International Journal of Economic Sciences and Applied Research

Volume 4 Issue 2 August 2011

Print ISSN: 1791-5120 Online ISSN: 1791-3373

International Journal of Economic Sciences and Applied Research

Contents

Volume 4 Issue 2 August 2011

A Region's Basic Image as a Measure of its Attractiveness

Vasilis Angelis[1] and Katerina Dimaki[2]

Abstract

*The growth or decline of a region depends on its power to pull and retain both business and the right blend of people to run them. This pulling power depends on what we call the **"Image"** of the region, a variable which expresses the region's present state of development and future prospects and may be defined as a function of a multitude of factors; economic, social and environmental ones. The Image of a region may prove a very useful tool for planning purposes, since it doesn't only give an early diagnosis of any possible changes, sometimes discontinuous, in the region's pattern of development, but it also indicates the reasons for those changes. Hence, it may be used as the basis for designing appropriate measures to assist a region's development. The objective of this paper is to define a region's image, based on an analysis of the business and the residential location process, to identify, through literature, the factors needed to quantify this image and finally, to suggest ways of measuring these factors.*

Keywords: Region's Image, Region's Attractiveness, Regional Development

JEL Classification: C02, C65, R58

1. Introduction

History has taught us that regional development is a complex process. It results from the balanced presence of tangible and intangible elements mainly originated from the economic and social spheres. Over the past years, a large number of regional growth theories have been developed and a number of models have been built in an effort to describe, explain and eventually predict regional development trends (Pike et al., 2006; Stimson et al., 2006; Capello, 2007; Capello et al., 2008). These models may be classified in various ways on the basis of certain characteristics. A number of such classifications is given below. The first classification is based on the theoretical perspective applied

[1] Quantitative Methods Laboratory, Department of Business Administration, University of the Aegean, 8, Michalon str., 82100 Chios, Greece, v.angelis@aegean.gr

[2] Department of Statistics, Athens University of Economics and Business, 76, Patission str., 10434 Athens, Greece, dimaki@aueb.gr

to the analysis of regional growth and distinguishes between regional growth and local development models (Stilwell, 1972; Stimson et al., 2006). The regional growth models treat a region as a portion of a national system and its purpose is to explain the aggregate growth rate of income and employment. The local development models treat a region as a set of individual economic factors and its purpose is to identify all the tangible and intangible factors of the growth process. A second classification is made on the basis of the role played by space in the various theories and distinguishes between models where the role of space is passive or active respectively (Capello et al., 2008). The first group of models considers space as a mere physical container of growth. The second group considers space as a resource in itself, being the source of increasing returns in the form of agglomeration economies and territorial externalities, and therefore of local development. Furthermore, a region's location is considered to play a key role in its development. A third classification is based on the way regional disparities are treated and distinguishes between equilibrium and disequilibrium models (Holland, 1976; Pike et al., 2006). The equilibrium models consider that regions will converge regardless of any intervention, due to the causal mechanisms of growth that move regions towards equilibrium. The disequilibrium models consider that regional growth disparities persist and are reproduced over time since they are inherent in the process of a region's development. State intervention is needed in order to reduce disparities and assist the less developed regions. Finally, a fourth classification is based on the assumptions made for the potential trend of regional development and distinguishes between linear and non-linear models (Capello, 2007). The linear models may provide approximate replications of short and medium-run changes, but they fail to interpret long-term developments characterized by structural shifts of an irregular nature. The non-linear models allow for a change in a system's dynamics generated by even small perturbations in structural forms (Pike et al., 2006; Capello et al., 2008).

The present paper introduces a region's Basic Image, a variable which expresses a region's state of development and its future prospects. Furthermore, the factors affecting this variable are defined and ways of measuring them are suggested. Finally, this variable is used as the basis for the building of a model calculating regional growth. Regarding the above mentioned classifications the proposed model

- implicitly assumes that regional disparities persist and are reproduced over time; hence a mechanism is needed to reduce and/or alleviate them
- treats space as a resource and considers that a region's location plays a crucial role in its development
- is a local development model as it aims to identify all the tangible and intangible elements of the growth process
- is a non-linear model

Following this brief introduction, section 2 describes a region as a socio-economic unit, section 3 outlines the process of both business and residential location and section 4 introduces and defines the concept of a region's Image. Section 5 presents the basic properties characterising a region's Basic Image, while section 6 develops a mathematical model for its estimation. Finally, section 7 applies this model to the case of the thirteen

Greek regions, whereas Section 8 summarizes the conclusions and makes suggestions for further research.

2. The region as a socio-economic unit

The growth and decline of a region depends on their ability to attract and retain business and people. The realisation that places compete for investment has expanded in recent years to encompass competition among places for the attention of investors and workers (Malecki, 2004). As people and businesses become more mobile, they will move towards attractive places and evacuate unattractive places. The shrinking of time and distance in the global marketplace means that developments in other parts of the world can impact the fortunes of a place once thought to be competitive. This raises fundamental questions about what places can do, not only to survive but, also, to prosper. Places must routinely reassess whether they are meeting the needs of their citizens and businesses. Each place must be continually involved in a process considering the benefits and attractions to be provided to its inhabitants and the ways in which it can help them find and create new value. A region is successful when it manages to meet the needs of its inhabitants and potential movers and, also, to maximize its efficient social and economic functioning in accordance with whatever goals have been established (Ashworth and Voogd 1988; 1990). The success of a region depends on its capability to attract and keep firms with stable or increasing market shares, whilst maintaining stable or increasing standards of living for those who participate in it (Storper, 1997); in other words, when the region is able to generate high profits for its businesses and high standard conditions for its residents (Bristow, 2010).

On the basis of all the above, one can say that a region should ensure livability, investibility and visitability (Kotler et al., 1999) and, in doing so, it performs a number of functions: economic, social and environmental. These functions, however, are not always compatible; on the contrary, the idea of a potential conflict between them often appears in urban literature. A region is a place for work and social interaction. Thus, working and living must be compatible; however, factories often make the living environment unpleasant or even unbearable. Regions are growing mightily in population, wealth and geographic extent but with potentially adverse social and environmental consequences. A region is an environment in which people live, invest, and share ideas and spaces. They meet people and receive information; they send their children to schools and meet entrepreneurs at parties, seminars, restaurants and sporting grounds. People, including entrepreneurs, are embedded in this environment (Boschma and Lambooy, 1999).

Concluding, it could be said that a region, as a socioeconomic system, has to achieve not only a satisfactory economic performance but also a number of other basic social objectives. If these are not met, then, over the longer term, a conflict would arise and the situation would almost certainly not be sustainable (Llewellyn, 1996; Lovering, 2001; Bristow, 2005). Many examples of that inherent conflict between the economic and social development of a region may be provided. They illustrate the basic fact that much of the region's advantages stem from its infrastructure, which, however, detracts from its attractions or, in general from, its social and environmental dimensions. The construction of a motorway in a region, for example, may improve its transportation infrastructure but, at the same time, it may cause a deterioration of its environmental conditions. Similarly, excessive use of the land available for business purposes may improve the business development infrastructure of the region but, at the same time, it may restrict the land available for houses, open spaces and recreational grounds. Those were two examples of a rather direct conflict between the factors affecting the economic development potential of a region and those affecting the level of its environmental conditions and social amenities respectively. More generally, however, a kind of indirect conflict between them may also be detected. The improvement of both sets of factors depends largely on the amount of expenditure the region is prepared to place on each one of them. Nevertheless, given that the budget of a region is always limited, a conflict of priorities is bound to develop.

3. The process of business and residential location

Having discussed the nature of a region as a socio-economic unit, it is time now to study the location process of both business units and employees. Mobility within a state is largely a voluntary process. Hence, the growth and decline of a region depends on its ability to attract and retain business activities and the right blend of people to run them (Bristow, 2005). It is, therefore, important to understand how business firms make investment and location decisions. A region should base its business attraction plans on an assessment of its economic conditions and locational characteristics. Furthermore, an accurate and frequent updating of the prevailing economic conditions, cost factors and quality-of-life features provide an understanding of how well this region compares with others. As a rule, business firms rate places as potential sites after considering various factors that constitute the overall local business climate of any given place. Some of those factors can be measured in more-or-less objective terms; these include location accessibility, economic stability, costs, property value, infrastructure, incentives schemes and programs, financial resources, local support services and networks. Others are not easily measured, as they represent more subjective characteristics of a given place; these include business culture, local entrepreneurial culture, personnel, management, professional and workforce competences, availability of specialised suppliers, quality of the local living or social environment, quality of public education, quality of health services etc. These factors, as well as their relative importance, keep changing over time. Factors representing more subjective characteristics have become increasingly more important for location decisions. Furthermore, the content of various factors has changed over time. In the case of accessibility, for example, the emphasis has been shifted from physical accessibility to communication accessibility. Similarly, in the case of labour, the emphasis has shifted from low-cost, unskilled labour, to quality, high-skilled labour. Furthermore, in the case of housing, education and health facilities, the emphasis has shifted from the availability of those facilities to their quality. Finally, environmental considerations, such as clean air and facility compliance with stronger air, water and chemical and waste disposal regulations, have also grown in importance.

A healthy environment has become a powerful stimulus of capital flow and investors are increasingly critical of the environment they choose to establish new projects; in many cases potential investments and the subsequent economic benefits have been lost, simply because the quality of the area was not enough (Kotler et al., 1999).

Therefore, it can be argued that regarding the movement and location of business units, although traditional factors such as location, accessibility to market and natural resources, transportation facilities, land availability, labour, capital and infrastructure remain important, a number of factors representing more subjective characteristics, as well as a number of environmental factors, have also appeared. However, every region must provide some basic standards of services to attract and retain people, business and visitors. Further analysis of the process of locational choice suggests that a distinction must be made between a list of factors or requirements that is seen by the investors of a firm as a minimum for all locations and those additional factors which may tip the balance between one alternative location or another. However, it is only after the basic requirements have been satisfied that the additional factors are considered. In other words, the process of business location appears to be a two-stage process, whereby the final choice is made from a small group of possible locations all of which satisfy a set of basic criteria (Malecki, 2004). Places not only try to attract businesses and investors, but they also attempt to shape a policy towards attracting and keeping residents as part of building a viable community. Places seek to appeal to certain groups, including the wealthy, young families, workers with special or relevant skills, professionals, managers, technicians, senior officials, administrators and connected families. The free movement of goods and people on the internal market and the opening of new markets have led to more intensive competition for talented persons. Local access to intellectual capital represents one of the most important factors in place development (Kotler et al., 1999).

People attraction is likely to become an even more important strategic component in place competition in the years ahead. As a rule, employees rate places as potential sites after considering various factors. Although job availability/quality and job/pay prospects still remain important, a number of other factors have also appeared. These factors include infrastructure, quality of life, housing options and quality, educational services quality, health services quality, access to daycare centres, competitive social security costs and conditions, a positive attitude to newcomers, and relocation services that include efforts to find job opportunities for family partners. This last factor is gaining in importance. Since in most families today both adults are working, an unhappy partner could discourage the move. Furthermore, as in the case of business, the factors affecting the movement of people, their content and their relative importance keep changing over time (Kotler et al., 1999).

Concluding, it could be suggested that regarding the movement and location of households, although job availability and quality, as well as employment earnings, remain among the most important factors – underlining, therefore, the importance of a strong business activities presence in the region – a new set of factors representing more subjective characteristics are becoming increasingly more significant. Employees look not only for reward and job satisfaction, but, also, for intellectual, social and cultural interests.

Although the two-stage process described in the case of business location is not explicit in the case of people, the existence of certain minimum standards is beyond doubt. In many cases, the people who would generally consider moving were not prepared to move to certain particular areas (Bristow, 2010). More importantly, no specific reasons could be given for this attitude; most of the people justified it in terms of "general dislike of the area" or "lack of appeal" for them. In other words, they were not convinced, in their own minds that although the sums were right, the atmosphere was also right for them. Hence, as in the case of business units location, people are also prepared to consider moving only to areas satisfying a set of basic criteria (Burgess, 1982).

On the basis of all the above, there seem to be a set of "attraction" factors, common for both investors and employees. Those factors include economic stability, economic viability, location, accessibility, land availability, infrastructure, financial resources, housing, health and education facilities of high quality. Furthermore, they may be divided into two basic sets. The first set contains factors related to the economic function of the region, such as economic stability, economic viability, location, accessibility, land availability and infrastructure, whereas the second set contains factors related to the social function of the region, such as housing, health and education facilities. Finally, there exist a number of factors related to the environmental function of the region, including clean air and water supply, as well as pollution control. For the purposes of this paper and in order to keep our model as simple as possible, environmental factors will be grouped together with social factors. However, as the environmental dimension becomes increasingly important, those factors should form a distinct third set.

4. The Concept of a Region's Image

As it has already been mentioned in section 2, the growth or decline of a region depends on its power to "pull" and retain both business activities and the right blend of people to run them; this pulling power depends on what we call the Image of the region. The term image is currently used in a variety of contexts. Image is a sum of beliefs, ideas and impressions. It is the total impression an entity makes on the minds of people and exerts a powerful influence on the way people perceive things and react to them (Dowling, 1998; Dichter, 1985). Marketing literature suggests that image is important in this process and identifies different types, including projected and received place images (Kotler et al., 1993). Projected place images can be conceived as the ideas and impressions of a place that are available for people's consideration. This type of images reach the potential mover by an image transmission or diffusion process through various channels of communication, which themselves can alter the character of the message. The received place images are formed from the interaction between these projected messages and the movers' own needs, motivations, prior knowledge, experience, preferences, and other personal characteristics. In this way, potential movers create their own unique representations or mental constructs, resulting in their own personal images of place (Ashworth and Voogd, 1990; Gartner, 1993; Bramwell and Rawding, 1996).

In this paper, image is defined in a slightly different way, as a function of objectively measured factors, which influence the movement of both business units and people. It is clear that a region's image, based on objectively measured factors, may be improved through marketing and promotion activities. Nevertheless, it is believed that the impact of those activities on the region's image is temporary and limited and the only lasting effect is the objective improvement of the various attributes of this image. Competition among places involves the improvement in the attributes that make it possible to attract and keep investment and migrants – that is, to become 'sticky places' (Markunsen, 1996; Malecki, 2004).

Different people hold quite different images of the same place. Because a region may be related to a number of potential movers' groups that have a different type of interaction with it, each of these groups is likely to have a different image of the particular region. Hence, a region does not have an image, but multiple images (Dowling, 1998). Based on the above, it can be argued that, at each point in time, the region "sends out" its Image and, depending on its impact on the people (both employers and employees), the region may be considered attractive or non attractive. One may also argue that since people "receiving" the image of the region belong to various distinct groups (i.e. employers, unskilled workers, skilled workers etc.) and are sensitive to different factors, the impact of the region's Image on the members of each particular group will be different (Kotler et al., 1999).

Whilst this argument is plausible, the evidence presented in section 2 suggests that all groups of potential movers react similarly to a basic set of factors; more precisely, a set of minimum standards, largely common to all groups, must be satisfied if the region is to be considered as a potential choice by any of them. Every community must provide some basic standards of services to attract and retain people, business and visitors. Admittedly, no uniform standards exist. Hence, every region, in order to be/remain attractive, should determine the standards pertaining each time and try to meet them (Kotler et al., 1999).

To reconcile these two views the concept of a region's Image is refined by introducing the following two concepts: the Basic Image and the Specific Image.

The **Basic Image** of a given region measures the degree to which the region satisfies a set of basic criteria, common for all movers. A region satisfying those criteria is considered by all potential movers worthy of a closer examination and a potential final choice.

The **Specific Image** of a given region, as perceived by a particular group of potential movers, measures the degree to which movers who belong to that particular group consider the region as their best final choice. However, although this Specific Image is a function of specific factors appealing mainly to members of that group, it is primarily a function of the Basic Image.

The remainder of this paper will focus on the definition and study of a region's Basic Image. This is a rather abstract concept which expresses the actual state of the region; a physically realisable measure for the Basic image is difficult to find. What may be measured more easily is the net change of a region's population due to migration during each time period. Such a change, however, is of very little importance as a measure of the real state of the region. The perception and reaction times to any change in the state of a region's Basic Image are different for the various groups of potential movers and are particularly long for certain vulnerable minorities, who lack real choice in place to live and work. Hence, the measurable changes of the region's population due to migration may be generally considered as the delayed and considerably smoothed consequence of changes in the Basic Image.

The study of the mechanisms governing the shaping and the changes of a region's Basic Image is a task of imperative importance. Apart from simplifying the analysis of a region's behaviour, the Basic Image, as an overall measure of its attractiveness and performance, has the following two advantages:

i. It gives an early warning of any potential danger of decline.
ii. It gives the "true" picture of the region and helps decision makes to detect the causes and not only the symptoms of any existing problems.

An early and correct diagnosis of a problem is perhaps the biggest step towards its solution. In the case of regional development, however, the seeds of decay are usually planted during a period of prosperity and no action is taken against them until it is too late. Ironically, the very state of being an attractive place may unleash forces that ultimately unravel the attractiveness of a place. Many places experience a period of growth, followed by a period of decline, and the fluctuations may be repeated several times. Therefore, a monitoring device, which will alert us at the first sight of danger, is a tool of great importance (Kotler et al., 1999).

We have, so far, introduced the Basic Image of a region, as a measure of the degree to which a region satisfies a set of criteria common for all movers. A region satisfying this general set of criteria will be considered as a potential final choice for both people and business activities. Hence, a region's Basic Image will be a function of the factors which influence the movement of both people and business units. Mobility within a state is essentially a voluntary process and state intervention may only be negative in the sense that it can stop or influence movement but it can not direct it. Hence, any attempt to improve or sustain the attractiveness of a region must be directed towards providing the framework within which this voluntary process can flourish. The Basic Image, as defined, may be the basis for such a framework and the factors affecting it will be the prime targets for improvement. The factors affecting the Basic Image, as presented in the previous section, include economic stability, economic viability, location, accessibility, land availability, infrastructure, financial resources, housing, health and education facilities of high quality. Furthermore, they have been divided into two sets according to whether they express the economic or the social function of the region. The factors of the first set (*Accessibility to Centers of Influence, Land Availability, Financial Conditions*) provide a measure of the region's economic development prospects. This measure is referred to as **Economic Indicator.** Similarly, the factors of the second set (*Housing Conditions, Environmental Conditions, Social Conditions*) provide a measure of a region's social profile. This measure is referred to as **Social Indicator.** Hence,

$$\text{Basic Image} = \varphi\left(\text{Economic Indicator, Social Indicator}\right)$$

Ways for the quantification of those Indicators will be presented in section 6.

The expression of the Basic Image as a function of those two Indicators is not accidental; on the contrary, it is consistent with the concept of a region as a socio-economic unit. The main advantage of such an expression is that it may be used to underline and, eventually, describe the potential conflict between the economic and social functions of a region in the course of development (Llewellyn, 1996; Lovering, 2001; Bristow, 2005).

Concluding, it should be mentioned that the growth of a region may be expressed both in absolute or relative terms. In the latter and most interesting case, the development pattern of a given region is compared to that of a hypothetical region, which is referred to as the "typical" region and expresses, as far as possible, an average of the main regions of a similar type to that under study. In this paper, we shall be looking at the relative development patterns of a region. Hence, all the factors affecting its images (Basic and Specific) should be expressed in relative terms, as compared to the corresponding values of the "typical" region.

5. Properties of the Basic Image

Let us now move a step further and concentrate on the problem of the theoretical shape of the graph of the Basic Image. It is reminded that the Basic Image has been defined as a function of two potentially conflicting indicators and, hence, its graph must be a three-dimensional one. In order to get a first feeling of the shape of that graph we start by stating the following simple observations describing the way in which the two indicators operate.

i. The higher the Economic Indicator of a region, the more attractive its Basic Image.
ii. The lower the Social Indicator of a region, the less attractive its Basic Image,
iii. If the Economic Indicator of a region is continuously increasing but, at the same time, its Social Indicator is continuously decreasing, the Basic Image of the region may be either attractive or non attractive and sudden changes in its state may be expected.

Observation (iii) is the most interesting because it implies that the graph we want to draw may be discontinuous.

The general mathematical theory of discontinuous and divergent behaviour from continuous underlying forces is called Catastrophe Theory (Thom, 1975; Zeeman, 1973). The theory is derived from Topology and is based upon some new theorems in the geometry of many dimensions, which classify the ways in which discontinuities may occur, in terms of a few archetypal forms called elementary catastrophes (Poston and Stewart, 1996). Although the underlying mathematics is difficult and the proofs of the theorems involved complicated, the elementary catastrophes themselves are relatively easy to understand and can be used effectively, even by non-experts in the subject (Angelis and Dimopoulou, 1991). Catastrophe theory was developed and popularized in the early 1970's. After a period of criticism, it is now well established and widely applied (Rosser, 2007). Today, the theory is very much alive and numerous nonlinear phenomena that exhibit discontinuous jumps in behaviour have been modeled by using the theory, for instance in chemistry (e.g Wales, 2001), in physics (e.g. Aerts, 2003), in psychology (e.g. Van der Mass et al., 2003) in clinical studies (e.g. Smerz and Guastello, 2008) and in the social sciences (e.g. Smith et al., 2005; Dou and Ghose, 2006; Huang, 2008).

Table 1 summarizes the elementary catastrophes in the case where a process is expressed through one behaviour variable depending on one up to four control variables.

In the case of a process, for example, where the behaviour depends on two control variables, it is sufficient to know that a theorem exists giving the qualitative shape of a 3-dimensional surface, which shows all possible ways in which a discontinuity in the behaviour may occur. The two control variables are usually referred to as normal and splitting factor respectively.

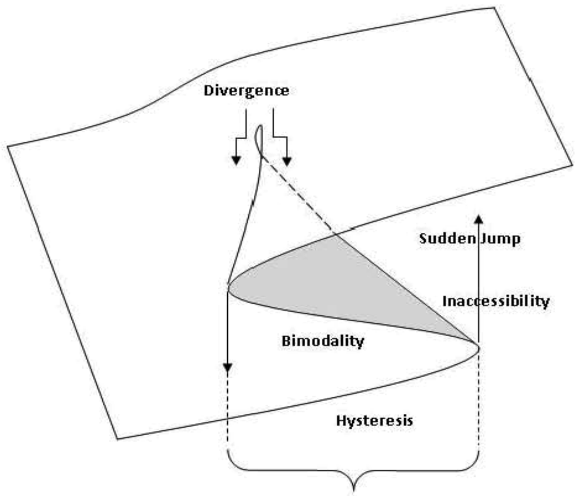
Table 1: Some Elementary Catastrophes

Number of Behaviour Variables	Number of Control Variables	Type of Catastrophe
1	1	Fold
1	2	Cusp
1	3	Swallowtail
1	4	Butterfly

Elementary Catastrophes have characteristic invariant properties and, often, even wave flags (Gilmore, 1993) to gain our attention. Figure 1 illustrates graphically those five properties for the case of cusp catastrophe. Further details about them are given below.

- ***Modality*** arises when, for some combinations of values of the control parameters, there are two or more possible stable values for the state variable.
- ***Sudden Jumps*** take place when a small change in the values of the control parameters may result in a large change in the value of the state variable, as the system jumps from one local minimum to another.
- ***Hysteresis*** occurs whenever a physical process is not strictly reversible. That is, when the jump from one local minimum to another does not occur over the same point in the control parameter space as the reciprocal jump.
- ***Divergence*** arises when small changes in initial values of the control parameters lead to large changes in the final value of the state variable.
- ***Inaccessibility*** means that the physical system has an equilibrium state which is unstable.

Figure 1: The five Characteristic Properties of the Cusp Catastrophe Graph



Returning to the present case, our intention is to show that the process of shaping a region's Basic Image may be modeled in terms of a cusp catastrophe. The first step towards this direction will be to show that at least some of the five invariant properties characterizing phenomena that may be described by the cusp catastrophe (i.e. bimodality, sudden transitions, hysteresis, divergence and inaccessibility) are present in our particular case.

Camagni (2002) notes that regions, unlike nations, more or less can go out of business, becoming so depleted by outmigration that they have, at a long-run, competitive disadvantage. The European continent, with its many competing communities, regions and nations, is now experiencing extreme economic turbulence, where two basic dimensions may be identified. First, every place is subject to internal growth and decline cycles. Second, every place may experience external shocks and forces beyond its control. A large number of European places are experiencing problems, but some more than others (Kotler et al., 1999). The situations fall along a continuum. At the most desperate extreme are places that are dying or are chronically depressed. Many such places have emerged in Europe, as a result of recent decades of economic crises and industrial restructuring. These depressed places lack even the internal resources to launch recovery. There are also acutely depressed places that nonetheless have some potential for revival. While their debts and problems keep worsening, these places possess assets that could support a turnaround should the right leadership and vision emerge. Other places have boom and bust characteristics. These places, as a result of their mix of industries and growth companies, are highly sensitive to business cycle movements. In many cases, in order to survive, those places have shifted their focus from a declining business sector, to a more promising one. On the brighter side of Europe, we find many places that have undergone healthy transformations. These places have devised effective plans to create new conditions that improve their attractiveness and ensure a turbulence free transition. Finally, some places deserve the title of favoured few, as they enjoy a strong position and continue to attract businesspeople, new residents and visitors (Kotler et al., 1999).

The concept of **modality** and **sudden transitions** in the development of a region is certainly not novel in literature. Many places experience a period of growth followed by a period of decline and the fluctuations may be repeated several times (Camagni, 2002). The growth period inevitably ends because growth lays the seeds of its own destruction. The decline period will also end, but for different reasons. The processes underlying growth and decline dynamics can occur independently of the business cycle stage. However, these processes may be accelerated by sudden changes in the economic climate (Kotler et al., 1999). Boschma and Lambooy (1999) writing about the industrial areas of the 1970's and 1980's, mentioned that had often showed long periods of economic growth, before they declined or even collapsed. Their position became vulnerable due to developments like technological change or the increasing opportunities to shift production to other regions or countries with cheap labour. Within a decade or so, several regions lost many jobs in mature business activities, like textiles, steel making, coal mining and shipbuilding. This was something quite unexpected, because, traditionally, regional economists focus their attention on the positive impacts of agglomeration economies on regional development.

Hysteresis is a characteristic property of the development of a region and is reflected in the delays observed before any sudden changes in the Basic Image of the region take place. For an attractive region, in the process of decline, those delays extend its stay on the attractive side and they are due to the strong attachment to an area displayed by both business units and people. Some business activities are maintained in districts where the original reasons for their development are no longer significant, or even no longer exist. This phenomenon is sometimes referred to as "industrial inertia". While the main reasons for the business units attachment to any area are economic, in the case of people they are essentially psychological. The strength of those psychological factors has been reflected on the poor results of various government schemes aiming to transfer unemployed workers and their families from depressed to more prosperous areas. This finding verifies the existence of strong "residential inertia". Boschma and Lambooy (1999) argue that the poor ability of old agglomerations to learn, innovate and adapt is explained in terms of socio-cultural factors rather than purely economic factors.

Divergence is usually recognisable in the case of two competing regions, especially in a period of rationalisation of their main business activities. Some regions lack the resources to launch a recovery, whereas others have the potential for revival (Kotler et al., 1999). The line between successful and open regions and old industrial, insular, inward-looking industrial districts can be very thin. Some agglomerations, when confronted with catastrophic changes in their regional specialisation, have displayed a strong vitality, like the Boston region and the Birmingham region. On the other hand, similar agglomerations, like the Ruhr area, the Manchester area, the Liverpool area and the Newcastle upon Tyne area, have followed a quite different trend of no adaptation and, therefore, decline (Boschma and Lambooy, 1999).

The idea of **inaccessibility**, although never expressed in this explicit way, is not novel in literature. The loss of confidence of the community in the future of a region leads to accelerated immigration, rapid shifts in investment and physical neglect. The more sudden the loss of confidence, the more rapid the decline. A region in decline enters a vicious circle as the problems mount and at the same time the region's financial resources and consequently its ability to face those problems decrease. The blight spreads at an accelerating rate and acts as a negative multiplier reinforcing and speeding up the depression. Once a process of regional decline has set in, it becomes self-reinforcing through all kinds of sub indicator and accelerator mechanisms (Boschma and Lambooy, 1999). As a place begins to lose its attractiveness, forces are released that worsen the situation and the image of the place becomes further tarnished. The community raises taxes to maintain or improve the infrastructure and to meet social needs, but the higher taxes only accelerate the out-migration of resources. Unfortunately, the European maps depict numerous places of decaying examples (Kotler et al., 1999). The potential of high acceleration in the loss or gain of a region's attraction power, once it has entered a cycle of deprivation or prosperity, suggests that in such cases a range of values of its Basic Image representing neutrality may be generally considered as unstable and, therefore, practically unattainable.

6. Modeling a Region's Basic Image

6.1 The general form of the model

We have so far shown that the process of shaping a region's Basic Image has all the properties characterising phenomena which may be modeled in terms of Catastrophe Theory. Hence, we may now use Catastrophe Theory to estimate a region's Basic Image. It is reminded that the Basic Image of a region has been defined as a function of two conflicting indicators. Therefore, the appropriate elementary catastrophe is the cusp. Consequently, the value x, of a region's Basic Image, at each point in time, is given as a solution of the equation:

$$x^3 - bx - a = 0 \qquad (1)$$

with,

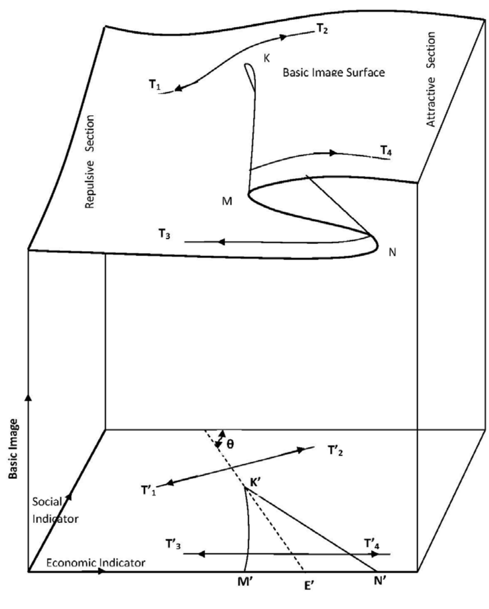
$$\begin{cases} a = m(\alpha - \alpha_0) + (\beta - \beta_0) \\ b = (\alpha - \alpha_0) - m(\beta - \beta_0) \end{cases} if\ m \le 1 \left(i.e.\ \theta \le \frac{\pi}{2} \right) \qquad \text{and}$$

$$\begin{cases} a = (\alpha - \alpha_0) + (1/m)(\beta - \beta_0) \\ b = (1/m)(\alpha - \alpha_0) - (\beta - \beta_0) \end{cases} if\ m > 1 \left(i.e.\ \theta > \frac{\pi}{2} \right)$$

Equation (1) is referred to as the **Basic Image Equation** and its graph is qualitatively equivalent to the Cusp Catastrophe Graph (Figure 2).

The variables *α, β* express the values of the two Indicators, while α_0, β_0, express the values of those two Indicators for the "typical" region. The point (α_0, β_0) corresponds to the vertex of the cusp, while $m = \tan\theta$ represents the slope of the cusp axis and expresses the relative weights attached to each one of the two indicators in defining the Basic Image. For the purposes of this work, the values of all Indicators lie in the interval [0,1], whereas the value of its Basic Image lies in the interval [-1,1]. The value of the "typical" region's Basic Image is 0. Hence, positive Basic Image indicates an attractive region that may be considered as a potential final choice by the various groups of prospective movers.

Figure 2: The Cusp Catastrophe graph in the case of Basic Image



The position of the cusp in Figure 2 is indicative. The trajectory of a region's Basic Image lies on the Basic Image surface. As long as the trajectory remains on the upper section of this surface, the area is attractive, whereas in case the trajectory moves on the lower part, the region becomes repulsive. T_1T_2 and T_3T_4 are typical trajectories of an area's Basic Image and $T_1^{'}T_2^{'}$, $T_3^{'}T_4^{'}$ are their projections on the two dimensional Control Space C. The line KM is the locus of breaking points for areas undergoing sudden loss of attractiveness while the line KN is the locus of turning points for regions going through a phase of sudden increase of attractiveness. $K^{'}M^{'}, K^{'}N^{'}$ are the projections of KM, KN on the Control Space and $K^{'}E^{'}$ is the projection on C of the cusp axis.

We have so far defined a region's i Basic Image as a function of a multitude of factors, grouped into two potentially conflicting indicators. A large variety of indicators, either simple or composite, quantifying the economic, social and environmental dimension of a region may be found in the relevant literature (Hammond et al., 1995; Freudenberg, 2003; Slavova, 2008; Eurostat Regional Yearbook, 2008). For the purposes of our model, those indicators are expressed as the geometric mean of several **Sub indicators**, each of which depends on a number of factors among those affecting the region's Basic Image. The use of this geometric mean is justified by the fact that each one of the Sub indicators affecting the respective indicator is considered to be critically important for this indicator's value. Consequently,

$$IND_i^h = \sqrt[m]{\prod_{j=1}^{m} SbI_{ij}^h},\ h = 1,\ 2;\ i = 1,\ 2,\ldots,n$$

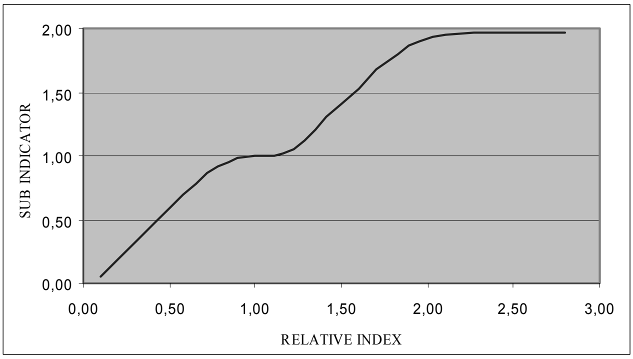
where, IND_i^h denotes the h^{th} Indicator of region i and SbI_{ij}^h denotes the j^{th} Sub indicator of region i, which is related to Indicator h. Each Sub indicator SbI_{ij}^h is defined as a non-linear function of a respective Relative Index RI_{ij}^h. This index is, in turn, a function of all variables, measured or estimated, affecting the Sub indicator and may be defined in the following two ways:

- The values of all variables, expressed in relative terms with respect to the typical region, are used to obtain directly the Relative Index RI_{ij}^h, $h = 1,\ 2$, $i = 1,\ 2,\ldots,n$, $j = 1,\ 2,\ldots,m$.
- The variables are classified into various sets, depending on the specific component of the Sub indicator they affect. The values of all variables which belong to every set, expressed in relative terms with respect to the typical region, are used to obtain directly the respective Relative Sub indices $RSI_{ij_k}^h$, $h = 1,\ 2$, $i = 1,\ 2,\ldots,n$, $j = 1,\ 2,\ldots,m$, $k = 1,\ 2,\ldots,r$. Finally, those Sub Indices are combined so as to give Relative Index:

$$RI_{ij}^h = \frac{\sum_{k=1}^{r} w_k RSI_{ij_k}^h}{\sum_{k=1}^{r} w_k},\ h = 1,\ 2,\ i = 1,\ 2,\ldots,n,\ j = 1,\ 2,\ldots,m,$$

where, w_k, $k = 1,\ 2,\ldots,r$ are weights indicating the relative importance attached to each Sub index in defining the respective Relative Index.

Figure 3: An indicative Transformation of a Relative Index into the corresponding Sub indicator



Once the Relative Index has been defined, its values are transformed so as to obtain the corresponding values of Sub indicator SbI_{ij}^{h}. For the purposes of the model, this transformation has been based on available data but, also, on assumptions consistent with generally accepted views expressed in literature. In the case of a given sub indicator SbI_{ij}^{h}, depending on a Relative Index RI_{ij}^{h}, a simple relationship has been used. As long as the value of the Relative Index RI_{ij}^{h} is close to 1, the value of the corresponding sub indicator remains also close to 1, i.e. close to the typical region's value, thus indicating a limited variation of the sub indicator's influence on the region's Image. However, as the value of the Relative Index RI_{ij}^{h} becomes substantially greater or lower than 1, in other words substantially better or worse than the typical region's value, SbI_{ij}^{h} also increases rapidly, indicating its strong influence on the region's development. An indicative transformation is shown in Figure 3. The prospective user of the model, however, may easily modify this transformation if his underlying set of assumptions is different.

Note that each Sub indicator and the respective Relative Index summarize the same aspect of a region's development. The transformation used works as a standardization process and it is needed to ensure that:

- all sub indicators have the same range. For the purposes of the model, all sub indicators have the same range values –usually [0,2]; hence, the range of their product is $[0, 2^n]$ and, consequently, the range of IND_i^h, $h=1, 2;\ i=1, 2, \ldots, n$ is also [0,2]. In certain cases, however, the dominance of a particular sub indicator needs to be emphasized. This may be done by increasing its range. In such a case, the range of the remaining sub indicators must be modified, so that the range of their product remains the same i.e. $[0,\ 2^n]$.
- the effect of changes in the values of variables on the values of the respective sub indicators follow the same pattern for all sub indicators.

6.2 Estimation of the model parameters for the case under study

Returning to the present case, it is reminded that the factors affecting a region's Basic Image may be allocated into two sets, according to whether they express the economic or the social aspect of the region. The factors of the first set provide a measure of the region's economic development prospects. This measure is referred to as the **Economic Indicator.** Similarly, the factors of the second set provide a measure of the region's social profile. This measure is referred to as the **Social Indicator.** Furthermore, each of those Indicators is expressed as the geometric mean of several Sub indicators as shown below:

$IND_i^1 = \sqrt[3]{\prod_{j=1}^{3} SbI_{ij}^1},\ i = 1, 2, \ldots, n$		$IND_i^2 = \sqrt[3]{\prod_{j=1}^{3} SbI_{ij}^2},\ i = 1, 2, \ldots, n$	
where			
IND_i^1 :	**The Economic Indicator of region i**	IND_i^2 :	**The Social Indicator of region i**
SbI_{i1}^1 :	The Location Sub indicator of region i	SbI_{i1}^2 :	The Housing Conditions Sub indicator of region i
SbI_{i2}^1 :	The Land Availability Sub indicator of region i	SbI_{i2}^2 :	The Social Conditions Sub indicator of region i
SbI_{i3}^1 :	The Financial Conditions Sub indicator of region i	SbI_{i3}^2 :	The Environmental Conditions Sub indicator of region i

It should be noted that IND_i^1, IND_i^2 coincide with the coefficients a and b of the Basic Image Equation (equation (1)). A clear overview of the variables affecting a region's Basic Image and their conversion through Sub Indices, Relative Sub indices, Relative Indices and Sub-indicators into Indicators and, finally, into the region's Basic Image is given in Table 2. One may argue that some significant variables expressing the region's power to retain/attract movers belonging to various groups are missing from Table 2.

Table 2: Conversion of the variables affecting the Basic Image of region i

<table>
<tr><th colspan="6">INDICATORS, INDICES AND VARIABLES CONCERNING REGION i</th></tr>
<tr><th>Indicators of region i</th><th>Sub indicators of region i</th><th>Relative Indices of region i</th><th>Relative Sub indices of region i</th><th>Sub indices of region i</th><th>Variables</th></tr>
<tr><td rowspan="6">Economic Indicator $\left(IND_i^1\right)$</td><td rowspan="2">Location Sub indicator $\left(SbI_{i1}^1\right)$</td><td rowspan="2">Relative Location Index $\left(RI_{i1}^1\right)$</td><td></td><td></td><td>Size of Influence Centres</td></tr>
<tr><td></td><td></td><td>Distance/Cost from Influence Centres</td></tr>
<tr><td rowspan="2">Land Availability Sub indicator $\left(SbI_{i2}^1\right)$</td><td rowspan="2">Relative Land Availability Index $\left(RI_{i2}^1\right)$</td><td rowspan="2">―</td><td rowspan="2">―</td><td>Area</td></tr>
<tr><td>Population</td></tr>
<tr><td rowspan="2">Financial Conditions Sub indicator $\left(SbI_{i3}^1\right)$</td><td rowspan="2">Relative Financial Conditions Index $\left(RI_{i3}^1\right)$</td><td rowspan="2">―</td><td rowspan="2">―</td><td>Gross Domestic Product,</td></tr>
<tr><td>Population</td></tr>
<tr><td colspan="6"></td></tr>
<tr><td rowspan="14">Social Indicator $\left(IND_i^2\right)$</td><td rowspan="4">Housing Conditions Sub indicator $\left(SbI_{i1}^2\right)$</td><td rowspan="4">Relative Housing Conditions Index $\left(RI_{i1}^2\right)$</td><td rowspan="2">Relative Housing Availability Sub index $\left(RSI_{i1_1}^2\right)$</td><td rowspan="2">Housing Availability Sub index $\left(SI_{i1_1}^2\right)$</td><td>Total Number of Houses</td></tr>
<tr><td>Population</td></tr>
<tr><td rowspan="2">Relative Housing Quality Sub index $\left(RSI_{i1_2}^2\right)$</td><td rowspan="2">Housing Quality Sub index $\left(SI_{i1_2}^2\right)$</td><td>Number of New Houses</td></tr>
<tr><td>Total number of Houses</td></tr>
<tr><td rowspan="6">Social Conditions Sub indicator $\left(SbI_{i2}^2\right)$</td><td rowspan="6">Relative Social Conditions Index $\left(RI_{i2}^2\right)$</td><td rowspan="3">Relative Health Services Sub index $\left(RSI_{i2_1}^2\right)$</td><td rowspan="3">Health Services Sub index $\left(SI_{i2_1}^2\right)$</td><td>Number of Doctors</td></tr>
<tr><td>Number of Hospital Beds</td></tr>
<tr><td>Population</td></tr>
<tr><td rowspan="3">Relative Educational Services Sub index $\left(RSI_{i2_2}^2\right)$</td><td rowspan="3">Educational Services Sub index $\left(SI_{i2_2}^2\right)$</td><td>Number of Teachers</td></tr>
<tr><td>Number of Classrooms</td></tr>
<tr><td>Population</td></tr>
<tr><td rowspan="4">Environmental Conditions Sub indicator $\left(SbI_{i3}^2\right)$</td><td rowspan="4">Relative Environmental Conditions Index $\left(RI_{i3}^2\right)$</td><td rowspan="2">Relative Industrial Pollution Sub index $\left(RSI_{i3_1}^2\right)$</td><td rowspan="2">Industrial Pollution Sub index $\left(SI_{i3_1}^2\right)$</td><td>Industrial Electricity Consumption</td></tr>
<tr><td>Total Electricity Consumption</td></tr>
<tr><td rowspan="2">Relative Car Pollution Sub index $\left(RSI_{i3_2}^2\right)$</td><td rowspan="2">Car Pollution Sub index $\left(SI_{i3_2}^2\right)$</td><td>Number of Cars</td></tr>
<tr><td>Population</td></tr>
</table>

Such variables include labour availability/quality and financial incentives for investors, as well as job availability/quality, employment earnings and financial incentives for employees. This is a plausible argument but, on the other hand, it must be noted that those factors will be used in a next step, which is, however, beyond the scope of this paper, for the estimation of a region's Specific Images, as perceived by the various groups of potential movers. The Specific Images express the degree to which the members of each group consider the region as their best final choice. A physically realisable measure of those Images is the net change, over a period of time, in the number of the members of each group present in the region. As it has already been mentioned, those changes may be generally considered as the delayed and considerably smoothed consequences of the changes in the Basic Image. Hence, a region's Basic Image, as defined, gives a reliable overall estimate of the region's prospects of development and an early warning for any potential danger. Finally, it must be underlined that the choice of variables used for the estimation of a region's Basic Image depends, among other things, on the availability of data. In any way, however, our intention in this paper is to provide a generic framework for the estimation of a region's Basic Image. Within this framework, every researcher may make the appropriate modifications according to both his research requirements and the data availability.

All the Sub Indicators presented in Table 2 are defined below.

- **The Location Sub indicator**

Every business activity, in order to operate effectively and efficiently, requires access to sources of raw materials, commerce and service centres, as well as clusters of other industries. In other words, it requires access to what we may generally call "influence centres". An area, the location of which offers "influence centres", has a strong comparative advantage over its competitors in attracting industrial units.

The Location Sub indicator of region i, $\left(SbI_{i3}^{2}\right)$, is a non linear transformation of the Relative Location Index $\left(RI_{i3}^{2}\right)$, which expresses the region's relative position with respect to the various influence centers. Every region is generally surrounded by more than one influence centers. Hence, the Relative Location Index expresses the total influence exerted on region i by all influence centres. In other words, the Relative Location Index is the sum of r Relative Location Sub indices $\left(RSI_{i1_k}^{1}\right)$, $k = 1, 2, \ldots, r$, each one expressing the influence exerted on region i by the respective influence centre k. Hence,

$$RI_{i1}^{1} = \sum_{k=1}^{r} RSI_{i1_k}^{1}.$$

Furthermore, each of the Relative Location Sub indices is a function of:

- The influence centre's size, as defined by its Gross Domestic Product, expressed in relative terms.
- The region's accessibility to the given influence centre, which depends on
  - – the cost of transporting a unit quantity between region i and the given influence centre, expressed in relative terms.
  - – The degree of a region's spatial discontinuity, as defined by the transport modes available and their transportation capacity, expressed in relative terms.

- **The Land Availability Sub indicator**

Measuring land availability is a delicate subject. If the area available for business use is considered fixed, as in the case of an area surrounded by a clearly defined "green belt", then land availability, at any time, may be measured as the fraction of the area which is available for use. Generally, however, the area available for business use is allowed to expand in order to accommodate any further growth. Although expansion is not limitless, the measure presented above is meaningless in this case. A more suitable measure would be the density of population in the surrounding region. High density indicates a high degree of urbanization in the area and makes further expansion difficult. Local regulations on land use must also be taken into account, whenever it is necessary.

The Land Availability Sub indicator of region i, $\left(SbI_{i3}^{2}\right)$, is a non linear transformation of the Relative Land Availability Index $\left(RI_{i3}^{2}\right)$, which is defined as the inverse population density ratio for this region expressed in relative terms.

- **The Financial Conditions Sub indicator**

The term refers to the level of general economic conditions prevailing in the region and, somehow, reflects the standard of living of its inhabitants. The Financial Conditions Sub indicator of region i, $\left(SbI_{i3}^{1}\right)$, is a non linear transformation of the Relative Financial Conditions Index $\left(RI_{i3}^{1}\right)$, which is defined as the region's relative gross domestic product per capita expressed in relative terms.

- **The Housing Conditions Sub indicator**

The Housing Conditions Sub indicator of region i, $\left(SbI_{i1}^{2}\right)$, is a non linear transformation of the Relative Housing Conditions Index, $\left(RI_{i1}^{2}\right)$, which combines two aspects of the region's housing stock: availability and quality. Housing availability is expressed through the Relative Housing Availability Sub index $\left(RSI_{i1_1}^{2}\right)$, which is the ratio of the total number of houses available over the population, expressed in relative terms. Housing quality is expressed through the Relative Housing Quality Sub index $\left(RSI_{1_2 j}^{2}\right)$, which is the ratio of the number of new houses over the total number of houses, expressed in relative terms. Hence, on the basis of the above we have that:

$$RI_{i1}^{2} = \frac{w_1 RSI_{i1_1}^{2} + w_2 RSI_{i1_2}^{2}}{w_1 + w_2}$$, where w_k, $k = 1,2$ are the appropriate weights.

- **The Social Conditions Sub indicator**

The Social Conditions Sub indicator of region i, $\left(SbI_{i3}^{2}\right)$, is a non linear transformation of the Relative Social Conditions Index, which combines two aspects of the region's social profile, health services and educational services.

The level of health services is expressed through the Relative Health Services Sub index, $\left(RSI_{i2_1}^{2}\right)$, which is the weighted average of two ratios: the ratio of the number of doctors available in the region over its population, expressed in relative terms and the ratio of the number of hospital beds available in the region over its population, measured in relative terms.

Similarly the level of educational services is expressed through the relative Educational Services Sub index, $\left(RSI_{i3_1}^2\right)$, which is a weighted average of two ratios: the ratio of the number of teachers available in the region over its population and the ratio of the number of school classrooms available in the region over its population, both ratios expressed in relative terms. Hence, on the basis of the above we have that:

$$RI_{i2}^2 = \frac{w_3 RSI_{i2_1}^2 + w_4 RSI_{i2_2}^2}{w_3 + w_4}$$, where w_k, $k = 3,4$ are the appropriate weights.

- **The Environmental Conditions Sub indicator**

Environment is a unity wherein many elements interact but several of them may be distinguished; air pollution, water pollution, noise, solid waste disposal, and dereliction of land. For the purposes of the present work, two sources of environmental degradation are considered: excessive industrialization and heavy use of cars.

The Environmental Conditions Sub indicator of region i, $\left(SbI_{i3}^2\right)$, is a non linear transformation of the Relative Environmental Conditions Index, $\left(RI_{i3}^2\right)$, which combines two aspects of the region's environmental profile, industrial pollution (excessive industrialization) and car pollution (heavy use of cars).

The level of industrial pollution is expressed through the Relative Industrial Pollution Index, $\left(RSI_{i3_1}^2\right)$, which is the ratio of the total annual electrical consumption in the region over the electrical consumption for industrial uses only, expressed in relative terms.

The level of car pollution is expressed through the Relative Cars Pollution Index, $\left(RSI_{i3_2}^2\right)$, which is the ratio of the region's population over the total number of cars available, expressed in relative terms. Hence, on the basis of the above we have that:

$$RI_{i2}^2 = \frac{w_5 RSI_{i2_1}^2 + w_6 RSI_{i2_2}^2}{w_5 + w_6}$$, where w_k, $k = 5, 6$ are the appropriate weights.

7. Application of the Proposed Model

The methodology presented in the previous section is now used for the estimation of the Basic Image of the 13 regions of Greece (Figure 4). The required data have been drawn from the official site of the Hellenic Statistical Authority. The results are shown in Table 3, which gives the values of Economic Indicator, Social Indicator and Basic Image for all 13 regions of Greece for the year 2005. The values of Economic and Social Indicator for the typical region have also been calculated and found to be 0.45 and 0.52 respectively.

Figure 4: The 13 Greek Regions

1. Attica
2. Central Greece
3. Central Macedonia
4. Crete
5. East Macedonia and Thrace
6. Epirus
7. Ionian Islands
8. North Aegean
9. Peloponnese
10. South Aegean
11. Thessaly
12. West Greece
13. West Macedonia

Source: http://upload.wikimedia.org/wikipedia/commons/9/9d/GreeceNumberedPerepheries.png

Table 3: Basic Image values for the 13 Greek Regions

$m = 1,\ \alpha_0 = 0.45,\ \beta_0 = 0.52$

REGION	Economic Indicator	Social Indicator	BASIC IMAGE
1. ATTICA	0.6969	0.4748	0.7493
2. EAST MACEDONIA & THRAKI	0.4355	0.4436	-0.4954
3. CENTRAL MACEDONIA	0.5234	0.5083	0.4663
4. WEST MACEDONIA	0.4989	0.5655	0.4579
5. EPEIROS	0.4840	0.6169	0.4666
6. THESSALY	0.5283	0.4930	0.4646
7. IONIAN ISLANDS	0.3676	0.5732	-0.1755
8. WEST GREECE	0.5023	0.4841	0.3649
9. STEREA ELLADA	0.5673	0.4445	0.5222
10. PELOPONNISOS	0.5107	0.5158	0.4398
11. NORTH AEGEAN	0.2238	0.5485	-0.4405
12. SOUTH AEGEAN	0.2605	0.5754	-0.3588
13. CRETE	0.2801	0.5494	-0.3952

By looking at the results the following conclusions may be drawn:

- All mainland regions have positive Basic Image with the exception of East Macedonia and Thrace, whereas all island regions have negative Basic Image.
- East Macedonia and Thrace, the only mainland region with a negative Basic Image value, is a remote border region with poor accessibility, something which is reflected in the relatively low value of its Economic Indicator. Hence, any effort to improve its Basic Image should start from the improvement of its accessibility i.e. transportation infrastructure and means.
- The negative Basic Image value of all island regions is a natural consequence of their high spatial discontinuity which makes it extremely difficult for them to attract economic activities involving transportation of materials and goods. All efforts aiming to reduce their geographic discontinuity, through the improvement of transportation infrastructure and means, have limited results. Hence, a realistic alternative way to overcome the problem will be to bypass geographic discontinuity through one of the following measures or a combination of them:
  - Development of local business activities, not requiring extensive transportation of physical entities. The effectiveness of this measure, however, is questionable, as the potential markets for the local products are usually very limited.
  - Development of business activities for which unfavourable location is not necessarily a handicap. Tourism is such an activity, where geographical discontinuity may not be a problem but, on the contrary, in certain cases, a strong comparative advantage. The exclusive dependence of the region's development, however, on a single activity, such as tourism, is vulnerable to external factors and therefore risky.
  - Establishment of a communication network, where no discontinuity occurs. In this way, the regions will be able to attract or develop economic activities involving the production of intangible goods (financial services, computer software) locally, which, then, may be communicated to customers located elsewhere. The rapid development of Information and Communication Technologies (ICT) over the last years has made the third solution possible.

  The choice of the proper measure or combination of measures depends on the specific characteristics of the region given.

The estimation of a region's Basic Image for a given year gives a "snapshot" view of a region's development. A more interesting exercise however, would be to estimate the region's Basic Image for a number of years, to identify its respective trend and to study its changes. It must be noted that the way in which the Basic Image has been structured, allows the researcher to determine not only the changes in the region's Basic Image value, but also the causes of those changes. Going backwards from the Basic Image, through indicators, sub indicators, indices and sub indices to the variables, one can identify the real causes of the Basic Image changes. Hence, the Basic Image may prove a very useful managerial tool for both local authorities and business firms. The local authorities may use the Basic

Image in order to monitor the development of the various regions, get an early warning of any potential problems they may face and take the necessary measures to prevent them. The business firms on the other hand, may use the Basic Image in order to follow the development of various regions, assess their potential for future growth and take the proper location and investment decisions.

8. Conclusions and Suggestions for Further Research

A region's development depends on its ability to retain existing business activities and attract new ones. This ability depends on what we call the Image of a region and it is a measure expressing the region's current state of development and its future prospects. The paper introduced the concept of a region's Basic Image, developed a mathematical model for its estimation, applied the model to the case of the thirteen regions of Greece and presented the results. The Basic Image gives a "true" picture of a region's development, an early warning of any future problems. Furthermore, its structure allows a researcher to identify not only the changes in the Basic Image values, but also the causes of those changes and, hence, take the necessary measures. Consequently, the Basic Image may prove to be a very useful managerial tool, which can help the authorities to improve the region's attractiveness and future prospects of development. The application results seem logical and expected. They show that the proposed model expresses a region's attractiveness in a realistic way, in the sense that it quantifies the region's appeal to the full range of its existing and potential business units and employees.

A special note, however, should be made for the island regions. As it has been mentioned in the previous section, all islands have negative Basic Image values expressing their difficulty in attracting economic activities involving heavy transportation of raw materials and finished goods. However, a number of alternative measures for overcoming this problem have been proposed, opening up new prospects for those regions' growth. Hence, the Basic Image should be redefined so as to take into account these prospects and the effects of measures taken in this direction. This redefinition of a region's basic image is one of the main areas for further research.

The Basic Image, as defined so far, has left out a number of important variables, endogenous or exogenous. Hence, another area of further research would be to redefine a region's Basic Image, so as to include some of those variables. A first set of such variables may be those related to the region's environmental conditions and could define a third indicator, which may be referred to as Environmental Indicator. A second set of variables may be those related to the prevailing socio-economic environment in which the region operates and could define a fourth indicator, which may be referred to as General Economic Climate Indicator. As it has been seen in section 5, in the case of three or four indicator the most appropriate elementary catastrophes are the Swallowtail and the Butterfly catastrophes respectively. Hence, our task will be to examine how those elementary catastrophes may be used to model the enriched Basic Image.

References

Aerts, D., Czachor, M., Gabora, L., Kuna, M., Posiewnik, A., Pykacz, J., Syty, M., 2003, 'Quantum morphogenesis: a variation on Thom's catastrophe theory', *Physical Review E,* 67, pp. 1-13.

Angelis, V. and Dimopoulou, M., 1991, 'A Decision Support System for the Location of Industrial Units. Theoretical framework and Applications', *Studies in Locational Analysis*, 3, pp. 83-99.

Ashworth, C. J. and Voogd, H., 1988, 'Marketing the City: Concepts, Processes and Dutch Applications', *Town Planning Review*, 59, pp. 65-79.

Ashworth, C. J. and Voogd, H., 1990, *Selling the City*, Belhaven Press, London.

Boschma, R. and Lambooy, J., 1999, 'Why do old industrial regions decline? An exploration of potential adjustment strategies', *European RSA Congress, Proceedings,* August 23-27, Dublin, Ireland.

Bramwell, B. and Rawding, L., 1996, 'Tourism Marketing Images of Industrial Cities', *Annals of Tourism Research*, 23, 1, pp. 201-221.

Bristow, G., 2005, 'Everyone's a winner: problematising the discourse of regional competitiveness', *Journal of Economic Geography*, 5, pp. 285-304.

Bristow, G., 2010, *Critical Reflections on Regional Competitiveness: Theory, policy and practice*, Routledge, London.

Burgess, J. A., 1982, 'Selling places: Environmental images for the Executive', *Regional Studies*, 16, 1, pp. 1-17.

Camagni, R., 2002, 'On the concept of territorial competitiveness sound or misleading?', *Urban Studies*, 39, pp. 2395-2411.

Capello, R., 2007, *Regional Economics,* Routledge, London.

Capello, R., Camagni, R., Chizzolini, B., Fratesi, U., 2008, *Modelling Regional Scenarios for the Enlarged Europe*, Springer, Berlin.

Dichter, E., 1985, 'What's An Image', *The Journal of Consumer Marketing*, 2, pp. 75-81.

Dou, E. and Ghose, W., 2006, 'A dynamic nonlinear model of online retail competition using cusp catastrophe theory', *Journal of Business Research*, 59, pp. 838-848.

Dowling, G. R., 1998, 'Measuring Corporate Images: A Review of Alternative Approaches', *Journal of Business Research*, 17, pp. 27-37.

Eurostat Regional Yearbook 2008, *Office for Official Publications of the European Communities*, Luxemburg.

Freudenberg, M., 2003, 'Composite Indicators of Country Performance: A Critical Assessment, OECD Science', *Technology and Industry Working Papers*, 2003/16, OECD Publishing.

Gartner, W., 1993, 'Image Formation Process', *Journal of Travel and Tourism Marketing*, 2, pp. 191-215.

Gilmore, R., 1993, *Catastrophe Theory for scientists and engineers.* Wiley, New York.

Hammond, A., Adriaanse, A., Rodenburg, E., Bryant, D., Woodwart, R., 1995, 'Environmental Indicators: A Systematic Appraoch to Measuring and Reporting on

Environmental Policy Performance in the Context of Sustainable Development', *World Resources Institute*.

Huang, Y. K., 2008, 'The study of customer segmentation examined by catastrophe model', in Olya, M. and Uda, R. (eds) *Towards Sustainable Society on Ubiquitous Networks,* Springer, Boston, pp. 37-48.

Holland, S., 1976, *Capital versus the Regions,* Macmillan, London.

Kotler, P., Haider, D. H., Irving, R., 1993, *Marketing Places: Attracting Investment, Industry and Tourism to Cities, States and Nations*, Free Press, New York.

Kotler, P., Asplund, C., Rein, I., Haider, D. H., 1999, *Marketing Places Europe*, Prentice Hall, London.

Lever, W. F. and Turok, I., 1999, 'Competitive cities: introduction to the review', *Urban Studies*, 36, pp. 791-793.

Llewellyn, J., 1996, 'Tackling Europe's competitiveness', *Oxford Review of Economic Policy,* 12, pp. 87-96.

Lovering, J., 2001, 'The Coming Regional Crisis (And How To Avoid It)', *Regional Studies*, 35, 4, pp. 349-354.

Malecki, E., 2004, 'Jockeying for Position: What It Means and Why It Matters to Regional Development Policy When Places Compete', *Regional Studies*, 38, 9, pp. 1101-1120.

Markusen, A., 1996, 'Sticky places in slippery space: a typology of industrial districts', *Economic Geography*, 72, pp. 293-313.

Pike, A., Rodriguez-Pose, A., Tomaney, J., 2006, *Local and Regional Development*, Routledge, London.

Poston, T. and Stewart, I., 1996, *Catastrophe Theory and its Applications,* Dover, New York.

Rosser, J. B., 2007, 'The rise and fall of catastrophe theory applications in economics: Was the baby thrown out with bathwater?', *Journal of Economic Dynamics and Control,* 10, pp. 3255-3280.

Slavova, T., 2008, 'A rank order and efficiency evaluation of the EU regions in a social framework', *Empirica*, 35, pp. 339-367.

Smerz, K. E. and Guastello, S. J., 2008, 'Cusp catastrophe model for binge drinking in college population', *Nonlinear Dynamics, Psychology and Life Sciences*, 12, pp. 205-224.

Smith, M., Lancioni, R. A., Oliva, T. A., 2005, 'The effects of management inertia on the supply chain performance of produce-to stock firms', *Industrial Marketing Management*, 24, pp. 614-628.

Stilwell, F., 1972, *Regional Economic Policy,* Macmillan, London.

Stimson, R. J., Stough, R. R., Roberts, B. H., 2006, *Regional Economic Development,* Springer, Berlin.

Storper, M., 1997, *The Regional World*, Guilford Press, New York.

Thom, R., 1975, *Structural Stability and Morphogenesis: An Outline of a General Theory of Models.* Addison-Wesley, Reading M.A.

Van der Mass, H. L. J., Kolsteib, R., Van der Pligt, J., 2003, 'Sudden transitions in attitudes', *Sociological Methods and Research*, 32, pp. 395-417.

Wales, D. J., 2001, 'A microscopic basis for the global appearance of energy landscapes', *Science,* 293, pp. 602-604.

Zeeman, E. C., 1973, 'Applications of Catastrophe Theory', *International Conference on Manifolds*, Tokyo, Japan, pp. 11-26.

The Impact of Mining and Services Industries on the Structural Change of Australia

Ikhlaas Gurrib[1]

Abstract

Both services and mining industries gained relatively more importance in the structural changes occurring, particularly in recent years, where the rate of structural change has increased. The study aims at defragmenting the mining and services industries, and analyses their impacts on structural change. VAR model, where activity is measured in terms of output, shows social and business services to have more forecasting abilities than other variables. On the hand, VAR model where activity is measured in terms of investment shows the mining industry to have relatively less Mean Absolute Errors forecasts. Results deteriorate as one moves from one step ahead to twelve step ahead forecasts, suggesting longer periods of one step ahead forecasting should be used to avoid cyclical fluctuations in activity variables such as output or investment for industries.

Keywords: Australia, mining, services, VAR, structural change

JEL classification: L16, N15, O11

1. Introduction and Background to Study

Structural change and economic development have always been part of the economics literature since Smith (1776). More recent research includes Silva and Teixeira (2008) and Krüger (2008) who survey previous literature on structural change at different aggregation levels. Structural change plays a critical role in the process of economic growth in raising standards of living for any economy. In the Australian context, the Productivity Commission is chartered by the government to facilitate adjustment within the economy by those individuals, firms and regions affected by economic change, by providing accurate information about structural change in Australia since the 1970s (Productivity Commission, 1998).

The framework of this study is based upon an extension of the traditional three-sector hypothesis, i.e. primary (agriculture and mining), secondary (manufacturing and

[1] Fellow of Financial Institute Services of Australasia (FINSIA); *Assistant Professor* at the Canadian University of Dubai, Sheikh Zayed Road, Dubai, UAE, PO Box 117781, ikhlaas@cud.ac.ae

construction), and tertiary sectors (private services) mostly affect economic growth, measured in terms of employment or output. Earlier contributions supporting this hypothesis are Fisher (1952), Wolfe (1955) and Fourastié (1969). Neoclassical growth as Solow (1956) as well as new growth theory like Lucas (1988) disregard such hypothesis by assuming sectoral composition to be constant such that there is no structural change. More recent research have tried to complement structural change with formal growth theory (e.g. Bonatti and Felice, 2008), but in these models causality between economic growth to structural change was simply assumed.

The aim of this study is to look at the structural change in Australia since the 1960s, with particular reference to the impact of the service and mining industries, where the country has experienced three mining booms over the 50 year period. The traditional three-sector hypothesis mentioned above will be extended to a four-sector hypothesis based upon using the United Nations (UN) International Standard Industrial Classification (ISIC) for agriculture, mining and market and non-market services industries. The study is broken down further to each Australian State level to allow an analysis of the trend in activity flowing from each state. This latter part of the study, which looks at the forecasting ability of activity in the service and mining industries, is beneficial especially due to recent regulatory barriers being lifted in Western Australia to allow for uranium mining and increasing iron ore trade with China. The rest of the paper looks at some literature review pertinent to the study, followed by some emphasis on data and research methodologies. Research findings are presented before giving some concluding remarks.

2. Literature Review

In an attempt not to duplicate the already replete literature on structural change, only those relevant to the study are reviewed. Meckl (2002) and Kongsamut et al. (2001) show that balanced growth is possible consistent with significant changes in the sectoral composition. Structural change theory literature is broadly divided into two arms, i.e. preference changes in demand and sectoral specific productivity gains. For the former arm, societies consume products and services according to their personal preferences and hence spend their income in a certain distribution for goods of the primary, secondary and tertiary sector. The distribution varies with increasing income. Supported by Engel's law, the share of basic needs as primary products decreases as income increases. Fourastié (1969) links this transfer of preferences with a hierarchy of needs associated with different saturation levels for the goods of the three sectors. He even postulates that the demand for goods and services in the tertiary goods will never drench. Rowthorn and Ramaswami (1999) support these views by providing evidence of different income elasticities among the three core sectors, where income elasticity in the service sector is above one. For the latter arm, the productivity hypothesis indicates that different levels of technical progress among sectors will shift the sector shares. Technical progress results in better production process, which means less input of employment is needed to produce one unit of output, which in turn reduces prices of goods and services produced. Baumol (1967) and Meckl (2002) provide evidence that different productivity increases are responsible for economic development. Fourastié (1969) argue preference changes in demand is the driving force of structural change. Dietrich and Kruger's (2008) analysis of Germany's structural change provides additional support to the demand side perspective. No matter what perspective is taken, higher growth rates of any economy GDP in one period should lead, via rising income and changing demand to more structural change or to a higher speed of structural change in the following periods, measured either in terms of output or employment of core economic sectors.

Echevarria (1997) investigates the relations between sectoral structures and economic growth, and find a hump shaped relationship between economic growth (measured as growth rate of Real GDP per capita) and sectoral composition. The author provides views that sectoral composition plays a major role for the growth rate of GDP. Aiginger (2001) studies the relationship between economic dynamics and structural change of production by using the norm of absolute values (NAV) and a disaggregation of 23 sectors or 99 industries, and finds structural change has a more significant impact on growth than vice versa. Productivity Commission (1998) looks at the manufacturing sector and its structural change for 15 OECD countries, and finds considerable differences in both sector and industry specific growth rates. This paper contributes to existing literature by defragmenting the mining and services industries of Australia, and analyses how they affect the structural change in the Australian economy.

3. Data and Research Methodology

For the purpose of this study, structural change is measured by using a traditional method called the coefficient of compositional structural change or structural change index (SCI)[1]. The SCI for output can be defined as half the sum of the absolute value of the differences in value-added shares over a time period, and is calculated as follows:

$$SCI = \frac{1}{2}\sum_{t=1}^{N}\left(\left|\phi_{i,t} - \phi_{i,t-x}\right|\right) \tag{1}$$

where, and, $\phi_{i,t}$ and $\phi_{i,t-x}$ represent each industry's share of total value-added at time t and t-x. Absolute values guarantee negative and positive changes in industry shares do not annul each other when summed up across industries. The amount of structural change equals exactly the share of the movements of the sectors as a percentage of the whole economy. If the structure remains unchanged, the indicator is equal to zero and if all sectors change at its most, which means the whole economy has a total change, then the index is equal to unity (Dietrich, 2009).

[1] Earlier literature also called it the Norm of Absolute Values (NAV) Michaely Index (Michaely, 1962) or Stoikov Index (Stoikov, 1966).

Before conducting any structural change analysis, it is important that the data are comparable. Further, the SCI results are sensitive to the level of industry aggregation, price movements and time period, which need consideration before interpretation. Information on value-added, employment and investment for Australia within the mining sector was sourced from the International Economic Data Bank (IEDB) and cross checked out with the Australian Bureau of Statistics (ABS). While SCIs can be biased by the level of industry aggregation, results from De Laine, Lee and Woodridge (1997) find a similar result on Australian employment data for 11 and 54 industry disaggregations. To allow the possibility of future comparative studies among different countries, the United Nations (UN) International Standard Industrial Classification (ISIC) is adopted for agriculture, mining and market and non-market services industries. More importantly, five subdivision level industries (ISIC C 10-14) have been implemented for the Mining Industry[2], Business Services subdivided into 5 sublevels (ISIC K 70-74), and Social Services into 3 sublevels (ISIC 75, 80 and 85)[3].

Due to the sensitiveness of price movements over time, the distinction between using current and constant prices, when examining the changing share of output/investment accounted for by various industries, becomes important. Resources are transferred from one industry to another due to changes in demand and supply, which in turn are influenced by changes in relative prices and non-prices factors like quality, competition and government policies. By using current prices, SCIs captures the impact of all changes on the total value of goods and services produced (Productivity Commission, 1998). Current price data are preferred to constant price data, due to the sensitiveness of the base year when using constant prices, which requires rebasing of data series[4]. In choosing the time period for comparison, year-to-year comparisons have a tendency to exhibit significant variability which suggests the influence of temporary and cyclical fluctuations in activity. To ensure the effect of longer term changes in output and investment shares between industries, five years window periods were chosen. Five is substituted for x in equation 1 above. In line with the Reserve Bank of Australia (RBA) which identified three mining booms since the 1960s (see BIS 2010) the data is set from 1960 to 2010[5]. When conducting the same SCI analysis at State levels, Gross State Product (GSP) as opposed to Gross Domestic Product (GDP) is used. Due to little or no activity in the mining sector of New South Wales (NSW), it is combined with Australian Capital Territory (ACT).

[2] See, ISIC (2010) for ISIC classification codes.

[3] Only Business Services and Social Services are analysed due to their relative significance over other kinds of services.

[4] Clark, Geer and Underhill (1996) provide a good example of the need to rebase constant prices in the Australian agricultural sector.

[5] These booms were the 1960s/early 1970s mining and energy boom, the late 1970s/early 1980s energy boom, and the current (2010) mineral and energy boom.

4. Research Findings

4.1 Activity by Industry

Compared to the 19th century where Australia's agriculture accounted for about one third of output, the 20th century witnessed significant activity for the manufacturing and service sector. As it can be observed in Figure 1 below, by 1950, employment in the manufacturing sector rose to roughly 25 percent, compared to 15 percent at the start of the century. While the share of manufacturing has continued being dropped since the beginning of the mining boom in 1960s, service industries have emerged strongly to reach above 80 percent in 2010. A possible explanation to the decline of manufacturing and surge of service can be explained since in the 1950s services were closely linked to manufacturing, with whole trade and transport supporting the production and distribution of secondary goods. From then, the share of distribution services has dropped continually with the declining importance of manufacturing. On the other hand, the fastest growing service industries in recent years have been the business services, the social services and the personal services, as observed in Figure 2. Although not reported here, it is important to note that service industries tend to be more labour intensive than other sectors, with 85 percent of the Australian work force, but only 70 percent of investment.

More relevant to the study is the share of nominal output of the mining industry which has witnessed considerable volatility since the 1960s, but has drifted upwards to around 8 percent in the 2009-2010 period as shown in Figure 3. The first two peaks in the Investment Figure can be attributed to the early 1970s and early 1980s mining booms as reviewed by BIS (2010). Interestingly, investment in the mining industry in the 2009-2010 period has well exceeded those from the two previous mining booms, reaching 19 percent. The mining booms in recent years have also had positive effects on other industries in Australia. Figure 2 supports strong demand for mining related construction. Further, there has also been a shift in the composition of the manufacturing industry towards mining related manufacturing, and away from import competing manufactures. In spite of the increase in investment share, employment in the mining sector remains subdued, reflecting more concentration on capital intensive methods compared to the service sector.

While the service sector dominates the share of output and employment, Australia's exports are still led by the mining industry as shown in Figure 4, where mining exports were around half of the country's export revenues. It can also be observed that the shares of manufacturing and services exports rose through the 1990s, but declined in recent years. This can be attributed to the strengthening of commodity prices, and the demand of China for more Australian iron ore and other minerals.

4.2 Activity by State

Since activity is based on employment, a look at the Australian population movement across states is vital. From Figure 5, there has been a shift from South Eastern states like

Victoria, New South Wales (NSW), Australian Capital Territory (ACT), towards Western Australia (WA) and Queensland (QLD). The population in Victoria was the highest in the 1850s due to the gold rush, but since then it has dropped significantly due to the 1890 depression and migration towards Western Australia where mining booms occurred (Blainey, 1963). Compared to the 19th century, it can be seen that changes in the population shares towards WA and QLD have occurred smoothly with current mining trends being less labour intensive than their predecessors. Western Australia and Queensland have grown relatively strong, with their combined share of output and employment rising to over 30 percent recently, compared to 20 percent in the 1960s. Sturdy population growth has contributed to faster output growth in these states relative to the national average over the past two decades. These resource-rich states contributed more to the nation's investment as shown in Table 1.

4.3 Structural Change

Based on equation (1), SCI is calculated to measure the change of the share of different industries in total nominal output, employment and investment over five years. The same calculation is replicated to measure the change in economic activity across Australian states. These different indices help us to identify periods of high rates of structural change and periods of relative stability over the past 50 years. Results are reported in Figure 6.

Findings suggest that structural change was high from the late 1960s through to the late 1970s, from the late 1980s through to mid 1990s, and more so in recent times. The earlier structural changes can be attributed to investment booms in mining in the late 1960s and business services in the late 1980s. These led to rising output shares in these industries in the coming years. Resources have moved from agriculture and manufacturing towards more appealing industries, i.e. mining and services. In early 1990s, increase in structural change by industry and state were associated mostly with the recession, in which there was a sharp contraction in manufacturing output and employment. Although not reported here, Victoria and South Australia were the most affected due to their relatively large manufacturing industries. The mid 1960s and early 200s both experienced periods of solid economic growth and price stability.

The rate of structural change appears to have increased in recent years, partly led by the mining boom in WA and QLD. This can be noticed in the measure of nominal output, reflecting the sharp rises in commodity prices. The measure of investment across states also supports the relative increases in rate of structural change, although by industry, the measure remains well below the peak of the late 1980s[6]. Importantly too, structural change in employment has not picked up to the same extent as investment. This is backed by the fact that the mining industry is still a relatively more capital intensive industry, hence, a lower share of the total Australian work force. Although not reported here, changes in real output SCI remained low compared to investment SCI, suggesting the unavoidable lag

[6] This period was associated with the commercial property boom in Australia.

between investment and real output[7]. Moreover, the rate of structural change across all states has been the highest since the 1960s on most measures. This reflects that the current mining boom is larger and more geographically concentrated than the previous boom in the late 1960s and early 1980s. Nonetheless, the strong growth in WA and QLD has been quite broad across industries.

4.4 Factors driving Structural Change

4.4.1 Economic Development

Australian policy changes over the last 50 years are one of the major causes of structural change. These policy changes include deregulation of a broad category of service industries and a reduction in trade protectionist barriers imposed on goods producing industries. Forsyth (2000) shows that these policies have increased the relative share of services in Australia by enhancing competition. Particularly, the deregulation of the finance industry in the 1980s and the start of compulsory superannuation gave a big boost to the growth of funds management and banking. Reforms also include cuts in the protectionist trade barriers previously imposed in the manufacturing industry. Since the early 1970s, these barriers were gradually wound back. Productivity Commission (2010) support that the effective rate of assistance to manufacturing has dropped from 35 to 5 percent in the last 40 years. The agricultural sector has also witnessed drops in tax concessions and subsidies, which were particularly allocated during droughts or low commodity prices. Lower trade barriers have given households access to lower imported goods, hence benefiting from the comparative advantage of East Asia manufacturing goods. This is consistent to Dwyer and Fabo (2001) who also find that the removal of these barriers have led to the manufacturing sector being increasingly export oriented than before.

4.4.2 Increase in demand for services

With increases in Australia real income per capita, consumer demand for services has increased the share of services in the Australian economy. In fact, the share of consumption spent on services has risen from 40 to 60 percent over the last 50 years, reflected by upward trends in educational, health, and recreational expenditures. Similar trends of increasing share of services in output can be found in other economies as shown in Connolly and Lewis (2010). Most importantly, the increase of the share of services has coincided with increasing labour share as well. For instance, the increase in the proportion of working couples has led to more demand for services such as child care, pre schooling, house maintenance and aged care. These services were previously classified as households and thus not production. Similarly, more health services were demanded as a result of increase

[7] Real output is based upon choosing a base year for constant prices. The choice of base year affects the results as supported by Clark, Geer and Underhill (1996) and is not further pursued in that paper.

in prolonged existence. The service industries provide much more flexible working hours than traditional manufacturing or agricultural production, hence, its preference for the working class population.

4.4.3 East Asia Emerging Markets

Although Australia manufacturing and agricultural shares of the total output produced have decreased partly because of the shift towards services and mining industries, this was also due to the emergence of East Asia as fierce competitors of manufacturing goods globally. Having one of the lowest regional labour costs, East Asian economies' share of global manufacturing have more than doubled since the last 40 years. This process started by Japan in the 1960s, followed more recently by newly industrialized countries like China. While Australia share of global manufacturing fell consequently, China and Japan's need for mining resources as part of their production process meant a much more significant rise in Australia's exports, particularly in the iron ore and coal exports. On top of economic reforms mentioned earlier, adequate Australian government policies have also played a vital role in the rise of its mining industry. For example, to encourage mining activity, an embargo imposed on iron ore exports was lifted in the 1960s, together with the introduction of bulk carriers. This kept transportation costs at lower levels than previously held. A more recent important policy change is the removal of the 'no mining policy' in Western Australia in 2008 on the uranium industry. While almost 100 percent of the uranium is exported, such policy changes added value to the structural change of the economy towards the mining industry.

4.4 Robustness testing

From the above analysis so far, it can be observed that there has been structural changes from the traditional agricultural and manufacturing industries towards the services and mining industries. The increase in the rate of structural changes can also be found in more recent periods than earlier ones. However, it is also important to understand whether the mining or the services industry is leading those structural shifts in the Australian economy. While existing literature is full of granger causality testing, the latter allows only for *one step ahead* prediction in bivariate settings[8]. If all the above variables are considered simultaneously, there might be more benefits in the predictive power of the resultant equation. A vector autoregressive (VAR) model provides a solution to both of these issues[9] and can be generalized in matrix notation[10] as follows:

[8] See, Erdil and Yetkiner (2004) for review of granger causality tests between structural change and other variables.

[9] See, Webb (1988) on the forecasting reliabilities of VAR models.

[10] While (2) is for a 2 variable multivariate setting with two variables y_t and z_t, it can be expanded to accommodate for all the different independent variables used in traditional Granger causality tests.

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a^y \\ a^z \end{bmatrix} = \begin{bmatrix} \beta^y & \gamma^y \\ \beta^z & \gamma^z \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} = \begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix} \tag{2}$$

where $a'\ s$ and $\beta'\ s$ are parameters. The epsilons are white noise, i.e,
E (ε_t^i) = 0,
Var (ε_t^i) = σ^2,
Covar (ε_t^i , ε_s^i) = 0, where i, j = y, z and i ≠ j, t ≠ s.

To ensure that the vector of endogenous variables are stationary in our VAR models, stationarity testing using Augmented Dickey Fuller (ADF) tests is carried out to avoid spurious regression. All variables in the VAR models are found to be stationary in their levels. The optimal lag length is obtained by minimizing the Schwarz Information Criteria for the two VAR models and results are provided in Table 2. While there is no exact date for the start of the recent boom in Australia, 2007 is chosen as the time where the mining boom severely tested the productive capacity and flexibility of the Australian economy as per BIS (2010). This is supported by earlier findings in Figure 6, where SCI indices for both nominal output and investment increased significantly, at states and industry levels. Similar to Gurrib and Ahmad (2010), each VAR model is estimated using data through to January 2007 and forecasts were computed for each month through September 2008. The forecasts for February 2007 were compared with actual data and the resulting one-step ahead errors were recorded; forecasts for February were used for two-step ahead errors; and similarly, forecast errors up to twelve steps ahead were calculated. The process was then updated for one month, with the model estimated through February 2007 and forecasts made until February 2008. This process or estimation and forecasting was repeated each month through July 2010. Results of the forecast errors are reported in Table 3.

Findings support that both VAR models performed better than a naïve model with a no change forecast. In fact, values for Theil inequality statistics were less than unity, indicating that the VAR models forecasted outperformed a naïve no change forecast. The VAR 1 model which includes the structural change index at nominal output level ($SCIO_t$), shows that social services and business services had more significant forecasting abilities on Australia's nominal output due to their lower Mean Absolute Error (MAE) forecasts than in the VAR 2 model. On the other hand, the VAR 2 model which includes the structural change index at investment level ($SCII_t$), shows the mining industry had more forecasting ability than the services industry due to its lower MAE than VAR 1 model. These two findings support earlier results that social services and business services had some significant impact on the volatility of structural change over the last 50 years. Results also support the mining industry has gained more momentum in recent years due to the recent ongoing boom. However, in most instances, the relative accuracy decreases with the forecast window as we move from one step to twelve step ahead forecasts, suggesting that a longer time period forecast (say three years) compared to one month forecast is of more value, due to the influence of temporary and cyclical fluctuations in activity which can obscure the effects of longer term changes in output and investment shares between the two industries.

5. Conclusion

Structural change had a tendency to witness volatility over the last 50 years in Australia. Noticeable shifts from the traditional agricultural and manufacturing industries have been observed, with the services and mining industries capturing most of those structural shifts in the economy. These movements occurred not only due to Australia policy changes in regards of deregulation and removal of protectionism, rising demand for services and more trade with emerging markets, but also due to three mining booms that occurred in the last fifty years. The rate of structural change has increased, particularly in recent times, due to the ongoing mining boom where states like Western Australia and Queensland had larger shares of output and investment relative to others. While one Vector Autoregressive model supports social and business services to have some forecasting abilities when structural change is measured in terms of nominal output, the other VAR model suggests the mining industry have lower Mean Absolute Error forecasts, when structural change is measured in terms of activity in investment. Forecasts tend to deteriorate as data are regressed from one to twelve step ahead months. This suggests a need of longer forecast period for each step to avoid cyclical fluctuations when measuring structural change.

References

Aiginger, K., 2001, 'Speed of Change and Growth of Manufacturing, Peneder, M., Aiginger, K. and Marterbauer, M., *Structural Change and Economic Growth*, Austrian Institute of Economic Research, pp. 53-86.

Baumol, W. J., 1967, 'Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis', *The American Economic Review*, 57, pp. 415–426.

BIS, 2010, 'Mining Booms and the Australian Economy', addressed by Mr Ric Battellino, Deputy Governor of the Reserve Bank of Australia, to the Sydney Institute, Sydney, 23rd February 2010.

Blainey G., 1963, 'The Rush that Never Ended', Melbourne University Press, Melbourne.

Bonatti, L. and Felice, G., 2008, 'Endogenous Growth and Changing Sectoral Composition in Advanced Economies', *Structural Change and Economic Dynamics*, 19, pp. 109-131.

Clark, C., Geer, T. and Underhill, B., 1996, 'The Changing of Australian Manufacturing', *Industry Commission Staff Information Paper*, AGPS, Canberra.

Connolly, E. and Lewis, J., 2010, 'Structural Change in the Australian Economy', *RBA Bulletin*, September Quarter.

De Laine, C., Lee, K. and Woodridge, G., 1997, 'Microeconomic Reform and Structural Change in Employment', *Industry Commission Staff Research Paper* no. 6, AGPS, Canberra.

Dietrich, A., 2009, 'Does Growth Cause Structural Change, or Is it the Other Way Round? A Dynamic Panel Data Analyses for Seven OECD Countries', *Jena Economic Research Papers* 2009-034, Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

Dietrich, A., Krüger, J. J., 2008, 'Long-Run Sectoral Development – Time Series Evidence for the German Economy', *Jena Economic Research Papers*, no. 013/2008.

Dwyer J. and Fabo, J., 2001, 'The Manufacturing Sector: Adapting to Structural Change', *RBA Bulletin*, March, pp. 1-9.

Echevarria, C., 1997, 'Changes in Sectoral Composition Associated with Economic Growth', *International Economic Review*, 38, pp. 431-452.

Erdil, E. and Yetkiner, I. H., 2004, 'A Panel Data Approach for Income-Health Causality', FNU-47.

Fisher, A.G.B., 1952, 'A Note on Tertiary Production', *Economic Journal*, 62, pp. 820 834.

Fourastié, J., 1969, 'Le Grand Espoir du XXe Siècle: Progrès Technique, Progrès Économique, Progrès Social', 1949, Paris: Presses Universitaires de France.

Gurrib, I. and Ahmad S., 2010, 'Saudi Arabia's Inflation Agenda: A Vector Autoregressive Framework', *International Journal of Trade, Economics and Finance*, 1, 1, Singapore.

ISIC, 2010, 'ISIC Classification codes', http://www.top500.de/lexikon/nace_code_c.php, (accessed 10/10/2010).

Kongsamut, P., Rebelo, S., and Xie, D., 2001, 'Beyond Balanced Growth', *Review of Economic Studies*, 68, pp. 869-882.

Krüger, J. J., 2008, 'Productivity and Structural Change : A Review of the Literature', *Journal of Economic Surveys*, 22, 2, pp. 330-363.

Lucas, R. E., 1988, 'On The Mechanics of Economic Development', *Journal of Monetary Economics*, 22, pp. 3-42.

Meckl, J., 2002, 'Structural Change and Generalized Balanced Growth', *Journal of Economics*, 77, 3, 241-266.

Michaely, M., 1962, 'Concentration in International Trade', Amsterdam, North Holland.

Productivity Commission, 1998, 'Aspects of Structural Change in Australia', *Research Paper*, December.

Rowthorn, R., Ramaswami, 1999, 'Growth, Trade and Deindustrialization', *IMF Staff Papers*, 46, 1, pp. 18-40.

Silva, E. G. and Teixeira, A. A. C., 2008, 'Surveying Structural Change: Seminal Contributions and a Bibliometric Account', *Structural Change and Economic Dynamics*, 19, pp. 273-300.

Smith, A., 1776, 'An Inquiry into the Nature and Causes of the Wealth of Nations', Oxford, Clarendon Press, Campbell, R.H., Skinner, A.S., Todd, W.B. (Eds.)

Solow, R. M., 1956, 'A Contribution to the Theory of Economic Growth', *The Quarterly Journal of Economics*, 70, 1, pp. 65-94.

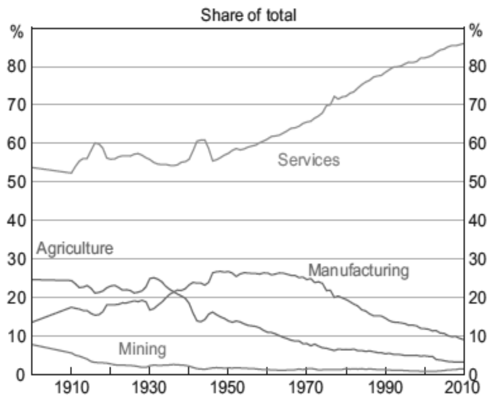
Stoikov, V., 1966, 'Some Determinants of the Level of Frictional Unemployment: A Comparative Study', *International Labour Review*, 93, pp. 530-549.

Webb, R. H., 1988, 'Commodity Prices as Predictors of Aggregate Price Change', *Economic Review*, Federal Reserve Bank of Richmond, pp. 3-11.

Wolfe, M., 1955, 'The Concept of Economic Sectors', *Quarterly Journal of Economics*, 69, pp. 402-420.

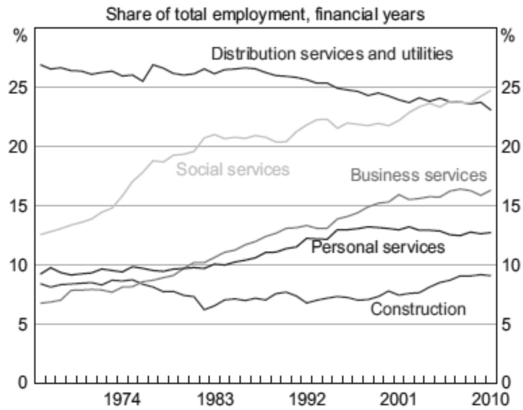
Appendix

Figure 1:
Employment by Industry



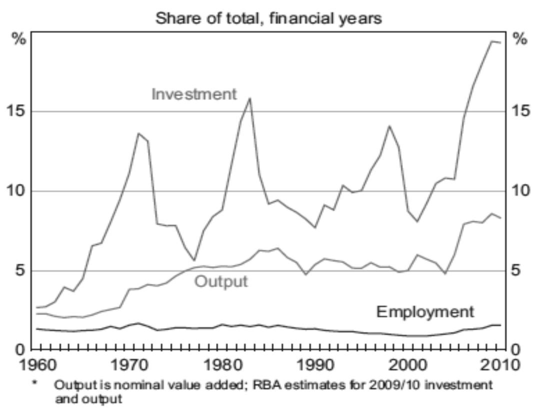
Source: ABS and RBA.

Figure 2:
Employment in the Service



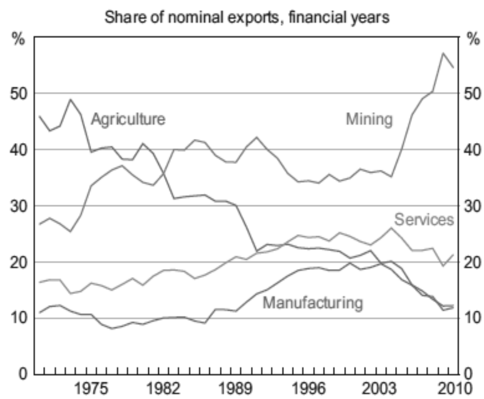
Source: ABS and RBA.

Figure 3:
Activity in the Mining Industry



Source: ABS and RBA.

Figure 4:
Australian Exports by Industry



Source: ABS and RBA.

Figure 5: Australian Population by State

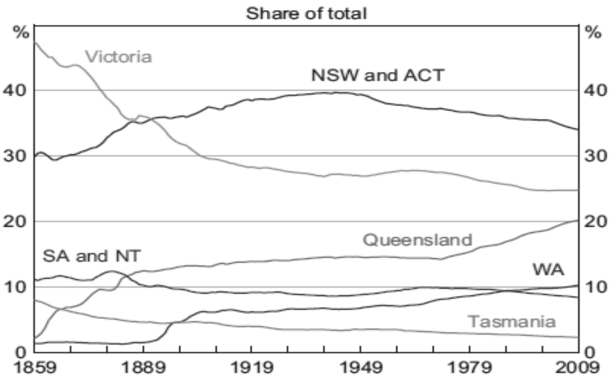
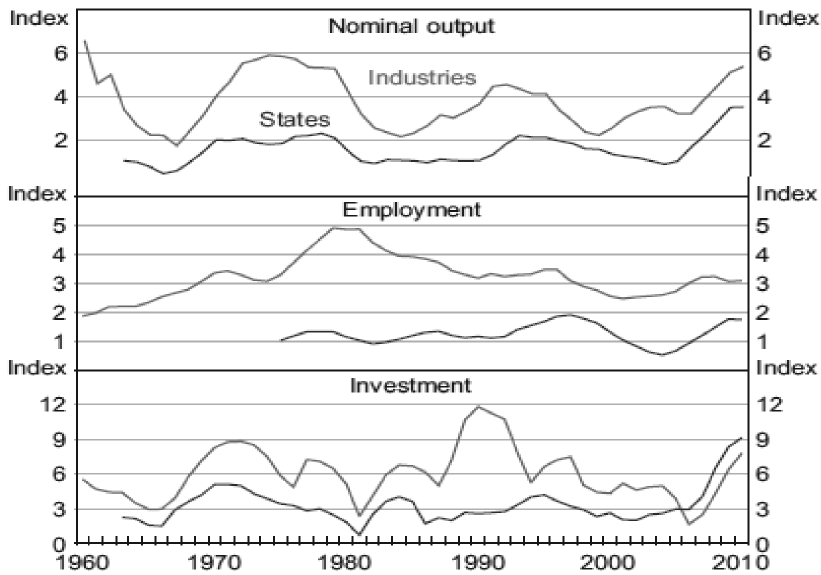


Figure 6: Structural Change Indices



Note: SCIs are calculated as half the sum of the absolute five-year change in five-year average industry shares or state shares. Data is set from 1960 to 2010. Output refers to value added by industry and gross state product (GSP). Investment includes business and public investment, but excludes dwelling investment and cultivated biological resources.

Table 1: Activity Share of States (per cent)

	NSW and ACT	Queensland	SA and NT	Tasmania	Victoria	WA
Output[(a)]						
– 1960s	38	13	9	3	32	6
– 1980s	35	15	9	3	30	9
– 2000s	35	18	8	2	25	12
Employment						
– 1960s[(b)]	38	14	10	3	28	8
– 1980s	36	16	10	3	27	9
– 2000s	35	20	8	2	25	10
Investment						
– 1960s	39	13	9	4	28	7
– 1980s	36	18	9	2	24	10
– 2000s	32	21	8	2	24	14

(a) Nominal gross state product
(b) 1966–1969

Source: ABS and RBA.

Table 2: Lag length of VAR models

This table reports the optimal lags for VAR models. VAR 1 model includes the structural change index ($SCIO_t$), mining industry activity variable (MIN_t), business services ($BSERV_t$) and social services ($SSERV_t$). VAR 2 model includes the structural change index ($SCII_t$), mining industry activity variable (MIN_t), business services ($BSERV_t$) and social services ($SSERV_t$). ($SCIO_t$) measures the change in the share of the different industries in total nominal output, while ($SCII_t$) measures the change in the share of the different industries in total investment. ($SCIO_t$) and ($SCII_t$) are calculated as half the sum of the absolute five year change in the five year average industry shares. Number of lags is optimized by minimizing the Schwarz Information Criteria (SIC).

VAR 1

Dependent variable	Independent Variable $SCIO_t$	MIN_t	$BSERV_t$	$SSERV_t$
$SCIO_t$	1	1	1	1
MIN_t	1	1	_	1
$BSERV_t$	1	2	1	1
$SSERV_t$	1	2	_	1

VAR 2

Dependent variable	Independent Variable $SCIIt$	MIN_t	$BSERV_t$	$SSERV_t$
$SCIIt$	1	1	1	1
MIN_t	2	1	2	_
$BSERV_t$	1	1	1	_
$SSERV_t$	1	1	1	1

Table 3: VAR Forecast Error Statistics

This table reports the Mean Absolute Error forecast (MAE) and Theil Inequality U statistics for VAR 1 and VAR 2 models. VAR 1 model includes the structural change index ($SCIO_t$), mining industry activity variable (MIN_t) and business services ($BSERV_t$) and social services ($SSERV_t$). VAR 2 model includes the structural change index ($SCII_t$), mining industry activity variable (MIN_t) and business services ($BSERV_t$) and social services ($SSERV_t$). $SCIO_t$ measures the change in the share of the different industries in total nominal output, while $SCII_t$ measures the change in the share of the different industries in total investment. $SCIO_t$ and $SCII_t$ are calculated as half the sum of the absolute five-year change in five-year average industry shares. Data is set from January 1960 to July 2010. These data are based on the International Standard Industrial Classification (ISIC). Results are reported for 1-step, 6-step and 12-step ahead forecasts.

VAR 1	**1-step**		**6-step**		**12-step**	
Variables	MAE	Theil U	MAE	Theil U	MAE	Theil U
SCIO	1.52	0.85	1.31	0.82	1.74	0.91
MIN	2.3	0.95	2.54	0.94	2.36	0.98
BSERV	1.76	0.82	1.81	0.84	1.92	0.89
SSERV	1.3	0.81	1.51	0.83	1.68	0.86

VAR 2	**1-step**		**6-step**		**12-step**	
Variables	MAE	Theil U	MAE	Theil U	MAE	Theil U
SCII	1.66	0.87	1.76	0.92	1.95	0.95
MIN	0.93	0.71	1.02	0.76	1.12	0.78
BSERV	2.12	0.94	2.27	0.95	2.43	0.98
SSERV	1.83	0.91	1.95	0.95	2.11	0.94

Note: MAE (Mean Absolute Error) has been annualized and reported in percentage points

The Dynamics of Relationship between exports and economic growth in India

P. K. Mishra

Abstract

In this era of open economy, nations are concerned with increasing the quality of life of their citizens. And, the quality of life mainly comes from the macro-economic prosperity. Thus, fast growth of gross domestic Product has become the most important objective of any economy. There are various approaches to achieve this target of which one strategy is to promote exports of the country. At this juncture, an important issue immediately breaks the minds of economists and researchers, that is, whether export promotion leads to higher economic growth or economic growth promotes exports. Thus, this paper is an attempt to reinvestigate the dynamics of the relationship between exports and economic growth for India over the period 1970 to 2009. Applying popular time series econometric techniques of cointegration and vector error correction estimation, the study provides the evidence of stationarity of time series variables, existence of long-run equilibrium relation between them, and finally, the rejection of export-led growth hypothesis for India by the Granger causality test based on vector error correction model estimation.

Keywords: Export-led Growth Hypothesis, India, GDP, Granger Causality, Error Correction Model

JEL Classification Code: C22, C32, F43

1. Introduction

In this era of open economy, nations are concerned with increasing the quality of life of their citizens. And, the quality of life mainly comes from the macro-economic prosperity. Thus, increasing Gross domestic Product is the most important objective of any economy. There are different approaches to achieve this target of which one possibility is to promote exports. At this juncture, an important issue immediately cracks the minds of economists and researchers, that is, whether export promotion leads to higher economic growth or economic growth promotes exports growth. Thus, economists came up with different views at different times and the literature puts forward a debate for researchers and policy

[1] Assistant Professor of Economics, Siksha O Anusandhan University, Khandagiri, Bhubaneswar, Orissa, India-751030, e-mail: pkmishra1974@gmail.com

makers since the last few decades. One school of thought argues in favour of export-led growth hypothesis while the other school advocates for growth-driven export hypothesis. In addition, the existing literature also provides the evidence that export promotion leads to economic growth and economic growth leads to export promotion, i.e., the bi-directional causality between exports and economic growth.

The Export-led growth hypothesis generally reflects the relationship between exports and economic growth. The proponents of such hypothesis argue that export promotion through policies such as export subsidies or exchange rate depreciation will increase economic growth. The substance of the neo-classical arguments underlying the export-led growth hypothesis is that competition in international markets promotes economies of scale and increases efficiency by concentrating resources in sectors in which the country has a comparative advantage. These positive externalities promote economic growth (Bhagwati, 1978; Balassa, 1978; Krueger, 1978; Feder, 1982; Krueger, 1990; Vohra, 2001; Ullah et al., 2009). On the contrary, the argument that economic growth promotes export growth stands on the idea that gains in productivity give rise to comparative advantages in certain sectors that lead naturally to export growth. Also, countries with high growth rates and relatively low absorption rates must necessarily export the excess output (Arnade and Vasavada, 1995; Fosu, 1996; Thornton, 1996; Henriques and Sadorsky, 1996; Sharma and Panagiotidis, 2005). In addition, some studies demonstrate that there exists a bi-directional relationship between these variables such that export causes economic growth and economic growth causes export (Dutt and Ghosh, 1994; Thornton, 1997; Shan and Sun, 1998a; Shan and Sun, 1998b; Khalafalla and Webb, 2001).

It is due to such contradicting evidences about the dynamic relation between exports and economic growth that many developing countries are still in dilemma whether to open up their economies to promote international trade or whether they should concentrate on economic activities that will promote international trade. Today, there has been much worldwide debate about Doha Development Agenda, Trade for Aid discussion, etc. and a good number of researchers and policy makers believe that developing countries can achieve economic growth through free market while others believe that developing countries should protect their industries from imported goods and promote their economic activities which will lead to the economic growth.

Now, it is believed that the rapid growth of China and India is mainly due to the expansion of their exports. "The success of China and India largely caused by both the export-led growth and access to technology through globalization" (Stiglitz, 2007). Exports imply access to the global market and permit increased production while trade encourages efficient allocation of resources, and trade contributes to economic growth by generating long-run gains (Easterly, 2007). Thus, India can be an interesting case study of the export and economic growth relationship.

It is against this backdrop that the paper attempts to revisit the issue of the relationship between growth of exports and economic growth in case of India for the period 1970 to 2009. The rest of the paper is organised as follows: Section 2 reviews the theoretical as well as empirical literature; Section 3 discusses the data and methodology; Section 4 makes the empirical analysis; and section 5 concludes.

2. Literature Review

The argument concerning the role of exports as one of the main deterministic factors of economic growth is not new. It goes back to the classical economic theories by Adam Smith and David Ricardo, who argued that international trade plays an important role in economic growth and that there are economic gains from specialisation.

The argument of the neo-classical economists is that competition in international market promotes economies of scale and increases efficiency by concentrating resources in sectors in which the country has a comparative advantage. These positive externalities promote economic growth.

These theoretical arguments regarding exports-economic growth nexus have been empirically verified by economists and researchers at different times. A number of studies including Jung and Marshall (1985), Chow (1987), Darrat (1987), Hsiao (1987), Bahmani-Oskooee et al (1991), Kugler (1991), Dodaro (1993), Van den Berg and Schmidt (1994), Greenaway and Sapsford (1994), and Islam (1998) have had adopted time series analysis for exploring the causal liaison between exports growth and output growth. Using the Granger (1969), Sims (1972) and Hsiao (1987) causality procedures, these studies failed to provide an unvarying conclusion about the export-led growth hypothesis. However, these time series studies were not free from disparagement. Although standard Granger or Sims tests are only valid if the original time series are not cointegrated, none of these studies checked the cointegrating properties of the time-series variables involved. When two or more time series variables are cointegrated, inferences based on traditional time-series modelling techniques will be misleading, as pointed out by Granger (1988), this is because traditional causality tests would miss some of the "forecastability", hence, reach incorrect conclusions about causality. Moreover, all the studies reviewed above used growth of Gross Domestic Product (GDP) and that of exports which are akin to first differencing and filter out long-run information. In order to alleviate such occurrences, cointegration and error-correction models have been recommended to combine the short-term as well as long-run information. Bahmani-Oskooee and Alse (1993) took all these issues into account and employed quarterly instead of annual data for the nine countries studied. The study found strong empirical support for two-way causality between exports growth and GDP growth in eight out of nine countries.

Darrat (1986) worked on four Asian countries, (Hong Kong, South Korea, Singapore, and Taiwan) and found no evidence of unidirectional causality from exports to economic growth in all the four economies. In the case of Taiwan, however, the study detected unidirectional causality from economic growth to export growth.

Kim (1993) has examined the major trends of key macroeconomic variables in South Korea and Chile and correlated them to export performance. Kim identified exports as a major source of economic growth and provided the evidence of the validity of the claim that an open and trade-oriented economy is not only the best guarantee for long-term economic growth, but it lightens the initial impacts of external shocks. Kim, further, mentioned that there are factors other than trade which increase economic growth.

Erfani (1999) examined the causal relationship between economic performance and exports over the period of 1965 to 1995 for several developing countries in Asia and Latin America. The results showed the significant positive relationship between exports and economic growth. This study provides the evidence of export-led growth hypothesis.

Vohra (2001) showed the relationship between the exports and economic growth in India, Pakistan, Philippines, Malaysia, and Thailand for the period 1973 to 1993. The empirical results indicated that when a country has achieved some level of economic development then the exports have a positive and significant impact on economic growth. The study also showed the importance of liberal market policies by pursuing export expansion strategies, and by attracting foreign investments.

Subasat (2002) investigated the empirical linkages between exports and economic growth. The study suggested that the more export-oriented countries like middle-income countries grow faster than the relatively less export-oriented countries. The study further showed that export promotion does not have any significant impact on economic growth for low and high income countries.

Amavilah (2003) determined the role of exports in economic growth by analyzing Namibia's data from 1968 to 1992. Results explained the general importance of exports, but the study finds no discernible sign of accelerated growth due to exports.

Lin (2003) stated that 10 per cent increase in exports cause 1 per cent increase in GDP in the 1990s in China on the basis of new proposed estimation method, when both direct and indirect contributions are considered.

Shirazi et al (2004) studied the short-run and long-run relationship among real exports, real imports, and economic growth on the basis of co-integration and multivariate Granger causality test as developed by Toda and Yamamoto (1995) for the period 1960 to 2003.This study showed a long-run relationship among imports, exports, and economic growth and found unidirectional causality from exports to output. But, it did not find any significant causality between imports and exports.

Thurayia (2004) studied the relationship between exports and economic growth experience in Saudi Arabia and Sudan. Results showed that the growth rate in total exports in Saudi Arabia had an active role in achieving economic growth while it had a weak influence in Sudan. The results of cointegration and error correction models showed a positive effect of exports on GDP in the short- and long- run, which confirms the validity of the hypothesis of export-led growth in Saudi Arabia, and Sudan.

Mah (2005) studied the long-run causality between exports and economic growth for China with the help of the significance of error correction term, EC_{t-1} . This study indicates that export expansion is insufficient to explain the patterns of real economic growth.

Tang (2006) stated that there is no long-run relationship among exports, real Gross Domestic product, and imports. This study further shows no short- and long-run causality between export expansion and economic growth in China on the basis of Granger causality test while economic growth does Granger-cause imports in the short-run.

Jordaan (2007) analyzed the causality between exports and GDP of Namibia for the period 1970 to 2005. The export-led growth hypothesis is tested through Granger causality and cointegration models. It tests whether there is unidirectional or bi-directional causality between exports and GDP. The results revealed that exports Granger-cause GDP and GDP per capita, and suggested that the export-led growth strategy through various incentives has a positive influence on growth.

Rangasamy (2008) examined the exports and economic growth relationship for South Africa, and provides the evidence that the unidirectional Granger causality runs from exports to economic growth.

Pazim (2009) tested the validity of export-led growth hypothesis in three countries by using panel data analysis. And, it is concluded that there exists no significant relationship between the size on national income and amount of exports for these countries on the basis of one-way random effect model. The panel unit root test shows that the process for both GDP and exports at first glance is not stationary, while the panel co-integration test indicates that there is no co-integration relationship between the exports and economic growth for these countries.

Ullah et al (2009) re-investigated the export-led growth hypothesis using time series econometric techniques over the period of 1970 to 2008 for Pakistan. The results reveal that export expansion leads to economic growth.

Elbeydi, Hamuda and Gazda (2010) investigated the relationship between exports and economic growth for Libya for the period 1980 to 2007. The findings indicate that there exists a long-run bi-directional causality between exports and income growth, and thus, the export promotion policy contributes to the economic growth of Libya.

The study of the dynamics of the relation between growth of exports and economic growth has been addressed by a number of researches in the context of India. Nandi and Biswas (1991) found the evidence of unidirectional causality from growth of exports to economic growth. This study does not test for stationarity and conduct Sims causality test on the levels of the income and export variables. Given that the levels of the income and export variables are usually non-stationary, the results are unreliable.

Sharma and Dhakal (1994) offer some evidence of the export-led growth hypothesis for India, but the empirical evidence offered by it is unreliable. The study concludes that the income and export series for India are non-stationary using the Phillip-Perron test. It tests for causality, but does not test for cointegration. However, the correct application of Granger tests requires the identification of a possible cointegrating relationship.

Bhat (1995) re-examines the exports-economic growth nexus for India, and finds evidence of bi-directional causality between growth of exports and economic growth. Xu (1996) confirms rejection of the export-led growth hypothesis for India. Ghatak and Price (1997) conclude that growth of exports is caused by output growth in India. Dhawan and Biswal (1999) examine the same issue for the period 1961 to 1993, and find that growth in GDP causes growth in exports while causality from exports to GDP appears to be a short-run phenomenon.

Nidugala (2000) finds that exports had a crucial role in influencing GDP growth in the 1980s. Anwar and Sampath (2000) examine the export-led growth hypothesis for 97 countries (including India, Pakistan and Sri Lanka) for the period 1960 to 1992. They found the evidence of unidirectional causality in the case of Pakistan and Sri Lanka, and no causality in the case of India. However, Kemal et al (2002) finds a positive association between exports and economic growth for India as well as for other economies of South Asia.

In case of India, Chandra (2000; 2002) found bi-directional causal relationship between growth of exports and GDP growth which is a short-run causal relation, as cointegration between growth of exports and GDP growth was not found.

Sharma and Panagiotidis (2004) test the export-led growth hypothesis in the context of India, and the results strengthen the arguments against the export-led growth hypothesis for the case of India.

Raju and Kurien (2005) analyzed the relationship between exports and economic growth in India over the pre-liberalization period 1960-1992, and found strong support for unidirectional causality from exports to economic growth using Granger causality regressions based on stationary variables, with and without an error-correction term.

Dash (2009) analyzes the causal relationship between growth of exports and economic growth in India for the post-liberalization period 1992-2007, and the results indicate that there exists a long-run relationship between output and exports, and it is unidirectional, running from growth of exports to output growth.

It is, therefore, clear from the above literature review that the evidence regarding exports-economic growth nexus is rather ambiguous and mixed. A number of studies support the export-led economic growth while others do not. Furthermore, studies on this issue in the context of India are only a few, and again provide mixed evidences. Also, the literature lacks studies including the period of recent global financial crisis. Therefore, this paper is an attempt to re-investigate the exports-economic growth nexus for India considering the period of recent global financial downsizing. This study shall provide the useful information helpful to policy makers. It can serve as a reference to subsequent research works on the issue 'exports-economic growth nexus' in the context of India.

3. Data and Methodology

The objective of this paper is to investigate the dynamics of the relationship between exports and economic growth in India using the annual data for the period 1970 to 2009. In this study, the variables are Total Exports by India (EX) and Economic Growth (EG). Total Exports by India is the sum of oil, and non-oil exports expressed in crores of rupees. And, the real Gross Domestic Product (GDP) is used as the proxy for economic growth in India. All necessary data for the sample period are obtained from the Handbook of Statistics on Indian Economy published by Reserve Bank of India[1]. All the variables are taken in their natural logarithms to avoid the problems of heteroscedasticity.

The estimation methodology employed in this study is the cointegration and error

[1] See, the web data at: http://www.rbi.org.in/scripts/AnnualPublications.aspx?head=Handbook of Statistics on Indian Economy

correction modeling technique. The entire estimation procedure consists of three steps: first, unit root test; second, cointegration test; third, the error correction model estimation.

3.1 Unit Root Test

The econometric methodology first examines the stationarity properties of each time series of consideration. The present study uses Augmented Dickey-Fuller (ADF) unit root test to examine the stationarity of the data series. It consists of running a regression of the first difference of the series against the series lagged once, lagged difference terms and optionally, a constant and a time trend. This can be expressed as follows:

$$\Delta Y_t = \alpha_0 + \alpha_1 t + \alpha_2 Y_{t-1} + \sum_{j=1}^{p} \alpha_j \Delta Y_{t-j} + \varepsilon_t \ . \tag{1}$$

The additional lagged terms are included to ensure that the errors are uncorrelated. In this ADF procedure, the test for a unit root is conducted on the coefficient of Y_{t-1} in the regression. If the coefficient is significantly different from zero, then the hypothesis that Y_t contains a unit root is rejected. Rejection of the null hypothesis implies stationarity. Precisely, the null hypothesis is that the variable Y_t is a non-stationary series ($H_0 : \alpha_2 = 0$) and is rejected when α_2 is significantly negative ($H_a : \alpha_2 < 0$). If the calculated value of ADF statistic is higher than McKinnon's critical values, then the null hypothesis (H_0) is not rejected and the series is non-stationary or not integrated of order zero, I(0). Alternatively, rejection of the null hypothesis implies stationarity. Failure to reject the null hypothesis leads to conducting the test on the difference of the series, so further differencing is conducted until stationarity is reached and the null hypothesis is rejected. If the time series (variables) are non-stationary in their levels, they can be integrated with I(1), when their first differences are stationary.

3.2 Cointegration Test

Once the unit roots are confirmed for data series, the next step is to examine whether there exists a long-run equilibrium relationship among the variables. This calls for cointegration analysis which is significant so as to avoid the risk of spurious regression. Cointegration analysis is important because if two non-stationary variables are cointegrated, a Vector Auto-Regression (VAR) model in the first difference is mis-specified due to the effects of a common trend. If cointegration relationship is identified, the model should include residuals from the vectors (lagged one period) in the dynamic VECM system. In this stage, Johansen's cointegration test is used to identify cointegrating relationship among the variables. The Johansen method applies the maximum likelihood procedure to determine the presence of cointegrated vectors in non-stationary time series. The testing hypothesis is the null of non-cointegration against the alternative of existence of cointegration using the Johansen maximum likelihood procedure.

In the Johansen framework, the first step is the estimation of an unrestricted, closed

p^{th} order VAR in k variables. The VAR model as considered in this study is:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + ... + A_p Y_{t-p} + BX_t + \varepsilon_t \qquad (2)$$

Where Y_t is a k -vector of non-stationary I(1) endogenous variables, X_t is a d -vector of exogenous deterministic variables, $A_1, ..., A_p$ *and* B are matrices of coefficients to be estimated, and ε_t is a vector of innovations that may be contemporaneously correlated but are uncorrelated with their own lagged values and uncorrelated with all of the right-hand side variables.

Since most economic time series are non-stationary, the above stated VAR model is generally estimated in its first-difference form as:

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + BX_t + \varepsilon_t \qquad (3)$$

$$\text{Where, } \Pi = \sum_{i=1}^{p} A_i - I, \quad and \quad \Gamma_i = -\sum_{j=i+1}^{p} A_j$$

Granger's representation theorem asserts that if the coefficient matrix Π has reduced rank $r < k$, then there exist $k \times r$ matrices α *and* β each with rank r such that $\Pi = \alpha\beta'$ *and* $\beta' Y_t$ is I(0). r is the number of co-integrating relations (the *co-integrating rank*) and each column of β is the co-integrating vector. α is the matrix of error correction parameters that measure the speed of adjustments in ΔY_t.

The Johansen approach to cointegration test is based on two test statistics, viz., the trace test statistic, and the maximum eigenvalue test statistic.

3.2.1 Trace Test Statistic

The trace test statistic can be specified as: $\tau_{trace} = -T \sum_{i=r+1}^{k} \log(1-\lambda_i)$, where λ_i is the ith largest eigenvalue of matrix Π and T is the number of observations. In the trace test, the null hypothesis is that the number of distinct cointegrating vector(s) is less than or equal to the number of cointegration relations (r).

3.2.2 Maximum Eigenvalue Test

The maximum eigenvalue test examines the null hypothesis of exactly r cointegrating relations against the alternative of r + 1 cointegrating relations with the test statistic: $\tau_{\max} = -T \log(1-\lambda_{r+1})$, where λ_{r+1} is the $(r+1)^{th}$ largest squared eigenvalue. In the trace test, the null hypothesis of $r = 0$ is tested against the alternative of r + 1 cointegrating vectors.

It is well known that Johansen's cointegration test is very sensitive to the choice of lag length. So, at first a VAR model is fitted to the time series data in order to find an appropriate lag structure. The Akaie Information Criterion (AIC), Schwarz Criterion (SC) and the Likelihood Ratio (LR) test are used to select the number of lags required in the cointegration test.

3.3 Vector Error Correction Model (VECM)

Once the cointegration is confirmed to exist between variables, then the third step entails the construction of error correction mechanism to model dynamic relationship. The purpose of the error correction model is to indicate the speed of adjustment from the short-run equilibrium to the long-run equilibrium state.

A Vector Error Correction Model (VECM) is a restricted VAR designed for use with non-stationary series that are known to be cointegrated. Once the equilibrium conditions are imposed, the VECM describes how the examined model is adjusting in each time period towards its long-run equilibrium state. Since the variables are supposed to be cointegrated, then in the short-run, deviations from this long-run equilibrium will feedback on the changes in the dependent variables in order to force their movements towards the long-run equilibrium state. Hence, the cointegrated vectors from which the error correction terms are derived are each indicating an independent direction where a stable meaningful long-run equilibrium state exists.

The VECM has cointegration relations built into the specification so that it restricts the long-run behaviour of the endogenous variables to converge on their cointegrating relationship while allowing for short-run adjustment dynamics. The cointegration term is known as the error correction term since the deviation from long-run equilibrium is corrected gradually through a series of partial short-run adjustments. The dynamic specification of the VECM allows the deletion of the insignificant variables, while the error correction term is retained. The size of the error correction term indicates the speed of adjustment of any disequilibrium towards a long-run equilibrium state.

In this study the error correction model as suggested by Hendry (1995) has been used. The general form of the VECM is as follows:

$$\Delta X_t = \alpha_0 + \lambda_1 EC^1_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta X_{t-i} + \sum_{j=1}^{n} \alpha_j \Delta \mathrm{Y}_{t-j} + \varepsilon_{1t} \quad (4)$$

$$\Delta \mathrm{Y}_t = \beta_0 + \lambda_2 EC^2_{t-1} + \sum_{i=1}^{m} \beta_i \Delta \mathrm{Y}_{t-i} + \sum_{j=1}^{n} \beta_j \Delta X_{t-j} + \varepsilon_{2t} \quad (5)$$

Where Δ is the first difference operator; EC_{t-1} is the error correction term lagged one period; λ is the short-run coefficient of the error correction term ($-1 < \lambda < 0$); and ε is the white noise. The error correction coefficient (λ) is very important in this error correction estimation as the greater co-efficient indicates higher speed of adjustment of the model from the short-run to the long-run.

The error correction term represents the long-run relationship. A negative and significant coefficient of the error correction term indicates the presence of long-run causal relationship. If both the coefficients of error correction terms in both the equations are significant, this will suggest the bi-directional causality. If only λ_1 is negative and significant, this will suggest a unidirectional causality from Y to X, implying that Y drives X towards long-run equilibrium, but not the other way around. Similarly, if λ_2 is negative and significant, this will suggest a unidirectional causality from X to Y, implying that X drives Y towards long-run equilibrium but not the other way around.

On the other hand, the lagged terms of ΔX_t and ΔY_t appeared as explanatory variables, indicating a short-run cause and effect relationship between the two variables. Thus, if the lagged coefficients of ΔX_t appear to be significant in the regression of ΔY_t, this will mean that X causes Y. Similarly, if the lagged coefficients of ΔY_t appear to be significant in the regression of ΔX_t, this will mean that Y causes X.

4. Empirical Analysis

At the outset, the Pearson's correlation coefficient between exports and real GDP is calculated over the sample period, and its significance is tested by the t-test. The value of Pearson's correlation coefficient (r) between these two time series over the sample period is 0.99. It shows that exports and real GDP are positively related in India and that a very high degree of correlation is evident between them. To test whether this value of 'r' shows a significant relationship between the two time series, student's t-test is used. The null hypothesis of the test is r = 0 against the alternative of $r \neq 0$. Since the t-statistic at 37 degrees of freedom is 42.68, and the critical t-value at 5 per cent level of significance is less than it, the null hypothesis is rejected. So, it can be said that the correlation between exports and real GDP is statistically significant. Correlation, however, does not say anything about long-run relationship, and thus, leaves unsettled the debate concerning the long-run relationship between exports and real economic growth measured by real gross domestic product.

Now, it is required to determine the order of integration for each of the two variables used in the analysis along with their stationarity tests. The Augmented Dickey-Fuller unit root test has been used for this purpose and, the results of such test are reported in Table 1.

Table 1: Results of Augmented Dickey-Fuller Unit Root Test

Variables in their First Differences with intercept	ADF Statistic	Critical Values	Decision
LEX_t	**-4.94**	**At 1% : -4.22** **At 5% : -3.53** **At 10% : -3.20**	**Reject Null hypothesis of no unit root**
$LGDP_t$	**-4.51**	**At 1% : -4.22** **At 5% : -3.53** **At 10% : -3.20**	**Reject Null hypothesis of no unit root**

Source: Author's own Calculation

It is clear from Table 1 that the null hypothesis of no unit roots for both the time series are rejected at their first differences since the ADF test statistic values are less than the critical values at 10, 5 and 1 per cent levels of significances. Thus, the variables are stationary and integrated into the same order, i.e., I(1).

In the next step, the cointegration between the stationary variables has been tested by the Johansen's Trace and Maximum Eigenvalue tests. The results of these tests are shown in Table 2.

Table 2: Results of Johansen's Cointegration Test

Hypothesized Number of Cointegrating Equations	Eigen Value	Trace Statistics	Critical Value at 5% (p-value)	Maximum Eigen statistics	Critical Value at 5% (p-value)
None*	0.432	26.465	20.261(0.006)	20.421	15.892(0.009)
At Most 1	0.154	6.044	9.164(0.187)	6.044	9.164(0.187)

Source: Author's own Calculation

* denotes rejection of the hypothesis at the 0.05 level

The Trace test indicates the existence of two cointegrating equations at 5 per cent level of significance. And, the maximum eigenvalue test makes the confirmation of this result. Thus, the two variables of the study have long-run equilibrium relationship between them. But in the short-run there may be deviations from this equilibrium, and it is required to verify whether such disequilibrium converges on the long-run equilibrium or not. Thus, Vector Error Correction Model is used to generate such short-run dynamics. Error correction mechanism provides a means whereby a proportion of the disequilibrium is corrected in the next period. So, error correction mechanism is a means to reconcile the short-run and long-run behaviour.

The estimation of a Vector Error Correction Model (VECM) requires selection of an appropriate lag length. The number of lags in the model is determined according to the Schwarz Information Criterion (SIC). The lag length that minimizes the SIC is 1. Then, an error correction model with the computed t-values of the regression coefficients is estimated and the results are reported in Table 3.

The estimated coefficient of error-correction term (EC^{1}_{t-1}) in the LEX equation is statistically significant and has a negative sign, which confirms that there is not only any problem in the long-run equilibrium relation between the independent and dependent variables at 5 per cent level of significance, but its relative value (-0.167) for India shows the rate of convergence to the equilibrium state per year. Precisely, the speed of adjustment of any disequilibrium towards a long-run equilibrium is that about 16.7 per cent of the disequilibrium in exports is corrected each year. Furthermore, the negative and statistically significant value of error correction coefficient indicates the existence of a long-run causality between the variables of the study. And, this causality is unidirectional in our model being running from the real GDP to exports. In other words, the changes in exports can be explained by real GDP.

Table 3: Estimates for VECM Regression

$$\Delta LEX_t = 0.081 - 0.167EC^1_{t-1} + 0.205\Delta LEX_{t-1} + 0.411\Delta LGNP_{t-1}$$
$$\Delta LGDP_t = 0.082 - 0.0143EC^2_{t-1} + 0.103\Delta LEX_{t-1} + 0.216\Delta LGDP_{t-1}$$

Independent Variable	ΔLEX_t	$\Delta LGDP_t$
Constant **[t-statistic]** **(p-value)**	**0.081** [1.591] **(0.120)**	**0.082** [4.014] **(0.00027)**
EC_{t-1} **[t-statistic]** **(p-value)**	$EC^1{}_{t-1}$ **= -0.167** [-1.836] **(0.074)**	$EC^2{}_{t-1}$ **= 0.0143** [-0.393] **(0.696)**
ΔLEX_{t-1} **[t-statistic]** **(p-value)**	**0.205** [1.242] **(0.222)**	**0.103** [1.557] **(0.127)**
$\Delta LGDP_{t-1}$ **[t-statistic]** **(p-value)**	**0.411** [1.010] **(0.319)**	**0.216** [1.325] **(0.193)**

Source: Author's own Calculation

The existence of Cointegration implies the existence of Granger causality at least in one direction (Granger, 1988). The long-run causality test from the VECM indicates that causality runs from real GDP to exports, since the coefficient of the error term in LEX equation is statistically significant and negatively based on standard t-test which means that the error term ($EC^1{}_{t-1}$) contributes to explain the changes in exports. However, the coefficient of the error correction term in the GDP equation is negative, but statistically insignificant which means that the error correction term ($EC^2{}_{t-1}$) does not contribute to explain the changes in real GDP. Therefore, there is unidirectional causality running from real GDP to exports.

The coefficients of the first difference of LEX and LGDP lagged one period in LEX equation in Table 3 are statistically insignificant which indicate the absence of short-run causality from real GDP to exports based on VECM estimates. In order to confirm the result of the short-run causality between the ΔLEX and the ΔLGDP based on VECM estimates, a standard Granger causality test is also performed based on the F-value.

Table 4: Results of Granger Causality Test

Null Hypothesis	F-Statistic	Probability	Decision
ΔLEX does not Granger Cause ΔLGDP	**2.522**	**0.096**	**Accept**
ΔLGDP does not Granger Cause ΔLEX	**1.324**	**0.280**	**Accept**

Source: Author's own Calculation
(Number of lags = 2)

The results in Table 4 indicate that real GDP does not Granger-cause the exports at 5 per cent level of significance. This result supports the previous result obtained from VECM that there is no short-run causality at 5 per cent level of significance. Based on this causality tests, changes in the real GDP cause changes in exports in the long-run, but not in the short-run.

5. Summary and Conclusion

In this paper, the relationship between exports and economic growth in a developing country like India has been investigated using popular time series methodologies. The data properties are analyzed to determine the stationarity of time series using the Augmented Dickey-Fuller unit root test which indicates that the two series are I(1). The results of the Cointegration test based on Johansen's procedure indicate the existence of the Cointegration between exports and real GDP. Therefore, the two variables of the study have a long-run equilibrium relationship between them, although they may be in disequilibrium in the short-run. The vector error correction model based on VAR indicates that about 16.7 per cent of disequilibrium is corrected every year. In addition, the negative and significant error correction term in LEX equation supports the existence of a long-run equilibrium relationship between real GDP and exports. Furthermore, the estimates of the VECM indicate the existence of a unidirectional causality running from real GDP to exports. The Granger causality test indicates that there is a causal relationship running from GDP to exports in the long-run, but not in the short-run.

The results of the empirical analysis lead to the conclusion that both exports and economic growth are related to past deviations (error correction terms) from the empirical long-run relationship. It implies that all variables in the system have a tendency to quickly revert back to their equilibrium relationship. This means that any increase in real GDP would have a positive impact on the growth of exports in the long-run. In other words, India provides the evidence of growth-driven exports over the sample period. This finding, thus, rejects the export-led growth hypothesis in India and corroborates the studies of Xu (1996), Ghatak and Price (1997), and Dhawan and Biswal (1999). Moreover, this study finds just the opposite result of Dash (2009); that it is exports that have promoted growth in India in the period 1992 to 2007, not the reverse. Thus, leaving the issue unsettled may be due to the inclusion of a relatively small time horizon in Dash (2009) and about a quite long sample period (about 40 years) in this study.

The fact is that India's economy is mostly dependent on its large domestic market with external trade accounting for only 20 per cent of the country's GDP[2]. In 2008, India accounted for 1.45 per cent of global merchandise trade and 2.8 per cent of global commercial services export. This supports the finding that the trend in India is not exports-led growth but rather growth-led exports.

In 2008-09, global financial meltdown and economic recession in developed economies were the major contributors in India's economic slowdown. Due to this, India's exports declined by 29.2 per cent in June 2009[3]. This steep decline was because countries were hit the hardest by the global recession, such as the United States and members of the European Union, account for more than 60 per cent of Indian exports. The later part of 2009 to April 2010, there has been a remarkable surge in India's exports. In April 2010, exports by India reported an increase by 36.2 per cent higher than the level in April 2009. Recently it is further reported that India's export increased by 23 per cent year-on-year basis in September 2010[4]. All these support the empirical evidence that the long-term trend may not be exports-led growth in India. The policy implication of such empirical evidence may be that the Government of India and other policy-planning bodies should devise prudential norms and policies to make the macro-economic fundamentals of the country strong enough to absorb the external shocks thereby achieving a fast growth of real economic variables to ensure a noticeable surge in the country's exports. In this direction, increasing domestic and foreign investments in key areas and ensuring price, interest rate, forex and political stabilities would go a long way.

However, further investigations are required to resolve such moot point using relatively longer time series, and incorporating in the study more number of macro-economic variables that governs the growth of exports by the country as well as the economic growth of the country.

Acknowledgement

I sincerely acknowledge the anonymous referee whose suggestions were instrumental in preparing this research paper.

[2] "Research and Markets: The Total Crude Oil Import in India", Reuters.com 2009-02-04, http://www.reuters.com/article/pressRelease/idUS196866+04-Feb-2009+BW20090204

[3] Chaturvedi, Neelabh (2009-08-27), "India Cuts Export Aim, Seeks Solace in New Markets" Online.wsj.com, http://online.wsj.com/article/SB125135625520662979.html?mod=googlenews_wsj

[4] "Indian Economy Overview", IBEF, July 2010, http://www.ibef.org/economy/economyoverview.aspx

References

Amavilah, Voxi Heinrich, 2003, 'Exports and Economic Growth in Namibia, 1968-1992', *Economic Working Paper Series*.

Anwar, M., and R. Sampath, 2000, 'Exports and Economic Growth', *Indian Economic Journal* 47, 3, pp. 79-88.

Arnade. C and U. Vasavada, 1995, 'Causality between Productivity and Exports in Agriculture: Evidence from Asia and Latin America', *Journal of Agricultural Economics,* 46, pp. 174-186.

Bahmani-Oskooee, M., Mohtadi, H. and Shabsigh, G., 1991, 'Exports, Growth and Causality in LDCs: A Re-examination', *Journal of Development Economics*, 36, pp. 405-15.

Bahmani-Oskooee, M. and Alse, J., 1993, 'Export Growth and Economic Growth: an Application of Cointegration and Error-Correction Modeling', *Journal of Developing Areas*, 27, pp. 535-42.

Balassa, B., 1978, 'Exports and Economic Growth: Further Evidence', *Journal of Development Economics*, 5, pp. 181-189.

Bhagwati, J., 1978, 'Foreign Trade Regimes and Economic Development: Anatomy and Consequences of Exchange Control Regime', *Working Paper Series, NBER*, New York.

Bhat, S., 1995, 'Export and Economic Growth in India', *Artha Vijana* 37, 44, pp. 350-358.

Chandra, R., 2000, 'The Impact of Trade Policy on Growth in India', Unpublished Ph.D Thesis, Department of Economics, University of Strathclyde, Glasgow

Chandra, R., 2002, 'Export Growth and Economic Growth: An Investigation of Causality in India', *Indian Economic Journal*, 49, pp. 64-73.

Chow, P.C.Y., 1987, 'Causality between Export Growth and Industrial Performance: Evidence from the NIC's', *Journal of Development Economics*, 26, pp. 55-63.

Darrat, A., F., 1986, 'Trade and Development: The Asian Experience', *Cato Journal*, 6, 2, pp. 695-699.

Darrat, A.F., 1987, 'Are Exports an Engine of Growth? Another Look at the Evidence', *Applied Economics,* 19, pp. 277-83.

Dash, R. K., 2009, 'Revisited Export-Led Growth Hypothesis: An Empirical Study on India', *South Asia Economic Journal*, 10, 2, pp. 305-324.

Dhawan, U. and Biswal, B., 1999, 'Re-examining Export-led Growth Hypothesis: A Multivariate Cointegration Analysis for India', *Applied Economics,* 31, pp. 525-530.

Dickey, D. A. and Fuller, W. A., 1979, 'Distribution of the Estimators for Autoregressive Time Series with a Unit Root', *Journal of the American Statistical Association*, 74, 366, pp. 427-431.

Dodaro, S., 1993, 'Export and Growth: A Reconsideration of Causality', *Journal of Developing Areas*, 27, pp. 227-44.

Dutt, S. D. and D. Ghosh, 1994, 'An Empirical Investigation of the Export Growth – Economic Growth Relationship', *Applied Economics Letters*, 1, pp. 44-48.

Easterly, W. R., 2007, 'Free Market and Economic Development', *Presentation International Symposium, on the Poverty Reduction and Beyond Development Strategies for Low Income Countries.*

Engle, R. F. and Granger, C. W., 1987, 'Co-integration and Error Correction: Representation, Estimation and Testing' *Econometrica,* 55, 2, pp. 251-276.

Erfani, G.R., 1999, 'Export and Economic Growth in Developing Countries', *International Advances in Economic Research*, 5, 1, pp. 147-148.

Feder, G., 1982, 'On Exports and Economic Growth', *Journal of Development Economics*, 12, pp. 59-73.

Fosu, A. K., 1990, 'Exports and Economics Growth: The African case', *World Development* 18, pp. 831-835.

Ghatak, Subrata, and Stephen Wheatley Price, 1997, 'Export Composition and Economic Growth: Cointegration and Causality Evidence for India', *Weltwirtschaftliches Archiv,* 133, 3, pp. 538-553.

Granger, C. W. J., 1969, 'Investigating Causal Relations by Econometric Models: Cross Spectral Methods', *Econometrica*, 37, pp. 424-438.

Granger, C. W. J., 1988, 'Some Recent Development in a Concept of Causality', *Journal of Econometrics*, 39, pp. 199-211.

Greenaway, D. and Sapsford, D., 1994, 'Exports, Growth, and Liberalization: An Evaluation', *Journal of Policy Modelling*, 16, pp. 165-186.

Hendry, D.F., 1995, *Dynamic Econometrics*, Oxford University Press, Oxford

Henriques, I. and Sadorsky, P., 1996, 'Export-Led Growth or Growth-Driven Exports?: The Canadian Case, *Canadian Journal of Economics*, XXIX, 3, pp. 540-555.

Hsiao, M.W., 1987, 'Tests of Causality and Exogeneity between Export Growth and Economic Growth', *Journal of Development Economics*, 18, pp. 143-159.

Islam, M.N., 1998, 'Export Expansion and Economic Growth: Testing For Cointegration and Causality', *Applied Economics*, 30, pp. 415-425.

Johansen, S., 1988, 'Statistical Analysis of Cointegration Vectors', *Journal of Economic Dynamics and Control* 12, pp. 231-254.

Johansen, S. and Juselius, K., 1990, 'Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money', *Oxford Bulletin of Economics and Statistics*, 52, pp. 169-211.

Jordaan, A. C. and Eita, J. H., 2007, 'Export and Economic Growth in Namibia: A Granger Causality Analysis', *South African Journal of Economics*, 75, 3, pp. 540-547.

Jung, W.S. and Marshall, P. J., 1985, 'Exports, Growth and Causality in Developing Countries', *Journal of Development Economics*, 18, pp. 1-12.

Kemal, A. R., Din, M., Qadir, U. L., Fernando and Colombage, S., 2002, 'Exports and Economic Growth in South Asia', *A Study prepared for the South Asia Network of Economic Research Institutes*.

Khalafalla, K. Y. and Webb, A. J., 2001, 'Export-Led Growth and Structural Change: Evidence from Malaysia', *Applied Economics*, 33, pp. 1703-1715.

Khaled, R.M. Elbeydi, Abdulbaset, M., Hamuda, Vladimir, Gazda, 2010, 'The Relationship between Export and Economic Growth in Libya Arab Jamahiriya', *Theoretical and Applied Economics,* XVII, 1, 542, pp. 69-76.

Kim, K. S., 1993, 'Economic Implications of NAFTA for East Asia—With Special Reference to US and East Asian Trade and Investment Flows', *paper presented at the Korea Economic Institute and Princeton University Symposium, Princeton University, September*.

Krueger, A., 1978, 'Foreign Trade Regimes and Economic Development: Liberalization Attempts and Consequence', *Working Paper Series, NBER*, New York.

Krueger, A., 1990, 'Asian Trade and Growth Lessons', *American Economic Association Papers and Proceedings*, 80, pp. 108-12.

Kugler, P., 1991, 'Growth, Exports and Cointegration: An Empirical Investigation', *Weltwirtschaftliches Archive,* 127, pp. 73-82.

Lin, Yo–Long, 2003, 'Exports, Terms of Trade, and Growth,' Master Thesis, National Chengchi University (in English).

Mah, J. S., 2005, 'Export Expansion, Economic Growth and Causality in China', *Applied Economics Letters*, 12, pp. 105-107.

Nandi, S. and Biswas, B., 1991, 'Export and Economic Growth in India: Empirical Evidence', *Indian Economic Journal*, pp. 53-59.

Nidugala, G. K., 2001, 'Exports and Economic Growth in India: An Empirical Investigation', *Indian Economic Journal*, 47, 3, pp. 67-78.

Pazim, K. H., 2009, 'Panel Data Analysis of 'Export-led Growth Hypothesis in BIMP-EAGA Countries', MPRA, paper 13264.

Raju, Sudhakar S. and Kurien, J., 2005, 'Exports and Economic Growth in India: Cointegration, Causality and Error-Correction Modeling: A Note', *Indian Journal of Economics and Business*, June,
http://findarticles.com/p/articles/mi_m1TSD/is_1_4/ai_n25122014/

Rangasamy, Logan, 2008, 'Exports and Economic Growth: The Case of South Africa', *Journal of International Development*, 21, 5, pp. 603-617.

Sharma, A., and Panagiotidis, T., 2005, 'An Analysis of Exports and Growth in India: Some Empirical Evidence', *Review of Development Economics*, 9, 2, pp. 232-248.

Shan, J. and Sun, F., 1998a, 'Export-Led Growth Hypothesis for Australia: An Empirical Reinvestigation', *Applied Economics Letters*, 5, pp. 423-428.

Shan, J. and Sun, F., 1998b, 'On the Export-Led Growth Hypothesis: The Econometric Evidence from China', *Applied Economics*, 30, pp. 1055-1065.

Sharma, Subhash C. and Dhakal, Dharmendra, 1994, 'Causal Analyses between Exports and Economic Growth in Developing Countries' *Applied Economics*, 26, 12, pp. 1145-1157.

Shirazi, N. S. and Manap, T.A.A., 2004, 'Export-Led Growth Hypothesis: Further Econometric Evidence from Pakistan', *Pakistan Development Review*, 43, pp. 563-581.

Sidiq Thurayia, 2004, 'The Relationship between Exports and Economic Growth Experience of Saudi Arabia and the Republic of Sudan', *Economic studies: The Scientific Series of the Saudi Economic Association,* V6, A11.

Sims, C.A., 1972, 'Money, Income and Causality' *American Economic Review,* 62, pp. 540-52.

Stiglitz. J. E., 2007, *Making Globalization Work for Developing Countries*. W.W Norton & Company, Inc.

Subasat, T., 2002, 'Does Export Promotion Increase Economic Growth? Some Cross-Section Evidence', *Development Policy Review*, 20, 3, pp. 333-349.

Tang, T. C., 2006, 'New Evidence on Export Expansion, Economic Growth and Causality in China', *Applied Economics Letters*, 13, pp. 801-803.

Thornton, J., 1996, 'Cointegration, Causality and Export-Led Growth in Mexico, 1895-1992', *Economics Letters,* 50, pp. 413-416.

Thornton, J., 1997, 'Export and Economic Growth: Evidence from 19th Century Europe', *Economics Letters,* 55, pp. 235-240.

Toda, H.Y. and Yamamoto, 1995, 'Statistical inference in Vector Autoregressions with Possibly Integrated Processes', *Journal of Econometrics*, 66, pp. 225-250.

Ullah, Zaman, Farooq and Javid., 2009, 'Cointegration and Causality between Exports and Economic Growth in Pakistan', *European Journal of Social Sciences,* 10, 2, pp. 264-272.

Van den Berg, H. and Schmidt, J. R., 1994, 'Foreign Trade and Economic Growth: Time Series Evidence from Latin America', *Journal of International Trade and Economic Development,* 3, pp. 249-68.

Vohra, R., 2001, 'Export and Economic Growth: Further Time Series Evidence from Less Developed Countries', *International Advances in Economic Research*, 7, 3, pp. 345-50.

Xu, X., 1996, 'On the Causality between Export Growth and GDP Growth: an Empirical Reinvestigation', *Review of International Economics,* 4, pp. 172-184.

The Gender impact in Earnings Inequality: Evidence from Sri Lanka

Thankom Arun[1] and Vani K. Borooah[2]

Abstract

This paper estimates an earnings function for Sri Lanka, followed by a decomposition analysis of male-female earnings suggest that the gender disparity in earnings largely represents 'discrimination' against women. The findings showed that irrespective of their "inferior" labour market attributes, men had average earnings that were considerably higher than the female average and that this could be attributed entirely to discrimination in favour of male earners.

Keywords: Inequality, Earnings, Gender, Sri Lanka

JEL codes: O53,R2,O29

1. Introduction

Despite the high levels of human development, the civil war has affected Sri Lanka's economic growth significantly. Sri Lanka embarked on the liberalisation policies since late 1970s, and this led to an increase in economic growth; however, this was not reflected on the distribution of income, which remained more or less unchanged (Dunham and Jayasuriya, 2000). In developing countries, Kuznets (1955) has identified the shift of population from traditional to modern activities as an important reason for inverted U relationship between inequality and development and found that developing countries had relatively greater inequality than developed countries. The variations in inequality reflect real differences across countries in participation in the modern sectors of the economy and indicate the importance of urbanisation and industrialisation in determining the extent of inequality. However, a recent work reveals reduced inequality trends in Latin America, which may be due to the narrowing of the earnings gap between skilled and to the poor (López-Calva and Claudia, 2010).

[1] Professor of Development Finance and Public Policy at the Lancashire Business School, University of Central Lancashire, email: tgarun@uclan.ac.uk.

[2] Professor of Applied Economics at the University of Ulster, Northern Ireland, email: vk.borooah@ulster.ac.uk.

Both authors are grateful to the Editor and two anonymous referees for valuable comments. Needless to say, the usual disclaimer applies.

In India, the distribution of income and the process of development based on a cross country analysis support the increase in relative inequality in the early stages of development followed by a decline in the later stages which is attributed to factors such as changes in inter-sectoral shifts in the structure of production, educational attainment and labour skills (Ahluwalia, 1976). In China, studies have highlighted powerful divergences in earnings among its provinces as well (Knight, Li and Zhao, 2001). It is also noted that in China, signs of wage discrimination against minorities and women are strong and the productive characteristics of workers are rewarded in the labour market (Knight and Song, 2003). The decomposition analysis of the rapid increase in mean earnings showed that unskilled market wage rose very little in real terms, the impetus came from the rising returns to education and the growing gap between the local private sector and other ownership factors. According to Cunningham and Jacobson (2003) the policies that attempt to equalise earnings-related characteristics across the population are more effective than any other type of targeting. However, the issues on trends in poverty and inequality receive a range of empirical results, mainly due to the ambiguous concepts, data problems, and the adequacy of income measures to gauge welfare and its distribution (Nelson, 1998). Some of the recent evidence shows that poverty has been reduced successfully and sustainably, where governments used policy interventions to facilitate employment centred structural transformations of their economies (UNRISD, 2010).

In Section 2 of this paper, we estimate an earnings function for Sri Lanka based on the Consumer Finances and Socio Economic Survey that excluded war-affected districts. Section 3 addresses the question of how greater equality by gender and race/ethnicity in distribution of earnings would affect earnings inequality. Section 4 summarises the main findings of the study.

This study builds on earlier work on the decomposition of income inequality in Sri Lanka (Glewwe, 1986; 1988). These studies showed that in the wake of the Sri Lankan government's implementation of policies in 1977 to remove governmental controls on the economy and to provide incentives for investment, inequality has increased. In this context, this study takes a fresh look at income inequality in Sri Lanka, based on a new data set.

2. Earnings Function for Sri Lanka

2.1 Basic Features of the Survey

The total earnings of the 7,826 earners in the Survey (i.e. persons with positive earnings) were defined as the sum of their earnings from their primary and subsidiary occupations. Table 1 (in Appendix) shows some of the salient features with respect to the earners in the Survey such as the average earnings were the lowest for Tamils.

The second feature is the gap in male-female earnings: males earned, on average, 41 per cent more than females. 86 per cent of Sinhalese earners lived in rural areas while 80 per cent of Tamil earners lived on tea estates; the most urbanised ethnic group was the Muslims, one-third of whom lived in urban areas. The place of residence, which was home to one out of five of Muslim earners, while only 3 per cent of the Sinhalese earners and 7 per cent of Tamil earners lived in the Colombo Municipal Area.

The table also reveals that the proportion of earners with higher educational qualifications was very low for Tamils (7 per cent) and the highest for the Sinhalese (44 per cent). The magnitude of high percentage of school drop outs among Tamils will be more significant if we include Tamil heartlands such as 'Jaffna' in the sample which has the highest school drop-out rates in the country (World Bank, 2000). In many of these areas, children have interrupted schooling due to displacement of families by the conflict and children are also lured (forcibly or not) to join the civil war.

2.2 Estimating an Earning Function

The earnings equations were estimated, first, over all the 7,826 earners; then, separately for male and female, earners; and lastly, separately for the Sinhalese earners, the Tamil earners and the Muslim earners (see, Table 2 in Appendix).

The estimation results point to the fact that *ceteris paribus* a move from the estates sector to the rural sector and urban sector would increase earnings by 30 and 40 per cent respectively. The fact that an overwhelming majority of Tamil earners worked on estates, provides an explanation for the lowest earnings of Tamils, compared to Sinhalese and Muslims. The effect of residence on earnings also made it felt through the various zones.

Earnings increased with the number of days worked in the week. The average numbers of weekly working days were, in terms of ethnicity, 4.9 for Sinhalese earners, 5.1 for Tamil and 5.0 for Muslim earners and, in terms of gender, 4.8 for men and 5.0 for women. Although an increase in age and in years of experience added to earnings, the effect of age was stronger than the effect of work experience. The results showed that a high education qualification, compared to no schooling, raised earnings by 33 per cent and this effect was stronger for male than for female earners (34 against 29 per cent) and stronger for Sinhalese than for Tamil earners (37 against 30 per cent). In employment, the earnings equation showed that working in the public sector and in the organised private sector added, on average, 10-11 per cent to earnings offered in the unorganised private sector.

Despite the details in table 1, the estimation results show that there was no 'ethnic effect' *per se* on earnings in Sri Lanka. After other non-ethnic factors had been controlled for, the coefficients on the Tamil and Muslim dummy variables were not significantly different from zero, whether in the all earners, male earners or female earners equations. The most important of the non-ethnic factors were controlling for the urban/rural/estates sectors and for education effects. Firstly, the vast majority of Tamil earners worked on estates were average earnings (485 rupees per week) were considerably lower than in the rural (854 rupees per week) or urban (1,250 rupees per week) sectors. Secondly, relatively few Tamil earners - compared to Sinhalese and Muslim earners - had high educational qualifications when such qualifications offered a considerable earnings premium.

3. The Decomposition of Male-Female Earnings

We are using the Oaxaca (1973) and Blinder (1973) decomposition methodology to examine the issue of gender difference in earnings represents 'discrimination' against women. The male and female earnings equations may be written as:

$$\log(W_F) = \mathbf{X}'_\mathbf{F}\boldsymbol{\beta}_\mathbf{F} \text{ and } \log(W_M) = \mathbf{X}'_\mathbf{M}\boldsymbol{\beta}_\mathbf{M} \quad (1)$$

where: W_F and W_M are, respectively, female and male earnings; $\mathbf{X_F}$ and $\mathbf{X_M}$ are vectors, respectively, of observations on explanatory variables for female and male earnings; $\boldsymbol{\beta}_\mathbf{F}$ and $\boldsymbol{\beta}_\mathbf{M}$ are coefficient vectors for the female and male earnings equations.

Alternatively, equation (1) may be written as:

$$\log(W_M) - \log(W_F) = \mathbf{X_M}'\boldsymbol{\beta}_\mathbf{M} - \mathbf{X_F}'\boldsymbol{\beta}_\mathbf{F} = (\boldsymbol{\beta}_\mathbf{M} - \boldsymbol{\beta}_\mathbf{F})'\mathbf{X_F} + (\mathbf{X_M} - \mathbf{X_F})'\boldsymbol{\beta}_\mathbf{M} \quad (2)$$

or as:

$$\log(W_M) - \log(W_F) = \mathbf{X_M}'\boldsymbol{\beta}_\mathbf{M} - \mathbf{X_F}'\boldsymbol{\beta}_\mathbf{F} = (\boldsymbol{\beta}_\mathbf{M} - \boldsymbol{\beta}_\mathbf{F})'\mathbf{X_M} + (\mathbf{X_M} - \mathbf{X_F})'\boldsymbol{\beta}_\mathbf{F} \quad (3)$$

The first term in equations (2) and (3) - which may be interpreted as the 'discrimination' component - measures the (log) difference in male and female earnings resulting from differences in their respective coefficient vectors ($\boldsymbol{\beta}_\mathbf{M}$-$\boldsymbol{\beta}_\mathbf{F}$): in equation (2) these differences are evaluated at $\mathbf{X_F}$, the observations relating to the female attribute vector; in equation (3) they are evaluated at $\mathbf{X_M}$, the observations for the male attribute vector. The second term in equations (2) and (3), above, measures the (log) difference in male and female earnings resulting from differences in their respective attribute vectors ($\mathbf{X_M}$-$\mathbf{X_F}$): in equation (2) these differences are evaluated using $\boldsymbol{\beta}_\mathbf{M}$, the male coefficient vector; in equation (3) they are evaluated using $\boldsymbol{\beta}_\mathbf{F}$, the female coefficient vector.

The observed difference between men and women in the logarithm of their earnings - $\log(earn^M / earn^F)$ - was 0.279 for all earners (See Table 3 in Appendix). Consequently, average male earnings were 28 per cent higher than average female earnings[1]. When, for all earners, female attributes were evaluated at male coefficients ('women were treated as men'), the log difference in earnings was predicted to be 0.314 which is higher than the observed sample difference. In other words, if women were treated 'fairly' - in that their earnings attributes were evaluated using male coefficients - then the average log earnings of women (6.551) would *exceed* that observed for men (6.516). However, the superior female attributes were translated into earnings using coefficients which were markedly inferior to those used for converting male attributes into earnings. As a result, female earners - notwithstanding their superior attributes - had average earnings which were considerably lower than the male average and this could be attributed *entirely* to discrimination against women earners.

A similar conclusion emerges when 'men were treated as women'. If male earnings attributes were evaluated at female coefficients then average log earnings for men (6.08) would be *lower* than that observed for women (6.237). To put it differently, men in the

Sri Lankan Survey had *inferior* earnings attributes compared to women. However, these inferior male attributes were translated into earnings using coefficients which were markedly superior to those used for converting female attributes into earnings. As a result, male earners - notwithstanding their inferior attributes - had average earnings which were considerably higher than the female average and, as has been argued, this fact could be attributed *entirely* to discrimination in favour of male earners.

4. Conclusions

This study estimated an earnings function for Sri Lanka, which explains the significant positive effects of urbanisation and education on earnings. Following this, the decomposition exercise of male-female earnings indicates the extent to which the gender disparity in earnings represents 'discrimination' against women. These findings provide greater insights into the misconception of the perceived notion of no significant gender inequality either in access to health and education services, or in economic welfare and income poverty levels (World Bank, 2000). The evidences showed that irrespective of inferior attributes, men had average earnings that were considerably higher than the female average that attributed *entirely* to discrimination in favour of male earners. The ethnic and gender imbalances in earnings raise wider policy questions that need to be addressed adequately in development strategies.

References

Ahluwalia, M., 1976, 'Income distribution and development. Some stylized facts', *American Economic Review*, Papers and Proceedings, 66, 2, pp. 128-135.

Cunningham, W. and Joyce, J., 2003, 'Earnings inequality within and across Gender, racial and ethnic groups in Latin America', Background paper for the World Bank LAC Report *Inequality and the State in Latin America and the Caribbean.*

Dunham, D. and Sisira, J., 2000, 'Equity, Growth and Insurrection: Liberalization and the welfare debate in Contemporary Sri Lanka', *Oxford Development Studies*, 28, 1, pp. 97-110.

Glewwe, P., 1986, 'The Distributionof Income in Sri Lanka in 1969-70 and 1980-81: A Decomposition Analysis', *Journal of Development Economics*, 24, pp. 255-274.

Glewwe, P., 1988, 'Economic Liberalisation and Income Inequality: Further Evidence on the Sri Lankan Experience', *Journal of Development Economics*, 28, pp. 233-246.

Knight, J. and Song, L. and Jia, H., 2001, 'Chinese rural migrants in urban enterprises: Three perspectives', *Journal of Development Studies*, 35, 3, pp. 73-104.

Knight, J., Song, L., 2003, 'Increasing urban wage inequality in China', *Economics of Transition*, 11, 4, pp. 596-619.

Kuznets, S., 1955, 'Economic Growth and income inequality', *American Economic Review*, 45, pp. 1-25.

López-Calva and Lustig, C., 2010, eds., '*Declining Inequality in Latin America A Decade of Progress?*' Brookings Institution Press and United Nations Development Programme.

Nelson, J., 1998, 'Poverty, Inequality, and Conf lict in Developing Countrie's', Project report on World Security, Rockfeller Brothers Fund.

UNRISD, 2010, Policy Brief of the Executive Committee of Economic and Social Affairs, New York.

World Bank, 2000, 'Sri Lanka: Recapturing missed opportunities', World Bank Country Report.

Appendix

Table 1: The Socio-Economic Position of Earners in Sri Lanka, by Ethnicity

	Sinhalese	***Tamil***	***Muslim***	***All Earners***
Number in sample	6,514	1,041	271	7,826
% of Total sample	83	13	4	100
Average earnings total (rupees)	913	571	919	868
Average earnngs (male)	994	647	956	955
Average earnngs (female)	727	480	682	679
%Rural	86	9	63	75
%Urban	13	11	34	14
%Estate	1	80	3	11
% female	30	46	14	32
Average age (yrs)	37	37	36	37
Zone of residence:				
Zone 1	47	14	20	42
Zone 2	13	2	12	12
Zone 3	37	77	46	42
Zone 4	3	7	22	4
Marital Status:				
% Married	67	71	70	67
% Single/widowed/divorced	33	29	30	33
Educational attainment:				
High (passed year 10 or above)	44	7	34	39
Moderate (passed up to year 10)	50	66	61	53
Low (no schooling)	6	27	5	8
Nature of Employment:				
% Regular employees	38	62	21	41
% Casual or Contract employees	62	38	79	59
Occupation:				
Professional, Managerial or Technical	11	3	12	10
Clerical	12	1	9	10
Sales	4	4	20	5
Service	8	5	11	8
Production	65	87	48	67
Sector of Employment:				
Public	13	11	34	14
Organised Private Sector	86	9	63	75
Unorganised Private Sector	1	80	3	11

Notes to Table 1:

Total earnings: sum of earnings from employment in primary and subsidiary occupations. *However, for all earners, earnings from subsidiary occupations were zero*.

Zone: Zone 1 (Colombo, Gampaha, Kalutara, Galle, Matara); Zone 2 (Hambantota, Monergala, ampara, Polonnarwa, Anuradhapura, Puttalam); Zone 3 (Kandy, Matale, Nuwara-Eliya, Badulla, Ratnapura, Kegalle, Kurunegala); Zone 4 (Colombo Municipal Area).

Table 2: Earnings Functions for Sri Lanka

	All Earners	*Male Earners*	*Female Earners*	*Sinhalese Earners*	*Tamil Earners*	*Muslim Earners*
Urban	0.400	0.513	0.262	0.371	0.077	0.667
	(8.1)	(8.1)	(3.5)	(3.8)	(0.9)	(2.1)
Rural	0.303	0.417	0.153	0.257	0.284	0.596
	(6.6)	(7.0)	(2.3)	(2.7)	(4.0)	(2.1)
Sex	-0.350	-	-	-0.394	-0.078	-0.392
	(22.1)			(22.1)	(2.4)	(2.9)
Age (years)	0.030	0.037	0.020	0.030	0.028	0.056
	(8.3)	(8.1)	(3.5)	(7.4)	(3.5)	(2.7)
Age sq	-0.0004	-0.0005	-0.0003	-0.0004	-0.0003	-0.0007
	(9.5)	(8.7)	(4.2)	(8.6)	(3.3)	(3.0)
Married	0.175	0.231	0.030	0.177	0.088	0.152
	(9.4)	(8.9)	(1.2)	(8.5)	(2.1)	(1.2)
Experience (years)	0.006	0.003	0.006	0.007	-0.0003	0.004
	(5.8)	(2.6)	(3.4)	(6.0)	(0.1)	(0.5)
Regular employee	0.340	0.308	0.399	0.374	0.127	0.544
	(16.2)	(11.1)	(13.4)	(15.8)	(2.9)	(3.4)
Days worked in week	0.133	0.145	0.109	0.131	0.159	0.105
	(30.4)	(26.6)	(15.5)	(27.2)	(15.1)	(3.6)
Tamil	-0.040	-0.045	-0.031	-	-	-
	(1.0)	(0.9)	(0.5)			
Muslim	0.024	0.005	-0.027	-	-	-
	(0.6)	(0.1)	(0.3)			
Zone 1	-0.002	-0.010	-0.015	0.052	-0.315	-0.047
	(0.1)	(0.1)	(0.2)	(1.1)	(3.2)	(0.3)
Zone 2	-0.174	-0.199	-0.122	-0.113	-0.336	-0.399
	(3.9)	(3.6)	(1.6)	(2.1)	(2.1)	(2.1)
Zone 3	-0.221	-0.256	-0.182	-0.161	-0.481	-0.334
	(5.3)	(5.0)	(2.6)	(3.2)	(4.8)	(2.0)
Higher education	0.328	0.342	0.293	0.371	0.296	0.130
	(10.6)	(8.2)	(6.4)	(9.7)	(3.4)	(0.6)
Middle education	0.136	0.152	0.117	0.176	0.136	0.116
	(5.1)	(4.0)	(3.2)	(5.1)	(3.7)	(0.6)
Clerical	-0.309	-0.349	-0.211	-0.319	-0.351	-0.376
	(10.0)	(8.2)	(4.9)	(9.9)	(2.1)	(1.9)

Sales	-0.396	-0.405	-0.480	-0.409	-0.678	-0.312
	(9.5)	(7.9)	(6.0)	(8.8)	(4.9)	(1.5)
Service	-0.489	-0.453	-0.490	-0.482	-0.899	-0.535
	(13.9)	(10.1)	(7.7)	(12.7)	(6.6)	(2.5)
Production	-0.444	-0.458	-0.410	-0.448	-0.578	-0.442
	(15.3)	(11.7)	(9.1)	(14.5)	(4.6)	(2.3)
Public Sector	0.098	-0.002	0.334	0.082	0.098	-0.175
	(3.8)	(0.1)	(7.9)	(2.9)	(1.5)	(0.9)
Organised Private Sector	0.108	0.043	0.237	0.118	0.067	0.085
	(5.0)	(1.5)	(7.2)	(4.9)	(1.2)	(0.6)
Intercept	5.391	4.756	5.055	5.392	5.512	5.046
	(51.1)	(36.8)	(30.9)	(37.9)	(23.6)	(8.4)

Notes to Table 2:
Dependent variable is log(total earnings); figures in parentheses are t-values.
Explanatory variables as defined in Table 1.
Reference categories: male; single; casual or contract employee; Sinhalese; zone 4; low educational attainment; professional, managerial, technical occupation; unorganised private sector.

Notes to the equations:
All earners equation: 7,826 observations; R^2 (adj)= 0.433.
Male earners equation: 5,338 observations; R^2 (adj)= 0.397.
Female earners equation: 2,488 observations; R^2 (adj)= 0.512.
Sinhalese earners equation: 6,514 observations; R^2 (adj)= 0.437.
Tamil earners equation: 1,041 observations; R^2 (adj)= 0.348.
Muslim earners equation: 271 observations; R^2 (adj)= 0.296.

Table 3: The Decomposition of Differences in Earnings Between Males and Females

All Earners					
Sample Average	***Women Treated as Men***			***Men Treated as Women***	
$\log(earn^M / earn^F)$	$\mathbf{X_F}'\hat{\boldsymbol{\beta}}_\mathbf{M} - \mathbf{X_F}'\hat{\boldsymbol{\beta}}_\mathbf{F}$	$\mathbf{X_M}'\hat{\boldsymbol{\beta}}_\mathbf{M} - \mathbf{X_F}'\hat{\boldsymbol{\beta}}_\mathbf{M}$		$\mathbf{X_M}'\hat{\boldsymbol{\beta}}_\mathbf{M} - \mathbf{X_M}'\hat{\boldsymbol{\beta}}_\mathbf{F}$	$\mathbf{X_M}'\hat{\boldsymbol{\beta}}_\mathbf{F} - \mathbf{X_F}'\hat{\boldsymbol{\beta}}_\mathbf{F}$
6.516-6.237=0.279	**6.551-6.237 = 0.314**	**6.516-6.551 = -0.035**		**6.516-6.08 = 0.436**	**6.08-6.237 = -0.157**

[i] $earn^M/earn^F$=exp(0.279)=1.276

Distribution of poverty and inequality indices for various groups in Greece using the bootstrap technique

Dimitra Aggelopoulou[1], Stavros Zografakis[2] and Panayiotis Sypsas[3]

Abstract

The present work estimates poverty and inequality indices in various groups that are known to have high contributions to poverty and inequality, such as farmers, pensioners and families with 3 or more children. Using the bootstrap technique, we estimate confidence intervals for these indices and examine if the decrease of poverty and inequality during the years 1998/99 and 2004/05 in these groups is statistically significant or within the boundary of the confidence interval.

For the purpose of the study, we use household income data from the last two Household Budget Surveys (HBS) which cover the entire population of Greece conducted in 1998/99 and 2004/05 by the National Statistical Service of Greece (NSSG). In regard to the methodological issues, we chose the individual as the unit of analysis and the "family equivalence scales" used by Eurostat.

Keywords: poverty, inequality.

JEL Classification: D31, I32

1. Introduction

During the last decades the concern for poverty and inequality increased. Many empirical studies used different indices to show the magnitude and the intensity of poverty and inequality and some of them explored the structure of poverty and inequality by decomposing them for various socio-economic groups. Most studies conclude that poverty and inequality decreased in Greece during the last decades, and others show that the overall inequality arise mainly from inequalities "within" the various socio-economic groups

[1] University of Patras, Department of Economics, University Campus, Rio, 265 00 Patras, Greece, daggelop@upatras.gr
[2] Agricultural University of Athens, Department of Agricultural Economics and Rural Development 75, Iera Odos – 118 55 Athens Greece, stazog@aua.gr
[3] University of Patras, Department of Economics, University Campus, Rio, 265 00 Patras, Greece, sypsas@upatras.gr

and not "between" them and this result holds regardless of the groups, the indices, the equivalence scales or the reference units that have been used (Mitrakos, 2004; Zografakis, Mitrakos, 2005; Mitrakos, 2008). These studies also point out that some groups have large contributions to overall poverty and inequality, such as households with farmers, pensioners or unemployed members, families with many children and especially those with household head have low educational level. The purpose of this study is to examine poverty and inequality within such groups, and more specifically to study the households that have at least one member with income from agricultural occupation, the households with three or more children, and the households whose head gets a pension from the Agricultural Insurance Organisation (OGA). We estimate poverty and inequality indices along with their confidence intervals for these groups, and test if the decrease of poverty and inequality during the years 1998/99 and 2004/05 in these groups is statistically significant or within the boundaries of the confidence interval.

In literature, many studies have shown similar results concerning vulnerable population groups for many countries. For example, Forster (1995) analyzed poverty across three different demographic groups showing that, in general, poverty among single parents is significantly higher than among families with three or more children which, in turn, have higher poverty levels than all non-elderly families. More recently, Förster και Pearson (2002) studied trends and driving forces of income distribution and poverty in OECD area. Concerning the changes in relative positions of specific social groups, they found that, in those countries where inequalities increased, this happened mostly among the working-age population, whilst there were fewer changes among the retirement-age population. Changes in income distribution generally favoured the prime-age and elderly age groups, particularly those around retirement age. Younger age groups lost ground, in particular those aged 18 to 25, reflecting delayed labour market entry. Similarly, poverty rates for the elderly fell in all but four countries, youth poverty rates increased, and child poverty rates increased slightly in a number of countries. Relative income levels of single parents and persons in workless households are very low and have worsened in a number of countries.

Tsakloglou and Papadopoulos (2001) provided a methodology for identifying population members at high risk of social exclusion using the data of the European Community Household Panel (ECHP) and highlighted similarities and differences across EU member states. They found that the highest levels of aggregate risk of social exclusion were observed in some southern countries (Portugal and Greece) and the UK and the lowest in northern and central European countries (Denmark, the Netherlands, Luxembourg and Germany). In almost all countries it was observed that the looser the links of the individual or the household with the labour market, the higher the risk of social exclusion in comparison with the rest of the population. Turning to similarities and differences across demographic groups, they found that in almost all countries children are facing a higher risk of social exclusion than the rest of the population. To a large extent, this risk is accounted by the higher than average risk of social exclusion facing children living in lone-parent households. The population share of persons living in lone-parent households varies a lot across the EU and although in all countries this group's relative risk of social exclusion is higher than average, it differs significantly across countries. At the other end of the demographic spectrum, older persons seem to face a risk of social exclusion substantially higher than average in only two southern countries (Greece and Portugal). As a result, Tsakloglou and Papadopoulos concluded that one-size-fits-all policies aimed at fighting social exclusion in Europe are not likely to have a significant impact in all countries.

Of great importance is the study of Tsakloglou (2000) who found considerable similarities regarding the level and the structure of poverty as well as the role and the impact of welfare state policies in the fight against poverty with the rest of the Mediterranean EU member-states. Comparisons were made between the four Mediterranean EU member-states and the three big EU countries (France, UK and Germany) with respect to the relative poverty risk and the contributions to aggregate poverty of eight non mutually exclusive high poverty-risk groups: households headed by farmers, unemployed, retired, all households headed by persons aged 65 or more, one-member households consisting of persons aged 65 or more, mono-parental households, households headed by females, and households headed by persons with no or low educational qualifications. Tsakloglou found that households headed by unemployed or retired persons, persons with low educational levels or females are high poverty-risk groups in all countries under examination, (are high poverty risk groups) with no significant differences between different countries other than the population shares of these groups. On the other hand, there are considerable cross-country differences for the group of households headed by farmers, unemployed persons and the group of mono-parental households. Tsakloglou found that in all Mediterranean countries (including France) the poverty risk of households headed by farmers is higher than that of the rest of the population. The opposite is observed in the UK and, especially, Germany[1]. In addition, even though households headed by unemployed persons are a high poverty-risk group in all countries under examination, the group's position *vis-à-vis* the rest of the population appears to be substantially worse in the more developed countries (Germany, UK, France). Finally, mono-parental households appear in Greece to be a low poverty-risk group[2]. The group's poverty-risk does not appear to be substantially higher than the national average in the other Mediterranean countries (including France), but it is quite high in the UK and, particularly, Germany. These results augurs well with the view

[1] However, due to their higher population share in the poorer countries, these households contribute more than 15% to aggregate poverty only in Greece and Portugal, whereas the corresponding contributions in Germany and the UK are negligible.

[2] Tsakloglou also notes that despite the lack of a comprehensive protection system for the unemployed, the poverty situation of this group is far less pronounced in Greece than in the rest of the countries. As a result of the discrepancy in the poverty risks and the fact that a considerable proportion of the unemployed in the Mediterranean countries (with the partial exception of Spain) consists of young persons living with their parents and wives of household heads, the contribution of the group to aggregate poverty appears to be substantially higher in the Northern than in the Southern EU countries.

of the authors who claim that there exists a distinct Southern European model of welfare and attribute the differences between these countries and the rest of the EU to a number of socioeconomic factors (Leibfreid, 1993; Ferrera, 1996; Gough, 1996). According to this view, in comparison with the rest of the EU, southern European countries are characterized by more "dualistic" economic structures, with relatively high employment in the agricultural sector and considerable size of the "hidden" economy.

For Greece, there are several empirical studies of the distribution of income and consumption but they use different statistical data, different indices and/or different methodology, so it is not easy to combine or compare their results. In some cases there can be contradictious results that can be explained by the different sensitivity of various indices in different income transfers or, similarly, the intersection of Lorenz curves of distributions in use. Older studies also considered income data unreliable due to extended tax evasion but also because, until recently, not everybody was obliged to submit a tax return. Other studies characterise several inequalities as justified if they are the result of normal functions of the market or unjustified if they are connected with specific interests of various social groups that force the government to achieve higher compensations or incomes (Athanasiou, 1984: pp. 56-64). More recent studies indicate education as a significant factor determinant of income inequalities (Mitrakos, Panopoulou, Tsakloglou, 2002; Mitrakos, 2004) and others conclude that inequalities "within" social groups contribute impressively to the overall inequality than inequalities "between" these groups (Zografakis, Mitrakos, 2005). So, policies attempting to eliminate poverty and inequalities within social groups are more effective than those that attempt the elimination of inequality between groups[3].

In brief, it is clear that, the choice of source and the level of analysis of data (personal or grouping data, etc.) as well as methodological and other choices, such as the reference unit of analysis (household or individual), the equivalence scales used, poverty and inequality indices, etc., are determinants for the results. Some characteristics of poverty and inequality remain valid over time, such as education and unemployment, but others like the residence in rural or urban region have stopped being significant factors determinants of income inequalities (Mitrakos and Tsakloglou, 1998; Tsakloglou, 2000).

2. Methodology

The present work uses household income micro-data from the last two Household Budget Surveys (HBS) (1998/99 and 2004/05) which cover the entire population of Greece conducted by the National Statistical Service of Greece (NSSG). The surveys cover all the non-institutional households of the country and their sampling fraction is 2/1000 (it consists approximately of 6,500 households and 17,500 members). The material contains detailed information about consumption expenditures (actual and imputed), incomes after taxes, social security contributions and transfer payments, socio-economic characteristics of the

[3] For a general review of poverty and inequality in Greece, see, Studies 55, KEPE, 2004, *Distribution, Redistribution and Poverty.*

households and their members as well as information on a number of housing amenities and consumer durables owned by the household[4].

We chose the individual as the unit of analysis and took into account the differences in size and composition of the household and the differences in needs between children and adults by using “family equivalence scales”. The equivalence scales used, are given by the Statistical Office of the European Union (EUROSTAT), and assign a weight of 1.0 to the household head, a weight of 0.5 to each of the remaining adults and a weight of 0.3 to each child (person aged up to 13) in the household[5]. In comparison with other sets of equivalence scales used in empirical distributional studies, these scales are in the middle of the range regarding the economies of scales they imply[6]. These weights are used to estimate the number of equivalent adults of the household. Finally, we use the distribution of total equivalent income per capita that is obtained by dividing total household income by the number of equivalent adults and assigning the result to every member of the household.

The next step is to choose the poverty and inequality indices. We consider someone poor when his income (or consumption) falls below a predefined level called “poverty line”. A poverty line can be defined as the necessary income to cover the basic needs for survival (food, clothing and residence) or as the necessary income to obtain the minimum socially acceptable level of living. In the first case we refer to “absolute” poverty where emphasis is given to economic deficiency, though in the last case, the “relative” poverty, emphasis is given to economic inequality. The relative approach defines poverty line as a fraction of the median income (or expenditure) in the society because the mean or median income reflects the usual level of living conditions, so this approach defines the minimum amount for a tolerable life in the society. The main difference of these two approaches is that the relative poverty line increases with the same rate as the average level of living conditions unlike absolute poverty line. For Greek economy, where large population groups have ensure the necessary goods to maintain and reproduce life, conventional poverty approach seems more appropriate[7]. In this study we used as poverty line a fraction (40%, 50%, 60% and 70%) of the median of equivalent income.

In this study, we use poverty percentage (P), which is the proportion of the population falling below the poverty line, and Foster index (F) of Foster, Greer and Thorebecke (1984).

---

[4] Some researchers believe that it is better to “clean” primary HBS data before using them (Tsakloglou, 1990; Tsakloglou, 1996; Mitrakos, 2000). These corrections (deflation, different time correspondence, zero incomes and outliers, sample weights, consumption and income definition) concern both consumption expenditures and income and intent to lead to a better approach of real welfare level of households or individuals through common poverty and inequality indices. These corrections also smooth primary data variations and so decrease total inequality level based on total expenditure. For this reason, in this study, no correction is made.

[5] This scale was first proposed by Haagenars et al. (1994) and was adopted in the late 1990s by EUROSTAT as “OECD-modified equivalence scale”.

[6] See, Buhmann et al. (1988); Coulter et al. (1992); Banks and Johnson (1994); Blundell and Lewbel (1991); Burkhauser et al. (1996); Mitrakos (2000).

[7] EUROSTAT also supports the use of the conventional poverty approach.

Index F satisfies all basic axioms set by Sen (1976) for poverty indices (focus axiom, monotonicity axiom, transfer axiom, symmetry axiom, mean independence axiom and independence of population size axiom) and is sensitive to the magnitude and intensity of poverty gap as well as the distribution of resources among poor people. The parameter α of index F, called the poverty aversion parameter, takes value α=1, that corresponds to poverty gap (the average distance of poor incomes from the poverty line as a fraction of the poverty line), and value α=2 that gives greater weight to the larger distances from the poverty line and so incorporates the social aversion for extreme poverty. On the contrary, index P violates some of the desired properties for a poverty index (symmetry, mean independence, population size independence, principle of transfers among poor people, focus axiom, and monotony) but provides a clear indication of the magnitude of poverty in population. In particular, index P does not take into account how much poor those people are, but careful use of it and many different poverty lines give a wide image of poverty.

Inequality describes the unequal distribution of wealth or income to the members of a population. It is the opposite of "equality" where everyone gets equal shares of whatever is to be shared (income, wealth, etc.). Inequality was first described using Lorenz curves, but in most cases Lorenz curves intersect and that prohibits the ranking of distributions, so we have to select the appropriate index to measure inequality. An inequality index is a statistic that incorporates the characteristics and the variance of a distribution and corresponds, directly or indirectly, to a social welfare function. It is known that since there is no single social welfare function accepted by all economists, there cannot be a unique inequality index considered to be the best of all, thus different indices can lead to different results. However, the different sensitivity of various indices to different types of transfers, allows us to choose which indices to use with regard to our research and the target groups of the population in which we study inequality.

Since the selection of inequality indices is, to a certain point, subjective, we chose six inequality indices that satisfy all the desired axioms (symmetry, population size independence, mean or scale independence, principle of transfers) and have different sensitivity to transfers. We use the variance of the logarithms (L), the mean logarithmic divergence (N) and Atkinson index with ε = 2, that are relatively more sensitive in transfers in the bottom of the distribution, Theil index (T) and Atkinson index with ε = 0.5 that are more sensitive in transfers in the upper edge of the distribution and Gini coefficient that is more sensitive in transfers around the center of the distribution (Lambert, 1993; Cowell, 1995). We also use the deciles shares of income distribution in order to calculate the S_{80}/S_{20} index that indicates the gap between the first and the last two deciles shares.

In order to construct the confidence intervals for these indices, we use the bootstrap technique[8]. Originally, the bootstrap was suggested by Efron (1979) as a method to derive an estimate for the standard error of arbitrary estimator. The method consists of generating a large number B of samples, called bootstrap samples, and calculate in each of them the estimator or the index we want. If the primary data consists of *n* independent units, it then

[8] See, Efron (1979); Efron and Tibshirani (1993); Davison and Hinkley (1997).

suffices to take a simple random sample of size *n*, *with replacement*, from the *n* units of data, to get one bootstrap sample. The technique is called a nonparametric bootstrap if nothing is assumed (like a parametric distribution) about the underlying process that generated the data. We only assume that the data in the original sample were "representative" and that sample size was moderately large. Confidence intervals can be computed in the usual way, using either a parametric or nonparametric bootstrap, as $\bar{x} \pm 2s_x$ for a 95% confidence interval, where $\bar{x}$ is the mean value and s_x is the standard deviation of the index we study of the B bootstrap samples. It should be noted that when enough bootstrap samples have been generated, not only the standard error but any aspect of the distribution of the estimator could be estimated. Since we are interested in testing? the null hypothesis that there isn't a difference in poverty and inequality indices between the years 1998/99 and 2004/05, we construct a 95% confidence interval for the difference of each measure in the two periods, as $(\bar{x}_2 - \bar{x}_1) \pm 1.96 \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$, where $\bar{x}_1, \bar{x}_2$ are the mean values of the index we study for the years 1998/99 and 2004/05, s_1, s_2 the corresponding standard deviations, calculated of the B bootstrap samples, and n_1, n_2 are the number of bootstrap samples respectively[9]. If the value of zero is not included in the confidence interval we then reject the null hypothesis and there is, indeed, a difference between the years 1998/99 and 2004/05.

3. Empirical results

Most studies conclude that poverty and inequality decreased in Greece during the last decades, and this is also the case we observe for the years 1998/99 and 2004/05. Table 1 presents poverty and inequality indices for these years according to the distribution of total equivalence income. We observe that poverty percentages have decreased about 3 percentage units, independently of what poverty line we use, showing an important decrease in poverty in general. Most important is the decrease of 3.7% for the poverty percentage using the 40% of median equivalent income as poverty line, showing that almost half of the poorest people in 1998/99 managed to improve their economic status in 2004/05. This result is enhanced by Foster et al. index with poverty aversion parameter a=2 that decreased almost 1 percentage unit showing a decrease in extreme poverty. In addition, Foster et al. index with poverty aversion parameter a=1 decreased almost 2 percentage units, showing that poor people are closer to the poverty line in the year 2004/05 than in 1998/99.

Inequality also decreased significantly during 1998/99 and 2004/05, about 2-7 percentage units, depending on what index we use. Indices with more sensitivity to transfers in the bottom of the distribution decreased the most: Atkinson index with ε = 2 decreased 5%, mean logarithmic divergence (N) decreased 4% and the variance of the logarithms (L) decreased more than 7%. On the other hand, Theil index (T) decreased about 5% Atkinson index with ε = 0.5 decreased more than 2% and Gini coefficient decreased almost 4%.

9 Since the number of bootstrap samples B is large enough (B=300 in our case), the mean values are considered normally distributed.

These decreases indicate that the whole distribution of 2004/05 is more equal than the distribution of 1998/99. Only index S80/S20 shows the smallest decrease (about 0.5%) indicating that the relative gap between the first and the last two deciles shares did not decrease considerably.

We want to examine poverty and inequality within some groups that have large contributions to overall poverty and inequality, such as households with farmers, pensioners and families with many children. Table 2 shows the percentage (%) of people living in households with specific characteristics in years 1998/99 and 2004/05. We notice that the percentage of people living in households with three or more children decreased about 2% indicating that families are getting smaller in size through years. The greater decrease, of more than 4%, is noticed on the percentage of people living in households that have at least one member with income from agricultural occupation. This is an indication that people do not want to work in the agriculture sector and select another occupation, which often differs from that of their parents. On the other hand, the percentage of people living in households whose head gets a pension from the Agricultural Insurance Organisation (OGA) increases more than 2 percentage units, showing that old farmers become pensioners.

Table 1: Poverty and inequality indices in HBS (1998/99 and 2004/05) using total equivalence income distribution

Poverty indices		**1998/99**	**2004/05**
Poverty percentage using as poverty line	40% median	7.7	4.0
	50% median	12.9	9.4
	60% median	19.9	16.2
	70% median	27.6	25.4
Foster et al. index, a = 1 (Poverty line = 60% median)		5.9	3.8
Foster et al. index, a = 2 (Poverty line = 60% median)		2.7	1.3
Inequality indices			
Atkinson Index (A, ε = 0.5)		9.23	7.20
Atkinson Index (A, ε = 2)		30.81	25.40
Theil Index (T)		20.15	15.15
Mean logarithmic divergence (N)		18.91	14.80
Gini Coefficient (G)		33.61	29.91
Variance of logarithms (L)		36.38	29.09
S80/S20		3.56	3.05

Source: Calculated from Household Budget Survey (HBS) micro-data 1994/95 and 2004/05.

Table 2: The percentage (%) of people living in households with specific characteristics in HBS (1998/99 and 2004/05)

Characteristics of households	1998/99	2004/05
Households with three or more children	5.2	3.3
Households that have at least one member with income from rural occupation	13.9	9.7
Households whose head gets a pension from the Agricultural Insurance Organisation (OGA)	5.3	7.6

Source: HBS micro-data 1994/95 and 2004/05

For these groups, we used the bootstrap technique to evaluate standard deviation for poverty and inequality indices for both years 1998/99 and 2004/05. First, we notice that the average bootstrap values for all indices, calculated over 300 bootstrap samples, approach the corresponding values from the initial samples. This is shown in tables 3 and 4 for poverty and inequality indices, respectively and indicates that the number of bootstrap samples is enough to estimate standard deviations[10]. Only index S80/S20 differs significantly for all groups and this can be justified by the bootstrap technique.

Table 3: Poverty indices in initial samples 1998/99 and 2004/05 and their estimates from 300 bootstrap samples with the use of various poverty lines for total equivalence income distribution

Poverty Index		Households with three or more children		Households that have at least one member with income from rural occupation		Households whose head gets a pension from the Agricultural Insurance Organisation (OGA)	
		1998/99	Bootstrap	1998/99	Bootstrap	1998/99	Bootstrap
Poverty percentage using as poverty line	40% median	10.6	10.3	6.0	6.2	4.3	3.5
	50% median	13.7	13.6	11.8	11.5	7.5	8.5
	60% median	20.4	19.6	17.3	17.2	20.4	19.7
	70% median	27.2	26.1	25.3	25.1	30.7	29.6
Foster et al. index, a = 1		7.3	7.0	4.7	4.8	3.7	3.6
Foster et al. index, a = 2		3.9	3.8	1.9	2.0	1.0	1.0

10 There is a direct connection between the size of the initial sample and the number of bootstrap samples in need, in order to trust the results. Our initial samples are sufficiently large in size, so 300 bootstrap samples are enough.

		2004/05	Bootstrap	2004/05	Bootstrap	2004/05	Bootstrap
Poverty percentage using as poverty line	40% median	6.8	6.6	4.6	4 .4	0.8	0.7
	50% median	11.9	10.8	7.5	7 .7	2.3	2.2
	60% median	15.9	15.8	15.1	14 .5	6.4	6.7
	70% median	24.1	23.7	24.8	24 .4	16.9	16.6
Foster et al. index, a = 1		4.8	4.8	3.3	3.2	0.9	0.9
Foster et al. index, a = 2		2.1	2.1	1.2	1.1	0.2	0.2

Source: Calculated from HBS micro-data 1994/95 and 2004/05 and the bootstrap technique.

Table 4: Inequality indices in initial samples 1998/99 and 2004/05 and their estimates from 300 bootstrap samples for total equivalence income distribution

Inequality Index	Households with three or more children		Households that have at least one member with income from rural occupation		Households whose head gets a pension from the Agricultural Insurance Organisation (OGA)	
	1998/99	**Bootstrap**	**1998/99**	**Bootstrap**	**1998/99**	**Bootstrap**
Atkinson Index (A, ε = 0.5)	10.04	9.96	11.21	11.07	10.76	10.67
Atkinson Index (A, ε = 2)	43.45	42.13	33.90	41.07	32.14	31.98
Theil Index (T)	20.29	20.15	25.61	25.28	23.91	23.67
Mean logarithmic divergence (N)	22.54	22.30	22.11	22.39	21.61	21.47
Gini Coefficient (G)	34.63	34.42	36.02	35.88	36.56	36.36
Variance of logarithms (L)	52.17	51.20	39.43	42.32	38.64	38.47
S80/S20	4.13	7.74	3.38	6.05	3.82	6.21
	2004/05	**Bootstrap**	**2004/05**	**Bootstrap**	**2004/05**	**Bootstrap**
Atkinson Index (A, ε = 0.5)	7.51	7.38	7.35	7.36	4.68	4.69
Atkinson Index (A, ε = 2)	28.36	27.69	24.75	24.75	15.71	15.78
Theil Index (T)	15.43	15.17	15.82	15.84	10.06	10.09
Mean logarithmic divergence (N)	15.89	15.58	14.83	14.85	9.18	9.22
Gini Coefficient (G)	30.20	29.73	29.96	29.93	23.77	23.80
Variance of logarithms (L)	33.04	32.28	28.09	28.12	16.87	16.96
S80/S20	3.28	6.17	5.44	5.35	2.35	4.16

Source: Calculated from HBS micro-data 1994/95 and 2004/05 and the bootstrap technique.

Finally, Tables 5 and 6 present the mean values of all poverty and inequality indices along with their standard deviation (in brackets) as calculated from the bootstrap technique. We should bear in mind that these results measure poverty and inequality within the groups of interest and not in the entire population. In general, we observe a decrease both in poverty and inequality indices for the years 1998/99 and 2004/05, and this result is enhanced if we consider the confidence intervals for the difference of means, as well. Table 7 presents the confidence intervals for the difference of means of each index in the two periods. First of all, we notice that no confidence interval includes the value of zero, so we reject the hypothesis that poverty and inequality remained unchanged during 1998/99 and 2004/05 within these groups. This result is consistent with the general feeling that all the members of the society improved their economic position during these years, probably due to the economic development that Olympic Games brought.

Table 5: Poverty indices and their standard deviations (in brackets) for specific social groups, using various poverty lines and total equivalence income distribution in HBS 1998/99 and 2004/05

Poverty Index		**Households with three or more children**		**Households that have at least one member with income from rural occupation**		**Households whose head gets a pension from the Agricultural Insurance Organisation (OGA)**	
		1998/99	**2004/05**	**1998/99**	**2004/05**	**1998/99**	**2004/05**
Poverty percent-age using as poverty line	40% median	10.3 (2.5)	6.6 (2.5)	6.2 (0.9)	4.4 (1.0)	3.5 (1.9)	0.7 (0.5)
	50% median	13.6 (2.6)	10.8 (3.0)	11.5 (1.2)	7.7 (1.4)	8.5 (2.1)	2.2 (0.7)
	60% median	19.6 (2.9)	15.8 (3.7)	17.2 (1.4)	14.5 (2.1)	19.7 (2.3)	6.7 (1.2)
	70% median	26.1 (3.4)	23.7 (4.1)	25.1 (1.5)	24.4 (2.3)	29.6 (3.1)	16.6 (1.5)
Foster et al. index, a = 1		7.0 (1.4)	4.8 (1.3)	4.8 (0.5)	3.2 (0.5)	3.6 (0.8)	0.9 (0.2)
Foster et al. index, a = 2		3.8 (1.0)	2.1 (0.8)	2.0 (0.3)	1.1 (0.2)	1.0 (0.3)	0.2 (0.1)

Source: Calculated from HBS micro-data 1994/95 and 2004/05.

Table 6: Inequality indices and their standard deviations (in brackets), using total equivalence income distribution in HBS (1998/99 and 2004/05)

Inequality Index	Households with three or more children		Households that have at least one member with income from rural occupation		Households whose head gets a pension from the Agricultural Insurance Organisation (OGA)	
	1998/99	2004/05	1998/99	2004/05	1998/99	2004/05
Atkinson Index (A, ε = 0.5)	9.96 (1.19)	7.38 (1.17)	11.07 (1.11)	7.36 (0.62)	10.76 (1.17)	4.69 (0.53)
Atkinson Index (A, ε = 2)	42.13 (5.78)	27.69 (3.67)	41.07 (6.90)	24.75 (1.53)	32.14 (2.18)	15.78 (1.29)
Theil Index (T)	20.15 (2.58)	15. 17 (2.65)	25.28 (3.04)	15.84 (1.53)	23.91 (3.18)	10.09 (1.30)
Mean logarithmic divergence (N)	22.30 (2.96)	15. 58 (2.48)	22.39 (2.15)	14.85 (1.20)	21.61 (2.20)	9.22 (0.96)
Gini Coefficient (G)	34.42 (2.01)	29.73 (2.36)	35.88 (1.67)	29.93 (1.22)	36.56 (1.79)	23.80 (1.18)
Variance of logarithms (L)	51.20 (8.60)	32.28 (5.07)	42.32 (4.39)	28.12 (2.00)	38.64 (3.02)	16.96 (1.44)
S80/S20	7.74 (1.25)	6.17 (0.92)	6.05 (0.33)	5.35 (0.28)	3.82 (0.33)	4.16 (0.17)

Source: Calculated from HBS micro-data 1994/95 and 2004/05.

Table 7: 95% Confidence Intervals for the difference of inequality and poverty indices between 1998/99 and 2004/05.

Index		**Households with three or more children**		**Households that have at least one member with income from rural occupation**		**Households whose head gets a pension from the Agricultural Insurance Organisation (OGA)**	
		Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Poverty percentage using as poverty line	40% median	-4.1	-3.2	-1.9	-1.6	-3.1	-2.6
	50% median	-3.2	-2.4	-4.1	-3.7	-6.6	-6.1
	60% median	-4.4	-3.3	-3.0	-2.4	-13.3	-12.7
	70% median	-3.0	-1.8	-1.1	-0.4	-13.4	-12.6
Foster et al. index, a = 1		-2.5	-2.1	-1.6	-1.5	-2.8	-2.6
Foster et al. index, a = 2		-1.9	-1.6	-0.9	-0.9	-0.8	-0.7
Atkinson Index (A, ε = 0.5)		-2.77	-2.39	-3.86	-3.57	-6.12	-5.83
Atkinson Index (A, ε = 2)		-15.22	-13.67	-17.11	-15.51	-16.48	-15.91
Theil Index (T)		-5.40	-4.56	-9.83	-9.06	-13.96	-13.18
Mean logarithmic divergence (N)		-7.16	-6.28	-7.82	-7.26	-12.52	-11.97
Gini Coefficient (G)		-5.03	-4.33	-6.19	-5.72	-12.80	-12.31
Variance of logarithms (L)		-20.05	-17.79	-14.74	-13.65	-21.88	-21.12
S80/S20		-1.75	-1.39	-0.75	-0.65	-2.10	-2.02

Source: Calculated from HBS micro-data 1994/95 and 2004/05 and the bootstrap technique.

More specifically, we notice that, the confidence interval of the difference of means is farther than the value of zero for Atkinson index with ε=2, the variance of logarithms and the mean logarithmic divergence. Since these indices are more sensitive to the lower bottom of the distribution that means that people in the lower bottom of the distribution improved substantially their economic status towards equality in all three groups we examined. On the other hand, Foster et al. index with a=2 and S80/S20 index have the smallest distance of the value of zero, indicating that the poorest people in all three groups have barely improved their economic status while the relative gap between the first and the last two deciles shares had the smaller decrement during these years.

People living in households with three or more children improved their performance to poverty more than people living in households that have at least one member with income from rural occupation, while the latter improved more their performance to the former to inequality, according to the distance of corresponding confidence intervals from the value of zero. We have to notice, though, that both these groups decreased their population shares through 1998/99 and 2004/05 so the alleviation from poverty and inequality their members felt through time was more intense.

The group that improved its economic position without doubt is people living in households whose head gets a pension from the Agricultural Insurance Organisation (OGA). Within this group poverty was reduced immensely, and inequality decreased even in half for most indices, so the confidence intervals for the difference of means are afar the value of zero. This result is consistent with applied policies during that period of time, which aimed to alleviate the poorer people, such as pensioners. In fact, pensions from OGA were indeed among the lowest pensions in Greece, and in most cases remain small. Our results show not only that the increase in these pensions seems to be adequate to alleviate more than half of poor pensioners and to provide them a better quality of life but also that these policies were well targeted since they almost eliminated extreme poverty within this group (Foster et al. index with poverty aversion parameter 2 = 0.2). Despite the increase in the population of this group by 2.3% between 1998/99 and 2004/05, these policies improved the economic status of all its members driving them over the poverty line, making their income distribution more equal and even reduced considerably the relative gap between the first and the last two deciles shares.

In conclusion, all three groups examined in this study, namely people living in households with three or more children, in households that have at least one member with income from rural occupation or in households whose head gets a pension from the Agricultural Insurance Organisation (OGA), have gained a better quality of life, but still experience considerable inequalities among them between 1998/99 and 2004/05. The first group of larger families have decreased inequality within the group but poverty remains high, the second group of households with members in rural sector are in better place, having reduced not only inequality but also some poverty indices and the last group of pensioners gained considerably in living conditions, reducing more than in half both poverty and inequality.

4. Conclusion

The purpose of this study was to examine poverty and inequality within the households that have at least one member with income from agricultural occupation, the households with three or more children, and the households whose head gets a pension from the Agricultural Insurance Organisation (OGA) through the years 1998/99 and 2004/05. We estimated a variety of poverty and inequality indices along with their confidence intervals for these groups, and constructed the confidence intervals for the difference of means for each index, using the bootstrap technique. Then we tested if the decrease of poverty and inequality is statistically significant and, in every case, we rejected the hypothesis that poverty and inequality remain unchanged through 1998/99 and 2004/05 in these groups.

People living in households with three or more children improved their performance to poverty more than people living in households that have at least one member with income from rural occupation, while the latter improved (more) their performance in relation to the former regarding inequality. The group of households whose head gets a pension from the Agricultural Insurance Organisation (OGA) experienced a significant turn towards equality and defeated poverty decisively. This result is consistent with applied policies during that period of time, who aimed to alleviate the poorer people, such as pensioners. Our results show not only that the increase in these pensions seems to be adequate to alleviate more than half of poor pensioners from OGA and to provide them with a better quality of life but also that these policies were well targeted since they almost eliminated extreme poverty within this group, despite the increase of the population of this group between 1998/99 and 2004/05.

Since Greece joined the European Economic and Monetary Union (EMU), Greek economy experienced a significant growth showing a 4.1% increment in real GDP versus the corresponding 1.9% for the rest of the Euro zone for the years 2000-2005. This explains the decrement in poverty and inequality we observed for the entire population, driving to the conclusion that even the poorest population groups could have gained from this economic growth. Our tests confirm that, at least to the three population groups we examined, this was the case.

References

Athanasiou, L., 1984, *The distribution of income in Greece*, Center of Planning and Economic Research, Scientific Studies, 6, Athens (in Greek).

Banks, J., Johnson P., 1994, 'Equivalence scale relativities revisiting', *The Economic Journal*, 104, pp. 233-257.

Blundell, R., Lewbel, A., 1991, 'The information content of equivalence scales', *Journal of Econometrics*, 50, pp. 49-68.

Buhmann, B., Rainwater, L., Schmaus, G., Smeeding, T., 1988, 'Equivalence scales, well-being, inequality and poverty: Estimates across ten countries using the LIS database', *Review of Income and Wealth,* 34, pp. 253-272.

Burkhauser, R., Smeeding, T., Merz, J., 1996, 'Relative inequality and poverty in Germany and the United States using alternative equivalence scales', *Review of Income and Wealth*, 42, 4, pp. 381-400.

Coulter, F., Cowell, F., Jenkins, St., 1992, 'Equivalence Scale Relativities and the Extent of Inequality and Poverty', *The Economic Journal*, 102, 414, pp. 1067-1082.

Cowell, F.A., 1995, *Measuring inequality*, (2nd edition), LSE Handbooks in Economics Series, London.

Davison, A. C., Hinkley, D. V., 1997, *Bootstrap methods and their application*, NY, Cambridge University Press.

Efron B., 1979, Bootstrap Methods: Another look at the jacknife, *The annals of statistics*, 7, 1, pp. 1-26.

Efron, B., Tibshirani R., 1993, *An introduction to the bootstrap*, New York: Chapman & Hall.

Ferrera, M., 1996, 'The "southern model" of welfare in social Europe', *Journal of European Social Policy*, 6, I, pp.17-37.

Forster, M., 1995, 'Measurement of Low Incomes and Poverty in A Perspective of International Comparisons', *OECD Labour Market and Social Policy Occasional Papers*, 14.

Förster, M., Pearson, M., 2002, 'Income distribution and poverty in the OECD area: Trends and Driving Forces', *OECD Economic Studies No. 34*, 2002/I

Foster, J., Greer, J., Thorebecke, E., 1984, 'A class of Decomposable Poverty Measures, *Econometrica*, 52, 3, pp. 761-766.

Gough, I., 1996, 'Social assistance in Southern Europe', *Southern European Society and Politics*, I,I, pp. 1-13.

Hagenaars, A., de Vos K., Zaidi M. A., 1994, *Poverty Statistics in the Late 1980s: Research Based on Micro-data*, Office for Official Publications of the European Communities. Luxembourg.

Kanellopoulos, C., Anastasakou, Z., Kotsi, A., Maniatis, T., Pachaki, K., 2004, *'Distribution, Redistribution and Poverty*', Center of Planning and Economic Research, Scientific Studies No 55, Athens (in Greek).

Lambert, P.J., 1993, *The distribution and redistribution of income: A mathematical analysis*, (2nd edition), Manchester University Press, Manchester.

Leibfreid, S., 1993, Towards a European welfare state? In C. Jones (ed.), *New Perspectives on the Welfare State in Europe*, London: Routledge.

Mitrakos, T., 2000, *Economic Inequalities in Greece: 'Measurement, Structure, Intertemporal changes and International comparisons',* Ph.D. Thesis, Economic University of Athens.

Mitrakos, T., 2004, 'Education and Economic Inequalities', *Economic Bulletin of Bank of Greece*, 23, pp. 33-55.

Mitrakos, T., 2008, 'Child poverty: recent developments and determinants', *Economic Bulletin of Bank of Greece*, 30, pp. 67-100.

Mitrakos, T., Panopoulou G., Tsakloglou P., 2002, 'Economic Inequalities and poverty in Greece: Quantitative dimensions of the problem', in T. Pakos (ed.) *Society of 2/3*, pp. 255-78, Panteion University of Athens (in Greek).

Mitrakos,, T., Tsakloglou P., 2000, Changes in aggregate inequality and poverty in Greece after the restoration of democracy, *Meletes Ekonomikis Politikis,* 5, pp. 1–52, (in Greek).

Sen, A., 1976, 'Poverty: An Ordinal Approach to Measurement', *Econometrica,* 44, pp. 219-231.

Tsakloglou, P., 1990, 'Aspects of poverty in Greece', *Review of Income and Wealth,* 36, 4, pp. 381-402.

Tsakloglou, P., 1996, 'Ghanges in inequality in Greece in the 1970s and the 1980s', in E. Palmer, B. Gustafson and P. Gottschalk (editors). *Changing patterns in the Distribution of Economic Welfare: What happened during the 1980s?*, Cambridge, Cambridge University Press.

Tsakloglou, P., 2000, 'Poverty and anti-poverty policies in Greece and a comparison with other mediterranean EU member states', in A. Mitsos and E. Mossialos (eds) *Contemporary Greece and Europe*, pp. 331-355, Ashgate, Aldershot etc., 2000.

Tsakloglou, P., Papadopoulos, F., 2001, Identifying Population Groups at High Risk of Social Exclusion: Evidence from the ECHP, *Discussion paper series IZA* DP No. 392.

Zografakis, S., Mitrakos, Th., 2006, 'Social indicators and immigration in Greece', *International and European Policy*, Papazissis publishers, 1, pp. 124-51, Athens (in Greek).

Compartmentalising Gold Prices

Rohnn Sanderson

Abstract

Deriving a functional form for a series of prices over time is difficult. It is common to assume some linearly estimable form for prediction purposes. While this can produce accurate short run forecasts it fails to identify longer trends and patterns that may exist in financial data. Particularly troublesome is the potential for chaotic behaviour which can look like standard autocorrelation. Also, components of a price's behaviour may not be linear or may be unable to be structured well in a stationary series. Recently, more research has been devoted to whether or not a series of prices exhibits deterministic behaviour, instead of some type of Brownian Motion (regular or fractal). This research suggests that some time series data may pass typical tests for randomness where randomness does not exist. Given the breadth of current research, the most logical and reasonable beginning assumption for modeling a time series is that data probably exhibit both deterministic and random components. This paper will make use of the techniques of spectral analysis and the Hurst Exponent to measure the level of long-run dependence in the price data of gold. This technique will allow for the separation and quantification of how large the deterministic and random components of gold prices are.

Keywords: Dynamic Systems, Hurst Exponent, Spectral Analysis, Industrial Organisation

JEL Classification: C5, G1, L1

1. Introduction

When modelling price movements, it is common to use a random walk framework. The random walk assumption limits modelling changes in prices over time primarily to using auto regressive and moving average processes. While this technique offers strong short-term forecasting, it cannot offer much of a description about how and why prices are changing over time, aside from correlation to past prices. Since ARIMA (Auto-Regressive Integrated Moving Average) models remove elements of long-term relationships in order to make the data stationary, we lose quantification of the long-run elements in a time series

[1] Assistant Professor Economics/Finance W.H Thompson School of Business, Brescia University, 717 Frederica St., Owensboro KY 42301, rohnn.sanderson@brescia.edu

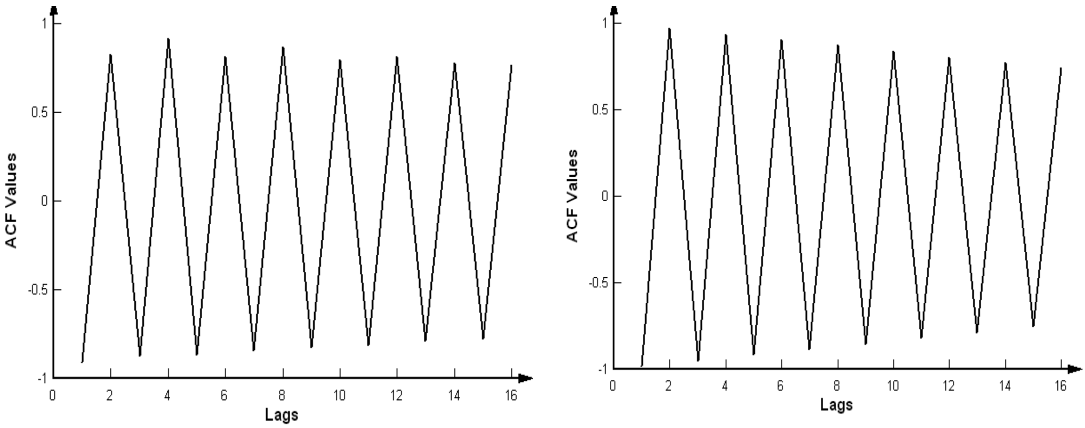
and move those elements into the error term. The complexity of modelling long-term relationships is compounded by the fact that if a dynamic economic system is non-linear, the system may be complex enough to pass standard tests of randomness and become misidentified (Baumol and Benhabib, 1989). The misidentification is usually due to the fact that deterministic processes have infinite variance and we tend to remove infinite variance from data, most commonly in the form of differencing. This critique is not novel, as many researchers have been concerned with measuring longer stationary cycles in dynamic economic systems (Baumol and Benhabib, 1989; Fama and French, 1988; Hsieh, 1991; Lo, 1991; Mayfield and Mizrach, 1992). One determinant of price path behaviour is long run dependence (i.e. memory/history). Highly persistent behaviour in economic systems can lead to events farther back in the history of the series that continue to have an influence on today's prices. The persistence of a series can be measured through the calculation of a Hurst Exponent (Hurst, 1951). While the Hurst Exponent does not definitively determine whether or not a system is linear or non-linear it does aid in our understanding of how the series will propagate in the future.

Gold, not unlike other financial instruments, is subject to memory (long-term cycles). Memory in a financial process implies that the history of prices partially dictates how prices fluctuate in the future. The presence of memory can mean that deterministic behaviour is present in a system. Deterministic behaviour is typically multiplicative in nature. Again this idea is not new to the literature, as many systems have been tested for deterministic and chaotic behaviour. Deterministic behaviour is just one component of the workings of gold prices, there is also a random component as well. This article will take advantage of the Hurst Exponent and a space-time regression in order to separate the deterministic from the random component of the changes in price. Throughout this article, deterministic components may be referred to as endogenous and random components may be referred to as exogenous. This is because the deterministic component represents intra-industry changes that have an effect on the market price, whereas the random component represents the effect on the market price from outside influences. Separation of the two components of the market price can allow us to see the extent to which market structure and macroeconomic changes affect price.

2. Analysis of Components of Gold Prices

Before we start with any data analysis let us get to the root of the long term memory problem. In Figure 1 there are ACF's for two different series: one is a deterministic chaotic logistic equation, the other is Regular Brownian Motion.

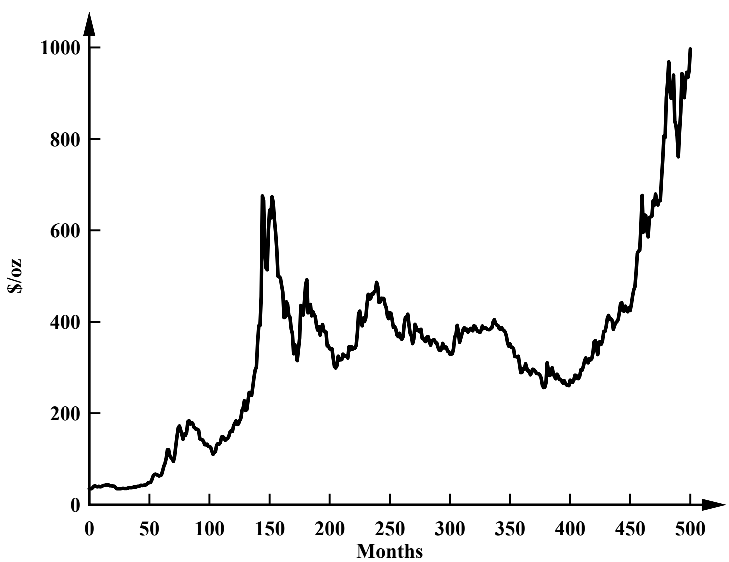
Figure 1: Comparison of a Deterministic and Random ACF



The first panel is the ACF for the chaotic logistic data and the second is the Regular Brownian Motion. Although there are slight differences, there is relatively little that distinguishes the two. In modelling we may see poor performance from an ARIMA model with chaotic data, but that is all. It should be kept in mind that there are many functional forms that can produce chaotic behaviour aside from a relatively simple logistic function. That is the motivation behind looking at the series in its entirety for memory prior to any modification.

The first dataset used in this study is the average monthly gold price per ounce from January of 1968 to October of 2009. The data are displayed in Figure 2 below.

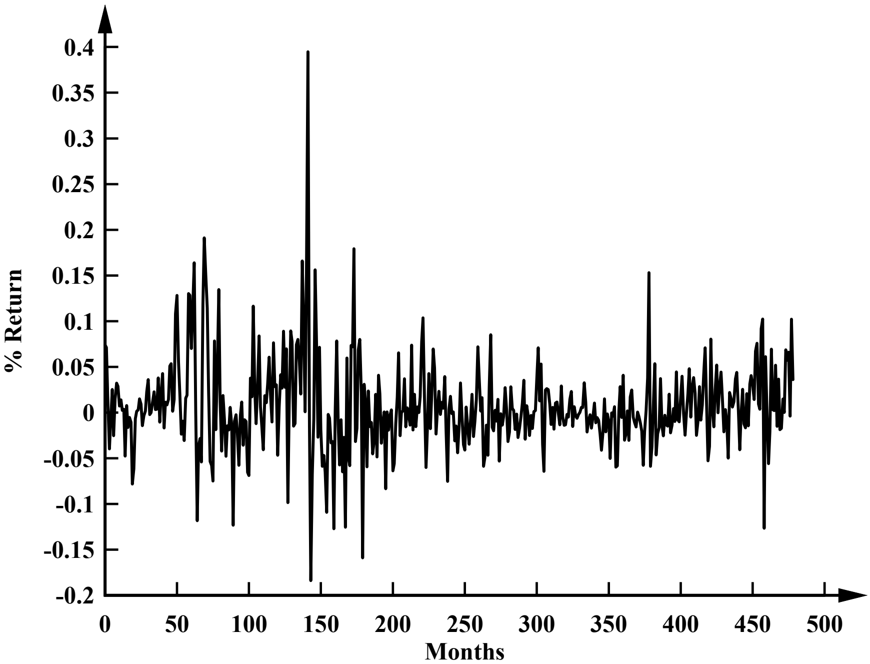
Figure 2: Average Price of Gold by Month (1968-2009)



Source: (Kitco, 2009)

As we would expect, there is a trend and some randomness in the series. Removing the trend by looking at the data of the percentage change in the price shows the stationary form of the data (Figure 3).

Figure 3: Average Monthly Return in Gold Price (1968-2009)



The existence of heteroskedasticity in the error term is also common in many financial series. A test of the squared residuals in both the stationary and non-stationary data shows the presence of heteroskedasticity in the error term. Although we will not do an ARCH model in this article for the sake of brevity, there are most definitely modeling techniques with an ARCH process that could model this behaviour. However we would still not be allowing for the possibility of an infinite variance process. The ACF and PACF plots as well as the non-stationarity of the data would suggest an ARIMA model of order (1,1,1). The results of the ARIMA(1,1,1) in Table 1 are as follows:

Table 1

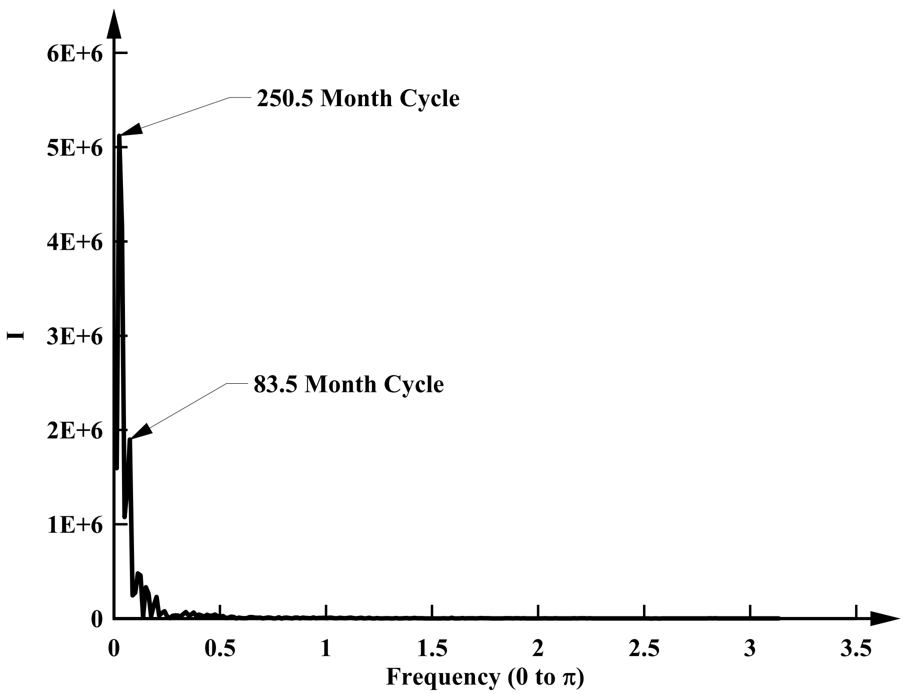
Variable	Coefficient	Std. Error	t-statistic	Significance
Constant	2.6864	1.5295	1.7564	[0.0797]
AR(1)	-.4718	0.1058	-4.4599	[0.0000]
MA(1)	0.7314	0.0834	8.7732	[0.0000[
R^2 = 98.63%	SE = 19.23			

Using a modelling approach such as ARIMA does not measure or investigate the existence of a deterministic long-run cycle with infinite variance. It should also be noted that the Hurst Exponent that we will use to estimate the amount of memory in the series is linked to the differencing parameter in an ARIFMIA model where the Hurst Exponent is equal to 1-d. Instead, we will start our analysis with the reverse question; is there memory or long-run cycles in the data?

To discover long-run cycles we want to impose as little of a functional form as possible and avoid averaging, differencing and the like. Although there are many directions that can be taken to accomplish this we will use a spectral analysis to test for the existence of long-run cycles, due to the acceptance of the technique (Clegg, 2005; Clegg, 2006; Sarker, 2007; Smith, 1992; Stone, Lewi, Landon and May, 1996). Then, the persistence of the memory in the system will be measured via the Hurst Exponent. This will provide a quantification of the level of the long-run effects.

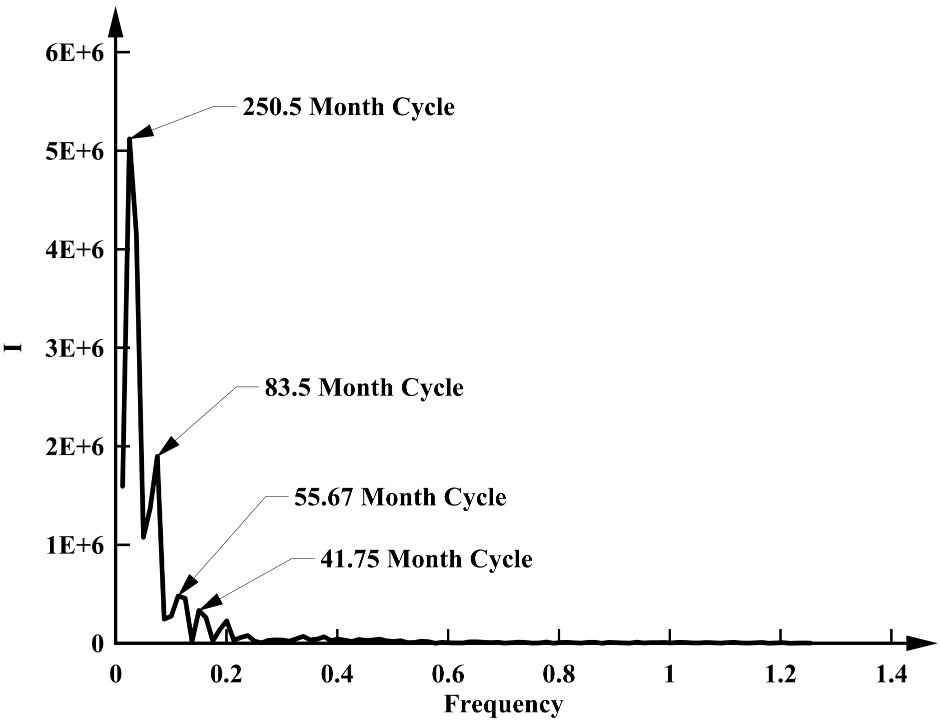
To determine if periodic components exist, a traditional spectral analysis will be used. We will not use a stationary series for this analysis because we want to allow for the possibility of infinite variance in the deterministic process. Instead, we will separate the deterministic and random components by their long term memory and their linear seperability via the ACF. In Figure 4 the only significant cycles at a 5% level are at 250.5 and 83.5 months. Over the rest of the frequency domain the periodicity falls off. The results suggest that the cycles in gold prices occur over very long intervals in time.

Figure 4: Full Spectrum Periodogram for Gold Price



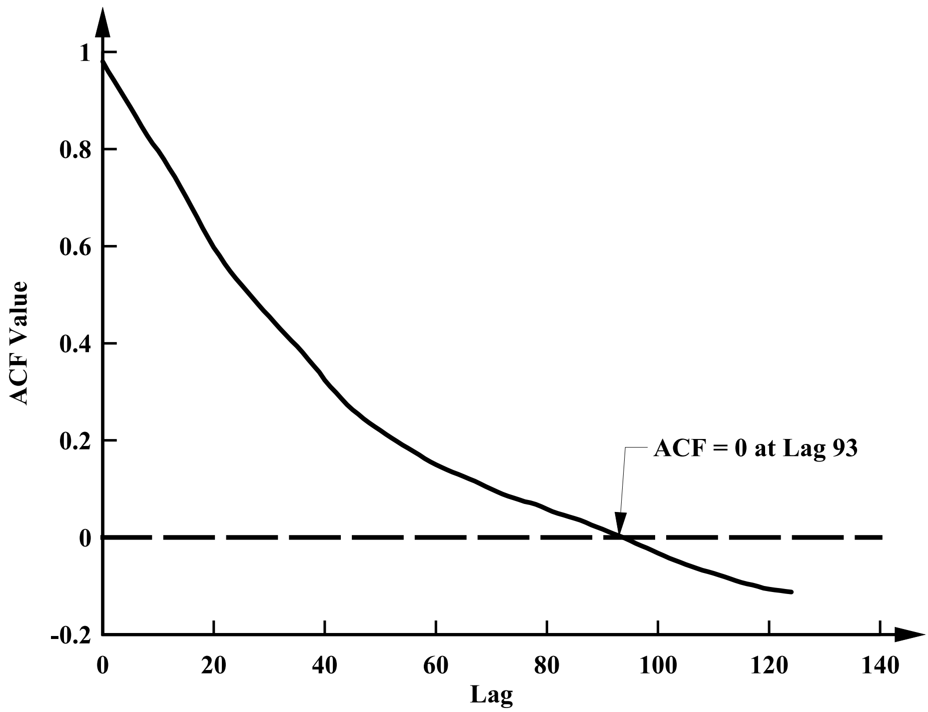
There are two more cycles that are significant at the 5% level that cannot be seen over the entire range of the frequency domain. To see all of the statistically significant cycles, the frequency window has been shortened to 1.4 (Figure 5). In addition to the previously mentioned cycles, there is also a 55.67 and a 41.75 month cycle that are significant at the 5 % level. These additional two cycles are still rather long and there is no statistically significant cycle under three years in length.

Figure 5: Shortened Periodogram for Gold Price



This demonstrates some of the problems with identifying patterns in financial data. Here we have a series that exhibits long cycles over time, which may suggest a certain amount of memory and deterministic behaviour in the system. If ARIMA modelling is used, these long-run cycles will be removed and we will not have a chance at identifying the potential for a portion of the series to have infinite variance. To identify the break between random and deterministic components in a linear fashion, we can measure the dependence through the autocorrelation function. Figure 6 displays the level of autocorrelation within the system, the ACF value does not reach zero until a lag of 93 months. This is where we will separate the data in the space-time regression by the components that have finite variance and the components that do not.

Figure 6: ACF of Gold Price Data



To measure the level of persistence in the system, the Hurst Exponent will be estimated via the periodogram method (this is to keep continuity with the previous spectral analysis). The results are shown in Table 2 and Figure 7 below. The Hurst Exponent is calculated from the regression equation results in Table 1 below.

Table 2

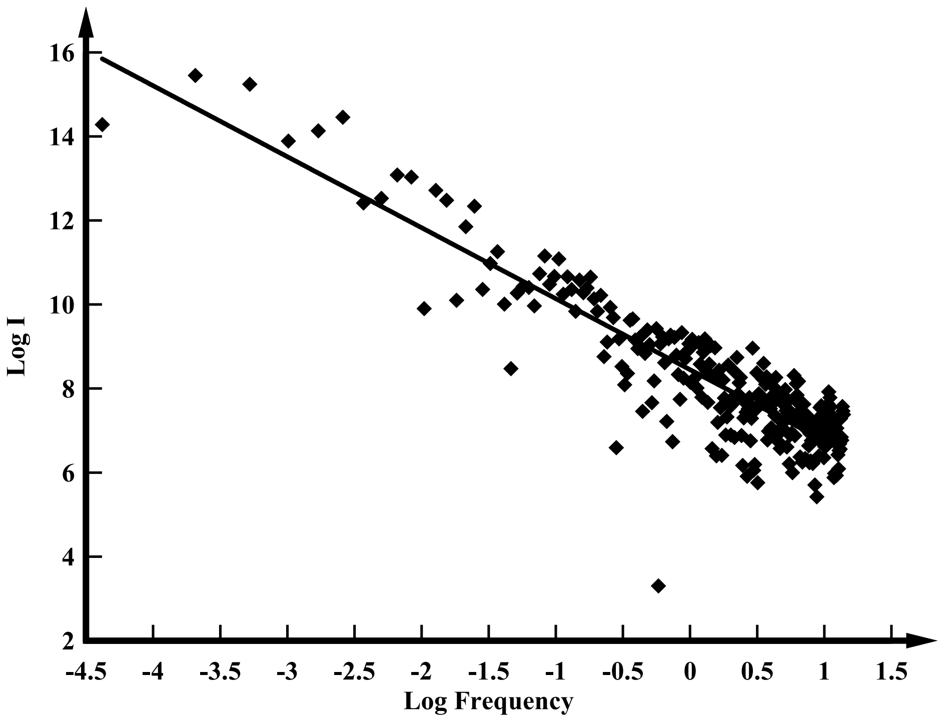
Variable	Coefficient	Std. Error	t-statistic	Significance
Constant	8.448	0.053	160.694	0.000
Log Frequency (α)	-1.689	0.054	-31.229	0.000
R^2 = 79.73%	SE = 0.82			

The Estimate of the Hurst Exponent (H) is:

$$H = {(1-a)}/{2} \qquad (1)$$

In this case, the point estimate of the Hurst Exponent is H= 1.345. Given a 95% confidence interval the Hurst Exponent has a range from 1.292 to 1.397. If the system is random (no memory) the Hurst Exponent would be equal to 0.5. This robust result confirms the presence of persistent memory in the system, meaning that history is causing some of the changes in price over time. This suggests that a portion of the structure of gold prices is deterministic.

Figure 7: Hurst Estimation of Periodogram Results

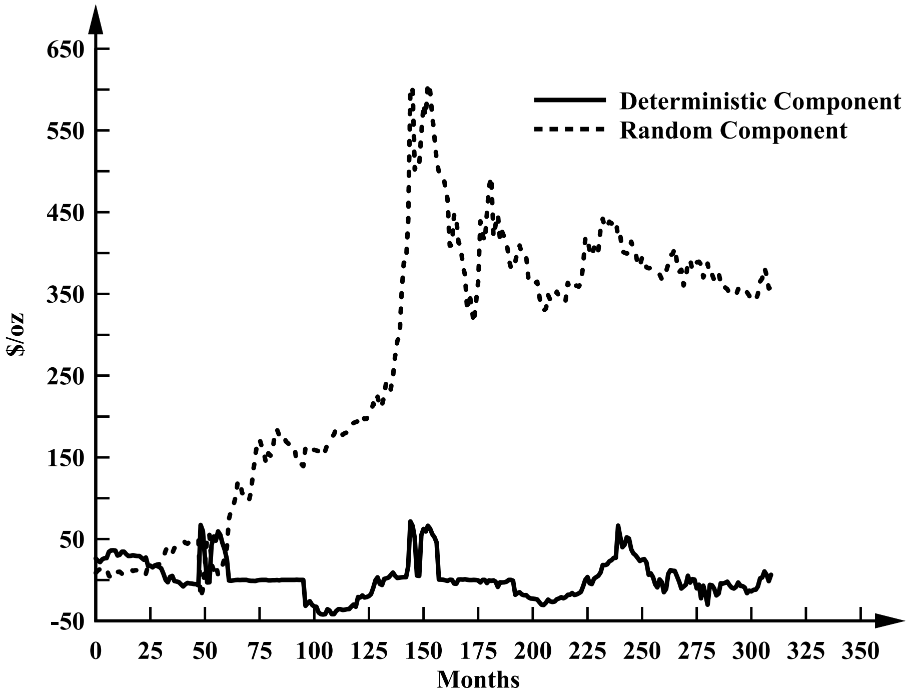


With the information gathered from the spectral analysis, autocorrelation function, and the Hurst Exponent, a space-time regression was performed in order to separate the deterministic from the random components. Since the space-time regression uses the memory information to separate components that are dimensionally independent, we can split the price data into two basic components, which sum to equal the entire signal. It should be retained that deterministic phenomena that are not dimensionally independent may have infinite variance, whereas a random phenomenon does not. Definitionally we will define the two together as the entire price, where:

$$Price_t = Deterministic_t + Random_t \tag{2}$$

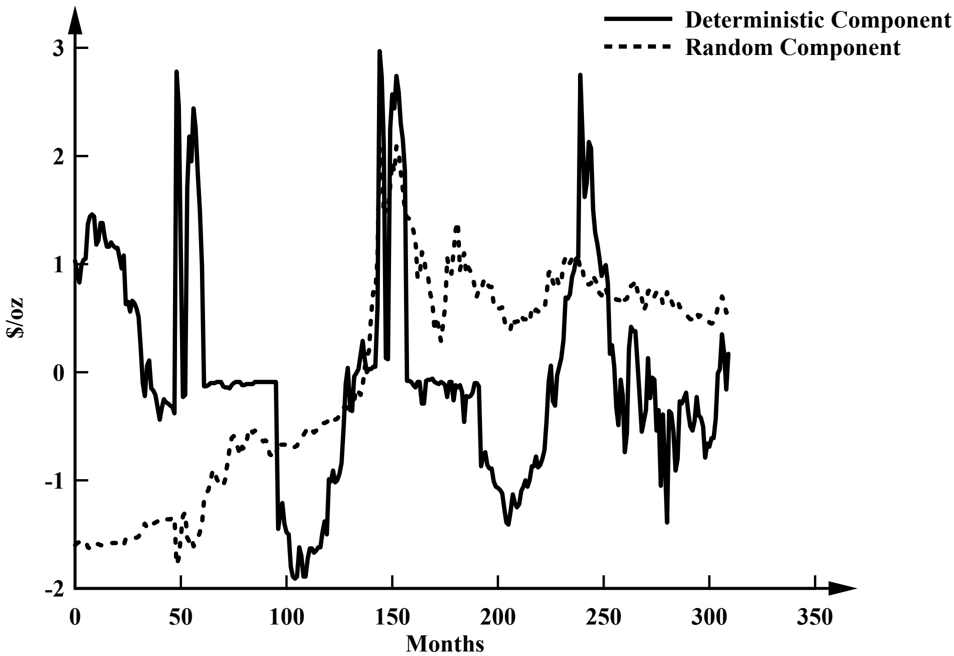
In Figure 8 below the two components of the price of gold can be seen.

Figure 8: Separated Components of Gold Price (1968-1993)



From Figure 8, it can be seen that the deterministic component is the smaller of the two components of gold price. The random (additively separable) component is the larger of the two components. This infers that most of the market price of gold is coming from exogenous events (outside of the gold industry) and that very little of the price of gold is determined by endogenous events. Thus the data suggests that the market structure of the gold industry has little impact on the market price for gold.

Figure 9: Normalized Separated Components of Gold Price (1968-1993)

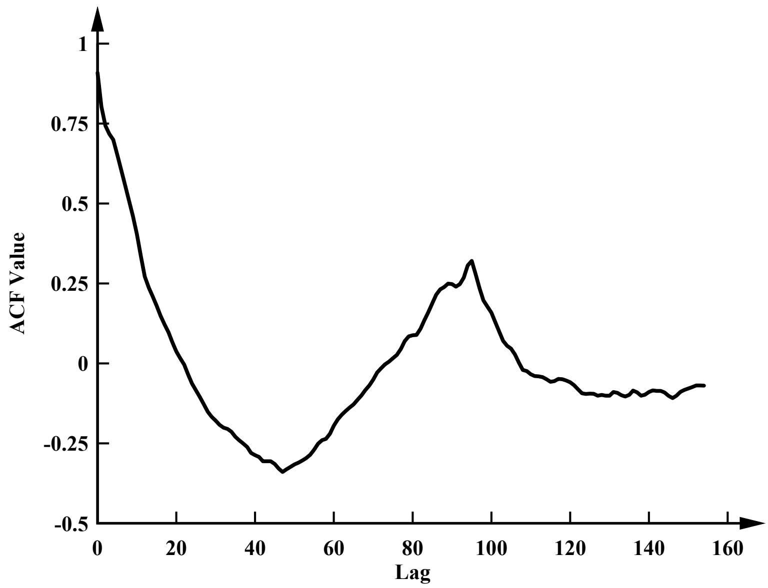


To better identify the behaviour of the two components, both series were normalized (Figure 9) to reduce the effects of scaling.

In Figure 9 it can be seen that the deterministic portion has cycling behaviour and that the additively separable component appears as if it is AR(1). Further investigation of these two components on an individual basis is necessary to determine their effect on market structure. That will not be done here as it is not the focus of this article.

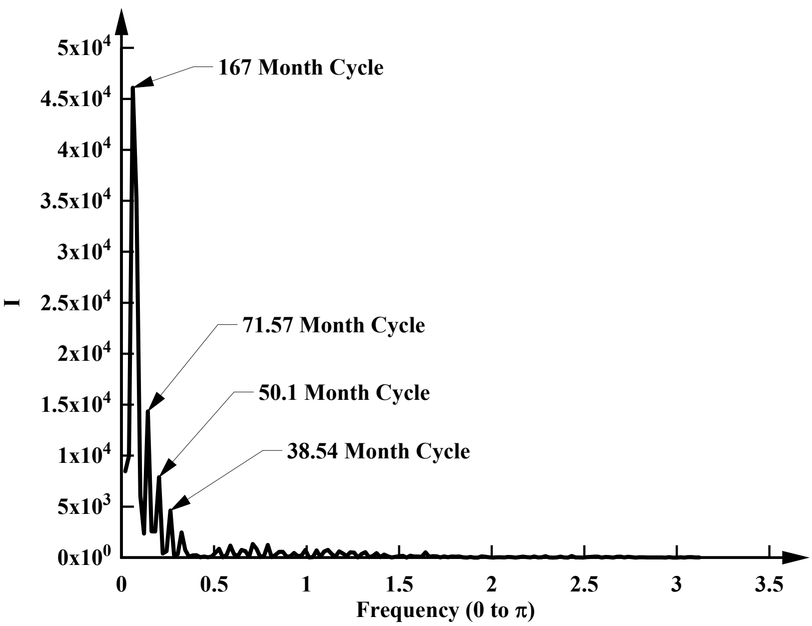
In Figure 10 below, the ACF plot of the deterministic component shows the cycling behaviour. It is important to note that we can now better identify the deterministic component.

Figure 10: ACF Plot of Deterministic Component



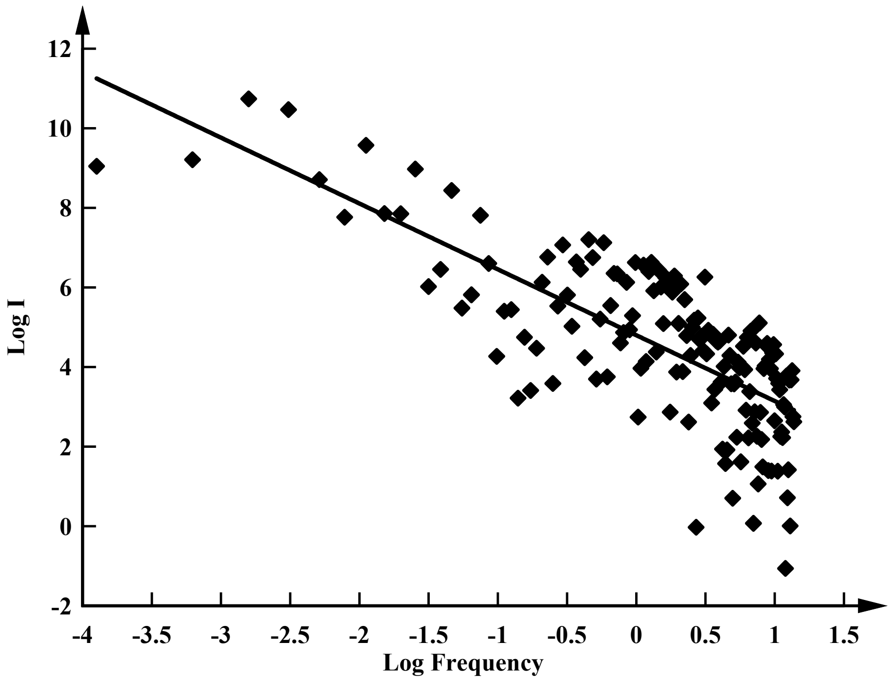
Performing a spectral analysis again on the deterministic component gives us the cycling of the deterministic behaviour (Figure 11).

Figure 11: Full Spectrum Periodogram of Deterministic Component



Now we can see four cycles that are significant at 167, 71.75, 50.1 and 38.54 months. This demonstrates that there is still long-run dependence in the system which is confirmed by another test of the Hurst Exponent (Figure 12).

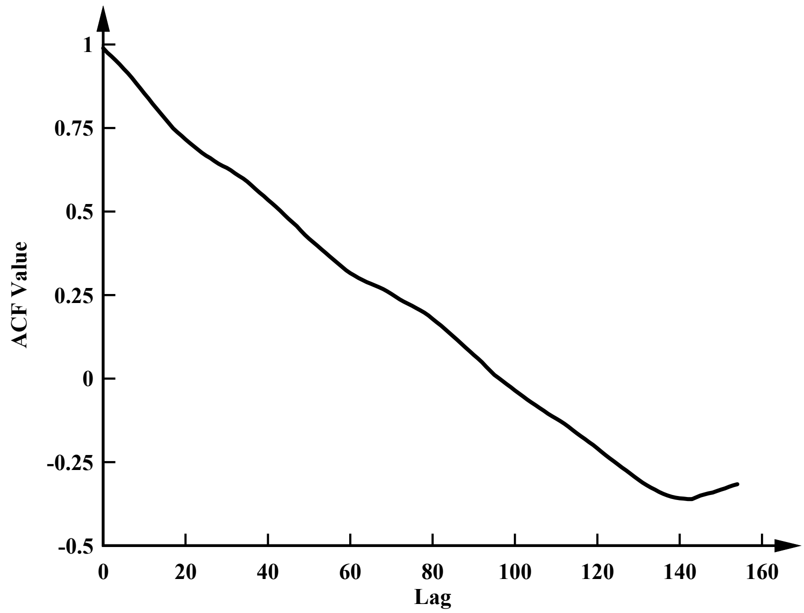
Figure 12: Hurst Estimation of Deterministic Periodogram Results



The Hurst Exponent is equal to 1.509 with a range of 1.414 to 1.604 at a 95% confidence interval, demonstrating that there is memory in the system. One conclusion that it might be drawn is that since there is memory in the endogenous portion of the gold price this displays that the industry itself is not perfectly competitive. If it was perfectly competitive, the endogenous component would attenuate to a flat signal over time. It can be concluded that although the endogenous component is small, market structure does play a role in the market price. However, as it was seen in Figure 8, the market structure impact is minimal in this case. The general result shows that changes in the market structure of the gold industry have very long-run impacts and that the market structure impacts have a small effect on the market price. From a theoretic standpoint this makes sense; although there are few sellers, there are many buyers. Therefore, it is the buyers of gold that are causing the large changes in the equilibrium price. In the case of gold prices, the exogenous component of the price has the greatest affect. Investigation of the exogenous component of the price of gold in more detail is necessary, but it is outside of the scope of this article.

In Figure 13 below, the additively separable (exogenous) component shows a series that has some autoregressive components. It should be noticed that this plot looks similar to the ACF plot of the entire series, which again reinforces the large difference in the magnitudes of the two components.

Figure 13: ACF of Random Component

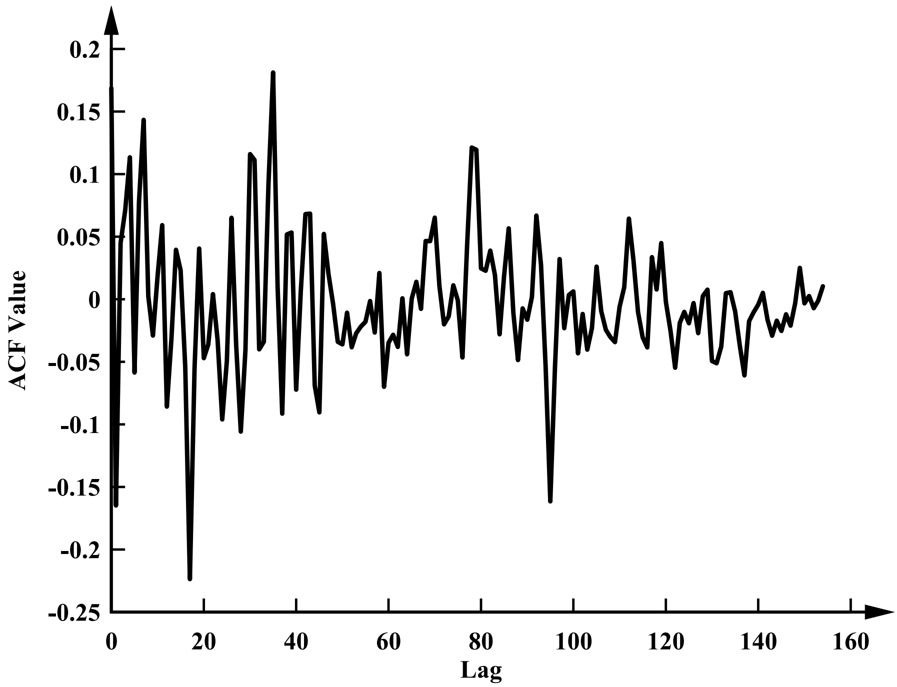


To account for the non-stationarity of the exogenous component, an AR(1) regression with a trend was performed (Table 3). This makes the exogenous data stationary, as it can be seen in Figure 14. Bias has also been removed from the estimators because the endogenous component has been removed.

Table 3: AR(1) Model of Random Component

Variable	Coefficient	Std. Error	t-statistic	Significance
Constant	47.919	78.146	0.613	0.540
Trend	1.279	0.401	3.194	0.002
AR(1)	0.981	0.011	88.121	0.000
R^2 = 98.70%	SE = 18.54			

Figure 14: ACF Plot of Residuals – AR(1) Model with Trend



The AR(1) model is commonly used for financial models of price movements over time. It is true that we could have just modelled the entire price of gold with an AR(1) model and would have obtained similar model results as to those in Table 1. However, the purpose of this technique is not about forecasting per se, it is about being able to compartmentalize prices in a way that helps determine cause and effect. In the case of gold, it was unknown a priori that the endogenous component of the price would be small. It is important to investigate these effects first before a modelling decision is made. For gold prices specifically, we have learned that there are cycles and they are very long. For forecasting purposes this may only be useful for longer time horizons. However, the market structure implications of the result are the most important. The small impact of the market structure tells us the changes in the market structure have little impact on market prices. This type of analysis is important to understand how much market structure changes will impact equilibrium prices. The small size of the endogenous component may not be the case for other commodities or precious metals; each one will need to be tested individually to better understand market structure impacts in those markets.

3. Conclusion

The price of gold has two major components, deterministic and random. In the case of gold prices the deterministic component is small relative to the random component. This suggests that industry structure has little effect on the price of gold. The preponderance of the results of the analysis concludes that external events (randomness) have the largest impact on the changing price of gold over time. This finding may have many important consequences.

For example, from an anti-trust standpoint this type of analysis can give better insights as to how mergers may affect an industry. The case of gold mergers will have very little influence on prices whereas the result may be different in other industries. It is important to note that there could be two industries with the same or similar HHI indices but with drastically different exogenous and endogenous signals that impact their respective markets differently.

From the analysis, we now know that external factors, such as business cycle events, will have a larger effect on price changes than that of intra-industry competition. There is very little that firms in the gold industry can do to alter market prices.

This paper should serve as just the beginning of a process of testing industries along these lines. More research needs to be done with this methodology on other industries to determine if there is true merit to the technique. In the appendix the same analysis is performed on two other commodities for comparison to the results on gold prices. Future extensions with respect to gold prices include determining supply and demand curves for both effects, which was not possible with this data set as the production numbers of gold have been historically unreliable. Further investigation as to what relationship exogenous and endogenous components may or may not have with the HHI and how much they vary with different industries is needed.

In terms of our understanding of long-memory processes, as well as deterministic and chaotic deterministic behaviour is concerned, further research needs to be done in economics and finance to better understand how and if we can use some of the techniques that have been developed in physics and the biological sciences. What we do know is that new dynamical system techniques are being further developed and further investigation of their validity and use in economics and finance is warranted as we continually strive to understand a really complex behaviour.

References

Baumol, W. J., and Benhabib, J., 1989, 'Chaos: significance, mechanism, and economic applications', *The Journal of Economic Perspectives*, 3, 1, pp. 77-105.

Clegg, R. G., 2006, 'A practical guide to measuring the hurst parameter', *International Journal of Simulation: Systems, Science & Technology*, 7, 2, pp. 3-14.

Clegg, R., 2005, 'A Practical Guide to Measuring the Hurst Parameter', UK Performance Engineering Workshop, Newcastle.

Fama, E. F. and French, K. R., 1988, 'Permanent and temporary components of stock prices', *Journal of Political Economy*, 96, pp. 246-273.

Hsieh, D. A., 1991, ‘Chaos and nonlinear dynamics: application to financial markets’, *The Journal of Finance*, 46, 5, pp. 1839-1877.

Hurst, H., 1951, ‘Long-term storage capacity of reservoirs’, *Transactions of the American Society of Civil Engineers*, 116, pp. 770-799.

‘Kitco - Gold Precious Metals - Buy Gold Sell Gold, Silver, Platinum - Charts, Graphs, Prices, Quotes, Gold Stocks, Mining Stocks, bullion dealers.’ Kitco - Gold Precious Metals - Buy Gold Sell Gold, Silver, Platinum - Charts, Graphs, Prices, Quotes, Gold Stocks, Mining Stocks, bullion dealers. N.p., n.d. Web. 18 Nov. 2009. <http://www.kitco.com/>.

Lo, A. W., 1991, ‘Long-term memory in stock market prices’, *Econometrica*, 59, 5, pp. 1279-1313.

Mayfield, E. S. and Mizrach, B., 1992, ‘On determining the dimension of real-time stock-price data’, *Journal of Business & Economic Statistics*, 10, 3, pp. 367-374.

Sarker, M.M.A., 2007, ‘Estimation of the self-similarity parameter in long memory processes’, *Journal of Mechanical Engineering*, 38, pp. 32-37.

Smith, L., 1992, ‘Identification and prediction of low dimensional dynamics’, *Physica D*, 58, pp. 50-76.

Stone, L., Giddy L. and May, R., 1996, ‘Detecting Time’s Arrow: A Method for Identifying Nonlinearity and Deterministic Chaos in Time-Series Data’, Proceedings: Biological Sciences, 263, 1376, pp. 1509-1513.

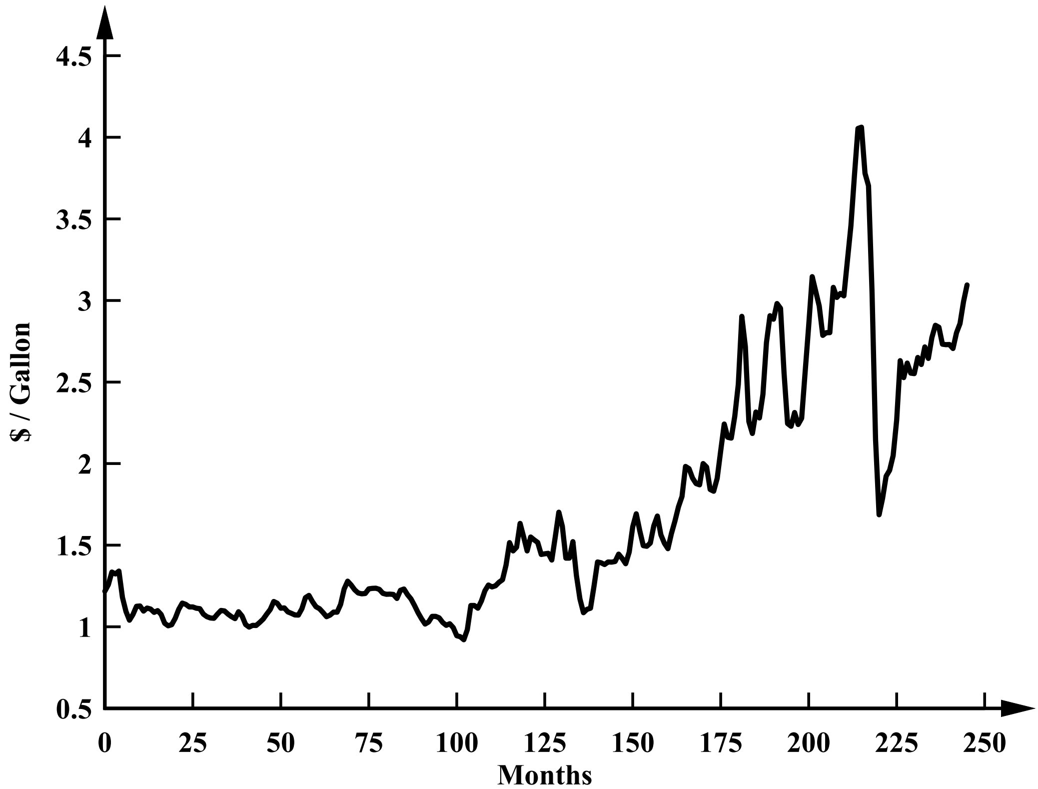
‘U.S. Price of Natural Gas Delivered to Residential Consumers (Dollars per Thousand Cubic Feet)’. U.S Energy Information Administration. N.p., n.d. Web. 21 Feb. 2011. <http://tonto.eia.gov/dnav/ng/hist/n3010us3m.htm>.

‘U.S. Regular All Formulations Retail Gasoline Prices (Dollars per Gallon)’. U.S Energy Information Administration. N.p., n.d. Web. 21 Feb. 2011. <tonto.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMM_EPMR_PTE_NUS_DPG&f=M>.

Appendix

In order to see how this technique works with other data, a similar analysis was performed on the average U.S regular formulation retail gas price from August 1990 to January 2011 (Figure 15).

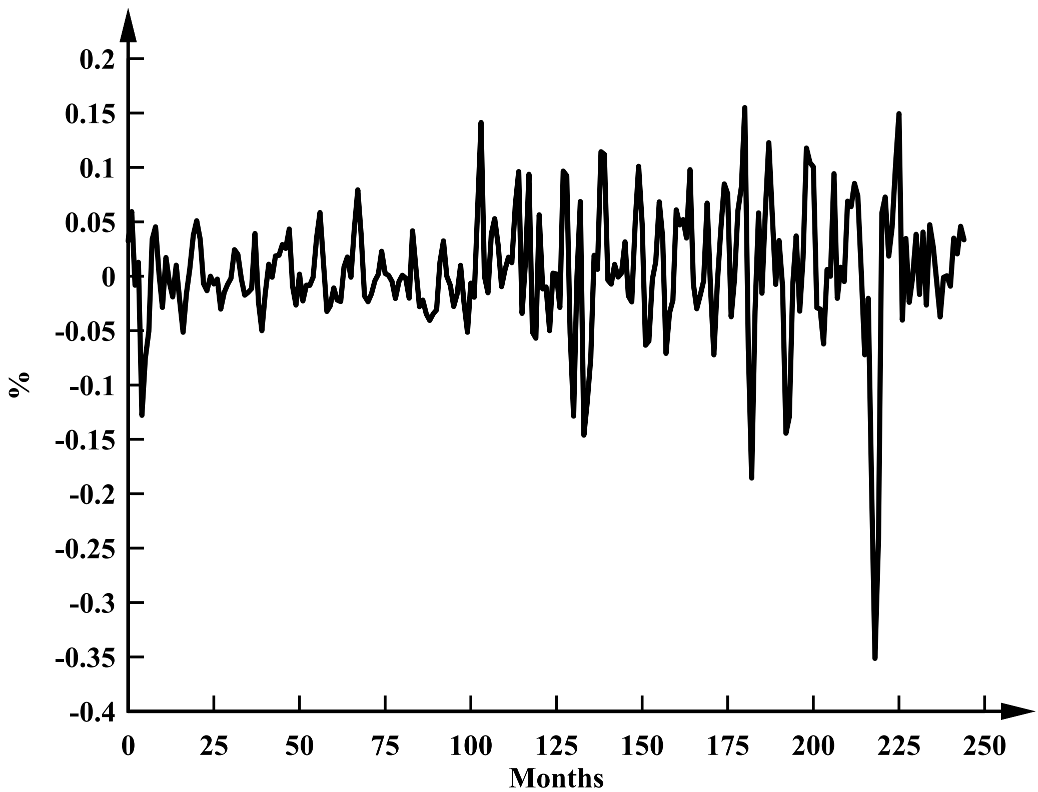
Figure 15: U.S Average Monthly Retail Gas Price (Aug 1990 – Jan 2011)



Source: (U.S Energy Information Administration, Feb-11)

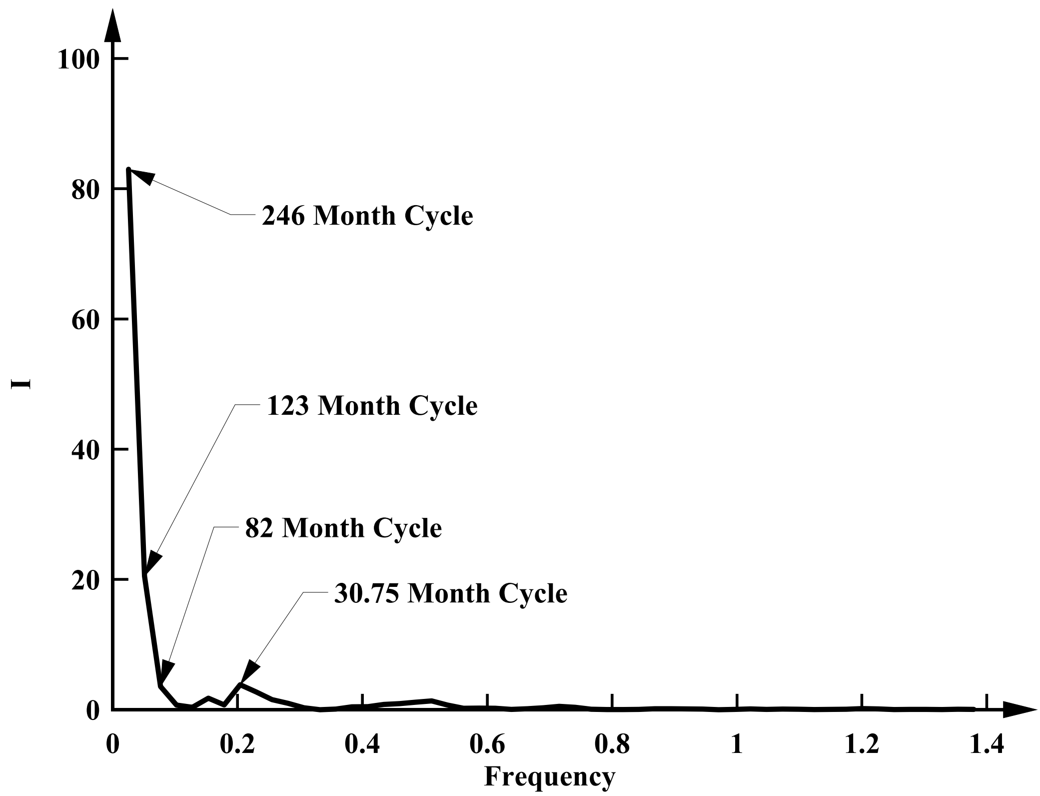
Again we can see similar data analysis problems such as the apparent heteroskedasticity in the data (Figure 16). This can be confirmed through the stationary plot below as well as with a t-test of the squared errors of the series.

Figure 16: Percentage Change in Average Monthly Retail Gas Price



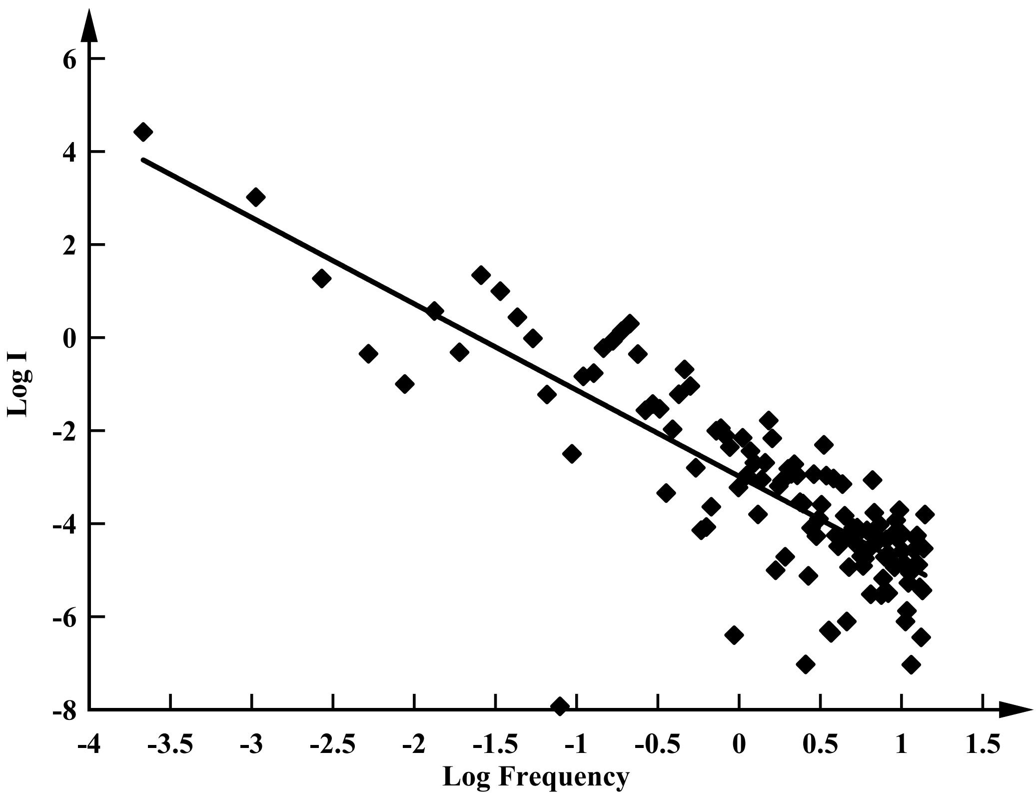
Again let us look at the original series to see what cycles may exist in the data. The following four cycles are significant at the 95% level. Again we see some long-run memory as the shortest cycle is 30.75 months (Figure 17).

Figure 17: Periodogram of Average Retail Gas Price



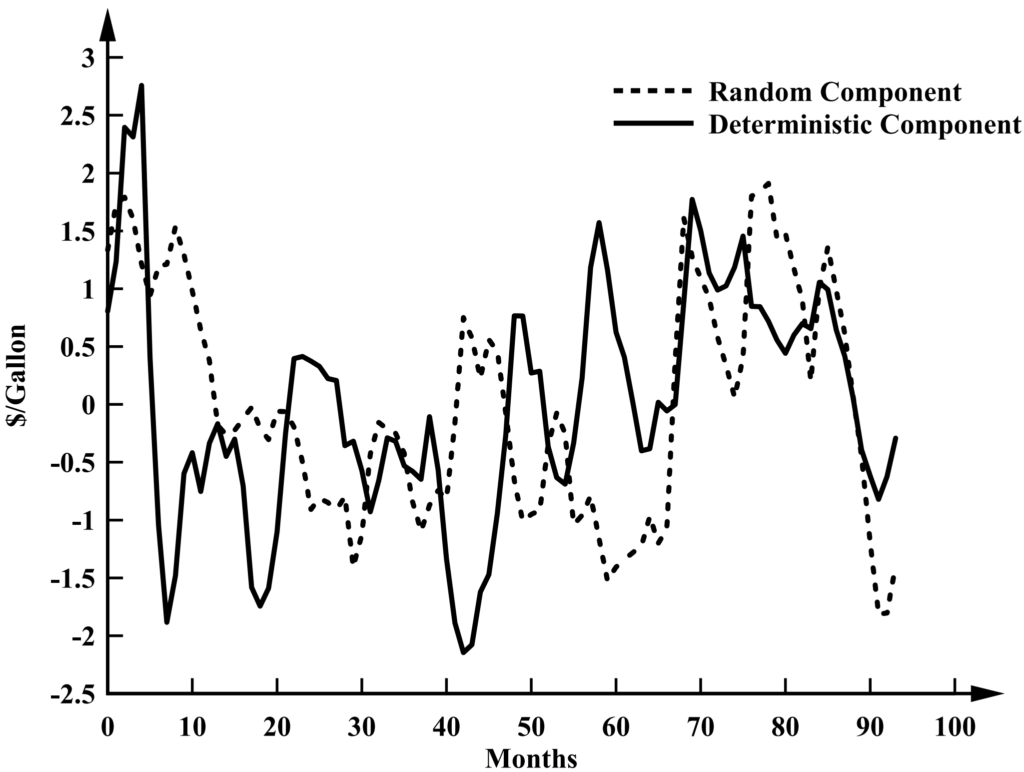
An estimation of the Hurst Exponent confirms that there is memory in the series with the Hurst Exponent being equal to 1.43 (Figure 18).

Figure 18: Hurst Estimation of Retail Gas Price



Performing the same analysis as before, the deterministic and random components of the series are separated and it can be seen below. As in the gold price data the random component is larger than the deterministic component so we will again look at the normalized data (Figure 19).

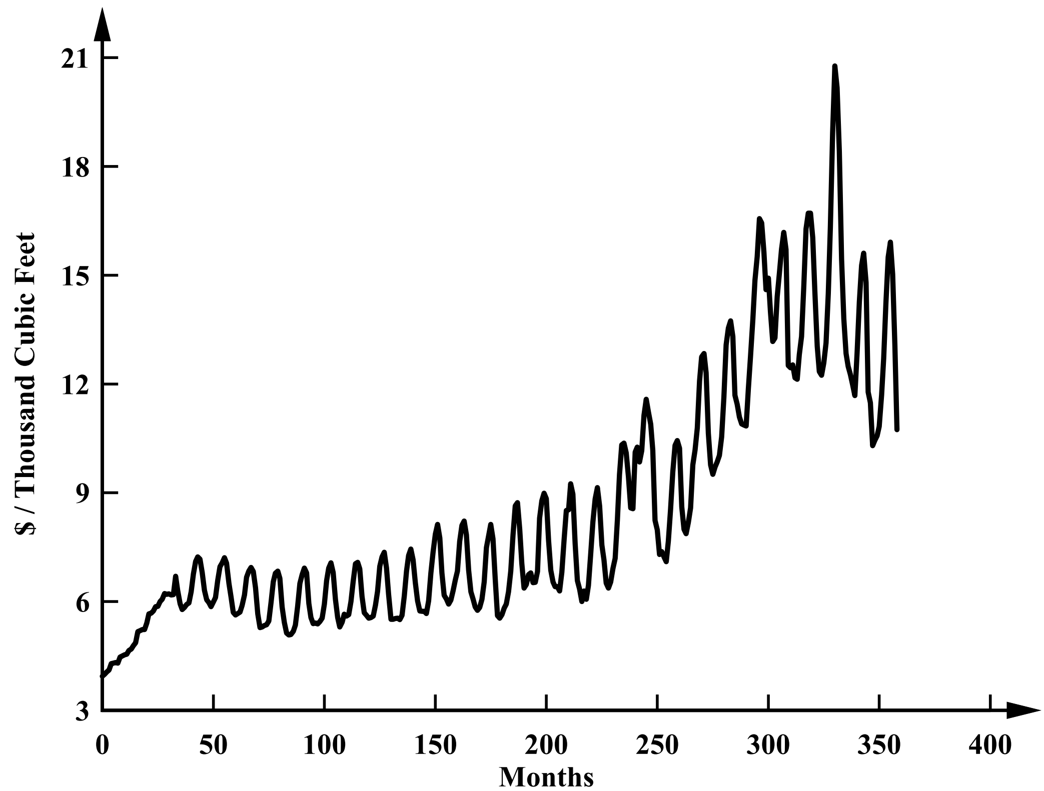
Figure 19: Normalized Separated Components of Gas Prices



In Figure 19, it looks as if both series may be random, again this could be a case where the deterministic component is chaotic. While we will not do so here, we could test the deterministic series for chaotic behaviour with tests as proposed by Stone, Landan and May (Stone et al., 1996). What is more of interest to the author of this article is that we need to allow for its existence when we model behaviour.

Finally, for one more look at methodology we will look at the residential natural gas price in the U.S. as seen in the graph below (Figure 20).

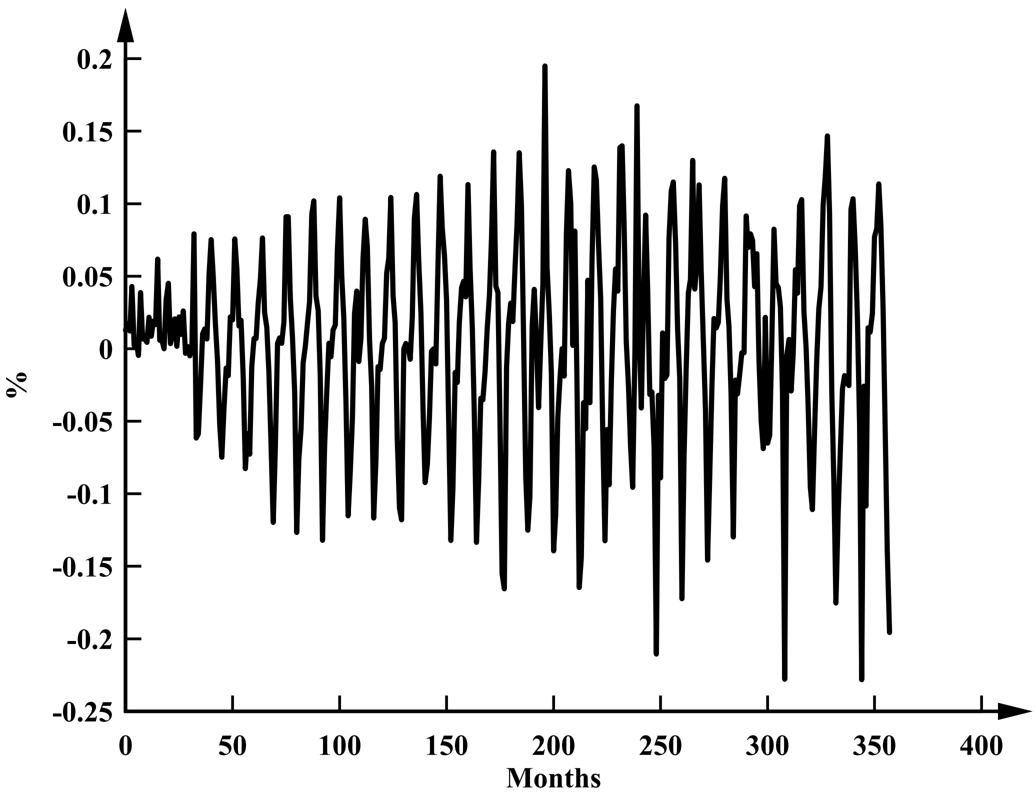
Figure 20: U.S Average Monthly Residential Natural Gas Price (Jan 1981 – Nov 2010)



Source: (U.S Energy Information Administration, Feb-11)

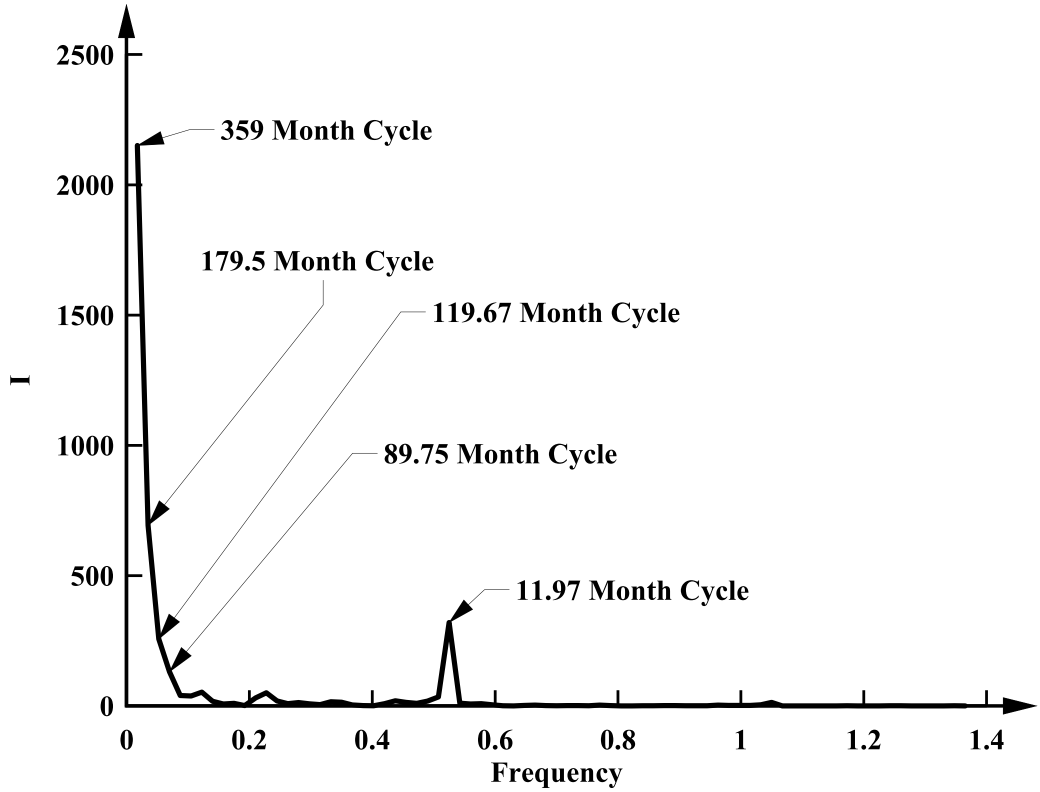
Again we see the same issues with heteroskedasticity, and as before we are confronted with the same methodological issues. Again we could remove the non-stationarity by differencing, but will still have the same methodological issues (Figure 21).

Figure 21: Percentage Change in Average Monthly Natural Gas Price



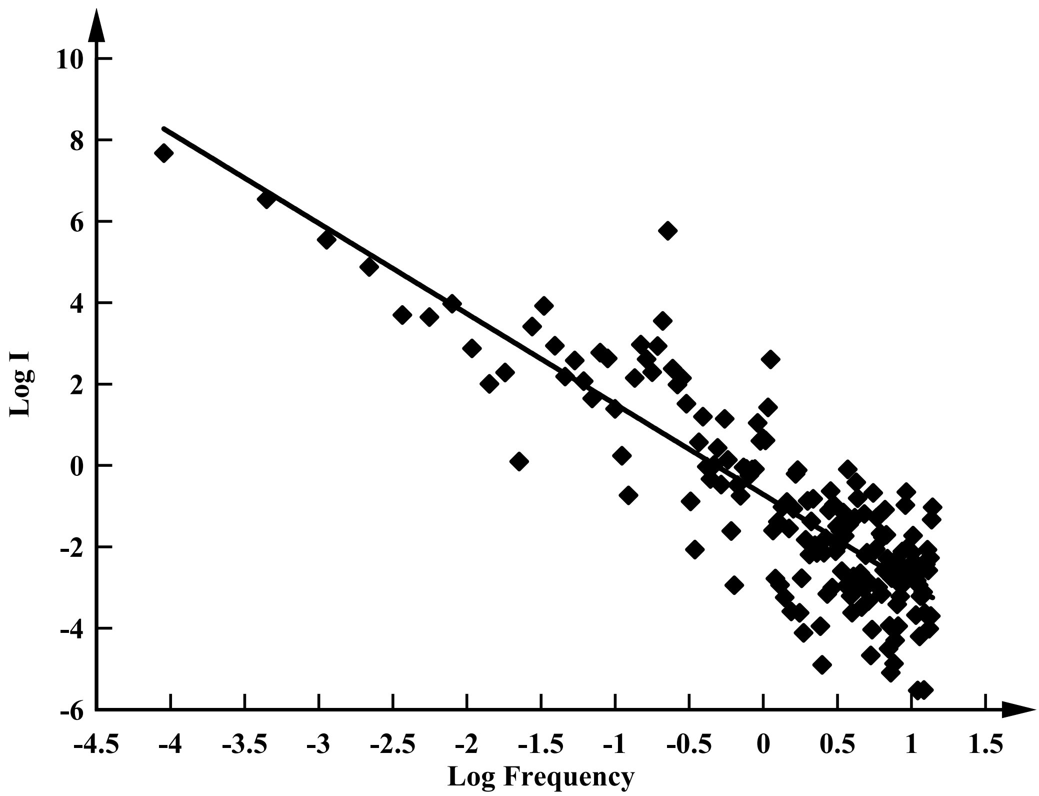
In the case of natural gas, there were five cycles that were significant at the 95% level, the shortest of them being 11.97 months (Figure 22). In this example, natural gas differs from the other two datasets because it does have a shorter cycle, but it is similar in that there is still long-term memory in the series.

Figure 22: Periodogram of Residential Natural Gas Price



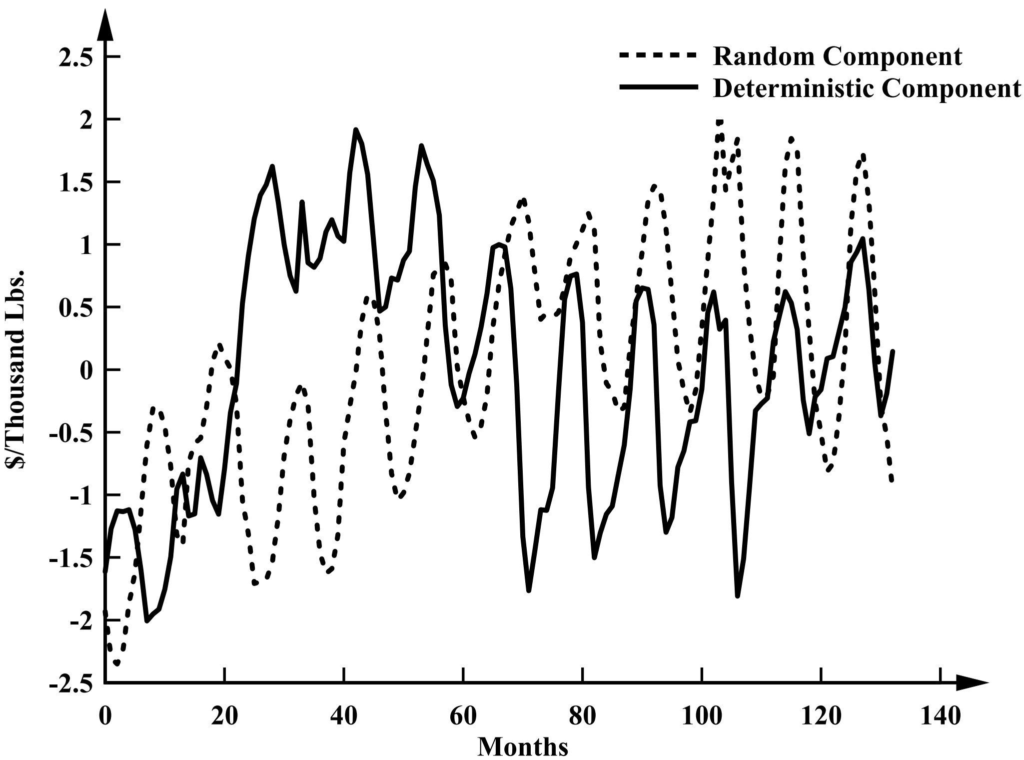
This is confirmed by the Hurst Exponent, which is estimated to be 0.81 (Figure 23), still showing persistence in the data, but at a lower level than the other two datasets.

Figure 23: Hurst Estimation of Residential Natural Gas Price



Performing the same analysis as before, we again find that the random component is larger than the deterministic component. Looking at the normalized data, we can see the behaviour of the two components. In this case, it appears as if the deterministic behaviour has a bit more regular cycling (Figure 24). This could be partially attributed to the lower level of persistence as measured by the Hurst Exponent. What we can see is that the deterministic behaviour in this series has been cycling on a more "regular" frequency than that of the other two series.

Figure 24: Normalized Separated Components of Natural Gas Price



Between all three datasets we can see some similarities and some differences. Why is the random component the largest in all three series? That is a good question that needs to be answered. We also need to ask the question of how prevalent is chaotic behaviour as well as how we can better model chaotic behaviour. These are important questions which hopefully will be answered with future research. A clear point is that we must first start by allowing for the existence of modelling deterministic infinite variance processes and possibly chaotic deterministic processes in order to discover if they are valid or not.

The Effect of stock market wealth on private consumption in Zimbabwe

Samuel Bindu[1], Lloyd Chigusiwa, D. Mazambani, L. Muchabaiwa, V. Mudavanhu

ABSTRACT

The study seeks to examine stock market wealth effects on private non-durable consumption for Zimbabwean households using quarterly data from 1994(1) to 2008(2). The bounds testing approach to cointegration is employed to test the long run relationship between stock market wealth and consumption. An autoregressive distributed lag model (ARDL) analysis is implemented to examine the relationship among the variables both in the short-run and the long run. The empirical findings suggest significant wealth effects for Zimbabwe, a developing country. This contradicts the commonly held view that LDCs should have insignificant wealth effect since the financial system is still underdeveloped. The dynamic short run error correction model also shows a speedy convergence to long run equilibrium.

Key words: wealth effects, consumption dynamics, income effect, convergence

JEL classification: E21, E44

1. Introduction

The study seeks to determine the extent to which household non durable consumption expenditure is affected by changes in stock market wealth for Zimbabwean households using an autoregressive distributed lag (ARDL) bounds test framework. Most studies on stock market wealth effects have been confined to developed countries which have well established financial markets, while very few are based on less developed countries (LDCs). The argument for underplaying the case for LDCs has been that since most of them have inefficient financial markets, stock market wealth effects cannot be expected to be significant. Above all, most researchers argue that since stock market wealth account for a small percentage of household wealth in LDCs, they are not important (Funke, 2004; Ahumada et al., 2003). However, following the revelation by African Stock Exchanges

[1] Bindura University of Science Education P. Bag 1020 Bindura Zimbabwe, tawandabindu@gmail.com

The views expressed here reflect the ideas of the authors and should not by any means taken to be those of Bindura University

Association (ASEA, 2007) that there has been a continuous rise in stock market wealth in household's wealth portfolios for most African countries over the past decade, the study seeks to establish whether we can obtain the same findings for developing countries as those of the developed countries.

Several reasons can be cited why the present study is important. Firstly, to the policymakers an understanding of the relationship between stock market wealth and consumption helps in ascertaining the speed with which shocks to the stock market are transmitted to the real sector through private consumption[1] (Macklem, 1994). In addition, an understanding of this relationship also gives a clearer picture to the policymakers as to the extent to which various stimulating interventions will affect the economy (Poterba, 2000). This is because the relationship between consumption and wealth determines the overall marginal propensity to consume, which in turn determines the value of the multiplier. Thus, ignoring such effects may result in unwanted policy reinforcement or the reversal.

The study of stock market wealth effects on consumption is also important to firms. It is apparent that firms should align their production levels to demand to avoid overproduction or underproductions. When equipped with knowledge of wealth effects, firms that produce consumer goods will be able to infer the corresponding change in their production level following a shock on the stock market (Singh, 1992).

1.1 Stock market wealth effects

Theoretically if households experience an increase in wealth, regardless of the source, they will experience an outward shift in their budget constraint. Such an increase in wealth can either be devoted to immediate consumption or it can be saved. If we assume that households have a finite life and do not leave behind bequest, it is certain that such an increase in wealth will be devoted to consumption at least within the household's lifespan. In such a case our concern will not be about the existence of wealth effects in general but the lag time it would take for changes in wealth to be devoted to consumption. However, when considering a more realistic household that leaves behind some bequest, households might save all or some of the gains in wealth for future generations. Thus wealth effects are concerned with the fraction of wealth increase that is devoted to current consumption.

It is important to note that household's wealth is composed of a wide range of assets. The major forms of household's wealth that are commonly held are stock market and housing wealth and the marginal propensities to consume (MPC) out of these two main types of wealth should be different (Pichette et al., 2003). For developing countries it is argued that stock market wealth is more reflective of household wealth for a number of reasons.

[1] This is because consumption spending of households that would have benefited from capital gains is one way in which such shocks are transmitted.

Firstly, for LDCs, most households do not own marketable houses. For example in Zimbabwe more than 65 percent of the population live in rural areas where houses are not easily traded due to customary laws and tradition (World Bank, 2010). Secondly, housing wealth is less liquid than stock market wealth, and the transaction costs associated with the former are usually higher than for the latter. Moreover, the indivisibility nature of housing wealth reduces its affordability to the general populace and as such households do not normally hold housing wealth for speculative purposes. Above all, households are likely to leave behind houses as bequest and as such do not normally consider housing wealth as redeemable (Singh, 1992). In LDCs stock market wealth is held for speculative purposes and thus gains are realised quickly than is the case for housing wealth.

1.2 The Zimbabwe Stock Exchange (ZSE): An Overview

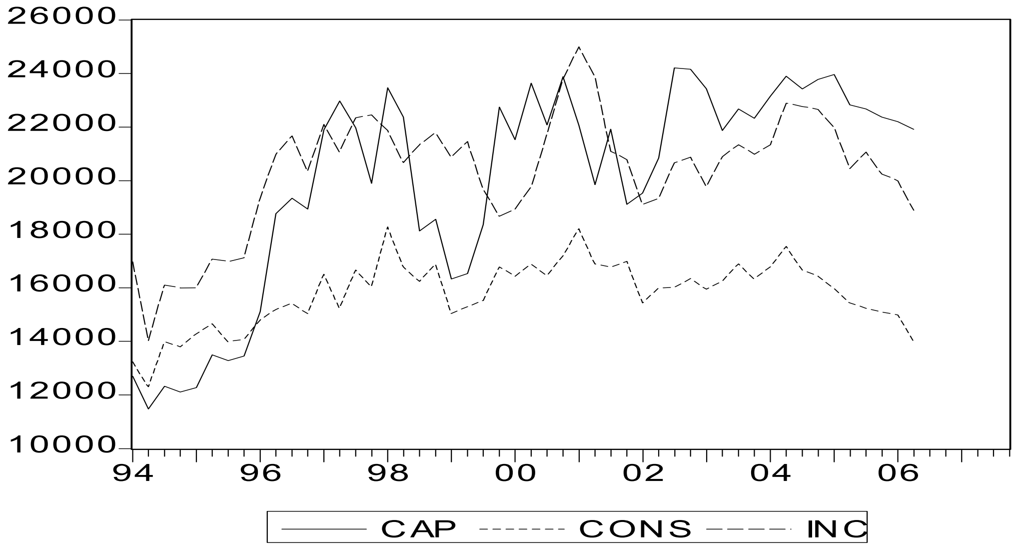
The ZSE has been in existence since 1946. Prior to 1993 the ZSE was highly regulated and its outcomes were not based on market fundamentals. After deregulation the ZSE has helped to mobilise investible funds and also provided an opportunity for insurance companies and pension funds to channel their portfolio investments to productive economic areas. This enabled pension funds to realise more returns on their investments which would ultimately accrue to the households for consumption. In March 1999, the ZSE was ranked sixth out of 33 emerging exchanges in the world (IFC, 1999). It was also ranked the second largest sophisticated stock market in the Southern African region after the Johannesburg Stock Exchange (Jefferis et al., 2000). Thus, compared to other African countries, the ZSE is relatively more efficient and as such provides a good ground for testing wealth effects for LDCs.

The ZSE has two indices, the Zimbabwe Industrial index and the Zimbabwe Mining index. There are 74 stocks listed on the Industrial index and 4 stocks listed on the Mining index. Trading in both indices is open to firms and individuals whether foreign or local. Stockholding across the two indices does not differ significantly and around 60 percent of stocks are owned by domestic households and firms (ZSE, 2008).

1.3 Trends in Zimbabwe stock prices and consumption

In order to give an elaborate analysis of the relationship between stock market wealth and household consumption expenditure on non-durables we shall analyse the trends in consumption and stock prices for the Zimbabwean case from 1994 to 2008. A graphical inspection of figure 1 highlights the likely marriage between the two series.

Figure 1. Trends in Consumption, Income and Stock Market Wealth



From observing figure 1 it can be established that periods of high stock market wealth (CAP) are accompanied by high consumption expenditure (CONS). It is also interesting to observe that income (INC) follow a similar pattern. For example, all the three series drop off after 1994 to reach a minimum in the second quarter of 1994. Stock market capitalisation seems to be the more volatile of the series, increasing and falling more frequently whereas the response of consumption seems to be more stable. The volatility in stock market wealth is because the financial markets are more responsive to external shocks than the goods market.

2. Related Literature

2.1 Theoretical literature

The discussion of wealth effects can either be based on the traditional consumption theories or on modern maximising theories. Following Carroll et al. (2004) who argued that the optimal behaviour of impatient consumers with labour income uncertainty is much better described by Friedman's original statement of the permanent income hypothesis than by the later explicit maximizing version, the study shall adopt the traditional consumption theories.

Keynes (1936) argued that households devote a certain percentage of their current income to consumption and the other percentage is then saved. Keynesian consumption function can be modelled as:

$$C_t = \alpha_0 + \beta Y_t + \varepsilon_t \quad (1)$$

where C_t is consumption and α_0 is autonomous consumption β is the marginal propensity to consume and Y_t is the current income and ε_t is the error term.

This theory is often criticised because it implicitly assumes that households only have human wealth and neglected non-human wealth such as equities and other financial wealth holdings. In addition, the theory overlooks the fact that incomes can be either permanent or transitory.

Friedman (1957) in his permanent income hypothesis (PIH) argued that the level of consumption depends on current and expected future income stream, that is,

$$C_t = \varphi Yp_t + \mu_t \tag{2}$$

where C_t is private household consumption in period t, Yp_t is household permanent income in period t and μ_t is the error term.

Friedman postulated that transitory income changes do not have an effect on household consumption but rather permanent changes.

The wealth effect on the household consumption expenditure was perhaps explicitly introduced by Ando and Modigliani (1963). They argued that households seek to smoothen their lifetime consumption to avoid large variance in their consumption in periods of high income and low income. Often, this effect has been analysed as the life-cycle hypothesis (LCH) in which private consumption is modelled as,

$$C_t = \eta Y_t + (\sigma - \gamma) A_t + \nu_t \tag{3}$$

Where A_t is the end period private wealth and γ is the rate of return on assets and ν_t is an error term. The rate of interest is important in explaining consumption in so far as it affects the value of wealth held in interest bearing assets. If interest rate goes up, household would expect to get more from their assets and as such may increase consumption.

2.2 The Empirical Literature Review

Early researches on wealth effects were mainly based on developed countries for example Mankiw et al. (1991); Starr-McCluer (1999); Maki et al. (2002); Lugwig et al. (2002). Most of these studies found positive and significant stock market wealth effects of between 2-9 percent. For the US wealth effects were found to be between 4-7 percent. Mankiw and Zeldes (1991) estimated micro level wealth effects for American households using OLS. They found a marginal propensity to consume out of stock market wealth of between 3 and 5 percent. They found the effect to be more pronounced for equity holders than for non-equity holders. Sousa (2008) found stock market wealth to be significantly smaller than non stock market wealth for USA. He also found that indirectly held stock market wealth has greater impact on consumption than direct property.

For Australia, Kohler et al. (2003) found a significant wealth effect of 7-9 percent. Similar to Mankiw et al. (1991), they also found out that in the wake of rising stock market wealth, households that hold stocks in their portfolios would increase their consumption more than those who do not hold equities. Their study, however, found the relationship between consumption and wealth to be unstable over time. Macklem (1994) found wealth effects of around 3.5cents per every dollar in the period of wealth change for Canadian households. He also found the effect to be lasting beyond the period of wealth change.

Very few studies based on emerging economies include Peltonen et al. (2009) who estimated wealth effects for 14 emerging economies. He found that a 10 percent rise in stock prices is associated with 0.29-0.35 percent rise in per capita consumption. Yet, they found that consumption reacts stronger to negative changes than it does to positive changes. Ahumada et al. (2003) also found significant but low wealth effect for Argentina. They used aggregate private consumption, national disposable income and stock market capitalisation macro level data without separating between equity holders from non-equity holders in their estimation of the wealth effects of an emerging country using data from Argentina. This finding was also in line with Funke (2004) who found small but significant wealth effects for 16 emerging economies for the period 1985-2000 ranging from 0.2-0.4 percent.

An important element to consider when studying wealth effects is the marked heterogeneity of stock ownership. Attanasio et al. (2002) used the UK Family Expenditure Survey (FES) data, to develop a method for separating the “likely” stockholders from non-stockholders and divided households into cohorts which depend on the likelihood of owning stocks in their wealth portfolios. They also found the consumption of stockholders was more highly correlated with stock market returns than the consumption of non-stockholders.

One shortcoming of their analysis is that they used a set of observable characteristics to construct a cohort of likely stockholders and likely non-stockholders, and regressed the average stock return against the average change in consumption of the two groups. However, those households who they predicted to own stocks might have other characteristics that are likely to affect their consumption behaviour. This indirect channel makes simple interpretation of the regression results more difficult (Grant et al. 2004).

However, the above studies are micro based in that they consider wealth effects on household level and not economy wide wealth effects. The micro based studies are criticised by macro based groups of researchers which include Bertaut (2002) who felt that not including those agents who do not hold stocks will only concentrate on the direct and immediate wealth effect and thus will estimate part of the effect and thus will misguide policy. Moreover, Paiella et al. (2004) argued that the micro based approach to wealth effects results in endogeneity problems. Endogeniety arises since households that hold stocks are usually those that are wealthier and as such are expected according to economic theory to have a low marginal propensity to consume and as such, it tends to underestimate the wealth effect. He further argued that the effect is further underestimated if the consumption on non-durables only is considered since wealthier households tend to expend even low percentage of their income on non-durables. In this study we shall follow the approach by Ahumada et al., (2003) and Macklem (1994) among others and estimate a stock market wealth effects using aggregate data for Zimbabwe. The approach is the most appropriate to the study because it is useful for policy purposes.

3. Methodology and results

3.1 Theoretical framework

Following Labbard et al. (2005) the aggregate consumption function can be modelled as:

$$C_t = f(Y_t, C_{t-i}, \gamma_t) \tag{4}$$

where C_t is household consumption, Y_t is total household disposable income, γ_t is the rate of interest in period t and C_{t-i} represents past consumption levels. Ideally from the basic Keynesian function (1936), disposable income and consumption are believed to have a positive relationship. Consumption also depends on past levels of itself C_{t-i} this idea was popularised by Friedman (1957) and Ando and Modigliani (1963) in the permanent income and the life cycle hypothesis respectively. γ_t is the rate of interest on savings and it takes into account interest elastic consumption that is mostly consumption of durable consumer goods. Interest rate also affects consumption by increasing the opportunity cost of consuming now thus a negative relationship is expected.

We shall split household income into human income (Y) and non-human income (W). Since the study is limited to non durable consumption which in most LDCs is not financed from borrowed funds we will not include the rate of interest in our analysis. Moreover, with the hyperinflationary environment over the past decade, there has been a negative real interest rate. Given the above mentioned, we shall therefore modify equation 4 to:

$$C_t = f(W_t, Yd_t, C_{t-i}) \tag{5}$$

where W_t is household wealth in period t, Yd_t is disposable income in period t.

3.2 Modelling Techniques

The study used an autoregressive distributed lag (ARDL) cointegration approach as pioneered by Pesaran et al. (1999), which they referred to as the Bounds Test Cointegration (BTC). The ARDL cointegration is more superior to the conventional cointegration approaches[2] in a number of ways.

Firstly, ARDL employs only a single reduced form equation thus avoiding estimation numerous equations some of which might not be important for the study in question (Pesaran et al., 1999). Furthermore, the ARDL approach also avoids stringent order of integration requirements as required by conventional cointegration procedures since it can be applied irrespective of whether the variables are I (0) or I (1) (Pesaran et al., 2001). Above all, given the low testing power of the unit root test, and the cyclical nature of most series, the ARDL is deemed superior to the conventional techniques. This is because the stationarity

---

[2] For more details on the conventional cointegration techniques see Engle et al. (1987); Johansen (1988).

tests result in biased results at times (Pesaran et al., 2001). However, it should be noted that in cases where some variables are I (2) or beyond, the ARDL will be inconsistent and will thus collapse (Pesaran et al., 2001).

3.2.1 ARDL modelling approach[3]

To illustrate the ARDL modelling approach, we followed the leads of Pesaran et al. (1997, 1999) as summarized in Choong et al. (2005), we apply the bounds test procedure by modelling equation 5 as general autoregressive (VAR) models of order *p* in G_t:

$$G_t = \alpha_0 + \sum_{i=1}^{p} \eta_i G_{t-i} + \varepsilon_t,\ t=1, 2, 3\ldots T \tag{6}$$

With α_0 representing a (*k+1*)-vector of intercepts (drift), and G_t is the vector of variables y_t and x_t respectively. Where y_t is the dependant variable consumption, (LCON) and x_t is a vector matrix representing a set of explanatory variables stock market capitalization and household disposable income $x_t = (LCAP_t, LINC_t)$. The short run unrestricted error correction (UECM) version of the ARDL model is given by:

$$DLCON_t = \alpha_0 + \sum_{i=1}^{p} \phi_i DLCON_{t-i} + \sum_{i=0}^{p} \beta_i DLCAP_{t-i} + \sum_{i=0}^{p} \delta_i DLINC_{t-i} + \lambda_1 LCON_{t-1}$$
$$+ \lambda_2 LCAP_{t-1} + \lambda_2 LINC_{t-1} + \varepsilon_t \tag{7}$$

Where D is the first difference operator, $\varphi_t, \beta_t, \delta_t$ are short run elasticities and λ_i are long run multipliers, ε_t is the stochastic error term and *p* is the optimal lag length for each variable which might not necessarily be equal for each variable. The first part of equation 7 with $\varphi_t, \beta_t, \delta_t$ represents short run dynamics of the model while the second part with λ_i represents the long run relationship. In estimating the UECM it is important to determine the optimal lag length *p*, which might as stated earlier be different for every variable depending on how sooner or later the variable applies its effect to the dependant variable.

3.2.2 Selection of optimal lag length

The selection of structural lags *p* is determined by using the SBC[4]. The SBC involves estimating $(p+1)^k$ equations and selecting the one with the largest SBC. Where *p* is the maximum number of lags to be used and *k* is the number of variables in the equation. It our model we take *p* = 4 since we are using quarterly data.[5] Since we have three variables in the model we took *k* = 3. Microfit 4.0 automatically selects the parsimonious equation from a total of 125 equations and found it to be an ARDL (1, 2, 1) model.

---

[3] Much of the material in this section follows the leads of Pesaran et al. (1999).

[4] It can also be selected using the general to specific modeling technique as argued by Narayan et al. (2004).

[5] For the rational of using *p* = 4 when using quarterly series see Pesaran et al. (1997).

3.2.3 Testing for stationarity

The Phillip- Peron (PP) test was applied in order to determine stationarity properties for each variable. Variables were first converted to logarithms before carrying out the test for stationarity. Table 1 reports the PP test for all the variables in their levels.

Table 1: Phillips-Peron test statistics for variables in levels and first differences

Variable		Intercept	Intercept and trend	Conclusion
Critical value	5%	-2.9215	-3.5025	
	1 %	-3.5682	-4.154	
LCON		-2.613	-2.0207	Non stationary
LCAP		-2.0622	-2.14295	Non stationary
LINC		-2.2444	-2.0479	Non stationary
DLCON		-10.6970***	-12.2216***	Stationary
DLCAP		-7.31087***	-7.51910***	Stationary
DLINC		-7.64374***	-8.08342***	Stationary

*** means significant at 1 percent level

Since the absolute values of the calculated PP statistic for the variables in levels do not exceed the critical values shown in table 1, we cannot reject the null hypothesis of non-stationarity and it can therefore be concluded that there exists a unit root even at 1 percent significance. We thus proceed to differentiate the series once. After differentiating the variables once, the stationarity test was redone to the first differences. The results of stationarity test in first differences are also reported in Table 1. This implies that all the variables are I (1) and as such the bounds test procedure can be carried out since no variable is I (2) or above. Having established that all the variables are I (1) and are permissible for ARDL cointegration we will proceed to carry out the cointegration test.

3.3 ARDL Cointegration test

After regression of equation 7, the ARDL test for cointegration can now be conducted using the Wald test (F-statistic) to discern the long-run relationship between the concerned variables[6]. The F test will be used to determine whether there is some kind of relationship between variables. Under the null hypothesis that there is no long run relationship between the variables, the cointegration test will be conducted.

According to the Pesaran et al. (2001), the lower bound critical values assumed that the explanatory variables are integrated into order zero, or I (0), while the upper bound critical values assumed that they are integrated into order one, or I (1). Therefore, if the

[6] The computed F-statistic value will be compared with the critical values tabulated in Pesaran et al. (1999).

computed *F*-statistic is smaller than the lower bound value and all our regressors are I (0), then the null hypothesis is not rejected. Conversely, if the computed *F*-statistic is greater than the upper bound value, then we do not reject the null hypothesis stated earlier. On the other hand, if the computed *F*-statistic falls between the lower and the upper bound then the result is inconclusive and thus the decision has to be based on the significance of the error correction term, (Pesaran et al., 2001). The results reported below show that the hypothesis of no cointegration cannot be accepted at the 5 percent marginal level.

Table 2: Cointegration test results

ARDL (1, 2, 1) selected basing on Schwarz Bayesian criterion

Test statistic	Value	Lags	Significance level	Bounds critical value[7]	
			%	I (0)	I (1)
F-statistic	4.439	2	1 5 10	4.614 3.272 2.676	5.966 4.306 3.586

From Table 2 it can be seen that the calculated F-statistic of F = 4.439 is higher than the critical for I (1) at the 5 percent level of significance using restricted intercept and no trend. This implies that the hypothesis of no cointegration cannot be accepted at the 5 percent marginal level. We therefore conclude that there exists evidence of a cointegrating relationship between the variables. We now proceed to estimate the long run relationship.

3.4 The Long Run Equation

The long run equation can be specified as;

$$LCON = \alpha + \sum_{i=1}^{p} \beta_i LCON_{t-i} + \sum_{t=0}^{q_1} \varphi_i LINC_{t-i} + \sum_{t=0}^{q_2} \delta_i LCAP_{t-i} + \varepsilon_t \quad (8)$$

The results for the parsimonious model ARDL (1, 2, 1) are reported in Table 3.

Table 3: Long run equation results

Dependant variable LCON

Variable	Coefficient	Std. Error	Prob.
C	-0.13308	0.94275	0.888
LCAP	0.048148*	0.02798	0.087
LINC	0.95533***	0.24966	0.000

*** and* denotes significance at the 1 percent and 10 percent level.

R-squared	0.84492	Adjusted R-squared	0.82223
D-W statistic	2.1283	Wald F (6, 41)	37.2310(0.001)

[7] Critical value include restricted intercept and no trend.

The empirical results, reported in Table 3, show that all the determinants of consumption used in this study are statistically significant at least at the 10 percent marginal level. The diagnostic tests show that the model is correctly specified and has the correct functional form. The diagnostics also indicate that the model does not have problems of serial correlation, multicolinearity and heteroscedasticity. The summary of diagnostics is reported in Appendix B.

The coefficient of stock market wealth (LCAP) of (0.0481) is positive and statistically significant at the 10 percent marginal level. This suggests that in the long run, an increase of one dollar in the value of stock market capitalization is associated with an increase of 4.8 cents in household consumption. On the other hand, the coefficient on LINC is 0.95533 that is positive and statistically significant. The coefficient also carries the expected positive sign. This implies that a change in household disposable income of 1 percent results in a 0.955 percent change in consumption.

The marginal propensity to consume out of stock market wealth of 0.048 tends to be consistent with findings of studies from the developed world. This MPC out of stock market wealth implies that there exists a long run relationship between consumption and stock market wealth in Zimbabwe. The finding might have been fostered by the fact that since households have fewer assets in their portfolios they pay particular attention to changes in stock wealth. Another reason is that given the falling incomes in Zimbabwe since 1999 to present day households are constrained by low-income levels and thus will expend any increase in income that will come their way.

It is also important to note that other determinants of consumption like income are also significant. A dollar increase in disposable income increases consumption by around 95cents and this relationship is statistically significant even at the 1 percent marginal level. This coefficient also carries a positive sign as expected which conforms to the simple Keynesian consumption function (1936) where current income is the major determinant of consumption. However, the marginal propensity to consume from income tends to be very large compared to other studies. The finding implies that households consume nearly 96 percent of the increase in income that they obtain. Most studies find an MPC of less than 60 percent. The falling incomes facing Zimbabwean households might be the reason for such a high MPC since households are not obtaining enough incomes to cover the basics.

3.5 The short run Error correction model

Though variables exhibit long run trend, it is possible that they might be in short run disequilibria. The short run equilibrium error correction term (ECM) is used to estimate how long it takes for the variables to converge to the long run equilibrium once they are in short run disequilibria. The ARDL short run error correction model is given below;

$$DLCON_t = a_0 + \sum_{t=1}^{p} b_i DLCON_{t-i} + \sum_{i=0}^{p} d_i DLINC_{t-i} + \sum_{i=0}^{p} f_i DLCAP_{t-i} + \eta ECM_{t-i} + \mu_t \qquad (9)$$

Where

$$ECM = LCON - \alpha - \sum_{i=1}^{p} \beta_i LCON_{t-i} - \sum_{i=o}^{p} \phi_i LINC_{t-i} - \sum_{i=0}^{p} \delta_i LCAP_{t-i} = \varepsilon_t$$

The results of the error correction model are reported in Table 4.

Table 4: Short run ECM results

Dependant variable DLCON

Regressor	Coefficient	Std. Error	Prob.
DLCON	-0.042913	0.29619	0.885
DLINC	0.56388***	0.097091	0.000
DLINC1	-0.30652**	0.095137	0.002
DLCAP	0.27429***	0.060432	0.000
ECM(-1)	-0.32247**	0.10579	0.004

, * Denotes significance at 5 percent and 1 percent marginal level respectively

R-squared	0.64952	Adj R-squared	0.59825
D-W statistic	2.1283	F (4, 43)	18.9959(0.000)

The dynamic error correction result shown above shows that the coefficient of DLCAP, DLINC and DLINC1 are statistically significant at the 5 percent marginal level. This means that changes in these variables are related to changes in consumption. From Table 4 it can be seen that a change in the value of capitalisation (DLCAP) has a statistically significant and positive effect of 0.27429 on the change in consumption (DLCON). It can also be seen from the table that a change in the household income (DLINC) has a significant effect on the change in consumption (DLCON).

The coefficient of ECM*t-1* is found to be reasonably large in magnitude and is statistically significant. It confirms a long run relationship between the variables. The coefficient of ECM term is -0.32247, which suggests a speedy adjustment process. More than 32 percent of the disequilibria of the previous quarter's shock adjust back to the long run equilibrium in the current quarter. The error correction term also carries a negative sign as expected. The negative sign confirms that the adjustment will be in the opposite direction of the disequilibria. This means that there is convergence to the equilibrium relationship, otherwise the series would diverge away from their long run equilibrium.

3.6 Data Sources and Problems

The study uses quarterly time series data covering the period from 1994(1) to 2007(2). The choice of the period is because the ZSE was deregulated in 1993; in order to maintain consistence and avoid unnecessary break in data the study shall start from 1994 after the ZSE was deregulated. However, for time series data the period is long enough to establish whether there exists the relationship between the variables. The study is based on secondary data extracted from existing publications and the data were obtained from the Central Statistical Office (CSO) and the Reserve Bank of Zimbabwe (RBZ).

Due to the hyperinflationary environment that characterised the economy during the period under study we converted the data to US dollars using the parallel market exchange rate obtained from Old Mutual Financial Holdings and Imara Asset Management. The parallel market exchange rate was preferred because we felt that it could mirror the value of the Zimbabwean dollar more accurately than official figures which tend to overvalue the local currency. We further deflated the US denominated data to constant 1990 US prices using the US consumer price index as reported in the Bureau of Labour Statistics. To obtain quarterly series for household income we used the Lisman and Sandee (1964) method which is discussed in Appendix A. Whilst every care was taken to maintain accuracy in manipulating the data, the possibility of errors and omissions cannot be completely rejected.

4. Conclusions and policy recommendations

Form the empirical results it can be concluded that stock market wealth has a long run relationship with household consumption. Thus, basing on these results the hypothesis that a change in stock market wealth has a positive effect on household, non durable consumption cannot be rejected. A change in stock market wealth of one dollar will result in a change in consumption of 4.8 cents. The marginal propensity to consume out of wealth has been found to be around 0.048. The results tend to be in line with the economic theory which postulates that a change in household permanent income has a positive effect on income. The dynamic error correction also validates the long run relationship since it is statistically significant there exists a positive and statistically significant relationship between consumption and stock market wealth. The dynamic error correction results indicate that there is a speedy error correction process implying that when variables are in short run disequilibria the system quickly corrects itself.

4.1 Policy Implications

The findings of this study imply that interventionist policies like monetary policy need to consider the agent's reactions as they experience a monetary shock before setting targets. In this case the wealth effects tend to reinforce the initial policy position and thus may produce unwanted results such as inflation and excessive unemployment of resources. The results of this study imply that discretionary policies in Zimbabwe need to provide for the effect of the stock market wealth before setting the level of intervention that is needed to achieve a certain specified target. For example, expansionary monetary policy aimed at increasing employment during recessions normally involves increasing money supply by a certain percentage so as to hit the intended target. The increase in money supply will see households holding excessive money balances above their requirements for transaction purposes. The households tend to offload the excess in the money market by way of buying bonds, thus, increasing their demand and pushing up their price. According to Keynes (1936) there exists a negative relation between bond prices and interest rates; therefore, interest rates tend to fall. The fall in rate of interest will in turn reduce the returns in the money market. Agents thus substitute away from the money market to the stock market; this will increase demand of equities and thus pushes up the price of equities. For agents who already hold stocks the increase in stock prices means increase in household's wealth and thus with the significant wealth effect that we have found empirically, agents increase consumption which further aggravates the initial stimuli.

The increased aggregated demand would see inventories falling and calls forth for producers to produce more, in the wake of near full employment this might be inflationary. In light of the foregoing, policymakers should consider the effect of stock market wealth effect on employment and output before they can calculate the necessary changes in their instruments required to hit the intended target, otherwise they will not realise the employment or inflation levels that they would want to achieve from economic models.

For contractionary monetary policy the initial effect will be to increase interest rates, the rising interest rates from initial policy stimuli attract agents in the money market away from the stock market. This substitution will exert a downward pressure on stock prices and thus reduce household wealth. The significant wealth effects that we have obtained mean agents will reduce consumption accordingly. This fall in consumption thus reduces aggregate demand resulting in accumulation of inventories, reduced production levels and a general deflation in the economy. In light of the policy reinforcement discussed above, it is important for policy makers to make adequate and correct calculation of the effect otherwise they will end up surpassing their targets thus causing unnecessary deflation or inflation.

4.2 Suggestions for Future Researches

Future researches can consider disaggregated wealth effects on non-durable consumption. Inclusion of housing wealth to the analysis of wealth effects in developing countries can help broaden the concepts discussed in this study.

There is also need to consider wealth effects in the VAR framework where we could include more than one cointegrating vector and run causality test to see whether there are multidirectional feedbacks in the model. Future studies could also consider wealth effects in the micro level excluding non-equity holders and see whether we can reach the same conclusion as we have obtained in this study.

Acknowledgements

The authors acknowledge the assistance they got from Dr. Mutenheri and Dr. Chifamba of the University of Zimbabwe, Department of Economics.

References

Ahumada, H., Garegnani, M. L., 2003, 'Wealth Effects in the Consumption Function of Argentina: 1980-2000', *Central Bank of Argentina publications*.

Ando, A. and Modigliani, F., 1963, 'The Life Cycle Hypothesis of Saving: Aggregate Implications and Tests', *American Economic Review*, 103, pp. 55-84.

ASEA, 2007, Third quarter review, 43.

Attanasio, O., Banks, J. and Tanner, S., 2002, 'Asset Holdings and Consumption Volatility' *Journal of Political Economy*, 110, 4, pp. 771-792.

Bertaut, C., 2002, 'Equity prices, household wealth, and consumption growth in foreign industrial countries: wealth effects in the 1990s', *Federal Reserve Board International Finance Discussion Paper*, 724.

Carroll, C., and Misuzu O., 2004, 'Estimating the Wealth Effect on Consumption' *Mimeo*, Johns Hopkins University.

Choong, C., Zulkornain Y. and Venus, L., 2005, 'Export led growth hypothesis in Malaysia. An investigation using bounds test', *Sunway academic journal*, 2, pp. 13-22.

Engle, R.F. and Granger, C., 1987, 'Co-integration and Error Correction: Representation, Estimation and Testing', *Econometrica*, 55, 2, pp. 251-76.

Friedman, M., 1957, 'The permanent income hypothesis', *A Theory of the Consumption Function*, Princeton, NJ, Princeton University Press, Chapter 3, pp. 25-30.

Funke, N., 2004, 'Is there a stock market wealth effect in emerging markets?', *Economic Letters,* 83, pp. 417-421.

Grant, C. and Peltonen, T., 2004, 'Income and wealth effects of Italian households', *European University Institute working paper.*

IFC, 2008, 'Emerging Stock Markets Fact book' Washington, D.C.

Jefferis, K. and Okeahalam, C., 2000, The impact of economic fundamentals on stock markets in southern Africa.

Johansen, S., 1988, 'Statistical Analysis of Cointegrating Vectors', *Journal of Economic Dynamics and Control*, 12, pp. 231-54.

Keynes, J.M., 1936, 'The General Theory of Employment, Interest and Money', London, Macmillan.

Kohler, M. and Dyornak, N., 2003, 'Housing wealth, stock market wealth and consumption: a panel data analysis for Australia', *Reserve Bank of Australia Discussion Paper*, 7.

Labhard, V., Stern, G., and Young, C., 2005, 'Wealth and consumption: an assessment of the international evidence', Bank *of England*.

Ludwig, A. and Slok, T., 2002, 'The impact of changes in stock prices and house prices in consumption in OECD countries', *IMF Working Paper*, 1.

Macklem, T., 1994, Wealth, Disposable Income and Consumption: Some Evidence for Canada. *Technical Report*, 74, Ottawa, Bank of Canada.

Maki, D. and Palumbo, M., 2001, 'Disentangling the Wealth effect: A Cohort Analysis of Household Saving in the 1990s', *Federal Reserve Board*.

Mankiw, G. and Zeldes, S., 1991, 'The Consumption of Stockholders and Non-stockholders' *Journal of Financial Economics*, 29, 1, pp. 97-112.

Narayan, P. and Smyth, R., 2004, 'Dead Man Walking: An Empirical Reassessment of the Deterrent Effect of Capital Punishment Using the Bounds Testing Approach to Cointegration' *American Law & Economics Association Annual Meetings*, 26, Berkeley Electronic Press.

Paiella, M., 2004, 'Does wealth affect consumption? Evidence for Italy', Bank of Italy Discussion paper, 510.

Peltonen, A., Sousa, R., and Vansteenkiste, I., 2009, Wealth effects in emerging market economies, *European Central Bank*, 1000.

Pesaran, H., Pesaran, B., 1997, *Working with Microfit 4.0. Interactive Econometric Analysis*, Oxford, Oxford university press.

Pesaran, H., Shin, Y., 1999, An ARDL modelling approach to Cointergration analysis', *Econometrics and economic theory in the 20th century*, Cambridge, Cambridge University Press.

Pesaran, H., Shin, Y. and Smith, R., 2001, 'Bounds testing approaches to the analysis of long run relationships', *Journal of Applied Econometrics*, 16, pp. 289-326.

Pichette, L. and Tremblay, D., 2003, 'Are Wealth Effects Important for Canada?', *Bank of Canada, working paper 30.*

Poterba, J., 2000, 'Stock market wealth and consumption', *Journal of Economic Perspectives*, 14, pp. 99-119.

Singh, A., 1992, The stock market wealth and economic development: Should Developing countries encourage stock markets?, *UNCTAD*, 49.

Sousa, R. M., 2008, 'Financial wealth, Housing wealth and Consumption', *International Research Journal of Finance and Economics*, 19, pp. 167-191.

Starr-McCluer, M., 1999, 'Stock Market Wealth and Consumer Spending', Mimeo, Federal Reserve Board of Governors.

World Bank, 2009, 'Zimbabwe Country Data' http://data.worldbank.org/country/zimbabwe (accessed 11 December 2010).

World Bank, 2010, 'Rural population (% of total population) in Zimbabwe http://www.tradingeconomics.com/zimbabwe/rural-population-percent-of-total-population-wb-data.html (accessed 15 December 2010).

Appendix

Appendix A: Lisman and Sandee (1964)

This method is employed to generate quarterly data from annual totals. The method assumes that the quarterly series of a variable Y_j depends linearly on three successive annual figures Y^*_{i-1}, Y^*_i, Y^*_{i+1} such that: $Y^*_j = a_{j,-1} Y^*_{i-1} + a_{j,0} Y_{1+a} + a_{j,+1} Y^*_{i+1}$. The twelve constants a_{jk} (j=1, 2, 3, 4: k=-1, 0, 1) are obtained by imposing twelve restrictions on the matrix of a_{jk}. In detail the method involves the following steps

Step 1

Starting from the first data point, group the observations in column vectors of order three for all successive years. For example if Y1, Y2 and Y3 are annual totals for a variable Y where 1,2and 3 are successive years the vector would be

$$Z=\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}, \quad L=\begin{bmatrix} 0.729 & 0.1982 & -0.0211 \\ -0.0103 & 0.3018 & -0.0415 \\ -0.0415 & 0.3018 & -0.0103 \\ -0.0211 & 0.1982 & 0.729 \end{bmatrix}$$

Step 2

The Lisman and Sandee interpretation matrix is given by matrix L above

The quarterly data for the year whose annual totals is in the middle position of the vector Z in step 1 is obtained by pre-multiplying the L matrix as in the example below:

$$\begin{bmatrix} Q_{12} \\ Q_{22} \\ Q_{32} \\ Q_{42} \end{bmatrix} = L.\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$$

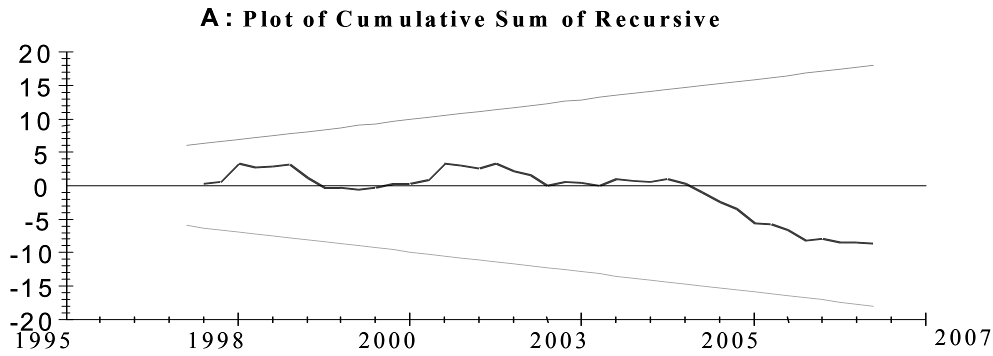
Where Q_{ij} is the ith quarter observation for year j (here j=2). This method has the nice property that the quarterly data it generates sum to exactly the annual total. The weakness however is that by the nature of the generation of the Z matrices, the method may introduce variability in the date generated which may not be present in the original series.

Appendix B:

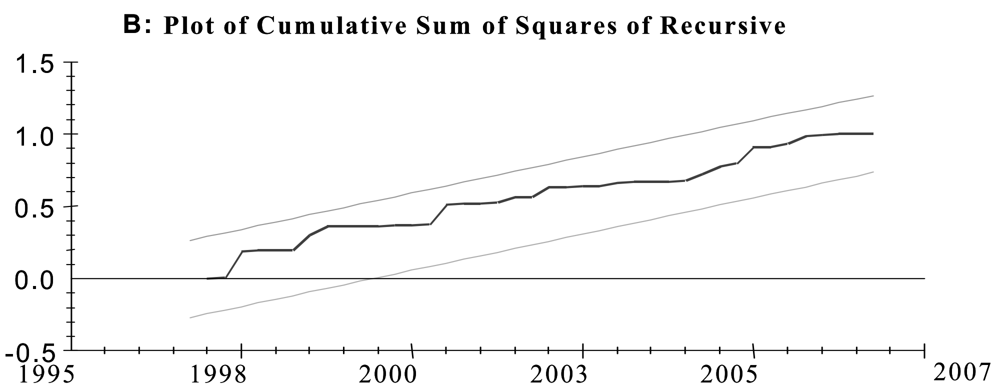
Table A1: Summary of diagnostics

Test	F-statistic	Probability
Specification error:		
Ramsey RESET test	0.39769	0.5321
Serial correlation:		
Breusch-Godfrey correlation LM test	0.32157	0.862
AR Cond. Heteroscedasticity:		
ARCH LM test	2.5824	0.1156

Stability Tests



The straight lines represent critical bounds at 5% significance



The straight lines represent critical bounds at 5% significance

Budget Deficit and Macroeconomics Fundamentals: The case of Azerbaijan

Kahnim Farajova[1]

Abstract

In recent years, the emergence of rising budget deficit is the main reason forcing economists to investigate the reasons for changes in fiscal balances. The purpose of the paper is to investigate the relationship between budget deficit and macroeconomic fundamentals using data from Azerbaijan. The empirical analysis applies ARDL Cointegration methodology in conjunction with Granger causality tests to provide evidence for both the long and short run dynamics between the variables involved in the analysis. Using the Error Correction specification, there was found evidence of long-run causality running from current account, real interest rate, GDP, inflation and exchange rate to budget deficit. There was also found evidence of short-run Granger causal effects running from current account and real interest rate towards budget deficit and a rather weak causal effect from inflation to budget deficit. However, there is no short – run causality running from interest rate to budget deficit.

Key words: Budget deficit, Fiscal policy, Cointegration methodology, Error correction model

JEL classification: E62, H62, C22

1. Introduction

Azerbaijan became an independent state in October 1991, following the dissolution of the USSR. The Azeri economy, geared to the demands of the Soviet GOSPLAN (*the state planning commission of the former Soviet Union or any of its constituent republics: it was responsible for coordination and development of the economy, social services, etc*) and part of CMEA (*Council for Mutual Economic Assistance*), faced external shocks with the demise of these structures, and by the onset of regional conflict over Nagorno – Karabakh. The break-up of trading links and payment mechanisms within the CMEA bloc, and more particularly within the Soviet Union was exacerbated by the deterioration in terms of trade for Azerbaijan as the formerly planned economies began to move towards transition at different speeds and with different rates of price adjustment. The start of the stabilization process was possible following the cessation of conflict Garabag in 1994. During this year

[1] Alexander Technological Educational Institute (ATEI) of Thessaloniki, 57400 Thessaloniki, P.O Box: 14 Greece, e-mail: xfarajova@gmail.com

a 30-year contract was signed between the State Oil Company of Azerbaijan Republic (SOCAR) and 13 oil companies, among them Amoco, BP, ExxonMobil, Lukoil.

In 1995, the Azeri authorities undertook an IMF-backed stabilisation plan. Both monetary and fiscal policies were tightened. These led to a sharp reduction in domestic inflation to 84,5 per cent in 1995 and to 4 per cent in 1997. Azerbaijan's high economic growth during 2006-08 was attributable to large and growing oil exports. GDP grew 41.7% in the first quarter of 2007, possibly the highest of any nation worldwide. Such rates cannot be sustained, however, despite reaching 26.4% in 2005, and over 36.6% (world highest) in 2006, they dropped to 11.6% in 2008. In 2009, economic growth dropped back to around 3% as oil prices moderated and growth in the construction sector cooled. The current global economic slowdown presents some challenges for the Azerbaijani economy as oil prices remain below their recent highs in mid-2008, highlighting Azerbaijan's reliance on energy exports and the ongoing difficulty diversifying its economy. In 2009, the government increasingly relied on financial transfers from the State Oil Fund to bridge its budget shortfalls. It should also be noted that besides important oil reserves Azerbaijan also has significant agronomic potential based on a wide variety of climatic zones.

On 18 June 2010, Azerbaijan's parliament, the Milli Majlis, has passed a bill amending the 2010 state budget. The amendments increase the revenue side of the budget by AZN 1.49bn ($1.855bn), bringing total revenue to AZN 11.505bn ($14.322bn). Expenditure is increased by AZN 1.11bn ($1.259bn), bringing it to AZN 12.275bn ($15.281bn).The increased revenue will be met by increased contributions from the State Oil Fund, which will go up by AZN 1bn to AZN 5.915bn, and from the Ministry of Taxes, which will contribute an additional AZN 490m, taking their total contribution to AZN 4.47bn.The amendments mean a reduction in the budget deficit from AZN 1.249bn to AZN 770.3m, or from 3.9% of GDP to 2.3%.

The main question is why such a country emerging as an important exporter of oil and natural gas and as a transport corridor between Europe and central Asia has budget deficit. Are fiscal deficits explained by a set of economic variables or do political factors bias fiscal policy towards deficit spending? In particular, do political systems and institutions have an inherent bias towards fiscal deficits or are these the result of the decisions of policymakers?

2. Theoretical Underpinnings

Theories of budget deficits run in two general directions. Some theories look on the effect of fiscal deficits on economic variables. Others look on the reverse direction, that is, what macroeconomic and fiscal variables (including budget rules and institutions) affect and determine fiscal deficits. This section gives a brief review of the theories of both –the effect of fiscal deficit on economic variables and the effect of macroeconomic variables on fiscal deficit.

2.1 Effect of Persistent Deficits on the Economy

How do persistent budget deficits and large government debt affect the economy?

Macroeconomic theory has divergent hypotheses regarding the implications of government deficits and debt on the economy. One strand of the literature contends that government debt reduces national saving which, in turn, crowds out capital accumulation. Thus, government debt hinders economic growth. Another strand of the literature implies the opposite: public debt does not influence national saving or capital accumulation. This view is based on the Ricardian equivalence theorem which asserts that it is only the quantity of government purchases, not whether such purchases are financed through between taxation or borrowing, which affects the economy. This implies that economic agents are indifferent between governments borrowing now or to a tax increase in the future. It has been shown empirically that this is not the case in the real world. In addition, when the permanent income hypothesis and the effect on consumption are considered, the Ricardian equivalence may not hold.

Barro's tax-smoothing theory states that what determines the deficit is the desire of the government to minimize distortions associated with raising taxes. The model implies that deficits and surpluses arise when the ratio of government purchases to output is expected to change. War and recession are times when the expected future ratio of government purchases to output is less than the current ratio. Consistent with the tax smoothing model, it has been observed that governments usually run deficits during these times. This implies that when national income is low, or government purchases are large, governments run deficits.

Roubini and Sachs (1988) find only partial evidence to support tax-smoothing, wherein tax rates are set over time to minimize the excess burden of taxation. They found a tendency for larger deficits in countries characterized by a short average tenure of government, the presence of many political parties in a ruling coalition and higher tax collection cost.

2.2 Effect of Macroeconomic Variables on Fiscal Balance

Inflation may affect budget deficits through various ways. The first way is through real tax revenues - inflationary conditions reduce the real tax revenues collected by government, thus, pushing towards budget deficits. The second way is via the effect on nominal interest rates. Inflation increases the nominal interest rates and consequently debt servicing, thus increasing the budget deficit. With these two factors in mind, it may be expected that inflation negatively affects fiscal balances.

However, inflation may positively affect fiscal stance by raising revenues via income tax 'bracket creep.' The US experience in the late 1970s was high federal tax receipts as a percentage of GDP in the face of high inflation rates (of approximately 10%). The explanation given by Saez (1999) and Auerbach (2000) was that the US income tax system at the time was not indexed for inflation (i.e. fixed in nominal terms), resulting in taxpayers near the top-end of a bracket to creep to the next bracket even if real income remained the same. Furthermore, if the tax system is designed to be elastic to changes in economic activity, it may be possible to have increased revenues with a boom and thus a positive influence on fiscal balance.

Easterly and Schmidt-Hebbel (1994) estimated the relationship between inflation and fiscal deficits. Across countries, the decision to print money to finance deficits (i.e. seignorage) would depend on the extent to which other means of financing are available. In their cross section estimation, they found no simple relationship between fiscal deficits leading to inflation. For case studies using time series data, revenue-maximizing inflation rates seem to rise with actual average inflation. In addition, money demand and inflation are nonlinearly related. It was found that money demand has decreasing semi-elasticity with respect to inflation. This implies that as inflation rises, money demand becomes less semi-elastic. They concluded that seignorage is unimportant as a steady state phenomenon, but it can be important as a temporary source of revenue in times of crisis. Furthermore, large surges of money creation are not closely linked to accelerated inflation. Though Easterly and Schmidt-Hebbel (1994) looked at how budget deficits affect inflation via seignorage, the opposite direction of this study, it is evident that the relationship of inflation and fiscal stance is not a simple one. The effect of inflation may be through various routes, thus making the actual relationship dependent on empirical evidence.

The level of development of the **financial market** is also believed to be related to fiscal performance. A more developed financial market would have more readily available forms of money to buy goods and services without incurring costs. The World Bank suggests that a more developed financial sector has increased flexibility in adjusting to macroeconomic shocks to prevent banking or financial crises. A measure of financial depth used by the World Bank is the ratio of liquid liabilities (i.e. broad money or M3) to GDP.

Another aspect of a financially deep economy is the link between banking openness and economic growth. Bayraktar and Wang (2006) found empirical evidence that banking sector openness may directly affect growth by improving the access to financial services and indirectly by improving the efficiency of financial intermediaries, both of which reduce the cost of financing and in turn, stimulate capital accumulation. Increased investments lead to economic growth and an improved fiscal performance, implying a positive relationship.

The literature on financial openness has also hinted at a positive relationship between financial depth and fiscal balance. Financial repression, as indicated by a less liquid banking sector, is practised by government either to finance its budget deficits or to direct its access of cheap credit to select industries, or both. Restrictive financial policy can be implemented in various ways: (1) imposing high nominal interest rate ceilings; (2) money creation (i.e. seignorage); and (3) imposing high reserve requirements. Denizer, Desai and Gueorguiev (1998) found evidence that the post-Communist governments in their study inhibit the development of financial institutions to ensure adequate flows of external capital to enterprise sectors rather than to finance deficits.

Other empirical evidence, however, has shown a negative relationship between fiscal deficit and financial market development. Woo (2001) examined the effect of financial depth on consolidated public sector deficit in developing countries. He found that an increase in financial depth is negatively associated with fiscal stance. He explained that a more liquid banking system can more easily finance fiscal deficits by issuing bonds without having to resort to inflationary finance. Aizenman and Noy (2003) found similar evidence that a budget surplus has a negative impact on financial openness for developing countries. That is, a bigger budget deficit will increase *de facto* financial openness. This was explained by evidence that developing economy engage in procyclical, rather than counter-cyclical, policy. In developing economies, financial crises tend to lead to recessions that in turn result in lower budget deficits because government reduces its spending. In addition, if the tax system is relatively inelastic to economic activity, an economic recession would lead to relatively higher tax revenues.

Turning to the open economy, most of the literature and studies about fiscal deficits and **exchange rates** have used fiscal stance as the independent variable. Easterly and Schimdt-Hebbel (1994) found robust relationships between the fiscal deficit, the trade deficit, and the real exchange rate. The fiscal deficit and the real exchange rate have a two-step relationship: the fiscal deficit and other determinants of investment and saving behaviour determine the external deficit, which then determines the real exchange rate consistent with clearing of the domestic goods market.

Long-term interest rate. A high interest rate worsens the overall budget balance via increasing interest expenditure on newly issued debt and on rolling debt. On the other hand, higher interest rates signal higher opportunity costs of bond market financing, possibly urging governments to improve the fiscal balance. Overall, however, the first effect is expected to dominate, thus producing a negative correlation between interest rates and budget balances.

An alternative measure could be interest expenditures as a percentage of GDP, on the ground that effects of high interest rates on fiscal policies depend on the prevailing debt level (e.g. Volkerink and De Haan, 2001 and Eschenbach and Schuknecht, 2002).

Short-term interest rate. In setting fiscal policy, the monetary policy stance may be an argument. The expected reaction, however, is ambiguous. High short-term interest rates to reduce inflationary pressures could be supported by fiscal policy or it could be countered, depending on policy preferences, views on the operation of the economy, and the allocation of tasks among policymakers. Modelling monetary policy by an interest rate, moreover, may capture other elements such as the cost of government financing, as described above when discussing long-term interest rates. This may be of particular importance in case of predominantly short-term financing or in case there is a strong link between short-term and long-term interest rates.

3. Methodological Issues

3.1 Concept of Stationarity

A stochastic process is said to be stationary if its mean and variance are constant over time and the value of the covariance between the two time periods depends only on the distance between the two time periods and not on the actual time at which the covariance is computed (Gujarati, 2003).When time series data are not stationary and are used in an econometric equation, there is the problem of spurious regression, which leads to unreliable results. In order to avoid this problem, it is necessary to investigate the time series data for their stationary properties.

The ADF test consists of estimating the following regression:

$$\Delta y_i = \beta_1 + \delta y_{i-1} + \sum_{j-1}^{p-1} a_j \Delta y_{i-j} + u_i \tag{3.1}$$

Where u_i is a pure white noise error term, $\Delta y_{i-1} = y_{i-1} - y_{i-2}$, and p is the class of autoregression. We test whether $\delta = 0$ (null hypothesis).

The ADF test with trend variable consists of estimating the following regression

$$\Delta y_i = \beta_1 + \beta_2 t + \delta y_{i-1} + \sum_{j=1}^{p-1} a_j \Delta y_{i-j} + u_i \tag{3.2}$$

Where t is the time or trend variable. The null hypothesis is $\delta = 0$ and if it is rejected, y_i is stationary around a deterministic trend.

The ADF test statistic is a modified *t* statistic. It has developed by Dickey and Fuller in cases where serial correlation exists and is conducted by adding the lagged values of the dependent variable Δy_i. The ADF test follows the same asymptotic distribution as the DF statistic, so the same critical values can be used. The power of the test to reject the null hypothesis decreases when the number of lags is increased. The rejection of the null hypothesis entails that the examined variable is stationary (Gujarati, 2003).

The testing procedure requires the estimation of the testing model and computes the t-value for the estimated coefficient β. Then we compare the calculated t ratio with the critical value τ from the Dickey - Fuller tables. If $t > \tau$, then the null hypothesis is rejected. If $t < \tau$, then the null hypothesis is accepted. In case the variables are found to be non-stationary we repeat the test using as depended variable the second difference $\Delta^2 y_{i-1}$ and so on, till we come to a stationary transformation of the original variable (Lazaridis, 2005).

3.2 The Concept of Cointegration

Many economic time series tend to change over time. However, this change may occur in: a) a stable or predictable way, in which case the mean and variance will be well defined, or b) in an unstable way, entailing that the mean and the variance will change over time. Unstable or non-stationary series can often be made stable or stationary if differentiated (*d*) one or more times, and are called integrated series of order *d*, [~*I* (*d*)]. Cointegration extends the univariate concept of integration to two or more series. Even in the case of two non-stationary variables, if a linear transformation of the variables is stationary, they are said to be cointegrated (more than two variables can generate more than one cointegrating vector). If cointegration is detected, the cointegrating equation defines the long-run relationship of the variables, but also, an error-correction model will exist to define both the short-run and long run behaviour of the variables.

If a long-run equilibrium relationship exists for a set of variables X, then it must be true that a cointegrating vector γ can be defined such that $Z_t = X_t\gamma \sim I(0)$.

That is, Z_t (the error term) is white noise representing random disturbances from a long-run equilibrium position; with the system again adjusting to the equilibrium. A direct test proposed for cointegration by Engle and Granger (1987) consists of a two-stage approach. First, each variable series is tested for stationary. In case the variables are found non stationary of the same order of integration, we proceed with the second stage. Next, the cointegrating vector is formed and then a test, that the errors (Z_t) are integrated of order zero, is performed. The Engle-Granger procedure has an advantage in the application of the least squares to identify and estimate a cointegrating vector. On the other hand, it has the disadvantage that it is efficient only for two variables. However, the Dickey-Fuller statistic used in testing has low power in distinguishing between unit roots and near unit roots. An alternative procedure for cointegration testing is the one of Johansen and Juselius (1990) which is more reliable for multivariate analysis. Nonetheless, large samples are needed for this test. The data are divided into a differentiated part and a levels' part. Under the assumption of $I(1)$ processes, the differentiated data are stationary. The technique of canonical correlations is used to find linear combinations of the data in levels. It is inferred that these linear combinations must be stationary. This procedure has the advantage of being able to identify more than one cointegrating vector. Finally, a relatively new approach used for cointegration testing is the ARDL cointegration - a single equation technique.

3.3 The Autoregressive Distributed Lag (ARDL) Cointegration Approach

The autoregressive distributed lag (ARDL) approach to cointegration, which has been chosen in this paper, is a relatively new technique for detecting possible long-run relationships among economic variables. The ARDL approach is considered to be a more efficient technique in determining cointegrating relationships in small samples, compared to the previously mentioned conventional techniques. An additional advantage of the ARDL approach is that it can be applied irrespective of the regressors' order of integration; therefore, it can be applied regardless of the stationary properties of the variables in the sample, thus allowing for statistical inferences on long-run estimates which are not possible under alternative cointegration techniques. Consequently, we are not concerned whether the applied series are *I(0) or I(1)*.

More particularly, in the first step the following unrestricted error correction (EC) version of the ARDL model is estimated for each of the examined variables:

$$DY_i = \alpha_0 + \sum_{i=1}^{p} b_i DY_{t-i} + \sum_{i=0}^{p} c_i DX_{t-i} + \delta_1 Y_{t-1} + \delta_2 X_{t-1} + e_i \qquad (3.3)$$

Where Y and X are the endogenous variables of the model; and D denotes first difference.

On the basis of equation (3.3), we perform bounds test for the presence of a long-run relationship between the variables. Actually, F-test is applied for the joint null hypothesis that the coefficients on the level variables are jointly equal to zero. According to the traditional approach and given that the testing statistic displays a non-standard distribution, we would have to take under consideration whether the variables are individually $I(0)$ or $I(1)$, the number of regressors and the existence of an intercept and/or a trend. Instead of the conventional F critical values, we may use two sets of critical value bounds for all classifications of the regressors into purely $I(1)$, purely $I(0)$ or mutually cointegrated. If the test statistic exceeds their respective upper critical values, it may be argued that there is evidence of a long-run equilibrium relationship. If the test statistic falls below the lower critical values, we cannot reject the null hypothesis of no cointegration. Finally, if the test statistic lies between the two bounds, then the test becomes inconclusive (Pesaran and Shin, 1999; Pesaran, et al., 2001).

The conditional long-run models can be extracted from the reduced form solution of equation (3.3), when the first-differenced variables jointly equal zero. The long-run coefficients of the EC models are estimated through the ARDL approach to cointegration and the use of OLS. The lag structure for the ARDL specification of the short-run dynamics, in this thesis, is determined by the Akaike's Information Criterion (AIC), in conjunction with the autocorrelation test. The corresponding EC specification is based on the implied ARDL specification, through a simple linear transformation (Pesaran and Pesaran, 1997).

The ARDL is based on a single-equation estimation method and requires the estimation of a fairly smaller number of parameters. Consequently, in case of small data samples, ARDL is considered a rather more suitable cointegration technique. In addition, the ARDL method avoids the problem of pre-testing for the order of integration of the individual variables which might be of crucial importance. Therefore, after the confirmation of a long run relationship between the variables, the derived from the ARDL analysis Error Correction (EC) model can be used to test for Granger-type causality. An advantage of using an EC specification to test for causality is that, on the one hand, it allows the testing of short-run causality through the lagged differenced explanatory variables and, on the other hand, the testing of long-run causality through the lagged EC term. A significant EC term confirms long-run causality from the explanatory variables to the dependent variable.

4. Data and Empirical Results

4.1 Data

The econometric method that is adopted to examine the relationship among the variables under examination is the ARDL cointegration technique in conjunction with Granger-Causality testing.

The data employed for the empirical analysis are annually covering the period 1992 to 2009 and are collected from the International Financial Statistics Yearbook (2000/2006) and World Bank database. Actually, the following variables are used: i) Interest Rate (we deduct inflation rate and used Real Interest Rate, denoted by RIR.) ii) Exchange Rate, denoted by ER. iii) Gross Domestic Product, denoted by Y iv) Inflation, denoted by P v) Current Account as a percentage of GDP, denoted by CA vi) Budget Deficit as a percentage of GDP denoted by BD. In this analysis ER and Y are used in logarithmic form and are denoted by LER and LY respectively.

The econometric analysis uses Microfit 4.0 developed by Pesaran and Pesaran (Pesaran and Pesaran, 1997).

4.2 Empirical Results

Before proceeding with the econometric estimations, it is required to investigate the integration properties of the used variables in order to avoid the problem of spurious regression. Consequently, the variables for their stationary properties are examined by means of the conventional Augmented Dickey-Fuller (ADF) test while the optimal ADF specification is determined by means of Akaike Information Criterion (AIC), the Schwarz Bayesian Criterion (SBC). The tests for all the variables in levels (BD, CA, P, LY, RIR and LER) as well as in first differences are presented in Tables 1 and 2. The results suggest that, all variables are non stationary in levels and stationary when tested in first difference form.

Table 1: ADF unit root test in levels

VARIABLE	LEVELS			
	Lag	With Intercept	Lag	With Intercept and Trend
BD	2	-2.4265	1	-1.2686
CA	0	-.94752	0	-1.8424
RIR	3	-2.8892	3	-1.5990
P	3	-2.3197	3	-1.2879
LY	0	-1.7757	0	-3.4704
LER	2	-2.3751	3	-1.5071

Note: 95% critical value for the augmented Dickey-Fuller statistic = -3.1004

Table 2: ADF unit root test in first differences

VARIABLE	FIRST DIFFERENCES			
	Lag	With Intercept	Lag	With Intercept and Trend
DBD	0	-4.9924	0	-5.3422
DCA	3	-6.6876	3	-7.5619
DRIR	2	-5.9034	2	-5.3513
DP	2	-5.0282	2	-4.3578
DLY	0	-5.7253	0	-5.4908
DLER	2	-1.9661	3	-4.1459

Note: 95% critical value for the augmented Dickey-Fuller statistic = -3.7921

Allthough the results provide evidence that all examined variables are integrated of order one I (1), we should accept this finding with caution given the limited number of available data. Following, the application of the ARDL Cointegration methodology proceeds in order to examine the dynamics of the investigated relationships. It is believed that this methodology is probably the most appropriate for the needs and limitations of the empirical analysis because of the following reasons:

1-While other methods require non-stationary variables of integration order I(1), like the traditional methodology proposed by Johansen, the ARDL method proposed by Pesaran has the advantage to avoid the problem of pre-testing for the order of integration of the individual series.

2-In addition, it is a single equation estimation technique and requires the estimation of a fairly smaller number of parameters compared to the Johansen's method. Consequently, ARDL proves to be more efficient when small data samples are available and thus, this method is adopted for the needs of the given empirical research.

In the first step, we proceed with the examination of the joint integration properties of the series using the ARDL cointegration methodology proposed by Pesaran and Shin (1999). Actually, we estimate the unrestricted error correction (EC) model (3.3) above, with DBD as the dependent variable and one of the other variables each time as independent and apply a F test on the group of the lagged level variables.

The F test results on the group of the lagged level variables are depicted in Table 3, below.

Table 3: Testing the existence of a long run relationship between variables

Dependent / Independent Variable	F-statistic	Intercept	Trend	Conclusion
BD/CA	22.4255	Yes	No	cointegration
BD/RIR	5.8721	Yes	No	Cointegration
BD/LY	8.2234	Yes	Yes	Cointegration
BD/P	5.6730	Yes	No	Cointegration
BD/LER	10.4014	Yes	No	Cointegration

Note: Critical values have been obtained from Pearan & Pesaran and are: 1) with constant 9.934 and 5.764 at the 95% level of significance and 2) with constant and trend 6.606 and 7.423 at the 95% level of significance.

We observe that the value of F-statistic exceeds the upper bound of the critical value bounds in all cases and consequently the tests suggests that there exist long-run equilibrium relationships between the BD and each one of the examined determinants CA, RIR, LY, P and LER with long-run causality running towards BD.

Having confirmed the existence of cointegration among the budget deficit variable and the tested determinants in a bi-variate framework, we proceed with the estimation of the appropriate ARDL models. Among a number of estimated alternative models, the most appropriate are selected according to the Akaike Information Criterion (AIC).

Next, the long-run coefficients from the implied ARDL model are estimated and are depicted in Table 4, below. All the independent variables (CA, RIR, LY, P, LER) bear statistically significant causal impacts on the dependent variable BD, in the long-run time horizon.

Table 4: Estimates of the long-run coefficients based on ARDL model

Estimated Long Run Coefficients using the ARDL Approach
selected based on Akaike Information Criterion.
Dependent variable is BD

Regressor	Coefficient	Standard Error	T-Ratio[Prob]
CA	4.2601	1.1607	3.6704[.004]
C	3.9477	1.0817	3.6495[.004]
RIR	-.0085576	.0035478	-2.4121[.031]
C	3.6593	1.6207	2.2578[.042]
LY	-3.5743	2.7225	-1.3128[.214]
C	-3.3194	5.3816	-.61681[.549]
t	1.2010	.79775	1.5055[.158]
P	.0075571	.0033579	2.2506[.042]
C	3.6098	1.6570	2.1785[.048]
ER	-22.2797	7.9950	-2.7867[.015]
C	-23.1119	6.8514	3.3733[.005]

More particularly, we can extract the following long-run econometric relationships:

$$\widehat{BD_t} = 3.947 + 4.2601\ CA_t \tag{4.1}$$

$$\widehat{BD_t} = 3.6593 - 0.00856\ RIR_t \tag{4.2}$$

$$\widehat{BD_t} = -3.3194 - 3.5743\ LY_t \tag{4.3}$$

$$\widehat{BD_t} = 3.6098 - 0.00756\ P_t \tag{4.4}$$

$$\widehat{BD_t} = 23.112 - 22.2797\ LER_t \tag{4.5}$$

In particular, according to the equations above, an increase in current account by 1%, increases budget deficit by 4.2%, increase in interest rate by 1% decreases budget deficit 0.008%, an increase in GDP by 1%, decreases budget deficit 3.6%, increase in inflation by 1% increases budget deficit 0.007% and increase in exchange rate by 1% decreases budget deficit 22.2%.

The stability of the coefficients of the EC models is tested by means of the CUSUM test. Figures 1, 2, and 3 below, confirm long-run structural stability for the model's coefficients.

Figure 1: Plot of cumulative sum of recursive residuals for BD on CA and RIR

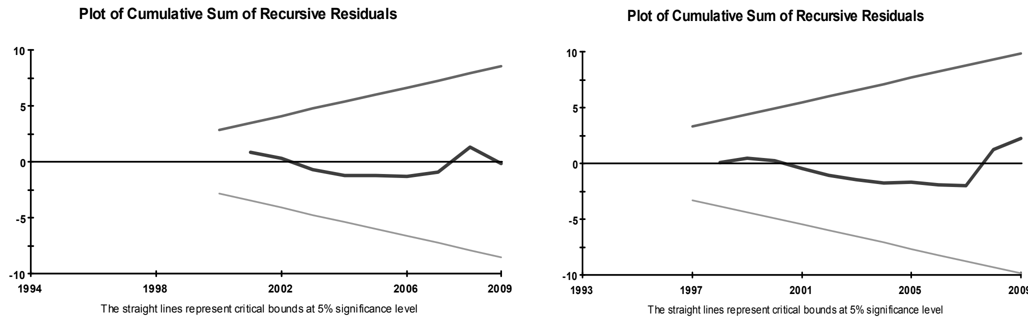


Figure 2: Plot of cumulative sum of squares of recursive residuals for BD on LY and P

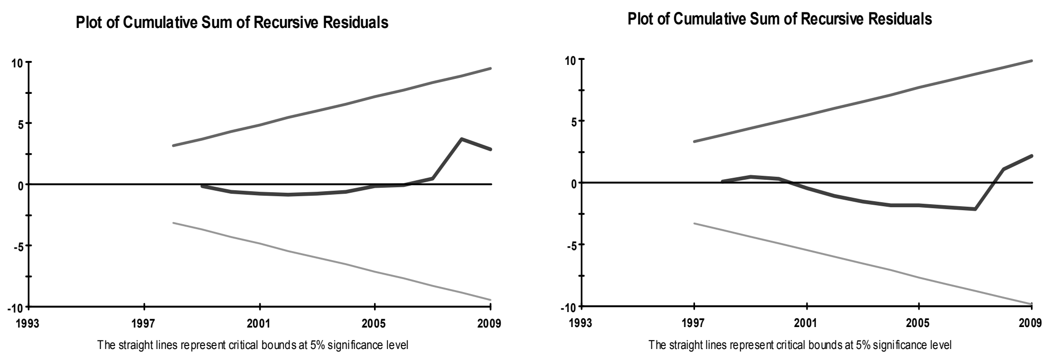
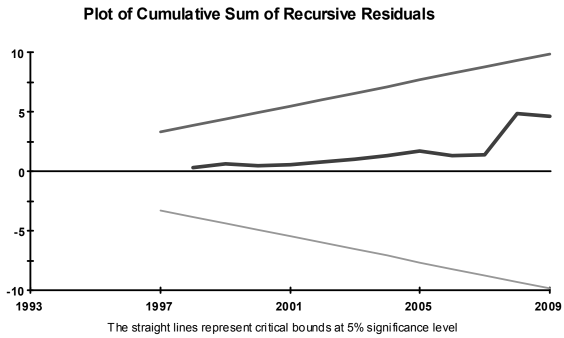


Figure 3: Plot of cumulative sum of squares of recursive residuals for BD on ER



Next, we proceed with the estimation of the corresponding Error Correction Models (ECM). The results are presented analytically in Table 5. The estimations are obtained from the following general form of equation:

$$DY_t = c + \sum a_j\, DY_{t-j} + \sum b_k\, DX_{t-k} + \delta\, ECM_{t-1} + \varepsilon_t \quad (4.6)$$

As we can notice in Table 5 below, the coefficients of the error correction terms in all estimated models were found negative and statistically significant at the 5% significance level. The larger the error correction coefficient (in absolute values), the faster the return to equilibrium, once shocked. Therefore, there is a statistically significant long-run impact from CA, RIR, LY, P and LER on BD.

Table 5: Causality test based on the Error correction Representation

Model	Wald statistic (lagged differences)	Coefficient of the error-correction term. (t-statistic/p-value)	Conclusion
CA →BD	59.0344(0.000)	-.59676 (- 4.6943 /0.001)	Both short and long run causality
RIR →BD	3.8706 (0.049)	-.94812 (- 3.4256 /0.004)	Both short and long-run causality
LY→BD	0.0014 (0.969)	-.84969 (- 3.2122 /0.005)	Long-run causality
P → BD	3.1142 (0.078)	-.95554 (- 3.3997 /0.009)	Long-run causality and weak short–run causality
LER →BD	2.3386 (0.126)	-.86592 (-3.6249 /0.003)	Long-run causality

In order to investigate the short-run dynamics among CA, RIR, LY, P and LER, Wald tests are performed. The results are depicted also in Table 5. We can see that for Current Account (CA) Wald statistics yields a value $\chi^2 = 59.0344$ (p- value 0.000), hence the above regressors are jointly statistically significant. Consequently, there is a short-run causal impact CA on BD. Regarding Real Interest Rate (RIR), the test gives a value $\chi^2 = 3.8706$ (p- value0.049), that is statistically significant and indicates a short- run causal effect on BD. GDP (LY) has no short –run causal effect on BD since $\chi^2= 0.0014$ (p – value 0.969). Wald statistic test yields a value $\chi^2= 3.1142$ (p-value 0.078) for Inflation (P) and it shows that Inflation has a rather weak short-run effect on BD. Finally, for the Exchange Rate (LER), we can extract from the Table 5 above, the value $\chi^2 = 2.3386$ (p-value 0.126) which is statistically insignificant, therefore, we conclude that there is no short-run causal impact from ER on BD.

Since the results were derived from bivariate relationships using a rather limited data sample we provide further evidence from variance decomposition analysis (Table 6), based on the above estimated error-correction models. The analysis traces the dynamics of a shock in any of the involved determinants of budget deficit over time. Actually, table 6 reports the percentage of the variance of the t-year ahead forecast that is attributable to each of the shocks for T=1, 3 and 5. One year ahead could be interpreted as the short-run, three years ahead as the medium-run and five years ahead as the long-run.

Table 6: Variance Decomposition Analysis

Period	CA	RIR	LY	P	LER
1	0.22	0.37	0.11	0.51	0.24
3	0.43	0.44	0.23	0.57	0.33
5	0.48	0.45	0.28	0.58	0.36

The reported results reveal that in the short-run, P with 51% and RIR with 37%, dominate in explaining the behaviour of budget deficits. However, in the medium and long-run horizon all determinants become very significant with LY to be rather weaker. Summarizing, the results are in line with the evidence obtained from the error-correction estimates and causality tests.

Bearing in mind the overall evidence, we can argue that an increase in current account, interest rate and inflation, cause Budget Deficit to increase as well and this should be carefully considered from the economic policy authorities. However, it should be mentioned that except the macro fundamentals referred above, there are other factors which may have an impact on budget deficit as well and are not considered in the present research.

Conclusions

In this paper, there was an attempt to investigate the determinants of budget deficit by developing an econometric model that would relate budget deficit in Azerbaijan (BD) to macroeconomic factors of the economy, namely, inflation (P), interest rates (RIR), current account as a percent of GDP (CA), exchange rate (LER), and Gross domestic Product (LY). First, unit root tests (ADF) were applied, in order to avoid any spurious regression and to ensure the reliability of the derived results. Next, the ARDL cointegrating technique was applied, being the most suitable for the given empirical analysis due to the limited number of the available data. Finally, the EC models were estimated in order to be examined for long-run and short-run Granger type causality running from the independent variables on the dependent ones. In sum, the application of the above mentioned techniques revealed that:

- Stationarity tests results suggest that, all variables are non stationary in levels and stationary when tested in first difference form
- Based on the ARDL method, it can be observed that the value of F-statistic exceeds the upper bound of the critical value bounds in all cases and consequently the tests suggest that there exist long-run equilibrium relationships between the BD and each one of the examined determinants CA, RIR, LY, P and LER with long-run causality running towards BD.
- The long-run ARDL estimates revealed the following: an increase in current account by 1% increases budget deficit 4.2%, an increase in interest rate by 1% decreases budget deficit by 0.008%, an increase in GDP by 1% decreases budget deficit by 3.6%, an increase in inflation by 1% increases budget deficit by 0.007% and an increase in exchange rate by 1% decreases budget deficit by 22.2%.
- Regarding the Error Correction specification, evidence of long-run causality running from CA, RIR, LY, P and ER to BD was found. There were also found evidence of short-run Granger causal effects running from CA and RIR towards BD and a rather weak causal effect from P to BD. However, there is no short – run causality running from LER on BD.

Nevertheless, there is a weak point in the present approach, along the paper, as there are other factors which may have an impact on budget deficit as well and are not considered in the given research. These factors may be macroeconomic factors like unemployment, money reserve, and government policy factors. Besides, the empirical analysis should be in a multivariate framework, but due to the limited number of available data it is not possible.

References

Diokno, B. E., 2006, 'Economic and Fiscal Policy Determinants of Public Deficits: The Philippine Case', *Discussion Paper*, 0702, School of Economics, University of the Philippines, February 2007.

Engle, F. R. and Granger, C. W., 1987, 'Co-integration and Error Correction: Representation, Estimation and Testing', *Econometrica*, 55, pp. 251-276.

Galli, E., Padovano, F., 2005, 'Sustainability and Determinants of Italian Public Deficits Before And After Maastricht', *CESIFO working paper*, 1391.

Gujarati, N. D., 2003, *Basic Econometrics*, 4th Edition, McGraw-Hill, New York.

Harvey, A.C., 1981, *The Econometric Analysis of Time Series*, Philip Allan, London, Table C, pp. 364-5.

Pesaran, M. H. and Pesaran, B., 1997, *Working with Microfit 4.0-Interactive Econometric Analysis*, Oxford University Press, Cambridge.

Pesaran, M. H. and Shin, Y., 1999, An Autoregressive Distributed Lag Approach to Cointegration Analysis', *In Econometric and Economic Theory in the Twentieth Century*, S. Strom (ed), Cambridge University Press, Cambridge.

Pesaran, M. H., Shin, Y. and Smith, R.J., 2001, 'Bounds Testing Approaches to the Analysis of Level Relationships', *Journal of Applied Econometrics*, 16, pp. 289-326.

Phillips, P.C.B. and Perron, P., 1998, Testing for a Unit Root in Time Series Regression', *Biometrika*, 75, pp. 355-46.

Roubini, N., Sachs, J., 1988, 'Political and Economic Determinants of Budget Deficits in The Industrial Democracies', *NBER Working Paper*, 2682.

Rupinder, S., Juhani, L., 1999, Azerbaijan: Recent Economic Developments and Policy Issues in Sustainability of Growth, *BOFIT Discussion Papers*, 5.

Tujulaand, M., Wolswijk, G., 'What Determines Fiscal Balances? An Empirical Investigation in Determinants of Changes in OECD Budget Balances', *ECB Working Paper*, 422.

About IJESAR

The **International Journal of Economic Sciences and Applied Research** is a blind peer-reviewed, scholarly journal devoted to publishing high-quality papers and sharing original economics research worldwide. It intends to serve as an outlet for theoretical and empirical research in all areas of economics. These include, but are not limited to the following:

- Economic Theory and Policy
- Micro- and Macro-economics
- International economics
- European economics
- Regional economics
- Public economics
- Financial economics
- Business Management
- Business Informatics
- Accounting and
- Mathematical methods for economics.

Papers in other sub-disciplines of economics (e.g., labour, public, money, industrial organization, health, environment and history) are also welcome if they contain international or cross-national dimensions in their scope and/or implications. The Journal especially encourages the submission of articles, which are empirical in nature, or deal with issues of open economy macroeconomics and international finance. Theoretical work submitted to the Journal should be original in its motivation or modelling structure. Empirical analysis should be based on a theoretical framework, and should be capable of replication. It is expected that all materials required for replication (including computer programs and data sets) should be available upon request to the authors.

Instructions for Authors

1. Submission of Papers

Please submit papers by e-mail to editassist@ijesar.org. If electronic submission is not possible, papers (and any other correspondence) should be sent to the IJESAR Editorial Manager:

Address:	TEI of Kavala, Agios Loukas, PC 65404, Kavala, GREECE
Phone:	(+30) 2510 462193
Fax:	(+30) 2510 462193
E-mail:	editassist@ijesar.org

2. Format of Papers

IJESAR accepts submissions in English only. The submission should include a cover page showing the author's name and affiliation(s), an address for correspondence (including email), and any acknowledgements. The title but not the author's name should appear on the first page of the manuscript, as IJESAR uses a blind peer-review process. Therefore, all information revealing the author's identity will be removed before we send the paper to the referees. Papers should be submitted in Microsoft Word electronically by e-mail as attachments to the Editor, to the e-mail address: editassist@ijesar.org.

A template is a set of styles and page layout settings that determine the appearance of a document. The use of the template is mandatory (see the webpage for the template file format).

3. Copyright

Your contribution must be original work that has not been published or submitted for publication elsewhere. The submission of a manuscript represents an offer to transfer the exclusive right of publication to the editor and publishers. The copyright to articles and figures published in this journal is reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without written permission. The transferred right of publication includes the right to process data related to the manuscript and to produce further duplications for commercial purposes. The authors retain the right to allow further publication after a minimum of one year. Every reprint must contain a reference to the original publication in "International Journal of Economic Sciences and Applied Research" including reference to the corresponding issue. The IJESAR subscribers may copy or download the paper from the network. Any distribution or publication of said paper by subscribers shall constitute an infringement of copyright.