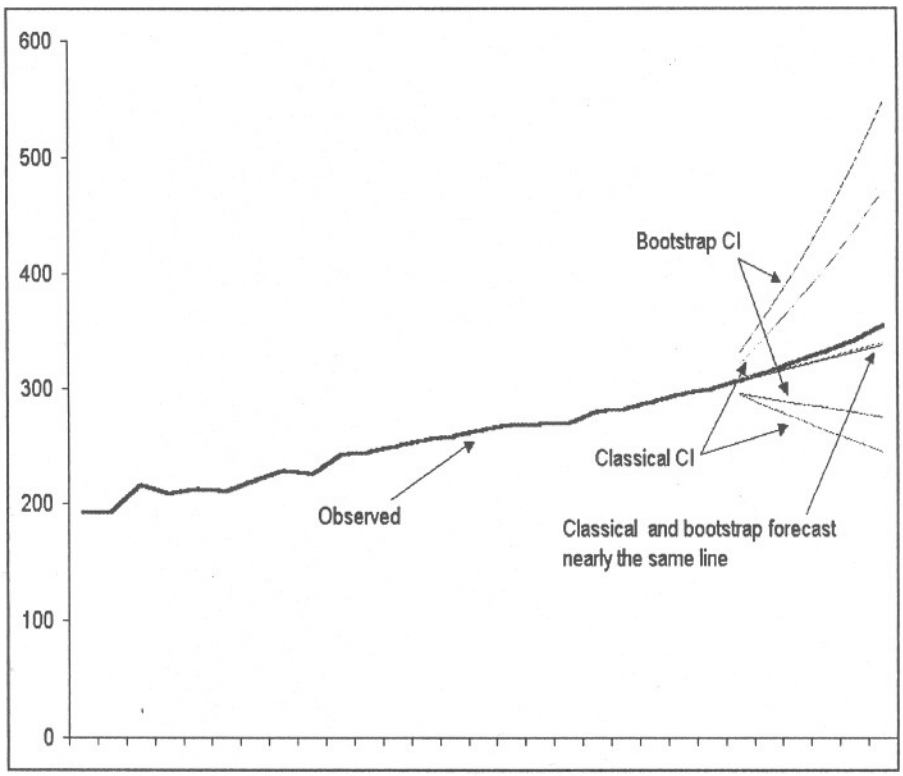


Figure 6.14: Forecasts with bootstrap and classical technique along with their 95% confidence interval for GDP per capita. Most of the forecasts lying on the same line so visually unidentifiable.

Table 6.7: Forecast of CO_2 emission using bootstrap. Classical estimator of θ ($\hat{\theta}^*_{classical}$) used for data generating process (DGP)

		Lower			Upper		
Time	Observed	95%	Mean	Median	95%	Residual 1	Residual 2
24	0.18200	0.14608	0.16297	0.16227	0.18580	0.01903	0.01973
25	0.18900	0.13918	0.17109	0.17023	0.22296	0.01791	0.01877
26	0.19200	0.13260	0.18005	0.17859	0.26755	0.01195	0.01341
27	0.18300	0.12634	0.18948	0.18780	0.32106	-0.00648	-0.00480
28	0.19800	0.12037	0.19869	0.19701	0.38527	-0.00068	0.00099
29	0.21000	0.11469	0.20849	0.20669	0.46232	0.00151	0.00331
					RMSFE	0.01205	0.01262

Mean and median indicates the means and the medians of the forecast replicates at each point of time

Table 6.8: Forecast of CO_2 emission using bootstrap. Bootstrapped mean estimator of θ ($\hat{\theta}^*_{mean}$) used for data generating process (DGP)

		Lower			Upper		
Time	Observed	95%	Mean	Median	95%	Residual 1	Residual 2
24	0.18200	0.14582	0.16101	0.16084	0.18360	0.02099	0.02116
25	0.18900	0.13893	0.16904	0.16874	0.22032	0.01996	0.02026
26	0.19200	0.13236	0.17789	0.17702	0.26439	0.01411	0.01498
27	0.18300	0.12611	0.18721	0.18615	0.31726	-0.00421	-0.00315
28	0.19800	0.12015	0.19631	0.19529	0.38072	0.00169	0.00271
29	0.21000	0.11448	0.20599	0.20487	0.45686	0.00401	0.00513
					RMSFE	0.01338	0.01370

Mean and median indicates the means and the medians of the forecast replicates at each point of time

Table 6.9: Forecast of CO_2 emission using bootstrap. Bootstrapped median estimator of θ ($\hat{\theta}^*_{median}$)used for data generating process (DGP)

		Lower			Upper		
Time	Observed	95%	Mean	Median	95%	Residual 1	Residual 2
24	0.18200	0.14582	0.16101	0.16084	0.18360	0.02099	0.02116
25	0.18900	0.13893	0.16904	0.16874	0.22032	0.01996	0.02026
26	0.19200	0.13236	0.17789	0.17702	0.26439	0.01411	0.01498
27	0.18300	0.12611	0.18721	0.18615	0.31726	-0.00421	-0.00315
28	0.19800	0.12015	0.19631	0.19529	0.38072	0.00169	0.00271
29	0.21000	0.11448	0.20599	0.20487	0.45686	0.00401	0.00513
					RMSFE	0.01338	0.01370

Mean and median indicates the means and the medians of the forecast replicates at each point of time

Table 6.10: Forecasts of GDP per capita using bootstrap. Classical estimator of θ ($\hat{\theta}^*_{classical}$) used for data generating process (DGP)

		Lower			Upper		
Time	Observed	95%	Mean	Median	95%	Residual 1	Residual 2
24	307	295.6351	306.0236	304.8374	331.6874	0.9764	2.16256
25	314	291.3380	312.6025	309.8174	366.6827	1.3975	4.18256
26	324	287.1333	318.6727	314.8322	405.3677	5.3273	9.16779
27	333	282.9630	325.2260	319.8944	448.1728	7.7739	13.1056
28	342	278.7780	331.4358	325.0198	495.4986	10.5642	16.9802
29	355	274.7396	337.9359	330.3155	547.6664	17.0641	24.6845
					RMSFE	9.078428	13.99778

Mean and median indicates the means and the medians of the forecast replicates at each point of time

Table 6.11: Forecast of GDP per capita using bootstrap. Bootstrapped mean estimator of θ ($\hat{\theta}^*_{mean}$) used for data generating process (DGP)

		Lower			Upper		
Time	Observed	95%	Mean	Median	95%	Residual 1	Residual 2
24	307	295.634	306.054	304.875	331.636	0.94637	2.125413
25	314	291.330	312.647	309.893	366.655	1.35319	4.107362
26	324	287.025	318.711	314.929	405.135	5.28835	9.071499
27	333	282.833	325.252	320.032	447.916	7.74755	12.96763
28	342	278.587	331.416	325.129	494.775	10.5837	16.87078
29	355	274.547	337.910	330.397	546.435	17.0899	24.60347
					RMSFE	9.0811277	13.91518

Mean and median indicates the means and the medians of the forecast replicates at each point of time

Table 6.12: Forecasts of GDP per capita using bootstrap. Bootstrapped median estimator of θ ($\hat{\theta}^*_{median}$) used for data generating process (DGP)

		Lower			Upper		
Time	Observed	95%	Mean	Median	95%	Residual 1	Residual 2
24	307	295.6872	306.0238	304.8746	331.6992	0.9762	2.1254
25	314	291.3859	312.5823	309.8926	366.8004	1.4176	4.1074
26	324	287.0416	318.5796	314.9285	405.5324	5.4204	9.0714
27	333	282.8994	325.1962	320.0324	448.4047	7.8038	12.9676
28	342	278.6236	331.3272	325.1292	494.7848	10.6728	16.8708
29	355	274.5699	337.9422	330.3965	546.4894	17.0578	24.6034
					RMSFE	9.112	13.9152

Mean and median indicates the means and the medians of the forecast replicates at each point of time

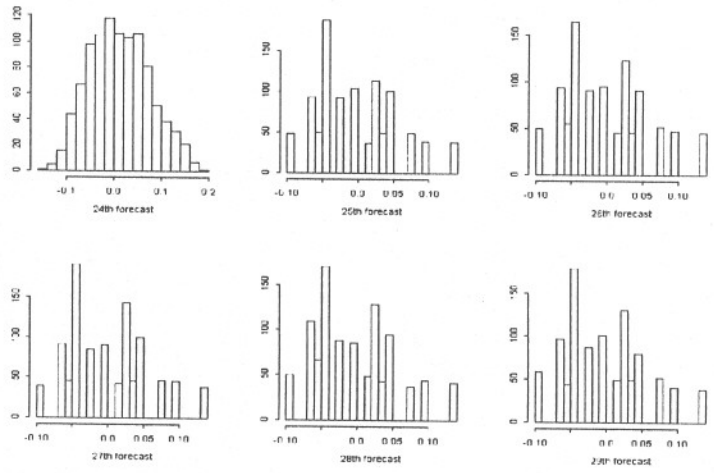


Figure 6.15: Histogram of the bootstrap forecast replicates for CO_2 emission in transformed metric. Classical estimate of θ is used for data generation process (DGP).

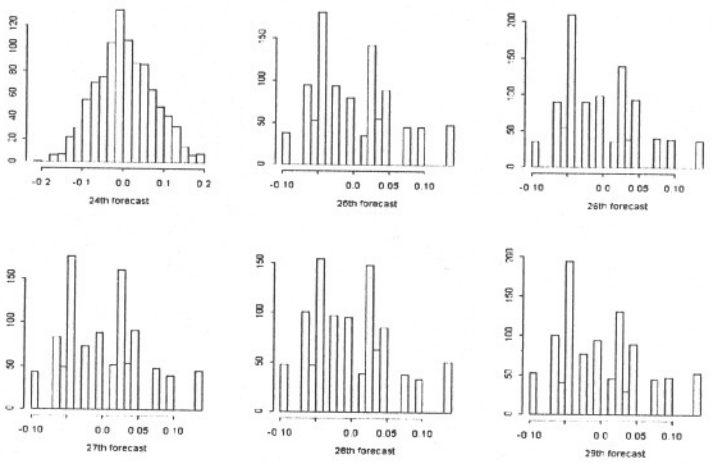


Figure 6.16: Histogram of the bootstrap forecast replicates for CO_2 emission in transformed metric. Mean bootstrapped estimate of θ is used for data generation process (DGP).

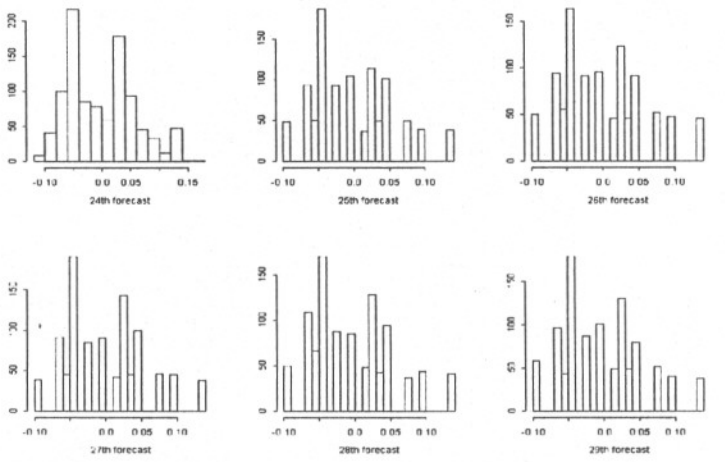


Figure 6.17: Histogram of the bootstrap forecast replicates for CO_2 emission in transformed metric. Median bootstrapped estimate of θ is used for data generation process (DGP).

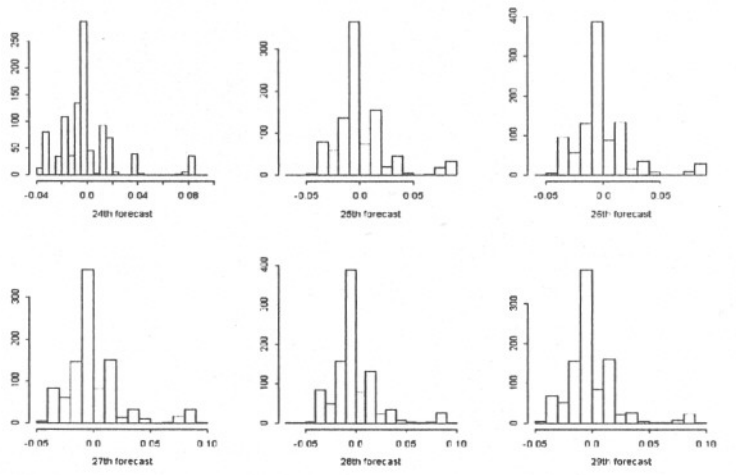


Figure 6.18: Histogram of the bootstrap forecast replicates for GDP per capita in transformed metric. Classical estimate of θ is used for data generation process (DGP).

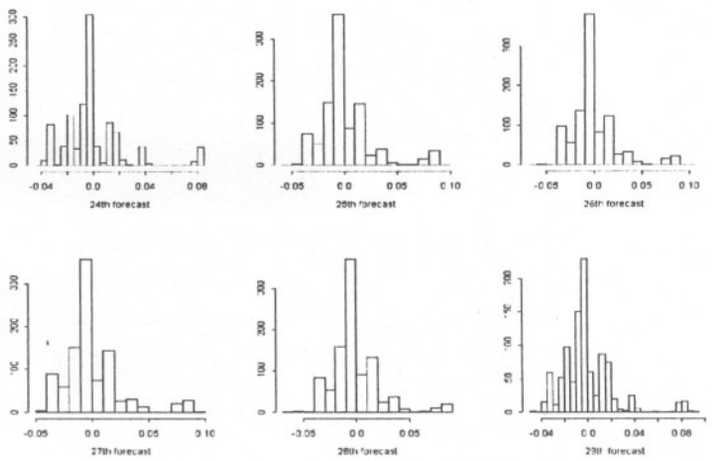


Figure 6.19: Histogram of the bootstrap forecast replicates for GDP per capita in transformed metric. Mean bootstrapped estimate of θ is used for data generation process (DGP).

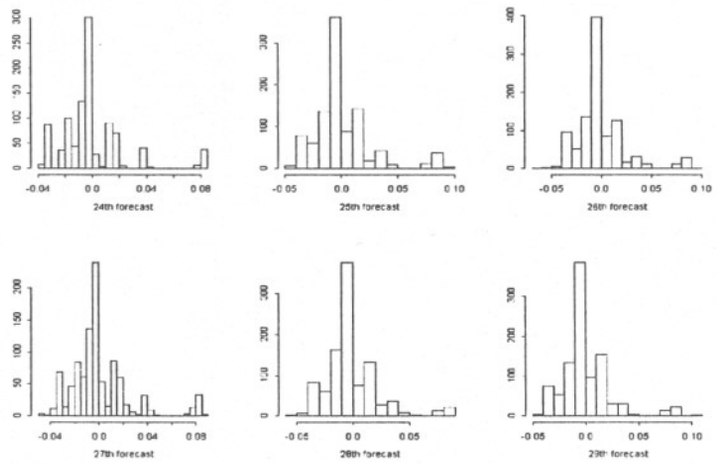


Figure 6.20: Histogram of the bootstrap forecast replicates for GDP per capita in transformed metric. Median bootstrapped estimate of θ is used for data generation process (DGP).

Chapter 7

Cointegration Analysis

Abstract

Since, CO_2 emission per capita and GDP per capita are I(1) process, we search whether there exists some cointegrating relation between CO_2 emission per capita and GDP per capita and squared GDP per capita. Classical method, bootstrap method and fast double bootstrap method is used to find the p-value of the test statistic for testing cointegrating relation. Classical method shows that these variables are trend stationary. The bootstrap methods show that the variables are non-stationary but not cointegrated. The result supports the results obtained from our previous studies.

7.1 Introduction

The bootstrap has become a standard tool for the econometric analysis. Roughly, the purpose of using the bootstrap methodology is to find the distributions of statistics whose asymptotic distributions are unknown or dependent upon nuisance parameters, and to obtain refinements of the asymptotic distributions that are closer to the finite sample distributions of the statistics. If properly implemented to pivotal statistics, the bootstrap simulations indeed provide better approximations to the finite sample distributions of the statistics than their asymptotic (Horowitz, 2002).

Small sample properties of tests on long-run coefficients in cointegrated system are still a matter of concern to applied econometricians. The asymptotic procedures proposed by Johansen (1991) was shown to suffer severe size distortion (see Bewley et al., 1994 and Li and Maddala, 1997). Two complimentary solutions has been proposed; I) applying Bertlett corrections to the test statistics so that the corrected statistic will follow a small sample distribution closer to the asymptotic one and thus bring actual sizes closer to the nominal sizes (Johansen 2000); and (II) trying to estimate the actual small sample distribution by the bootstrap. Again, since the statistical theories for the former are generally nonstandard and depend, often in a very complicated manner, upon various nuisance parameters bootstrap may be a very useful tool. Virtually a few work has been done on bootstrapping cointegrating regression. The bootstrap cointegrating regression was studied only by simulations as in Li and Maddala (1997). The bootstrap method, however, is used quite frequently and extensively by empirical researchers to approximate the distributions of the statistics in more general models with nonstationary time series. Chang, Park and Song (2002) tried to develop the bootstrap theory for cointegrating regressions. They used sieve bootstrap to resample the cointegrating regression. They showed that under the scheme, bootstrap become consistent for both the usual OLS and the efficient OLS by Saikkonen (1991) and Stock and Watson (1993). They also concluded that the bootstrap can thus be employed to correct biases in the estimated parameters, and to compute the critical values of the tests. With the bootstrap bias correction, the OLS estimator becomes asymptotically unbiased. Moreover, the OLS-based tests become asymptotically valid, if the bootstrap critical values are used.

Omtzigt and Fachin (2002) undertook a simulation based study to compare the bootstrap performance over the Bartlett correction in cointegration test for small samples. They used the bootstrap and the fast double bootstrap for testing cointegration parameter in the maximum likelihood framework. They have showed that the fast double bootstrap delivers superior size correction whereas the Bartlett cor-

rection leads to the least loss of power. They also concluded that all three techniques performs much better than the asymptotic tests and difference between them are small.

As we have seen from previous chapter 5 that CO_2 emission per capita and GDP per capita are both integrated of order I(1), So squared GDP per capita is also integrated of order I(1). Now we want to check whether there exists any long run equilibrium relation among these variables. Depending on the last comment of Omtzigt and Fachin (2002), we use bootstrap and fast double bootstrap test to check the cointegration.

The layout of the chapter is as follows; The next section 7.2 describes the maximum likelyhood estimation texhnique for cointegrated VAR model. Section 7.3 describes the method of testing cointegration in cointegrated VAR in both classical and bootstrap approach, section 7.4 illustrates the application of testing cointegration techniques and finally concludes.

7.2 MLE for Cointegrated VAR

Here we describe the Johansen's approach (1988, 1991), full information maximum likelihood estimation of a system characterized by exactly h cointegrating relations under the restriction that there is no deterministic time trend in any of the series. The constant term is restricted to lie in the cointegration space. The estimation techniques are summarized in the following steps.

Step 1: In this step we estimate the $(p-1)$th VAR of Δy_t and collect the OLS regressions in vector form as follows;

$$\Delta y_t = \hat{\Pi}_1 \Delta y_{t-1} + \hat{\Pi}_2 \Delta y_{t-2} + \ldots + \hat{\Pi}_{p-1} \Delta y_{t-p+1} + \hat{u}_t \tag{7.1}$$

$$1 = \hat{\omega}_1^{'} \Delta y_{t-1} + \hat{\omega}_2^{'} \Delta y_{t-2} + \ldots + \hat{\omega}_{p-1}^{'} \Delta y_{t-p+1} + \hat{w}_t \tag{7.2}$$

$$y_t = \hat{\chi}_1 \Delta y_{t-1} + \hat{\chi}_2 \Delta y_{t-2} + \ldots + \hat{\chi}_{p-1} \Delta y_{t-p+1} + \hat{v}_t \tag{7.3}$$

where Π_i and χ_i denotes the $n \times n$ OLS coefficient estimates. $\hat{u}_t$ and $\hat{w}_t$ and $\hat{v}_t$ denotes

the $n \times 1$ vector of OLS residuals.

Step 2: Next we estimate the canonical correlations $\hat{u}_t$ and $\bar{v}_t$ where

$$(\bar{v}_t)_{(n+1)\times 1} = \begin{bmatrix} \hat{w}_t \\ \hat{v}_t \end{bmatrix} \tag{7.4}$$

For this purpose we construct

$$\hat{\Sigma}_{VV} = 1/T \sum_{t=1}^{T} \bar{v}_t \bar{v}_t' \tag{7.5}$$

$$\hat{\Sigma}_{UU} = 1/T \sum_{t=1}^{T} \hat{u}_t \hat{u}_t' \tag{7.6}$$

$$\hat{\Sigma}_{UV} = 1/T \sum_{t=1}^{T} \hat{u}_t \bar{v}_t' \tag{7.7}$$

and find the eigen value of the $(n+1) \times (n+1)$ matrix

$$\hat{\Sigma}_{VV}^{-1} \hat{\Sigma}_{VU}' \hat{\Sigma}_{UU}^{-1} \hat{\Sigma}_{UV} \tag{7.8}$$

ordered as $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_{n+1}$. The maximum value achieved for the log likelihood function subject to the constrain that there are h cointegrating relations and no deterministic time trends is

$$\hat{\mathcal{L}}_h = -(TN/2)log(2\pi) - (TN/2) - (TN/2)log\left|\hat{\Sigma}_{du}\right| - (TN/2)\sum_{i=1}^{h} log(1-\lambda_i) \tag{7.9}$$

Step 3: The previous two steps provides all information needed to perform a likelihood ratio test of the number of the cointegrating relations. However, the likelihood estimates of the parameters also desired. This can be done in the following way.

Let, $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_h$ denotes the eigen vector of 7.8 associated with largest eigen values which forms the basis of the space of cointegrating relations. Johansen suggested normalizing these vectors $\hat{a}_i$ so that $\hat{a}_i' \hat{\Sigma}_{vv} \hat{a}_i = 1$. So Johansen's estimate is $\hat{a}_i = \bar{a}_i \div \sqrt{\bar{a}_i' \hat{\Sigma}_{vv} \bar{a}_i}$. Collecting the first h normalized vector in an $(n \times h)$ matrix $\hat{A}$, the MLE of δ and Π is given by

$$[\ \hat{\delta} \quad \hat{\Pi}\ ] = \hat{\Sigma}_{guv} \hat{A} \hat{A}' \tag{7.10}$$

The MLE of Φ_i is

$$\hat{\Phi}_i = \hat{\Pi}_i - \hat{\alpha}\hat{\omega}' - \hat{\Pi}\hat{\chi}_i \text{ for } i = 1, 2, \ldots, p-1$$

and the residuals

$$\hat{\varepsilon}_t = \hat{u}_t - \hat{\alpha}\hat{w}_t - \hat{\Pi}\hat{v}_t$$

7.3 Hypothesis Testing for Cointegration

Once we will get the unrestricted VAR model that satisfactorily describe the data, we can start the search of reduced-rank restriction, i.e., restrictions on the long-run parameter β, and finally restrictions on the short run adjustment parameter α and Φ. The first crucial steps is to discrinimate empirically between zero and non-zero eigen values when allowing sampling variation and then impose an appropriate cointegration rank restriction r on the Π matrix.

7.3.1 Cointegration Test in Classical Approach

A test for r cointegrating vectors can be based on the maximum likelihood approach proposed by Johansen (1988). The statistical problem is to derive the test procedure to discriminate between the λ_i for $i = 1, 2, \ldots, r$. The rank r is determined by a likelihood-ratio test procedure between the two hypothesis.

H_p: Rank= p, i.e., full-rank, so y_t is staionary

H_r: rank= $r < p$, i.e., r cointegrating relations

The test statistics is

$$S = \hat{\mathcal{L}}_{H_r/H_p} = -Tlog|(1-\lambda_{r+1})\ldots(1-\lambda_p)| = -T\sum_{i=r+1}^{p} log(1-\lambda_i) \qquad (7.11)$$

If $\lambda_{r+1} = \ldots = \lambda_p = 0$, the test statistic should be small enough to zero, which delivers the critical value under the null. The test is based on non-standard asymptotic distributions that have been simulated for the five cases that are already discussed.

There is an additional problem, in that H_r may be correctly accepted when $\lambda_r = 0$ even $\lambda_{r-1} = 0$; Therefore, if H_r is accepted, we conclude that there are at least $p - r$ unit roots, i.e., $p - r$ common trends in the process (but there can be more) corresponding to at most r stationary relations.

However, if the test statistic includes $log(1 - \lambda_r)$, which is not close to zero, so an outcome in excess of the critical value should be obtained, correctly rejecting the false null of fewer that r cointegration relations.

7.3.2 Cointegration Test in Bootstrap Approach

The general idea of the bootstrap test is to assess the value of the test statistic S obtained from the empirical analysis on the basis of the distribution of statistics S^* from suitably constructed pseudo data. $\mathcal{H}_0$ may be imposed when generating pseudo data (Eforn and Tibshirani (1993)) or the chosen DGP taken as the null hypothesis (as recommended by Hall (1992)). In both cases, $\mathcal{H}_0$ is true for the pseudo data. The proportion of S^* more extreme than S in the relevant direction is a natural estimate of the p-value of the test.

Bootstrap Test for Cointegration

With cointegrated VARs we want to test the hypothesis on the long run coefficients $\mathcal{H}_0 : \beta = \beta^0$. To perform bootstrap we followed Omtzigt and Fachin (2002) chosen approach. Estimating the unconstrained VAR, unconstrained parameters and a set of random noises, generating pseudo data using the unconstrained estimates of parameters and the set of random noises and testing $\mathcal{H}_0 : \beta = \beta^0$ on the original data and $\mathcal{H}_0^* : \beta = \hat{\beta}$ (where $\hat{\beta}$ is the unconstrained estimates of β) on the pseudo data.

Let us consider Θ the entire parameter set of the VAR. As we are interested in the test $\mathcal{H}_0 : \beta = \beta^0$, the test is implemented in the following bootstrap procedure.

1. Estimate VAR on data Y; for given cointegrating rank obtain unrestricted estimates $\hat{\theta}$, unrestricted residuals $\hat{\varepsilon}$, restricted estimates $\hat{\theta}_0$, restricted residuals

$\hat{\varepsilon}_0$ and test statistic S for the hypothesis $H_0 : \beta = \beta^0$

2. Construct pseudo data $Y^* = \phi(\hat{\theta}, \varepsilon^*)$, ε^* drawn at random with replacement from $\hat{\varepsilon}$ or NID using parametric bootstrap.

3. Estimate VAR on pseudo data Y^* and obtain $\hat{\theta}^*$, $\hat{\varepsilon}^*$, $\hat{\theta}_0^*$, $\hat{\varepsilon}_0^*$ and the test statistic S^* for the hypothesis $H_0^* : \beta = \hat{\beta}$

   Repeat (2)-(3) a large number of times

4. Compute bootstrap p-value: $p^* = prop(S^* > S)$

Fast Double Bootstrap Test for Cointegration

As discussed in the subsection 2.13.3 in the methodology chapter, we seen that double bootstrap may produce better result compared to simple bootstrap discussed above. The basic intuition is very simple. If the bootstrap estimate $p^* = prop(S^* > S)$ of true p-value of the test is distorted, we may get better estimate by replacing S with some $\tilde{S}$ chosen so to counterbalance the distortion. Now S by definition the p^*-th quantile of the distribution of S^*; hence, an obvious candidate for $\tilde{S}$ is the same quantile of the distribution of a second-level bootstrap distribution. If p^* is distorted downwards, such as a quantile will tend to larger than the true quantile S, and vice versa, thus delivering the desired effect.

The general structure of the fast double bootstrap is as follows; (Graphically presented in figure 7.1

1. Estimate VAR on data Y; for given cointegrating rank obtain unrestricted estimates $\hat{\theta}$, unrestricted residuals $\hat{\varepsilon}$, restricted estimates $\hat{\theta}_0$, restricted residuals $\hat{\varepsilon}_0$ and test statistic S for the hypothesis $H_0 : \beta = \beta^0$

2. Construct pseudo data $Y^* = \phi(\hat{\theta}, \varepsilon^*)$, ε^* drawn at random with replacement from $\hat{\varepsilon}$ or NID.

3. Estimate VAR on pseudo data Y^* and obtain $\hat{\theta}^*$, $\hat{\varepsilon}^*$, $\hat{\theta}_0^*$, $\hat{\varepsilon}_0^*$ and the test statistic S^* for the hypothesis $H_0^* : \beta = \hat{\beta}$

4. Construct second level pseudo data $Y^{**} = \phi(\hat{\theta}^*, \varepsilon^{**})$, ε^{**} drawn at random with replacement from $\hat{\varepsilon}^*$ or NID.

5. Estimate VAR on second-level pseudo data Y^{**} and obtain $\hat{\theta}^{**}$, $\hat{\varepsilon}^{**}$, $\hat{\theta}_0^{**}$, $\hat{\varepsilon}_0^{**}$ and test statistic S^{**} for the hypothesis $H_0^{**} : \beta = \hat{\beta}^*$

   Repeat (2)-(5) a large number of times

6. Compute bootstrap p-value: $p^* = prop(S^* > S)$

7. Compute fast double bootstrap p-value type 1: $p_1^{**} = prop(S^* > Q_{p^*}^{**})$ where $Q_{p^*}^{**}$ is the p^*-th quantile of the S^{**}.

8. Compute the fast double bootstrap p-value type 2; $p_2^{**} = 2p^* - prop(S^{**} > S)$.

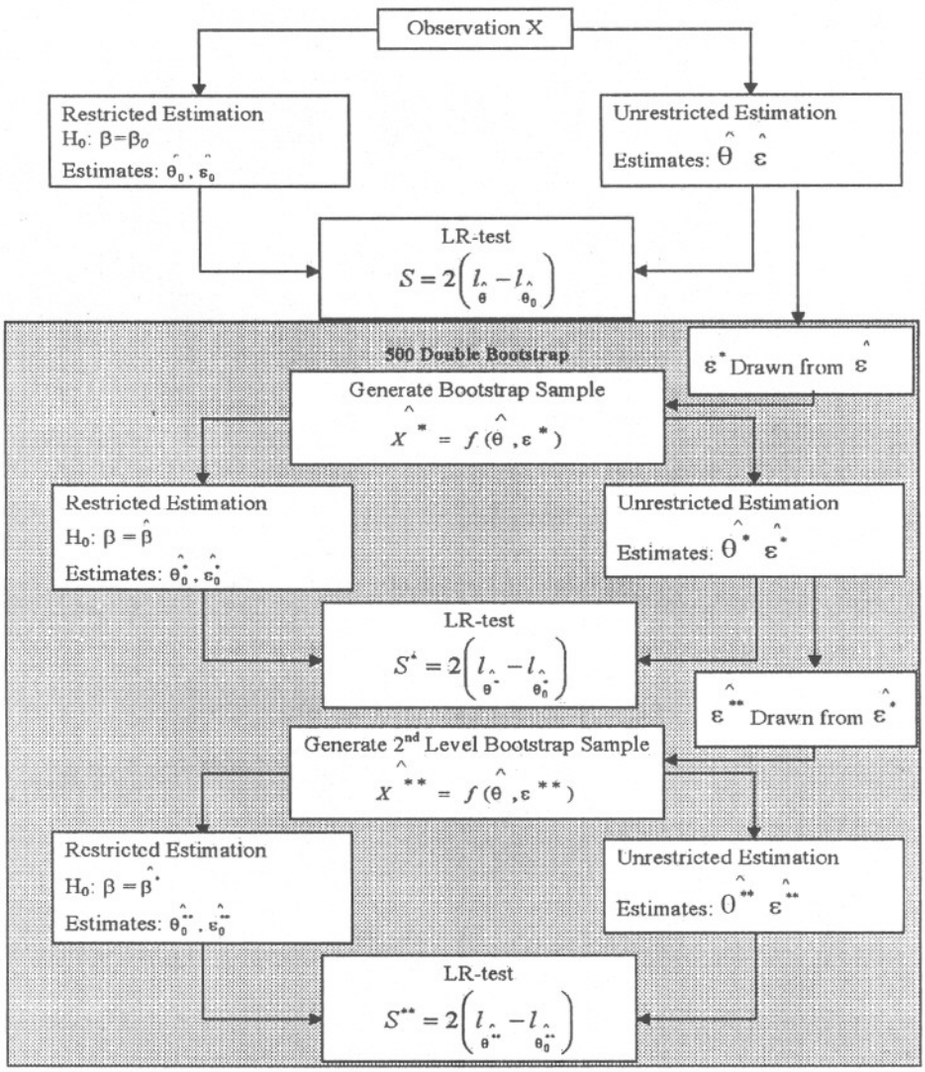
If for instance $p^* > p$, we can expect $pop(S^{**} > S) > p^*$, so that p_2^{**} will be closer to p than p^*. However p_2^{**} may not be greater that $2p^*$ and it may be negative. If the difference between the two p-value is sizeable neither of them should be trusted.

7.4 Analysis and Results

We want to test whether there is a cointegrating relationship between emission per capita and the first and second powers of GDP per capita for Bangladesh. Since the sample size is very small, we have used the methodology described above along with classical test.

7.4.1 Tentatively Estimated VAR

At the first step in the analysis, the unrestricted VAR(1) model with a constant term was estimated by OLS for the variables CO_2 emission per capita, GDP per



$p^* = prop(S^* > S)$, $p_1^{**} = prop(S^* > Q_{p^*}^{**})$ and $p_2^{**} = 2p^* - prop(S^{**} > S)$

Figure 7.1: Fast double bootstrap procedure for tests on the cointegration parameters

capita and Squared GDP per capita. Some descriptive statistics for the logs of the variables in level, differences, and for the residuals are reported in table 7.1. As in the previous chapter 5 we saw that both CO_2 per capita and GDP per capita were apparently non-stationary, but the empirical density was normal for these variables. On the other hand the change of the variables showed non-normality except CO_2 emission per capita. Normality was tested with Jarque-Bera test, as the test statistic distributed as $\chi^2(2)$ under the null. So due to either excess kurtosis of skewness, the test may produce ambiguous inference. Form table 7.1 we also have seen that the first differenced variables and their residuals are not significantly different from zero.

Table 7.1: Descriptive statistics of different time series under study

	ΔCO_2	ΔGDP	ΔGDP^2	LCO_2	LGDP	$LGDP^2$	$\hat{\varepsilon}_{CO_2}$	$\hat{\varepsilon}_{GDP}$	$\hat{\varepsilon}_{GDP^2}$
Mean	0.05	0.02	0.24	-2.18	5.57	31.08	0.00	0.00	0.00
t	4.12	4.53	4.67	-29.07	179.64	89.88	0.00	0.00	0.00
p-value	1.00	1.00	1.00	0.41	0.00	0.00	1.00	1.00	1.00
Median	0.04	0.02	0.23	-2.16	5.58	31.16	-0.01	0.00	0.02
Max.	0.18	0.11	1.13	-1.55	5.87	34.47	0.13	0.04	0.47
Min	-0.07	-0.03	-0.35	-2.81	5.27	27.76	-0.12	-0.04	-0.46
Stdev.	0.07	0.03	0.27	0.40	0.16	1.83	0.06	0.02	0.22
Skew	0.30	1.14	1.02	-0.08	0.01	0.05	-0.02	-0.10	-0.11
Kurt.	2.22	6.06	5.66	1.77	1.98	1.99	2.89	2.64	2.57
J-B stat	1.12	16.97	13.08	1.78	1.20	1.20	0.02	0.20	0.27
p-value	0.57	0.00	0.00	0.41	0.55	0.55	0.99	0.91	0.87
n	27	27	27	28	28	28	27	27	27

The graph of the residuals of the variables are shown in figure 7.2. We did not observe any countable situation which is far enough from $\pm 3\hat{\sigma}$. Figure 7.3 is the

autocorrelogram and histogram of the VAR residuals. There should be no significant autocorrelation left if the truncation after the first lag appropriate. Since the autocorrelation coefficients are very small, this seems to be the case. Furthermore, histograms have shown that the residuals may be normally distributed.

7.4.2 Stability and Unit Root Properties

Up to this point, we have discussed the VAR model as if it were stationary. However, the dynamic stability of the process can be checked by the companion matrix. The matrix was constructed from the estimated VAR. Estimations are presented in table 7.2. Stability depends on the eigen value of the companion matrix. For an n-dimensional VAR with p lags, there are np eigen values. For our trivariate process we have $3 \times 1 = 3$ roots, the moduli of which are

$$1.11,\ 0.906,\ 0.0846$$

We note that the system seems instable as one of the root is outside the unit circle, i.e., explosive root. But it is not much away from 1. Also we have another root nearer to 1. So we want to conclude that the process may be stable. This result also suggests the presence of the stachastic trend. Since there are two roots closer to unity for three variables, the series seems non-stationary and may be cointegrated.

When there are unit roots in the model, it is convenient to reformulate the VAR model in the error correction model. The presentation is discussed in chapter 2. From our study, Table 7.1 showed that $E(\Delta p_{i,t} = 0 \quad \text{for} \quad i = 1, 2, 3)$ was not rejected either of the case. (p_i is used against long variable names). Hence there is no evidence of linear deterministic trend in the variables, at least not over the sample period. So we think that, the cointegrated VAR model should be formulated as restricting the constant term to lie in the cointegration space. The only deterministic component in the model are the intercepts in any cointegrating relations, implying that some equilibrium means are different from zero.

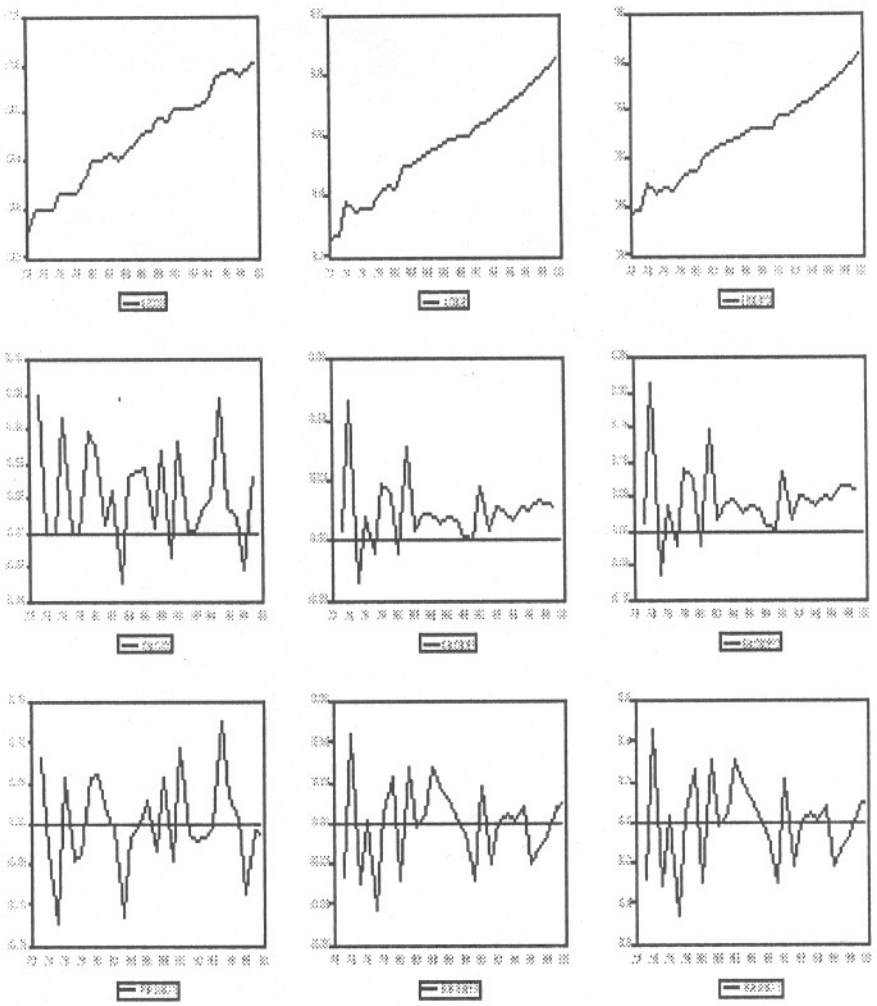


Figure 7.2: Graph of the log transformed variables, their differences and the VAR(1) residuals

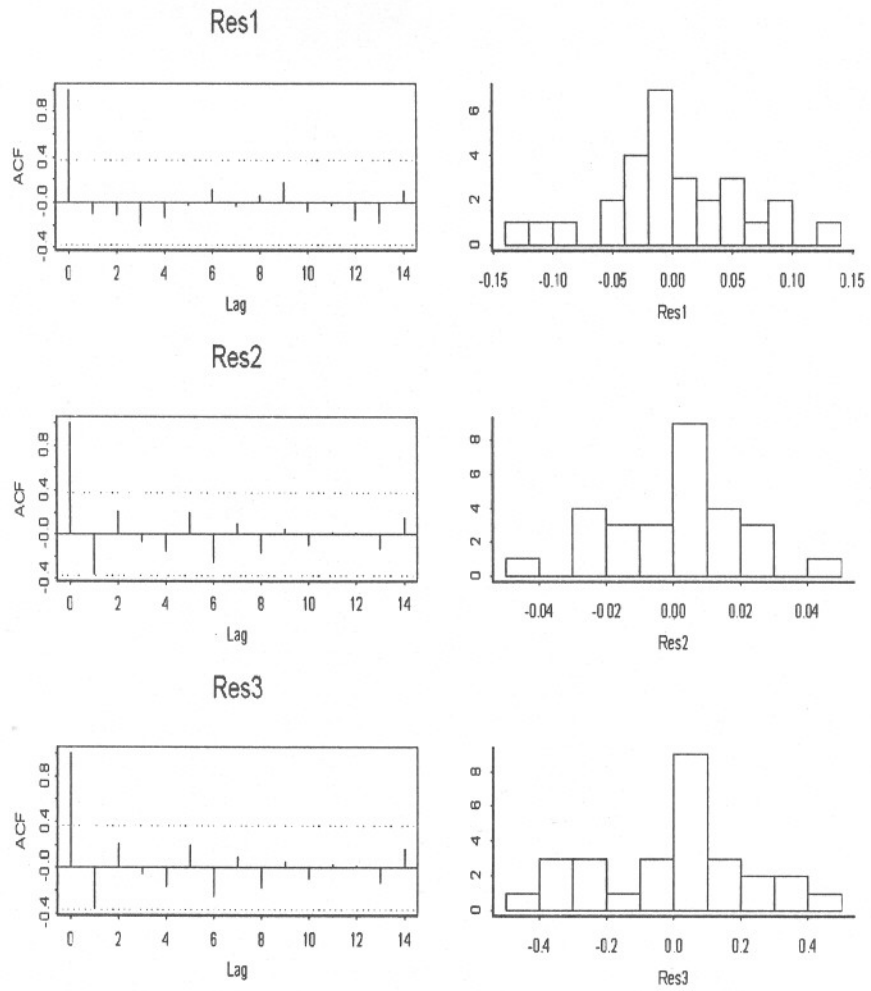


Figure 7.3: ACF plots and histograms of VAR(1) residuals

7.4.3 Test Results Using Classical Approach

To test the presence of cointegrating relations we estimated the system of equations 7.1, 7.2 and 7.3 using OLS for $y_t = (LCO_{2t}, LGDP_t, LGDP_t^2)'$. The regressions were estimated over 1972 to 1998. So the number of observation used for estimation was $T = 27$. The autoregressive order $p = 1$ was chosen using AIC. The sample variance-covariance matrix of the residuals were calculated from 7.5 through 7.7.

$$\hat{\Sigma}_{UU} = \begin{bmatrix} 0.0048322431 & 0.0005301099 & 0.005802559 \\ 0.0005301099 & 0.0007436179 & 0.008056441 \\ 0.0058025595 & 0.0080564411 & 0.087332531 \end{bmatrix} \tag{7.12}$$

$$\hat{\Sigma}_{VV} = \begin{bmatrix} 0.3186653 & -0.7591923 & 1.750866 & 9.622287 \\ -0.7591923 & 1.8486074 & -4.157156 & -22.767689 \\ 1.7508662 & -4.1571560 & 9.625579 & 52.931402 \\ 9.6222866 & -22.7676893 & 52.931402 & 291.250847 \end{bmatrix} \tag{7.13}$$

$$\hat{\Sigma}_{UV} = \begin{bmatrix} 0.02136993 & -0.04949809 & 0.11708052 & 0.6415721 \\ 0.00926036 & -0.02281532 & 0.05092345 & 0.2801825 \\ 0.10118424 & -0.24827699 & 0.55677856 & 3.0654734 \end{bmatrix} \tag{7.14}$$

The eigen value of the matrix in 7.8 are then

$$\hat{\lambda}_1 = 0.704$$
$$\hat{\lambda}_2 = 0.516$$
$$\hat{\lambda}_3 = 0.423$$
$$\hat{\lambda}_4 = 0.000$$

$$T \log(1 - \hat{\lambda}_1) = -32.85$$
$$T \log(1 - \hat{\lambda}_2) = -19.59$$
$$T \log(1 - \hat{\lambda}_3) = -14.83$$
$$T \log(1 - \hat{\lambda}_4) = 0.00$$

We want to test the null hypothesis of $r = 0$ cointegrating relation against the alternative of $r = 3$ cointegrating relation. The test statistic

$$S = -T \sum_{i=r+1}^{n} \log(1 - \hat{\lambda}_i) = 67.28 \tag{7.15}$$

The asymptotic distribution of S is tabulated in Hamilton(1994). The critical value of size 5% is 20.77 and 1% is 25.52. Since the estimated value of S 67.28 > 25.52 we may reject the null hypothesis and decide that all the series may be well described by trend stationary process.

7.4.4 Test Results Using Bootstrap Approach

As we came to know from several literature that bootstrap performs well for small sample, we used bootstrap technique to obtain more reliable p-value. Bootstrap and fast double bootstrap used for testing the cointegration

To perform we first obtained the unconstrained estimate of the model 2.39 The estimates are as follows;

Table 7.2: Model estimate of short-run and long-run effects

	$\hat{\Phi}$			$\hat{\Pi}$			
	$\Delta LCO_{2,t-1}$	$\Delta LGDP_{t-1}$	$\Delta LGDP^2_{t-1}$	α	$LCO_{2,t-1}$	$LGDP_{t-1}$	$LGDP^2_{t-1}$
ΔLCO_2	-0.3723	-15.504	1.423	0.461	0.428	1.319	-0.219
$\Delta LGDP$	0.0768	1.453	-0.183	-0.175	-0.172	-0.500	0.084
$\Delta LGDP^2$	0.8107	18.247	-2.216	-1.908	-1.860	-5.464	0.921

Also we obtained the residuals $\hat{\varepsilon}_t$ from the model. We resample $\hat{\varepsilon}_t$ and using the unconstrained model described in table 7.2 we generated the bootstrapped realization of the series. From those sample we calculated the test statistic S^*. 500 replicates are taken to draw the empirical distribution of S. The histogram of the bootstrap

replicates are shown in figure 7.5. The p-value for testing H_0; There are zero cointegrating relations against the alternative hypothesis that H_A; The series are trend stationary. The p-value obtained from bootstrap test is $p^* = 0.21$. We may accept the null hypothesis and conclude that the series are non-stationary but not cointegrated.

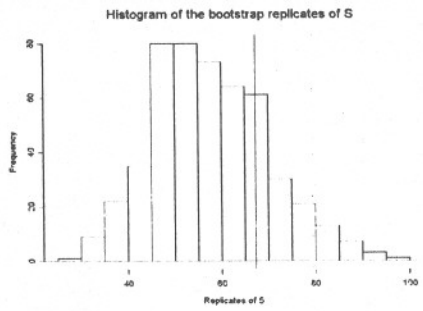


Figure 7.4: Histogram of the bootstrap replicates of S. Solid varticle line pointed the calculated value of S and the dotted varticle line pointed the 5% critical point

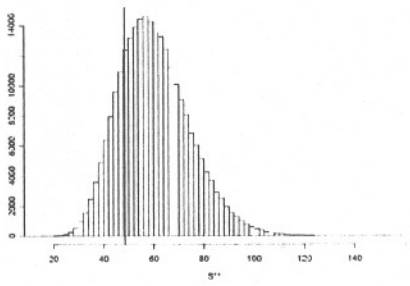


Figure 7.5: Histogram of the bootstrap replicates of S^{**}. Solid varticle line pointed the $p^* = 0.21$th quantile value of S^{**} and the dotted varticle line pointed the calculated value of S

In fast double bootstrap this model and $\hat{\varepsilon}_t$ were used for first step bootstrap. at second step the model is obtained by unconstrained estimate of each bootstrap realization of the first step bootstrap samples, i.e., $\hat{\theta}^*$ and $\hat{\varepsilon}_t^*$ is used for data gen-

eration. 500 replicates were taken to get the fast double bootstrap estimates of S. Same hypothesis is tested using double bootstrap. Five histogram of S^{**} are shown in figure 7.6. The histograms are little bit regular fashion. The double bootstrapped p-values are $p_1^{**} = 0.798$ and $p_2^{**} = 0.13$. Both the p-value suggested to accept the null hypothesis.

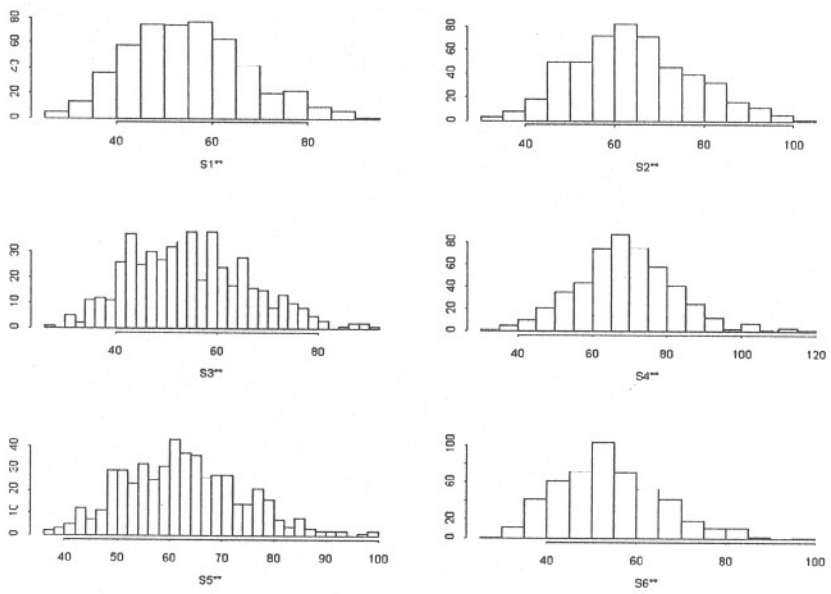


Figure 7.6: Histogram of the bootstrap replicates of S obtained from fast double bootstrap. Each histogram showed the empirical distribution of S^* obtained from first level of bootstrap

7.5 Conclusion

Cointegration analysis of the three series CO_2 emission per capita, GDP per capita and squared GDP per capita have studied. From analysis we see that, Though the variables are nonstationary, they are not cointegrated. There exist no long-run equi-

librium relationship among those variables. The test results also support our the result that we obtained from unit root test and the study of EKC.

This result is reasonable for Bangladesh point of view. The main part of GDP come from agriculture and manpower and foreign trade. Most of our trade is agriculture and manpower related. So it is clear that there does not exeist any short-run or longrun relationship between CO_2 emission per capita and GDP per capita.

Chapter 8

Conclusion

8.1 Summary of the Results

To study the relationship between development and CO_2 emission in Bangladesh, we study two time series variables mainly. GDP per capita used as the development variable and CO_2 emission per capita as the dependent variable. Secondary data is used. Since, the data series is of short length, we use bootstrap techniques.

- **Data Exloration.** At first data has been explored and analysed using visual inspection. Time series plot (in level and in first difference in different metric), scatter plot, phase diagram and box-plot have been used to explore data. Plots have shown that
  - CO_2 emission per capita and GDP per capita is trended over time. Although it is not clear whether the trend is stochastic or deterministic.
  - The rate of change of the variables corresponding to time are not centered to zero.
  - None of the variables create any attractor, i.e., have no stable state over the sample period. Both CO_2 emission per capita and GDP per capita are increasing at some constant rate.

- The decadal median of CO_2 emission per capita and the decadal variation of emission are increasing in consecutive decades.
- There may have any causal relationship from GDP per capita to CO_2 emission per capita
- There may exist a short-run as well as long-run equilibrium relationship. Plotting the transformed variables also support the hypothesis.

- **Environmental Kuznets Curve Fitting.** After visual inspection, the most familiar environment development relationship model EKC (Environmental Kuznets Curve) has been fitted for CO_2 emission of Bangladesh. Classical method has been used to estimate the parameters. To provide the valid inference on parameters, bootstrap technique has been used. Both classical and bootstrap result shows that;
  - Per capita CO_2 emission of Bangladesh is better explained by time trend rather than GDP per capita or, it's square.
  - The CO_2 emission of Bangladesh does not follow EKC in static sense.
  - The data have severe multicollinearity problem. So, it is not possible to get the real variance of the estimators. Consequently, we get the minimum number of significant coefficient though the overall regression is highly significant.
  - Such type of regression may produce spurious relationship as the variables are trending over time.
  - CO_2 emission of Bangladesh does not follow the EKC relation.

- **Test for Stoachastic Trend.** From exploratory data analysis and EKC study we have seen that all variables of the analysis are trended over time, it is necessary to check whether the trend is stochastic or deterministic. Significant stochastic trend is termed as unit root process. It is well known that the power

of conventional unit root tests such as ADF, PP, CADF, etc. are very low, and suffer severe size distortion problem in small sample time series. Bootstrap sometimes are helpful to get better estimate for small sample. So to decide, which test is better, a simulation based study has been done to extract the performance of bootstrap on classical ADF and CADF tests. Simulation shows the following results;

- For small sample like ($n = 30$), Bootstrapped Covariate Augmented Dickey-Fuller test (BCADF) performs relatively better than all other tests. Therefore, variables like GDP per capita and CO_2 emission per capita of Bangladesh where data are available from 1972-2000, i.e., only for 29 years it is wise to use BCADF test.
- The unit root test (both in classical and bootstrap approach) result shows that both CO_2 emission per capita and GDP per capita in Bangladesh are unit root processes.

- **Univeriate Time Series Modeling.** The ultimate goal of time series analysis is to find a suitable model that explains the relationship of development and CO_2 emission. At a conventional succession we fit ARIMA models for CO_2 emission per capita and GDP per capita. Both classical and bootstrapped approach have been used for parameter estimation, hypothesis testing and forecasting. We have found that both CO_2 emission per capita and GDP per capita of Bangladesh are better explained by ARIMA(0,1,1) process. The results can be summarized as follows;
  - The bootstrap method produces the same inference on parameter for both variables as the classical does.
  - The bootstrap method sometimes produces better forecasts and
  - Bootstrap forecast points histogram has a regularity among each other, which demands further study on bootstrap forecasts.

- **Cointegration:** ARIMA models are univariate time series model. So it does not describe any interrelationship among the study variables. Since our aim is to find the titled relationship, we search for some dynamic regression models. Since, CO_2 emission per capita and GDP per capita are I(1) process, we have searched whether there exists some cointegrating relation between CO_2 emission per capita and GDP per capita and squared GDP per capita as evidenced by the EKC of the variables. Classical p-value as well as bootstrapped method and fast double bootstrapped method have been used to find the bias corrected p-value of the test statistic for testing cointegrating relation. The results are as follows;
  - Classical test methods show that variables are trend stationary
  - Bootstrapped methods shows that the variables are non-stationary and not cointegrated.

8.2 Area of Further Research

There are some drawbacks of the thesis which is left for further research. First, in conjunction with GDP the variables, technological change, fuel consumption, deforestation, foreign trade, land use, vehicle use should be considered to study the relationship between development and CO_2 emission in Bangladesh. Another important drawback is that, most of the tests and inferences are asymptotic. We have tried to over come it by simulation and bootstrap. But most of the bootstrap techniques are mathematically not well established. For further study bootstrap techniques for small samples can be pursued.

8.3 Policy Implication

Our study shows that CO_2 is accumulating in the atmosphare and the emission is not statistically related to the GDP of the country. Perhaps, high traffic, large population and high rate of deforestation may be the causes of increasing CO_2 emission. On the other hand, it may happen that the relation has not been established due to labor intensive and agriculture dependent nature of GDP of Bangladesh. So the environmentalists and policy makers should take necessary research activities to find ways for the reduction of CO_2 accumulation in the atmosphare. Steps should also be taken to identify proper causes of increasing CO_2 emission and thereby take proper actions to control CO_2 emission in this regards will help the nation to have suitable, habitable and sustainable environment.

Bibliography

[1] Alberti M. and D. F. Layton (1996) "Indicators of Environmental Performance: Lessons from the Environmental Kuznets Curve," paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[2] Andreoni, J. and A. Levinson (2001) "The Simple Analytics of the Environmental Kuznets Curve," *Journal of Public Economics*, forthcoming.

[3] Andrews, D. W. K. (1991) "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

[4] Antle J. M. and G. Heidebrink (1995) "Environment and Development: Theory and International Evidence," *Economic Development and Cultural Change*, 43, 603-625.

[5] Arrow, K., Bolin, B., Costanza, R., Dasgupta, P., Folke, C., Holling, C. S., Jansson, B. O., Levin, S., Mäler, K.G., Perrings, C., and D. Pimentel (1995) "Economic Growth, Carrying Capacity, and the Environment," *Science*, 268, 520-521.

[6] Bartlett, B. (1994) "The High Cost of Turning Green," *Wall Street Journal*, Sep 14 1994, Sec. A, p 18 col. 3.

[7] Bartlett, M. S. (1946) "On the Theoritical Specification and Sampling Properties of Autocorrelated Time Series," *Journal of Royal Statistical Society*, Series B, 8, 27-41.

[8] Basawa, I. V., A. K. Mallik, W. P. McCormick, J. H. Reeves and R. L. Taylor (1991) "Bootstrapping unstable first-order autoregressive processes," *Annals of Statistics*, 19, 1098-1101.

[9] BBS, (1996) "Statistical Pocket Book of Bangladesh," Bangladesh Bureau of Statistics, Government of Bangladesh.

[10] BBS, (1997) "National Accounts Statistics of Bangladesh," Revised Estimates, 1989-90 to 1994-95, Bangladesh Bureau of Statistics, Government of Bangladesh.

[11] Becker N. (1996) "The implications of sustainable development on environmental quality: a theoretical note," forthcoming in Ecological Economics special edition on the environmental Kuznets curve.

[12] Beckerman W. (1992) "Economic Growth and the Environment: Whose Growth? Whose Environment?" *World Development* 20, 481-96.

[13] Beran, R. J. (1984) "Bootstrap Method in Statistics," *Jber. d. Dt. Math. Verein.* 86, 14-30.

[14] Beran, R. J. (1986) "Simulated Power Functions," *Annals of Statistics*, 14, 151-173.

[15] Beran, R. J. (1988) "Prepivoting Test Statistic: A Bootstrap View of Asymptotic Refinements," *Journal of American Statistical Association,* 83, 687-697.

[16] Bewley, R., D. Orden, M. Yang, and L. Fisher (1994) "Comparison of Box-Tiao and Johansen Canonical Estimator of Cointegration Vector in VEC(1) models," *Journal of Econometrics*, 64, 3-27.

[17] Bierens, H. J. (1999) "Unit Roots" *Companion in Theoritical Econometrics*, Blackwell Publisher.

[18] Birdsall, N. and D.Wheeler (1991) "Openness Reduces Industrial in Latin America: The Missing Pollution Haven Effect," Symposium of International Trade and the Environment, Washington, World Bank.

[19] Bollerslev, T. (1986) "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.

[20] Bongaarts, J. (1992) "Population Growth and Global Warming." *Population and Development Review,* 18, 299-319.

[21] Bose, A. (1988) "Edgeworth Correction by Bootstrap in Autoregressions," *Annals of Statistics,* 16, 1709-1722.

[22] Bose, A. (1990) "Bootstrap in Moving Average Models," *Annals. Institute of Statistical Mathematics*, 42: 753-768.

[23] Bousquet A. and Favardy P. (2000) "Does S. Kuznets' Belief Question the Environmental Kuznets Curves? Commissariat l'Energie Atomique, Universit des Sciences Sociales Toulouse. Email: bousquet@cict.fr. Universit de la Rochelle, Email: pfavard@cict.fr

[24] Bowman A.K. and Shenton, B.R. (1975) "Omnibus Test Contours for Departures for Normality Based on $\sqrt{b_1}$ and b_2," *Biometrica,* 64, 243-250.

[25] Box G.E.P and D.R. Cox (1964) "An Analysis of Transformation," *Journal of Royal Statistical Society*, Series B, 26, 211-252.

[26] Breidt, F.J., R.A. Davis, and W.T.M. Dunsmuir (1992) "On Backcasting in Linear Time Series Models," in *New Directions in Time Seties Analysis. Part I,* D. Brillinger et al., eds., Springer-Verlag, New York.

[27] Breidt, F.J., R.A. Davis, and W.T.M. Dunsmuir (1995) "Improved Bootstrap Prediction Intervals for Autoregressions," *Journal of Time Series Analysis*, 16, 177-200.

[28] Breusch, T. and A. Pagan (1979) "A Simple Test for Heteroskedasticity and Random Coefficient Variation," *Econometrica*, 47, 1287-1294.

[29] Bühlmann, P. (1997) Sieve Bootstrap for Time Series. Bernoulli 3, 123-148.

[30] Chang, I., G. C. Tiao and C. Chen (1988) "Estimation of Time Series Parameters in the Presence of Outliers," *Technometrics*, 30, 193-204.

[31] Chang, Y. and J.Y. Park (2003) "A sieve bootstrap for the test of a unit root" *Journal of Time Series Analysis*, 24 (4), 379-400.

[32] Chang, Y., Sickles Robin C. and Song, W. (2001) "Bootstrapping Unit Root Tests with Covariates," unpublished.

[33] Chang, Y., J. Y. Park and K. Song (2002) "Bootstrapping Cointegrating Regressions," Technical Report, Rice University.

[34] Chatterjee, S. (1986), "Bootstrapping ARMA Models: Some Simulations," *IEEE Transactions Systems, Man and Cybernetics*, SMC-16, 294-299.

[35] Chaudhuri, S. and A. S. Pfaff (1997) "Household Income, Fuel-choice and Indoor Air Quality: Microfoundations of an Environmental Kuznets Curve," mimeo, Columbia University.

[36] Chen, C., L. M. Liu and G. C. Tiao (1990) "Robust Estimation in Time Series: A Prametric Approach," Manuscript.

[37] Cropper M. and C. Griffiths (1994) "The interaction of population growth and environmental quality," *American Economic Review* 84, 250-254.

[38] CSIRO, Canberra, Australia. 1992. In: UNEP/UNSTAT Consultative Expert Group Meeting on Environmental and Sustainable Development Indicators. Geneva. Dec. 6-8, 1993.

[39] Daly, H.E. (1977) "Steady State Economics: The Economics of Bio-Physical Equilibrium and Moral Growth," San Francisco. W.H. Freeman.

[40] Datta, S. (1995) "Limit Theory and Bootstrap for Explosive and Partially Explosive Autoregression," *Stochastic Processes and their Applications,* 57,285-304.

[41] Davidson, R. and J. G. MacKinnon (2002a) "Bootstrap J Tests of nonnested Linear Regression Models," *Journal of Econometrics,* 109, 167-193.

[42] Davidson, R. and J. G. MacKinnon (2000b) "Bootstrap Tests: How Many Bootstraps?" *Econometric Review"* 19, 55-68.

[43] Davidson, R. and J. G. MacKinnon (1992) "A New Form of the Information Matrix Test," *Econometrica,* 60, 145-157.

[44] Day M. K. and Grafton Q. R. (2001) "Growth and the Environment in Canada: An Empirical Analysis," CJAE Manuscript No. 144, Department of Economics, University of Ottawa, Canada.

[45] De Bruyn S. M., J. C. J. M. van den Bergh, and J. B. Opschoor (1996) "Economic Growth and Patterns of Emissions: Reconsidering the Empirical Basis of Environmental Kuznets Curves," paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[46] De Wet, T., and J.W.J. van Wyk (1986) "Bootstrap Confidence Intervals for Regression Coefficients when the Residuals are Dependent," *Journal of Statistical Computation and Simulation,* 23, 317-327.

[47] Dickey, David N. and Wayne A Fuller (1979) "Distribution of the Estimators for Autoregressive Time Series With a Unit Root" *Journal of American Statistical Association* 74, 427-31.

[48] Dickey, David N. and Wayne A. Fuller (1981) "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root," *Econometrica*, 49, 1057-72.

[49] Dietz, T and E. A. Rosa (1994) "Rethinking the Environmental Impacts of Population, Affluence, and Technology," *Human Ecology Review*, Summer/Autumn. 1, 277-300.

[50] Dignon J. and S. Hameed (1989) "Global Emissions of Nitrogen and Sulfur Oxides From 1860 to 1980," *Journal of the Air Pollution Control Association*, 39, 180-186.

[51] Dollar J. (1992) "Outward Oriented Countries Really Do Grow More Rapidly: Evidence from 95 LDCs: 1976-85," *Economic Development and Cultural Change*, 40, 523-544.

[52] Efron, B. (1979) "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7, 1-26.

[53] Efron, B. (1982) "The Jackknife, The Bootstrap, and the Other Resampling Plans," CBMS Monograph No. 38, Society for Industrial and Applied Mathematics, Philadelphia.

[54] Efron, B. (1987) " Better Bootstrap Confidence Interval," *Journal of American Statistical Association*, 82, 171-185.

[55] Efron, B. and Tibshirani, R. J. (1986) "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy (with discussion)," *Statististical Science*, 1, 54-77.

[56] Efron, B. and R. J. Tibshirani, (1993) "An Introduction to the Bootstrap" Chapman and Hall, New York. Theory 11, 258-289.

[57] Engle, R. (1982) "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987-1008.

[58] Engleman, R. (1994) "Stabilizing the Atmosphere: Population, Consumption, and Greenhouse Gases," Washington, D.C.: Population Action International.

[59] Engleman, R. (1998) "Profiles in Carbon: An Update on Population, Consumption and Carbon Dioxide Emissions. Washington, D.C.: Population Action International.

[60] Fox, A. J. (1972) "Outliers in Time Series," *Journal of Royal Statistical Society*, Series B, 34, 350-363.

[61] Georgescu R. N. (1971) "The Entropy Law and The Economic Process," Harvard University Press.

[62] Goldfeld, S. M. and R. E. Quandt (1972) "Nonlinear Method in Econometrics," North Holland Publishing Company, Amsterdam, 93-94.

[63] Granger C.W.J. and P. Newbold (1977) "Forecasting Economic Time Series" ACADEMIC PRESS London.

[64] Grossman G. M. and A. B. Krueger (1991) "Environmental Impacts of a North American Free Trade Agreement," National Bureau of Economic Research Working Paper 3914, NBER, Cambridge MA.

[65] Grossman G. M. and A. B. Krueger (1994) "Economic Growth and the Environment," National Bureau of Economic Research Working Paper 4634, NBER, Cambridge MA.

[66] Grossman G. M. and A. B. Krueger (1995) "Economic growth and the environment," *Quarterly Journal of Economics,* 112, 353-378.

[67] Gupta S and P. Bhandari (1999) "An Effective Allocation Ariterion For CO_2 Emissions," *Energy Policy,* 1(1), pp. 727-736.

[68] Hadi, Ali S (1994) "A Modification of a Method for the Detection of Outliers in Multivariate Samples," *Journal of the Royal Statistical Society,* 56:B, 393-396.

[69] Hall, A. (1991) "Testing for Unit Root in Time Series with Pretest Data Based Model Selection," Department of Statistics, North Carolina State University, Mimeo.

[70] Hall, C. et al.(1986) "Energy and Resource Quality: The Ecology of the Economic Process," New York, John Wiley.

[71] Hall, P., J.L. Horowitz and B.Y. Jing (1995) "On Blocking Rules for the Bootstrap with Dependent Data," *Biomettika,* 82, 561-574.

[72] Hamilton D. J. (1994) "Time Series Analysis," Princeton University Press, Princeton, New Jersey.

[73] Hansen, B.E. (1995) "Rethinking the univariate approach to unit root testing: Using covariates to increase power," *Econometric Theory* 11, 1148-1171.

[74] Harbaugh T. W., Levinson A. & Wilson M. D. (2001) " Reexamining the Empirical Evidence for an Environmental Kuznets Curve," National Bureau of Economic Research harbaugh@oregon.uoregon.edu aml6@georgetown.edu

[75] Hendry, D. F. and K. Juselius (2000) " Explaining Cointegration Analysis: **Part I**," *Energy Journal,* 21, 1-42.

[76] Hendry, D. F. and K. Juselius (2001) " Explaining Cointegration Analysis: **Part II**," *Energy Journal,* 22, 75-120.

[77] Hilton, H. and A. Levinson (1998) "Factoring the Environmental Kuznets Curve: Evidence From Automotive Lead Emissions," *Journal of Environmental Economics and Management*, 35(3), 126-141.

[78] Hoaglin, C. David, F. Mosteller, J. W. Tukey (1985) "Exploring Data Tables, Trends, and Shapes," John Wiley & Sons, New York.

[79] Holtz E., Douglas and T. M. Selden (1995) "Stoking the fires? CO_2 Emissions and Economic Growth," *Journal of Public Economics*, 57(5), 85-101.

[80] Horvath R. J. (1996) "Environmental Quality, Throughput Intensity, and Economic Development: the Environmental Kuznets Curve Debate," paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[81] Horowitz, J. (2002) "The bootstrap. in J. Heckman and E. Leamer (Eds.)," *Handbook of Econometrics*, Vol. 5, Amsterdam. North-Holland.

[82] IBRD, World Development Report (1992) "Development and the Environment," Oxford University Press, New York.

[83] Imon, A.H.M.R. (2003) "Regression Residuals, Moments and Their Uses in Tests for Normality," *Communications Statististics -Theory and Mathods*, 32, 1021-1034.

[84] IEA, (International Energy Agency) (1997) "Key World Energy Statistics," Paris.

[85] Jaggers, K. and T. R. Gurr (1995) "Polity III: Regime Change and Political Authority, 1800-1994," 2nd ICPSR version. (Ann Arbor, MI: Inter-university Consortium 1995).

[86] Jarque, C. M. and Bera, A. K. (1987) "A test for Normality for Observations and Regression Residuals," *International Statistical Review*, 55, 163-172.

[87] Jenkins, G.M. and D.W. Watts (1968) "Spectral Analysis and Its Application," San Francisco, Holden-Day.

[88] Johansen, S. (1988) "Statistical Analysis of Cointegrating Vector," *Journal of Economic Dynamics and Control*, 12, 231-254.

[89] Johansen S. (1991) " Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive processes," *Econometrica*, 59, 1551-1580.

[90] Jones, L. E. and R. E. Manuelli (2000) "Endogenous Policy Choice: The Case of Pollution and Growth," *Review of Economic Dynamics.*

[91] Kaufmann R. K., B. Davidsdottir, S. Garnham, and P. Pauly (1996) "The Determinants of Atmospheric SO_2 Concentrations: Reconsidering the Environmental Kuznets Curve," Paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[92] Kilian, L. (1995) "Small-Sample Confidence Intervals for Impulse Response Functions," manuscript, Department of Economics, University of Pensylvania.

[93] Kilian, L. (1996a) "Impulse Response Analysis in Vector Autoregressions with Unknown Lag Order," manuscript, Department of Economics, University of Pennsylvania.

[94] Kiviet, J.F. (1984) "Bootstrap Inference in Lagged-Dependent Variable Models," University of Amsterdam Working Paper.

[95] Komen R., S. Gerking, and H. Folmer (1996) "Income and Environmental Protection: Empirical Evidence from OECD Countries," Department of Economics and Finance, University of Wyoming, Laramie WY.

[96] Kleiner, B., Martin, R. D. and D. J. Thompson (1979) "Robust Estimation ofPower Spectra," *Journal of the Royal Statistical Society*, Series B, 41, 3, 313-351.

[97] Kreiss, J.P., and J. Franke (1992). "Bootstrapping Stationary Autoregressive Moving-Average Models," *Journal of Time Series Analysis*, 13, 287-317.

[98] Kuik, O. and H. Verbruggen (1991) "Towards Sustainable Development Indicators" In Search of Indicators of Sustainable Development. ed. Kluwer Academic Publishers, Dordrecht.

[99] Künsch, H. R. (1989) "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics,* 17, 1217-1241.

[100] Kuznets S. (1955) "Economic Growth and Income Inequality," *American Economic Review*, 49, 1-28.

[101] Kuznets S. (1963) "Quantitative Aspects of The Economic Growth of Nations, VIII: the Distribution of Income by Size," *Economic Development and Cultural Change*, 11, 1-92.

[102] Leybourne, S. and P. Newbold (1999) "On the Size Properties of Phillips-Perron Tests," *Journal of Time Series Analysis* 20, 51-61.

[103] Li, H. and J.S. Maddala (1997) "Bootstrapping cointegrating regressions," *Journal of Econometrics*, 80, 297-384.

[104] Liddle B. T. (1996) "Environmental Kuznets Curves and Regional Pollution," Paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[105] Liu, L. M. and G. B. Hudak (1986) "The SCA Statistical System, Reference Manual for Forecasting and Time Series Analysis." Version III, Lincoln Center, Suit 106, 4513 Lincoln Avenue, Lisle, IL: Scientific Computing Associates.

[106] Liu, R.Y. and K. Singh (1992) "Moving Blocks Jackknife and Bootstrap Capture Weak Dependence," in *Exploting the Limits of Bootstrap,* R. LePage and L. Billard, eds. New York, John Wiley & Sons.

[107] Lopez R. (1994) "The Environment as a Factor of Production: The Effects of Economic Growth and Trade Liberalization," *Journal of Environmental Economics and Management*, 27, 163-184.

[108] Lucas R.E. B., D. Wheeler and H. Hettige (1992) "Economic Development, Environmental Regulation and the International Migration of Toxic Industrial Pollution: 1960-1988, P. Low (Ed.) *International Trade and the Environment,* World Bank Discussion Paper No. 159, Washington DC.

[109] Lutz, E. and S. El-Serafy (1989) "Environmental and Resource Accounting: An Overview," in Environmental Accounting for Sustainable Development, Yusuf J. Ahmad, Salah El-Serafy, and Ernest Lutz, eds. (The World Bank, Washington, D.C.).

[110] McCullough, B.D. (1994). "Bootstrapping Forecast Intervals: An Application to AR(p) Models," *Journal of Forecasting,* 13, 51-66.

[111] Maddala, G.S. and I.M. Kim (1998) "Unit Roots, Cointegration and Structural Change," Cambridge University Press, Cambridge.

[112] Martin, R. D. (1980) "Robust Methods for Time Series," Mimeographed. U. of Washington, Department of Statistics, Seattle, WA.

[113] Martin, R. D., Samarov, A. and W. Vandaele (1983) "Robust Method for ARIMA Models," in *Applied Time Series Analysis of Economic Data*, Proceedings of the Conference on Applied Time Series Analysis of Economic Data, October 13-15, 1981.

[114] Newey, W.K. and K.D. West (1994) "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631-653.

[115] Nicholls, D.F. and A.L. Pope (1988) "Bias in the Estimation of Multivariate Autoregressions, " *Australian Journal of Statistics,* 30A, 296-309.

[116] Omtzigt P. and S. Fachin (2002) "Bootstrapping and Bartlett Corrections in the Cointegrated VAR Model," Technical Paper No. 24, Università Dell'insubria Facoltà Di Economia, http://eco.uninsubria.it.

[117] O' Neill, C. Brian, F. L. MacKellar, W. Lutz (2001) "Population and Climate Change," Cambridge University Press.

[118] Panayotou T. (1993) "Empirical Tests and Policy Analysis of Environmental Degradation at Different Stages of Economic Development," Working Paper, No.WP238, Technology and Employment Programme, International Labor Office, Geneva.

[119] Paulsen, J., and D. TjFstheim (1985) "On the Estimation of Residual Variance and Orderin Autoregressive Time Series," *Journal of the Royal Statistical Society B*,47,216-228.

[120] Perman R. and Stern I. D. (1999) "The Environmental Kuznets Curve: Implications of Non-Stationarity," Center for Resource and Environmental Studies Ecological Economics Program No. 9901.

[121] Peters, E.E. (1991) "Chaos and Order In The Cpital Markets," John Wiley & Sons Inc, New York.

[122] Peters, S. C. and D. A. Freedman (1984) "Some Notes on The Bootstrap in Regression Problems," *Journal of Business and economic Statistics,* 2, 406-409.

[123] Pezzey J. C. V. (1989) "Economic Analysis of Sustainable Growth and Sustainable Development," Environment Department Working Paper No. 15, The World Bank, Washington DC.

[124] Phillips, P.C.B. (1987) "Toward a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74, 535-547.

[125] Phillips, P.C.B. and P. Perron (1988) "Testing for a Unit Root in Time Series Regression," *Biometrica* 75, 335-46.

[126] Politis, D. N. (2003) "The Impact of Bootstrap Methods on Time Series Analysis," *Statistical Science,* 18(2), 219-230.

[127] Politis, D. N., Romano, J. P. and Wolf, M. (1999) "Subsampling," Springer, New York.

[128] Pollak, R. A. and T. J. Wales (1992) "Demand System Specification and Estimation," Oxford University Press, New York.

[129] Pope, A.L. (1987) "Small-Sample Bias Problems in Time Series," unpublished MSC Thesis, Australian National University.

[130] Potscher, B.M. (1991) "Effects of Model Selection on Inference," *Econometric Theory,* 7, 163-185.

[131] Rayner (1990) "Bootstrapping p-Values and Power in the First-Order Autoregression: A Monte Carlo Investigation," *Journal of Business and Economic Statistics,* 8, 251-263.

[132] Reddy A. K. N. and J. Goldemberg (1990) "Energy for the developing world," Scientific American.

[133] Robert R. et al. (1989) "Wasting Assets: Natural Resources in the National Income Accounts," World Resources Institute, Washington, DC.

[134] Rock M. T. (1996) "Pollution Intensity of GDP and Trade Policy: Can The World Bank be wrong?" *World Development,* 24, 471-479.

[135] Rothman D. S. (1996) "Environmental Kuznets Curves - Real Progress or Passing the Buck? A Case for Consumption-Based Approaches," Paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[136] Runkle, D.E. (1987) "Vector Autoregressions and Reality," *Journal of Business and Economic Statistics,* 5, 437-442.

[137] Said, S.E. and D.A. Dickey (1984) "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika,* 71, 599-608.

[138] Saikkonen, P. (1991) "Asymptotically Efficient Estimation of Cointegration Regression," *Econometric Theory,* 7, 1-21.

[139] Sakamoto, Y., Ishiguro, M., and Kitagawa G. (1986) "Akaike Information Criterion Statistics," D. Reidel Publishing Company.

[140] Schwarz, G. (1978) "Estimating the Dimension of a Model", Annals of Statistics, 6, 461-464.

[141] Selden, T. and D. Song (1994) "Environmental Quality and Development: Is There a Kuznets Curve for Air Pollution Emissions?" *Journal of Environmental Economics and Management,* 27(9), 147-162.

[142] Selden, T. and D. Song (1995) "Neoclassical Growth, the *J* Curve for Abatement, and the Inverted U Curve for Pollution," *Journal of Environmental Economics and Management,* 29(9), 162-68.

[143] Shafik N. (1994) "Economic Development and Environmental Quality: An Econometric Analysis," *Oxford Economic Papers,* 46, 757-773.

[144] Shafik N. and S. Bandyopadhyay (1992) "Economic Growth and Environmental Quality: Time Series and Cross-Country Evidence," Background Paper for the World Development Report 1992, The World Bank, Washington DC.

[145] Shafik, N. (1994) "Economic Development and Environmental Quality: An Econometric Analysis," *Oxford Economic Papers,* 46(10), 201-227.

[146] Sharker, M.A.Y and M. Nasser (2004) "Prediction in Time Series by Bootstrapping: A Simulation Based Approach," Proceedings of 7th International Conference of Computer and Information Technology, BRAC University, Dhaka, Bangladesh, ISBN-984-32-1836-1.

[147] Shi, A. (2001) "Population Growth and Global Carbon Dioxide Emissions," Paper presented at IUSSP Conference in Brazil.

[148] Shukla V. and K. Parikh (1992) "The Environmental Consequences of Urban Growth: Cross-National Perspectives on Economic Development, Air Pollution, and City Size," *Urban Geography,* 12, 422-449.

[149] Smil, V. (1990) "Planetary Warming: Realities and Responses," *Population and Development Review* 16(1), 1-29.

[150] Stern D. I., M. S. Common, and E. B. Barbier (1994) "Economic Growth and Environmental Degradation: A Critique of the Environmental Kuznets Curve," Discussion Papers in Environmental Economics and Environmental Management 9409, EEEM, University of York, York.

[151] Stern D. I., M. S. Common, and E. B. Barbier (1996) Economic growth and environmental degradation: the environmental Kuznets curve and sustainable development, World Development, 24, 1151-1160.

[152] Stern, David I. (1998) "Progress on the environmental Kuznets curve?" *Environmental and Development Economics,* 3, 173-196.

[153] Stine R. A.(1987) "Estimating Properties of Autoregressive Forecasts," *Journal of the American Statistical Association,* 82, 1072-1078.

[154] Stock, J. H. and M. W. Watson (1993) "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated System," *Econometrica,* 61, 783-820.

[155] Stokey, N. L. (1998) "Are There Limits to Growth?" *International Economic Review,* 39(2), 1-31.

[156] Suri V. and D. Chapman (1996) "Economic Growth, Trade and the Environment: An Econometric Evaluation of the Environmental Kuznets Curve," Paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[157] Thombs, L.A., and W.R. Schucany (1990) "Bootstrap Prediction Intervals for Autoregressions," *Journal of the Ametican Statistical Association,* 85, 486-492.

[158] Torras M. and J. K. Boyce (1996) "Income, Inequality, and Environmental Quality: An International Cross-Sectional Analysis," paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[159] Tukey, W. John (1977) "Exploratory Data Analysis," Addison-Wesley Publishing Company California.

[160] Unruh G. C. and W. R. Moomaw (1996) "Kuznets Confounded. Evidence of Nonlinear, or Even Chaotic Systems Dynamics in Pollution Emissions Data," paper presented at the 4th biennial conference of the International Society for Ecological Economics, Boston University, Boston MA.

[161] Weisberg, S. (1980) "Applied Linear Regression," New York: Wiely

[162] Westbrook T. (1995) "An Empirical Examination of the Relation Between Carbon Dioxide,"

[163] WRI (1989) World Resources 1988-89, World Resources Institute, Washington DC.

[164] WRI (1991) World Resources 1988-89, World Resources Institute, Washington DC.

[165] Zapata, Hector O. and Alicia N. Rambaldi (1997) "Monte Carlo Evidence on Cointegration and Causation," *Oxford Bulletin of Economics and Statistics*, 59, 285-298.