Evaluating the Economic Growth Using Artificial Neural Networks and Panel Fixed Effects

Elmira Emsia

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Eastern Mediterranean University
October 2017
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Assoc. Prof. Dr. Ali Hakan Ulusoy
Acting Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Doctor of Philosophy in Economics.

Prof. Dr. Mehmet Balcılar
Chair, Department of Economics

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Doctor of philosophy in Economics.

Assoc. Prof. Dr. Çağay Coşkuner
Supervisor

Examining Committee

1. Prof. Dr. Tarkan Çavuşoğlu
2. Prof. Dr. Mustafa İsmihan
3. Prof. Dr. Özlem Önder
4. Prof. Dr. Sevin Uğural
5. Assoc. Prof. Dr. Çağay Coşkuner

ABSTRACT

This thesis uses a panel data to investigate the effects of eight macroeconomic variables on the evolution of growth rate of Gross Domestic Product per capita. The panel data consist of 23 years of observation for ten developed and ten developing countries. The years covered are from 1990 to 2012. The independent variables selected are: (i) initial GDP per capita (INIGDPPC) to account for the effect of convergence (ii) terms of trade (TOT), (iii) trade openness (OPEN), (iv) gross fixed capital formation (GFCF), (v) human capital (EDUC) measured as average years of schooling, (vi) inflation (INF), (vii) government size (GOVT) and (viii) population growth (POPUL). The thesis methodology is unique in combines cutting-edge data-driven models such as hybrid artificial neural network with genetic algorithm (ANN/GA) and fixed effect panel model. First, the impact of eight independent variables on growth is investigated and dominant variables are identified by using three data samples: developed countries only, developing countries only, and developed and developing countries together. Moreover the study uses three different data formatting for each sample: annual data, periodic data of 4 years overlapping and periodic data of 4 years non-overlapping. Second, two estimation methods are used to predict values of growth. This allows us to compare those forecasting methods with each other. The analysis indicates INIGDPPC, INF, GFCF, GOVT, EDUC, POPUL, TOT and OPEN variables have the statistically significant impact on growth in the panel regression. The INIGDPPC, POPUL, GOVT, and INF have negative and OPEN, EDUC and GFCF have positive statistically significant effects on the economic growth in developed and developing countries.

Moreover, the results obtained from the study have shown that the power of the hybrid ANN/GA method (combined the artificial neural network method and genetic algorithm) is more than Panel fixed effect estimation method in predicting the economic growth.

Keywords: Economic growth, Panel data, hybrid ANN-GA.

ÖZ

Bu tez, panel veri kullanarak, sekiz tane makroekonomik değişkenin kişi başı gayri safi yurt içi hasıla büyüme oranına etkisini inceler. Pnael very 23 yıldan; ve onu gelişmiş, onu da gelişmekte olan , toplam 20 ülkeden oluşmaktadır. Veri 1990 ile 2012 seneleri arasındaki yılları kapsamaktadır. Kullanılan 8 makroekonomik değişken şunlardır: (i) Kişi başı GSYİH başlangıç değeri (INIGDPPC), (ii) ticaret terimi (TOT), (iii) ticaret açıklığı (OPEN), (iv) yatırımlar (GFCF), (v) insan sermayesi (EDUC), (vi) enflasyon (INF), (vii) hükümet harcamaları büyüklüğü (GOVT), ve (viii) nüfus artış hızıdır (POPUL). Çalışma iki tane metodoloji kullanmaktadır: Biri genetic algoritma ile birleştirilmiş yapay neural network metotu, bir diğeri ise panel fixed effect metotudur. Calışma 3 ülke grubu ve 3 veri formatlaması kullanarak, toplamda 9 kez seçilen 8 makroekonomik değişkenin büyümeye etkisini inceledi. Ülke grupları: sadece gelişmiş ülkeler, sadece gelişmekte olan ülkeler, ve gelişmiş ve gelişmekte olan ükleler beraber olmak üzere 3 tane idi. Veri formatı ise yıllık veriler, 4 yıllık periyodik veri (yıllar örtüşüyor) ve 4 yıllık periyeodik veri (yıllar örtüşmüyor) şeklindeydi. Bu çalışma ayni zamanda ANN/GA metodu ile panel fıxed effect metodunu büyüme tahminleri alanındaki karşılaştırmasını yapmıştır. Sonuç olarak, INIGDPPC, INF, GFCF, GOVT, EDUC, POPUL, TOT and OPEN değişkenlerinin *istatistiksel* büyüme değerlerine etkisi olduğu görülmüştür. Hem gelişmiş hem de gelişmekte olan ülkelerde INIGDPPC, POPUL, GOVT, ve INF eksi bir etki, OPEN, EDUC ve GFCF ise artı bir etki yapmıştır.

Sonuçlar ayrıva ANN/GA metodunun panel fıxed effet metoduna gore daha güçlü bir metot olduğunu göstermiştir.

Anahtar Kelimeler: Ekonomik büyüme, panel veri, hibrit ANN/GA.

ACKNOWLEDGMENT

I should thank Assoc. Prof. Dr. Cagay Coskuner for his continuous support and guidance in the preparation of this study and I also like to thank my family who allowed me to travel all the way from Iran to Cyprus and supported me throughout my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Since the inception of economics as a discipline in social sciences, substantial differences in the living standards of nations, and the main determinant of this - substantial variation in the long-run economic growth rates of nations over the decades or even the centuries - has been a challenge for the policy-makers and researchers. In fact, diversity in long-run economic growth performances is one of the most debated and researched economic problems. Finding a robust and a lasting solution to the low level of economic growth can only be achieved through a proper identification of main determinants of growth, and then, using policies to improve on those determinants.

Indeed, high living standards and economic prosperities of several developed countries are mainly due to the fact that these countries have experienced high and sustained level of development and growth for several decades. These growth and development quite often have showed themselves in many areas of economic and social life such as education, technology, capital accumulation, infrastructure building, increased trade and increased output. Improvements on these areas and many other factors may be the driving force of economic growth and thus they need to be properly identified.

One may be stunned by the diversity in living standards and growth performances of nations, but yet one may still ask himself or herself why growth is important and whether it is as important for developed countries as it is for developing and less developed countries. In general, increased growth rates mean increase in per capita income for most families, and thus higher consumption and higher living standards. Usually it also means lower level of unemployment and better affordability of education and health services. As such growth is crucial for both developed and developing countries.

It is also no doubt that economic growth results in an increase in the wealth of a nation as a whole, which empowers such nation to fight poverty, reduce unemployment and solve other socio-economic and political problems. This is the reason why many countries of the world – developed and developing countries- consider a high and sustained level of economic growth as one of the main objective of macroeconomic policies.

While countries in Europe, North America as well as a few selected countries in Oceania and East Asia enjoy high per capita incomes and high living standards, billions of people in Africa, South and South East Asia and Latin America are achieving comparably lower living standards. Nevertheless economic growth matters for all people in both developed and developing countries. In less developed countries, lack of sufficient income means that millions of people go by without having adequate nutrition, while lack of revenues means that governments fail to provide adequate health and education services as well as electric power transportation and communication services. This, in turn, feeds into low labor productivity and low-income generation. The only way out of this vicious cycle is by increasing the GDP growth rates and sustaining it at such high levels.

Economic growth also matters for the people of developed countries through much the same way. It matters for the unemployed people, it matters for the buyers of health-care services, education services, houses and so on. That is to say it matters for the consumers, but equally so for the producers and businesses. In fact, recent disruptions in GDP growth rates and thus lower living standards – due to global financial crises between 2007 and 2011 – have shown that low growth rates may even lead to disruptions in institutional infrastructures as in the case of Brexit and recent United State (US) elections of 2016.

Furthermore, it is well known that human wants are unlimited, while the world population is increasing rapidly. Thus the ability to meet such increasing consumer demand depends on our ability to increase the world output, that is the global GDP, which is achieved through improving on national growth rates. The world population was just 2 billion in 1930 and only 4 billion in 1975. By 2000, it surpassed 6 billion and by 2050, it has been projected to be around 9.5 billion (US Census Bureau, 2015). Unless, such projected increase in the world population growth is accompanied by an advanced production techniques such as level of technology and increased capital accumulation, it may be difficult to maintain current level of living standards.

Therefore, identifying the determinants of economic growth and accurate in sample forecasting of real GDPs on a regular basis is crucial for economic policy-making which impacts both the current and future generations in all countries – developed and developing. It provides economists, policymakers, private institutions and businesses the information needed for a sound policy-making, planning as well as business investment.

This dissertation is in line with the vast literature in the profession of economic growth. It aims to determine the main factors which impacts the growth rate of real GDP per capita and then it uses cutting-edge techniques to forecast the growth rate of real GDP per capita. Another aim of the study is to make a comparison between in-sample forecasting methods in order to contribute to the literature through identifying "better" techniques in such real GDP in-sample forecasting.

More specifically, this dissertation uses a panel data of 20 countries and 23 years of observation from 1990 till 2012. Ten of these countries are developed countries (United Kingdom, Germany, Japan, Spain, Norway, New Zealand, France, Australia, Sweden, Greece) while the remaining ten are developing countries (India, Venezuela, Turkey, China, Nigeria, Iran, Russia, Ukraine, Pakistan and Brazil). The data is an annual data but formatted in three different ways: (i) yearly observation, (ii) periodic of 4 years in a non-overlapping way, (iii) periodic of 4 years in an overlapping way. The number of countries and years are limited because of limitation of Artificial Neural Networked (ANN) method.

Finally, the study uses two different estimation techniques. One is a more conventional panel fixed effect estimation technique while the other is an Artificial Neural Networked (ANN) with Genetic Algorithm (GA) method. These estimations are repeated for all countries together as well as for developed and developing countries separately. Moreover, as mentioned in the previous paragraph, we repeat each estimation for three different data formatting. This allows us to check thoroughly the robustness of the estimation and make a proper comparison of the two estimation techniques. All estimation techniques will be done in Eviews and Matlab software.

The living standard of a nation is best measured by real GDP per capita – GDP divided by the population size. Thus, an increase in the living standards would be calculated as the percentage change in real GDP per capita, and is named as economic growth. The real GDP is adjusted for changes in prices, thus it eliminates the impact of inflation. Therefore, economic growth must be calculated as the percentage change in real GDP, and not the nominal GDP. Thus, in this study, the real GDP per capita growth is used to proxy for actual level of the growth rate.

Furthermore, we employ eight macroeconomic variables as our explanatory variables, which have all been used commonly in empirical growth literature. These variables are gross fixed capital formation (GFCF), trade openness (OPEN), terms of trade (TOT), inflation (INF), human capital level that is proxied by average years of schooling (EDUC), government size (GOVT), population growth (POPUL) and initial GDP per capita (INIGDPPC) which stands to capture for the effects of convergence. As mentioned earlier, we refer to both empirical and theoretical literature in the area to identify these variables as our explanatory variables. Some of these literature may be listed as Mankiw, Romer and Weil (1992), Barro (1996), and Barro and Sala-i-Martin (2004). All of these data come from the following sources: the World Bank Database, Federal Reserve Broad Economics Database, Organization for Economic Co-operation and Development (OECD) National Accounts, and *United Nations Educational, Scientific and Cultural Organization* (UNESCO) Institute for Statistics.

As for the novelty of this study, as mentioned earlier, we make use of both the conventional panel fixed-effect method and the relatively new ANN/GA estimation technique on the same panel dataset and with the same explanatory variables; and the estimations are repeated for developed and developing countries together as well as separately and also for three different data formatting (annual, periodic overlapping and periodic non-overlapping). To the best of our knowledge, this study is the first to combine such methodologies and make such comparisons in empirical economic growth analysis. In fact, this is the only study, which uses ANN/GA method in "panel data "economic growth literature.

There are some studies that make use of both ANN and "time-serie" estimation techniques in economic growth literature with the aim of making comparison of these techniques. Most of these studies focused on a commonly used growth determinants and empirical methodologies. For instance, some of these studies are: Tkacz and Hu (1999), Heravi, Osborn and Birchenhal (2004), Binner et al. (2005), Sameti et.al (2013), Feng and Zhang (2014) and Sokolov et al. (2016). Most of these studies compared ANN method with ARIMA (Autoregressive Integrated Moving Average), AR (Autoregressive), and other time series linear in-sample forecasting models. They found that ANN method is better and more efficient than other time series in-sample forecasting models. We review briefly the articles related to the present study in the next chapter.

However, none of the existing literature on the topic has compared the panel fixed-effect model with a "panel data-based" ANN/GA method on the topic in a way, which has been investigated in this dissertation. In fact, one of the main objectives of this dissertation is to determine whether the predicting power of growth using ANN/GA provides better performance when compared to conventional panel method. Specifically, we apply the ANN/GA to forecast economic growth in developed, and developing countries, and we examine the forecast performance measures using the root mean squared error (RMSE) in the panel fixed effect and ANN/GA methods.

Conclusively, this dissertation attempts to answer the following research questions using new methods and through better application of the growth models: Why are some countries more economically buoyant and developed than others?

Which economic factors affect the growth of developed and developing countries? Which method can predict the economic growth better than others? Why and how?

1.1 Thesis Structure

The thesis comprises six chapters. Chapter 1 makes an introduction of the study, while in Chapter 2, we provide both theoretical and empirical literature review on the economic growth concept. Chapter 3 presents the empirical specification and gives information about the data used in this study. Chapter 4 highlights the methodology used in panel data fixed effect estimation technique and its results, while Chapter 5 presents and discusses the ANN/GA method and its results. Finally, in Chapter 6, we conclude the study and make suggestions for future research.

Chapter 2

LITERATURE REVIEW

Achieving economic growth has been one of the fundamental responsibilities of economists and policy-makers both in central and local governments. It is a crucial concept in the sense that it is the main determinant of the living standards of nations. It is also crucial in the sense that lower growth rates quite often may mean higher unemployment, other higher socio-economic problems as well as disrupted democracies and international relations.

Thus, this study aims to identify main determinants of economic growth and cast a light on better GDP estimation techniques with a hope of contributing not only to the vast economic growth literature but also to policy improvements geared towards faster economic growth. In line with this, in this chapter, we aim to present a review of both the growth theories and the empirical literature on economic growth.

In Section 2.1, let us present a brief on growth theories while in section 2.2, we present the literature review of the related empirical papers.

2.1 A Review of the Growth Theories

Economic growth is one of the most debated and researched topic in economics. It is done so both theoretically and empirically, and thus, there is a vast literature in the area. This dissertation is an empirical study; nevertheless, it has its roots embedded in theoretical studies. Thus, we find it useful to overview a brief theoretical literature, but not to over-extend it in order not to focus the reader away from the main purpose of this study, which is to determine empirically the main determinants of growth, and use this to identify a better GDP growth estimation technique.

Let us now look at some of these theoretical studies. In the first part, we will provide a brief history of growth theory as highlighted by R. Barro and X. Salai Martin (1995). In the parts that follow, we will provide a summary of some of these models.

2.1.1 Brief History of Growth Theory

One of the earliest and yet most comprehensive growth theory was Ramsey (1928). This study had utilized inter-temporal household optimization in explaining growth, and as such it was decades ahead of its time. In fact, the paper has only been widely accepted and credited after 1960 that is thirty years after its publication.

Another corner stone in growth theory is the Harrod-Domar model, which was synthesized from two separate studies: Harrod (1939) and Domar (1946). These studies were carried right after the Great Depression, and thus reflected the essence of its time by concluding that the capitalist system was inherently unstable. The authors achieved such results by mainly utilizing production functions where the inputs had little or no substitutability such as in Leontief Production Function.

However, the most fundamental growth theory emerged in 1956 as Solow-Swan model which was also a synthesis of two separate studies: Solow (1956) and Swan (1956). Its main assumption is a neoclassical production function with constant returns to scale but decreasing returns to each input. With a constant and exogenous saving rate, the model turns out to be simple to comprehend, and yet with a strong conclusions for the researchers and policy-makers. For example, one of these conclusions of this model is the conditional convergence, which simply states that all countries would achieve same income per person in the steady state on the condition that these countries have the same economic characteristics such as the level of technology and the rate of savings.

Another strong conclusion of the Solow-Swan model is that the continuation of growth of per capita income depends on continuous improvement in technology. However, the level of technology as well as the rate of savings in the model was exogenous, which turns out to be the main shortcoming of the model.

Cass (1965) and Koopmans (1965) combined Ramsey style consumer optimization with then-existing neoclassical growth model to formulate an endogenous saving rate to the model. This became to be known as Cass-Koopmans model. However, formulating the level of technology as endogenous was more difficult as it could involve increasing returns to scale, and thus, break-down of assumption of perfect competition.

Arrow (1962) and Sheshinski (1967) introduced "learning-by-doing" into the growth theory where the level of technology is improved through the decision of increasing production and/or investment. Romer (1986) and Lucas (1988) advanced these models as they included human capital to the growth theory. This would not achieve an endogenous technology level, but it would allow a continuous growth as human capital would include a non-diminishing returns to the input.

Finally, a "real" endogenous growth models emerged in 1990s through the works of Romer (1987 and 1990), Aghion and Howitt (1992), Grossman, and Helpman (1991), where explicit Research & Development (R & D) theories and imperfect competition models have been included.

These recent models are highly mathematical and in fact somewhat disconnected from the interest of empirical research. Thus, we will omit the recent models, but try to present more detailed framework for Harrod-Domar model, and Solow-Swan model, followed by a brief discussion of some endogenous growth models.

2.1.2 Harrod-Domar Model

Two famous economists Evsey Domar and Roy Harrod formulated Post-Keynesian theory of economic growth. Since these authors' independent works achieved very close results, their work has become to be known as the Harrod-Domar 's theory. Domar completed the Keynes theory in such a way that investment is a factor of production growth through a creation of production capacities. Domar's theory determines an investment growth rate that directly depends on the savings share in GDP inland and the average efficiency of investments. Then, investment should grow at this rate to ensure the growth of revenue. Therefore, he got good conclusion for the economic policy inland. The theory showed that with an investment growth balance between aggregate supply and aggregate demand can be provide.

The state can hold the balanced growth of investments and thus determine the productivity of capital through influencing the technological progress rate (savings share in GDP). Harrods's theory showed the ration of capital growth to the economic growth which is dedicate to the growth path. The theory also analyse the relationship between savings and income and showed that the expectations of entrepreneurs are the basis of the mechanism of balanced growth.

In Harrods's theory, the actual growth rate can be determined by the labour and capital productivities rates. If the real growth rate is consistent with the full utilization of existing capital resources, the economy will achieve a stable development. In his theory, the maximum probability of economic growth with full use of labor is called the natural rate. The stable equilibrium of the economic system is ensured by the equality of the guaranteed growth rate and the full employment. However, only with active state actions, such equality is maintained.

Over the time, combination of Harrod and Domar's works into a single theory was called Harrod-Domar model. This models basic conclusion is that under the technical conditions of production, marginal propensity to save determines economic growth, but the market dynamic equilibrium is unstable essentially and thus it needs purposeful interventions of the state to maintain the full employment.

The Harrod-Domar‘s theory limitation showed that, firstly, they need to have a prerequisite that economic growth depends linearly on the growth of investment. The model also assumes that there is no dependency between economic growth and the growth in labour demand. Finally, technological progress is not considered in the theory.

Moreover, Post-Keynesian historical setting was another limitation of the theory. This theory provided adequate and then-well-accepted explanation regarding the actual processes of economic growth when economic growth largely depended on a growth of production capacity utilization in the 1930s and the post-war period. However, as the production development in the 1950s till 1970s predominantly depended on qualitative and technological changes, the emphasis has shifted towards neoclassical theories of economic growth.

2.1.3 Neoclassical Solow-Swan Growth Theory

In the 1950s – 1960s, the first theories of neoclassical growth came about. As opposed to the Keynesian suggestion of state intervention, Robert Solow, together with the other neoclassical scholars, believed in competitive free-market system with minimal state role. They believed in allowing producers to achieve their growth potential by using of their available resources in a competitive market. The payments to these factors of production would then be determined through marginal productivity. That is, any production factor would earn an income according to its marginal.

Solow also suggested that a necessary condition for the economic equilibrium is determined by supply and demand equivalence, while the total supply is specified on the Cobb-Douglas production operation. Through such a production function, Solow model reveals interconnections between investments, workforce and technological progress as the sources of economic growth.

The key factor of this theory shows that the economic growth is determined by the savings rate: When the savings rate is high, the capital stock becomes larger, and so the production level can be greater and more. Other reason for the ongoing GDP growth in the stable economic condition was a population growth in the Solow's theory. However, Solow explains that if the growth of the population is not accompanied by an increase in investments, this would lead to a decrease in capital-labour ratio and results to lower income per person.

The other source of economic growth, in the model, is the progress of technology, which is the sole condition for the sustainable economic growth of welfare measured as GDP per capita. Dutt, and Ros (2008) explained "technical progress" as a qualitative change in the production such as the improvement of the organization, production scales growth or an increase in the educational level of workers and so on. In general, the Solow's theory points that the growth of technology as the main factor for the continuous growth of living standards. Thus, through his model, Solow, among all of his contemporaries, turns out to have a better perception of the economic growth, as the production efficiency deemed as the source of economic growth and social progress.

Furthermore, Robert Solow introduces the concept of “golden rule of savings” which is essentially the optimal level of savings that would maximize the per capita consumption in the steady state. The optimal saving rate or the golden-rule level of saving rate determines the optimal level of capital stock per capita, which maximizes the level of consumption in the steady state.

2.1.4 Theories of Endogenous Economic Growth

In the 1980s and 1990s, a new line of growth theories emerged, which reflects the impact of imperfect competition and the role of possible changes in the profit rate. In this theory, the scientific and technical progress has been considered as an endogenous factor created by internal reasons.

Paul Romer and Robert Lucas for the first time considered endogenous character of the most important technological innovations. They opined that human capital plays an important role in determining long-term economic growth. According to these theories, human capital can increase GDP growth by stimulating technology, invention and innovation. The endogenous theories are same as the Neo-classical ones but with significant differences in some part of assumptions and results.

According to the Solow model, the state with the support of economic policy instruments cannot provide the long run growth rate by influencing the savings rate (Romer, 1989a).

The theory of endogenous growth resolves the shortcoming of neoclassical theory by rejecting the marginal capital productivity diminishing. The assumption is that the impact of the scale of production through the entire economy can be concentrated and often focuses on the impact of external influences on the profitability of the investment. In the theories of endogenous growth, economic growth is not only originated from technological progress in the long term. Therefore, in the following the determinants of economic growth in the theories of endogenous growth are defined:

- The human capital quality depended on investment in human development such as schooling enrolment, health, and education;
- In the Imperfect competitive markets, government should protects from intellectual property rights;
- Government supports the innovation, new technology and science;
- The role of governments is to absorb new technology and create secure environments for investment.

Thus, the endogenous growth theories compared to the neo-classic ones support government intervention in the development process. The endogenous growth theories are divided into 2 groups.

The first group is theories of Romer (1989b) and Lucas (1988) which believes that human capital appears as an important determinant of economic growth. In fact, the inclusion of human capital in the production function distinguishes the theories of this group. Paul Romer names "knowledge" or "information" as the key factor in the endogenous growth theory that assumes the information is available to everyone to be used.

Romer believes that the total number of human capital is unchangeable over time, and only according to the function of consumer preferences, its distribution is possible between the research development activities and the production circle. The main idea of the Romer is that an exchange between today's use and knowledge that can be used to expand tomorrow.

In fact, Romer's idea is called "research technology," which creates "knowledge" from the past consumption; therefore, the economic growth is dependent on human capital values to acquire new technology. In fact, the new knowledge and idea affects the economic growth indirectly with the provision of human capital accumulation. This means that the human capital accumulation is essential for the economic growth of any country. Altogether, Romer in his theory implies that greater accumulation of human capital prepares the countries with higher economic growth rate.

In the theory of Robert Lucas, in contrast to the Romer's, accumulation of human capital is an outcome of optimization based on relative costs of alternative choices. The two choices are allocating time for: (i) contributing to current production and (ii) accumulation of human capital. In fact, it is the outcome of this optimization, which determines GDP growth rate. For example if a nation allocates less time for working and producing, this will lead to a reduction effect in the current production. At the same time, it will also increase the product output growth due to accelerated investment in human resources.

In the second group of theories, research and development activities are a main factor of growth. J.Grossman and E.Helpman describe the effect of endogenous high technical innovations on economic growth rates (UN, 2011). These authors also have indicated that subsidies for the introduction of new technology and R& D will boost the country's economy. This theory considers the possibility of inflow or outflow of capital to fund R & D.

Two of the followers of this group are P.Howitt and P.Aghion who believe in the endogenous technological progress theory, which accordingly economic growth is driven, by technological progress. Competition between firms results to technological progress by generating technological innovation, which brings new products and new technology used in a more effective production.

The main objective of the agents in research sector is to gain monopoly rents, which will allow the firms to pay for their costs, resulting from development and innovation activities. Intersectional movements of professionals between goods production and the R & D sector determine the rate of economic growth. Thus, endogenous growth theories as presented in the previous 3 or 4 paragraphs, formalize a link between economic growth rate and accumulation of human capital. All in all, these theories outline the reasons of differences in growth rates of different countries; the effectiveness of governments' technical, scientific, and industrial policies; and also the impact of trade openness and international finances on economic growth.

2.2 Empirical Literature of the Economic Growth

Let us now depart from the theories and use this section to provide a summary of some relatively old but well-known empirical papers as well as some recent ones. As mentioned earlier, there is a huge amount of empirical literature in the area. Thus, we try to present either only the papers, which are considered corner stones through their contributions to the literature, or the most recent papers in order to highlight the recent trends and results.

For example, one of the most pioneering studies who investigated the effect of different variables on economic growth in a cross-country framework is Barro (1991). Barro adopted a cross-sectional study of 100 countries, where the economic growth rate was calculated over the thirty years from 1960 to 1990. This paper not only provides a basis for empirical growth literature but also investigates about the concept of convergence. The results show that there is a negative relationship between political instability and growth. This is concluded to be via lowered property rights and lowered investment. Another finding is the support of conditional convergence. That is the countries with lower initial GDP per capita levels tend to grow faster.

Another essential study in the field is by Mankiw, Romer, and Weil (1992). Their empirical work used a modified Solow model and achieved remarkable results. Mankiw et al. found that the countries with different saving and population rate had the different level of income per capita. Moreover, they suggested that if they put human capital (education and training) into the Solow model, this model would produce superior empirical results. For their study, the authors used a sample of 98 non-oil-producing countries where subsamples included 22 OECD countries and 75 developing countries in 1960. They collected annual real income, government and private consumption, investment, number of labor, education, and population data and covered the period 1960-1985. The fundamental conclusion was that the accumulation of human capital has a larger positive impact on income per capita than the accumulation of other production factors, so that the authors were able to state that the "differences in saving, education, population growth, taxation and political stability" could explain vast cross-country differences in income per capita.

Nazrul Islam (1995) used dynamic panel data model for studying cross-country growth convergence in line with the previous works such as Solow (1956), Mankiw et al. (1992) and Barro et al. (2004). He believed that faster rate of convergence depends on the role of technological progress term as a determinant of the steady state level of income in cross-country. Islam's econometrically superior model concluded that a country can increase the economic growth in the long run by improving on the technological progress components which also have salutary effects on saving and population growth rates.

Let us now review some recent papers in the growth empirics. Since our study is not focused on any one specific explanatory variable, and rather is aiming to identify all main determinants of growth, we had to be very brief and selective. We will now present a few recent papers for each of our eight explanatory variables.

For example, Ilegbinosa et.al (2015) tested the effect of domestic investment and government expenditure on growth by using time series data in Nigeria between 1970 and 2013. They used multiple regression and cointegration method to analyze the sample data. They concluded that private investment had a positive effect on GDP growth, but government expenditure had a negative impact on growth.

Ssewamala, Nabunya, Ilic, Mukasa, Damulira (2015) also investigated the effects of private domestic investment and various governmental policies on growth. They did so by using random and fixed effects, and dynamic longitudinal techniques for 15 sub-Saharan African countries from 1980 to 2008. Their result showed that per capita income growth was positively influenced by government policies, which would increase the gross capital formation, encourage the human capital, and provide credits for the private sector.

AbuDalu, Ahmed, Almasaied, Elgazoli (2014) focused simultaneously on the impact of the real effective exchange rate, terms of trade, domestic money supply, domestic interest rate, and inflation on real GDP. They used "Autoregressive Distributed Lag (ARDL) co-integration" method in the ASEAN-5 countries. They found that real effective exchange rate has a positive effect on the growth rate from aggregated supply channel, but other variables have the negative impact on growth in the long-run. Their empirical results were also similar to the empirical studies which were done by Dimitris and Christopoulos (2004), David and Guillermo (2005), Justin, Byung, Lee, and Mark, (2005) and Julian and Jay (2008).

Klasen and Lawson (2007) investigated the relationship between population growth, economic growth and poverty by using Ugandan data from 1960 till 2000. They found that the poverty reduction and economic growth promotion by population growth is difficult in Uganda. That is, higher population growth has an inverse impact on growth per capita income. This is in line with Mankiw et al. (1992), Barro, et al. (2004), Furuoka (2005), Headey and Hodge (2009) and Brückner and Schwandt (2013).

In contrast, some literature provides evidence that the population growth has a positive impact on growth such as in Hernandez, Ortiz, Alejandre, and Cruz, 2017 as well as Thuku, Paul, and Almadi (2013). Indeed, Thuku, Paul, and Almadi (2013), by using an annual time series data during 1963-2009 and using Vector Auto Regression estimation method, finds that high economic growth and economic development was created by high population growth in Kenya.

On the other hand, there are also plenty of studies who find no relationship between GDP growth rate and population growth rate. For example, Dawson, Tiffin (1998) and Thornton (2001) did a similar study in India and seven Latin American countries. Both studies found out that population growth do not have a significant impact on real gross domestic product per capita.

As for trade openness as a determinant of economic growth, the results are also mixed results as it was the case in population variable. For example, Buigut, Soi, Koskei, and Kibet (2013) investigate the impact of gross capital formation, foreign direct investment, and openness on economic growth during 1960-2010 in Kenya.

They find that openness does not have the impact on GDP in developing countries. Inversely, Adhikary (2011) points out that trade openness has a negative effect on real GDP growth while the foreign direct investment and capital formation have a significant positive impacts.

Similarly, while Alcalá and Ciccone (2004), Tan (2012), Javed, Qaiser, Mushtaq, Sai-ullaha, Iqbal (2012), Samimi, Sadeghi, Sadeghi (2011), Kreinin (2006), and Wacziarg and Welch, (2008) find that high level of trade openness robustly increases economic growth rate, other studies such as Rodrik and Rodríguez (2000) shows that it is difficult to find any relationship between openness and GDP growth rate. Gries and Redlin, (2012) also produces results which are similar to those of Rodrik and Rodríguez (2000). In contrast, Adhikary (2011) Levine, Renelt (1992) and Rodrik (1992) are among the studies which conclude that trade openness leads to lower level of economic growth.

Hadass and Williamson (2003) examines the terms of trade effects to find that the global terms of trade shocks decrease the growth performance of developing countries relative to developed countries from 1870 to World War I. Sachs and Warner (1995, 2001) follows their study and indicates that the countries with greater resources grow more slowly than the countries with poor resources, implying that the terms of trade shock was more of a problem for natural resourced–based economies rather than for industrialized developed economies.

On the other hand, Blattman, Hwang, and Williamson, (1997) investigated the effect of terms of trade volatility on growth by using a panel data of 35 countries during1870-1939. They found that terms of trade have a statistically significant negative impact on growth for all countries in the sample.

Furthermore, authors such as Mendoza (1997), and Easterly, Kremer, Pritchett and Summers (1993) concludes that the relationship between terms of trade shocks and per capita GDP growth rate is positive.

Various studies in growth literature from 1950 until 1960 had showed that inflation had a positive impact on capital accumulation, and thus on economic growth. However, Fischer and Modigliani (1978) conclude that inflation resulting in as a taxation of the capital has a negative impact on income. Recent studies by Yabu, and Kessy (2015) and Gillman, (2009), produce similar results as in Fischer and Modigliani (1978).

Similarly, Fischer (1993) and De Gregorio (1993) both use panel regressions to conclude that there is a negative relationship between inflation and growth. Barro (1995) and Sala-i-Martin (1997) find that this relationship is nonlinear while Andres and Hernando (1997) also produces similar results. Ghosh and Philips (1998) and Gylfason (1991) can be listed, among many others, as some other studies with the conclusion that growth is negatively associated with inflation.

Easterly and Rebelo (1993) implied the government policy role is important for promoting economic growth. The government consumption expenditure on productive activities had a positive impact on growth while government consumption expenditure on unproductive activities had no impact on that. Similarly, Swamy (2015) determines that the relationship between the government and economic growth is positive.

On the other hand, Devarajan, Swaroop, and Zou, (1996) investigates the impact of government expenditure on growth by using a panel data from 1970 till 1990 for 43 developed and developing nations. They find that any increase in capital component of public expenditure has a negative effect on the economic growth.

Nenbee and Medee (2011) reaches to a similar conclusion that increases in federal government expenditure have either negative or no impact on growth in the short or long terms as they use a vector auto regression and vector error correction model on a Nigerian data from 1970 till 2008.

As mentioned earlier, Lucas (1988) and Romer (1990) are the pioneer works advocating that human capital has an important positive impact on long-run economic growth. The human capital can be considered knowledge, skills, and the ability of labor force in the labor market, but quite often it is proxied as the average level of education in a country (Barro, 1996). Thus, a report shows that some studies estimate that increasing average education in the population by one year would raise the level of output per capita by between 3 and 6 percent.[1]

Furthermore, Cohen D. and Soto M. (2007) use panel data estimation method to investigate the role of average years of schooling on growth. They conclude that human capital has a positive important role on growth in all of the High-Income, Middle-Income and Low-Income countries as well as for the following regions (Middle East and North Africa, Sub-Saharan Africa, Latin America and Caribbean, East Asia and Pacific, South Asia, and Eastern Europe and Central Asia) during the time period of 1960-2000. Their results are similar to those of Nehru Vikram, Swanson Eric, and Dubey Ashutosh, (1993).

2.3 Empirical Literature Review of ANN

In economics profession, it is quite common to use econometric estimation techniques to test for validity of economic theories as well as to forecast. Usually these methodologies can be grouped into two broad categories: (i) parametric

[1] Joint Report by the Economic Policy Committee (Quality of Public Finances) and the Directorate-General for Economic and Financial Affairs, (2010).

modeling which includes linear autoregressive and nonlinear Markov switching models and (ii) non-parametric techniques which includes kernel models, neural networks, and wavelet models. (Tong, 1990, and Pena, Tiao, and Tsay, 2003).

While parametric modeling is frequently used for economic forecasting and testing of theoretical analysis based on consistency, asymptotic properties and robustness of parameters, several problems may appear because of strong assumptions regarding model specification, estimation techniques and asymptotic properties of the estimated parameters. Non-parametric methods have overcome some of these problems by avoiding *a priori* specification of modeling approach and distribution of residuals. Recent high-speed computers help to overcome further such problems as they help to develop search algorithms from appropriate selection criteria (Becker, Chambers, and Wilks, 1988)

This thesis attempts to use both parametric conventional panel fixed-effect and non-parametric nonlinear artificial neural network (ANN) techniques in order to contribute to the literature in determining which technique is superior in such sample forecasting. One advantage of ANNs is that it can figure any nonlinear input/output. Indeed, Hoptroff, Bramson, and Hall, (1991) used ANNs to forecast trends in many UK macroeconomic variables, including the UK GDP, in order to predict turning points in the UK economy. Their results indicate that neural networks are comparable to the old method, and that they can be used to make predictions. Similarly, Kuan and White (1994) and Swanson and White (1997) introduced ANN as an application of econometrics and demonstrated that the use of neural networks to forecast macroeconomic variables is a better method than traditional method for nonlinear models. Tkacz and Hu (1999) have also confirmed this conclusion.

Among other papers which focus ANN, we can list a review paper by Yatchew (1998), as well as the works by Tkacz and Hu (1999); Blake (1999) who both use the neural networks to forecast Canadian GDP; and Ferrara, Guégan, and Rakotomarolahy (2010) who uses nearest neighbor and radial basis function methods to predict euro-area GDP. Nevertheless, the use of these techniques to forecast growth has been limited.

ANN can be weighted by the general algorithm using bit strings. In each test, prediction error is evaluated to measure a fitness value. The lower the error, the greater the fit, thereby yielding good weights. An empirical study by Koutmos and Booth (1995) proposed a “hybrid model” to investigate returns on developed-market stock exchanges. Other authors used the hybrid method to assess the relationship between the stock-price index and stock-price volumes.

Another example for the use of hybrid ANN model is the paper by Shi, Chen, and Xie, (2006). The authors considered a hybrid ANN model with genetic algorithm as an attempt to predict for China's GDP growth after the year 2000. In fact, they use not only an artificial neural network (ANN) trained with a genetic algorithm (GA), but also a model of overlapping generation (OLG) in order to predict trends in GDP. They find that a hybrid ANN/GA model can predict the economic growth better than the OLG model does.

Similar conclusions are reached by the papers: Samimi, Sadeghi, and Sadeghi, (2011); Demir, Shadmanov, Aydinli, and Eray (2015); Li, Xu, and Sun (2014); Chaudhuri, Ghosh, (2016); Heravi, Osborn and Birchenhal (2004); Feng and Zhang (2014); Binner et al. (2005); and Sameti et.al (2013). They all indicate that ANN which is trained by Genetic algorithms performs better and is more effective than both the linear and other nonlinear models. In line with these, Gjylapi, Proko, and Hyso, (2016) also finds that the GA progresses ANN method's efficiency compared with standard Feed- Forward Multilayer Perceptron Back Propagation Model.

As for panel-data application of ANN techniques, Giovanis (2008) compared in-sample forecasting performance of traditional panel regression with fixed effect and random effect, ARCH model, and ANN model for the greenhouse gas emission of 15 European Union countries from 1990 to 2004. Although, this paper is not about economic growth, it is an essential paper for our study as it is one of the rare panel data applications of ANN modeling. Giovanis concludes that ANN method could forecast greenhouse gas emissions far better than all traditional panel methods based on the results of the RMSE levels.

2.4 Overview of Present Study

Although the papers mentioned above indicate the superiority of ANN, especially ANN/GA models, several other papers claim the otherwise. Thus, quite often conventional econometric approaches are preferred in building prediction and in-sample forecasting models in almost all economic areas. For this reason, we would like to employ both conventional panel fixed effect and ANN/GA methods in estimating an equation for economic growth. The main purpose of the paper is then to compare the power of these methods in the prediction of growth rate in selected developed and developing countries by using a panel data for the period 1990 to 2012. By doing so, we also hope to contribute to the growth literature, as this is one of the rare studies which applies ANN/GA methods in GDP growth estimation; and moreover – and to the best of our knowledge – the only paper which does so in a panel data framework.

Chapter 3

EMPIRICAL SPECIFICATION AND DATA

As already mentioned, this thesis has three aims: (i) it aims to identify the main macroeconomic variables which affect the economic growth by using a panel data of twenty countries and twenty three years from 1990 to 2012; (ii) it applies both conventional panel fixed effect estimation technique and an ANN/GA model in order to assess which method is superior in GDP growth estimation; (iii) finally it aims to fill the gap in the literature by using ANN/GA in panel data framework in growth literature.

In this chapter, let us now first provide our empirical specification in Section 3.1. Then we will present information about our data in Section 3.2.

3.1 Empirical Specification

Based on the review of both theoretical and empirical papers in the growth literature we propose to use the following variables as our explanatory variables: the INIGDPPC, GFCF, GOVT, INF, POPUL, OPEN, TOT and average years of EDUC, while the GROWTH is used as the dependent variable.

The model specification is then presented as in Equation 1. In doing so, this paper is in line with Barro (1996) and Aydin et al. (2016) as well as with several other papers in the literature.

$$GROWTH_{it} = \alpha_0 + \alpha_1 INIGDPPC_{it} + \alpha_2\, GFCF_{it} + \alpha_3\, EDUC_{it} + \alpha_4 OPEN_{it} + \alpha_5 TOT_{it} + \alpha_7 INF_{it} + \alpha_8 POPUL_{it} + \varepsilon_{it} \quad \text{(Equation 1)}^2$$

[2]For more information about Equation (1) see Table 3.1

As for the theoretical expectation of the signs of these independent variables: Our first explanatory variable, the initial GDP per capita (INIGDPPC), stands to capture for the effects of conditional convergence. This concept states that a country with a lower initial per capita income will grow faster in order to catch up the higher income country provided that both countries have the same economic characteristics such as the level of technology, the rate of savings and the population growth rate. As such, a higher initial income per capita implies a lower growth rate; hence, the theoretically expected sign is negative for this variable.

The second explanatory variable is the gross fixed capital formation (GFCF) which captures the effects of investment on physical capital. The first insight is that GFCF should have a positive impact on economic growth since investment in physical capital would increase production, and thus accelerate the growth. On the other hand, physical capital enters into many production functions as an input with diminishing returns. If so, one can assume that more and more investment in physical capital would have lower and lower impact on the production, so that in the long-run there might be no relationship between investment and growth.

Given the short duration of our data as implied by only 23 years of observation, and given the inability of current investment rates to affect the levels of capital stocks substantially, we expect more of a positive impact from investment, rather than a no relationship as implied by diminishing returns to input. Thus our theoretically expected sign for gross fixed capital formation is positive.

Our third explanatory variable is human capital (EDUC) which is proxied by the average level of education attained in a country. Many recent theoretical papers, as well as several empirical papers show that human capital plays a positive role in economic growth through stimulating technological creation, invention and innovation, as well as facilitating the uptake and imitation of new technologies. Please see: Romer (1986); Lucas (1988); Romer (1990); Cohen and Soto (2007); Potančokováand Goujon (2014); and Mankiw, Romer, and Weil (1992) for more detailed discussion of human capital in growth literature. Based on such works, our theoretically expected sign for this variable is positive.

The next variable is trade openness (OPEN). This simply shows the size of the trade volume relative to the size of the economy. Fundamental trade theories such as Comparative Advantage Theory by David Ricardo (1817) and Heckscher-Ohlin Theory (based on the works of Eli Heckscher in 1919 and Bertil Ohlin in 1933) state that trade improves the welfare of a nation as a whole regardless of its income redistribution effect on sub-groups in a nation. In fact, one can argue that trade can increase welfare based on increased competition, specialization and economies of scales as well as though diffusion of technology and know-how. The works of Barro and Sala-i-Martin (2004), Gries and Redlin (2012) and Sokolov et al. (2016) confirm that trade openness positively influences economic growth. Based on these theories and empirics, our expected sign for trade openness is positive.

Terms of trade (TOT) is measured as export price divided by import prices. We would say that terms of trade are improving if we could export our products at a higher price or if we could buy our imports at a lower price. As such then, improvements in terms of trade are expected to increase per capita incomes through transfer of income from rest of the world into the domestic country. Studies such as Harberger (1950); Easterby, Kremer, Pritchett and Summers (1993); Mendoza (1997); and Barro and Sala-i-Martin (2004) are in line with this, especially confirming that TOT has a positive effect on growth in natural resource- and agriculture-based developing countries .

However, Sachs and Warner (1995, 2001), Hadass and Williamson (2003) and Kalumbu and Peyavali (2014) argued that improved terms of trade may deteriorate the economic growth. It means that TOT has a negative effect on growth. This may happen for a number of different reasons such as a decline in the competitiveness of other non-exporting sectors, the crowding out of human capital through the underinvestment in education, or as a result of corruption from the mismanagement of revenues from the natural resource sector. Based on these opposing arguments then, we have no expected sign for this variable.

Our sixth explanatory variable is the size of the government (GOVT) which is measured as the level of government final consumption expenditures relative to the size of the economy. Both theoretical and empirical papers are inconclusive about the effects of this variable. In general, though, to the extent that government expenditures reflect more effective stabilization policies, the effect of government size would be positive. To the extent that the government size reflects the size of the tax distortions, crowding-out effects, and/or other distortions in the market economy, the effect may be negative. For example, studies such as Devarajan et al. (1996), Nasiru (2012), and Medee, and Nenbee (2011) conclude that government size has a negative impact on growth. Based on these opposing arguments, we have no expected sign for this variable too.

As for inflation rate (INF), it is expected to have a negative impact on GDP per capita growth rate since the higher levels of inflation rate simply represents macroeconomic instability (Stanely Fischer, 1993). In fact, it is well documented that inflation reduces welfare and reduces incentives as it creates distortions in taxation system. Moreover, changes in relative prices lead to less-than optimal outcomes in cost-minimizing business decisions. “Menu costs” and “shoe-leather costs” are also well documented. In addition, unexpected inflation rates lead to random income redistribution between the borrowers and creditors as well as between the employers and employees; hence reducing incentives in economic decision further.

Finally, our last independent variable is population growth rate (POPUL). Although some theorists would argue that population growth provides incentives for R & D and investment decisions, more established theories would state that rapid population growth would reduce economic growth because rapid population growth diverts resources from productive sectors into efforts of raising children. Moreover, the larger is the population, the less are the resources and capital per person, leading to less productive societies. Hence, our expected sign for this variable is negative.

In Table 3.1 below, we present a summary of the expected signs for each of our eight explanatory variables.

Table 3.1: Expected Signs for Explanatory Variables

Explanatory Variable	Sign	Explanatory Variable	Sign
Initial GDP per capita	-	Terms of Trade	??
Gross Fixed Capital Formation	+	Government Size	??
Human Capital	+	Inflation Rate	-
Trade Openness	+	Population Growth Rate	-

In the next section, Section 3.2, we provide more details about our dependent and independent variables.

3.2 Data

This dissertation uses a panel data of 20 countries and 23 years of observation from 1990 till 2012. Ten of these countries are developed countries (United Kingdom, Germany, Japan, Spain, Norway, New Zealand, France, Australia, Greece) while the remaining ten are developing countries (India, Venezuela, Turkey, China, Nigeria, Iran, Russia, Ukraine, Pakistan and Brazil). The number of years and number of countries are limited because of limitation of ANN method, this model works better whith less than 500 observations. All of the data come from the following sources: World Bank Database, Federal Reserve Broad Economics Database, OECD National Accounts, and UNESCO Institute for Statistics. The data is an annual data but formatted in three different ways: (i) yearly observation, (ii) periodic of 4 years in a non-overlapping way, (iii) periodic of 4 years in an overlapping way. These various data formatting technique is based on the work of Checherita and Rother (2010), and is for checking the robustness of the models. Moreover, such data formatting allows us to capture the effects of conditional convergence better.

In this study, the real GDP per capita – as measured in constant US $- is extracted from the World Bank National Accounts dataset, and the growth rate is calculated as the percentage change from one year to the next.

Initial GDP per capita (INIGDPPC): This variable is included in the study to account for the effects of conditional convergence as implied by the Solow-Swan (1956) model. According to this concept, a poorer country with a lower initial GDP per capita grows faster and catches up with the richer country on the condition that both countries have the same economic characteristics such as saving rates and technology. It is extracted from the World Bank National Accounts dataset.

As said earlier, in this thesis we used three types of data formats so that: For the yearly observations, the initial GDP per capita is the GDP per capita of one year earlier; for the periodic data of 4 years, the initial GDP per capita is the first year observation of each group of four years both in overlapping and non-overlapping datasets.

Inflation Rate (INF): Inflation rate is the percentage change in the average prices from one year to the next. In this study, we use the consumer price index (CPI) to calculate the annual inflation rate as shown in below. The CPI figures are also extracted from the World Bank National Accounts dataset:

$$\mathrm{INF_t} = \frac{\mathrm{CPI_t} - \mathrm{CPI_{t-1}}}{\mathrm{CPI_{t-1}}}$$

where CPI_t is the consumer price index in time t and

$\mathrm{INF_t}$ is the inflation rate at time t.

CPI_{t-1} is the consumer price index of last year.

Gross Fixed Capital Formation (GFCF): In this study, GFCF is used as a measurement of level of physical investment in a country. More specifically, investment or GFCF is measured as a percentage of the GDP. This is in line with the previous literature such as Ilegbinosa et al. (2015). We obtain GFCF figures from the World Bank National Accounts dataset.

Human Capital (EDUC): Human capital of a country is often approximated by the level of education attained in that country. To this end, empirical studies may use average mean years of schooling; primary, secondary and tertiary school enrollment rate; adult literacy rate; as well as quality of education measurements such as specific test scores; and/or fraction of educated labor force. In this study, we used average years of schooling. Average or mean years of schooling of adults indicate the number of completed years of formal schooling received on average using citizenry population and age. These data were obtained from UNESCO Institute for statistic dataset.

Trade Openness (OPEN): It is measured by dividing the sum of export volume and import volume by the real GDP. We obtain the export and import volumes from OECD national accounts and Federal Reserve Broad Economic dataset while obtaining the real GDP figures from the World Bank dataset.

Term of Trade (TOT): The terms of trade is obtained by dividing the export value by import value, and then multiplying the result by 100. Thus, the formula for calculating the TOT is given as:

$TOT = (P_x / P_m)*100$ where

Px is the price of export

Pm is the price of import

These data are obtained from the World Bank and Federal Reserve Broad Economic dataset.

Government Size (GOVT): We obtain government consumption expenditure figures in constant US dollars from the World Bank dataset. Then we calculate the government size by dividing the government consumption expenditures by the total GDP values. In other words, we express the government expenditures as a percentage of the overall economic size.

Population Growth (POPUL): Population growth rate is the annual percentage increase in the number of resident people in a country in the given year. These data is collected from the World Bank dataset.

3.3 Descriptive Statistics

As mentioned previously, we used a balanced panel data of 20 countries over the period of 23 years. In this section, we present descriptive statistics for each variable and for selected countries in the sample in order to make readers familiar with the data and its measurement units as well as to highlight the variations among the countries. We report the descriptive statistics only for 10 countries out of 20 countries in the sample but the selection includes ten developed countries and ten developing countries in order to show the differences between measures of variables in both the developed and developing countries. These statistics are presented in Tables from Table 3.2 to Table 3.6 below.

As we can see from Table 3.2, the mean (average) growth rate over the 23 years (from 1990 to 2012) for China is 9.15%, which is the highest for the reported countries. This is followed by India who achieved an average growth rate of 4.63%. The lowest average growth rate was for Japan with a growth rate of 0.99%. However, the lowest (minimum) growth rate for any one year is for Turkey with an economic growth of negative 7.80%. On the contrary, the country with the maximum growth rate for any one year is Nigeria with a growth rate of 30.34%. The variability of data ranges from a standard deviation of 1.95 for the UK to standard deviation of 6.78 for Nigeria.

Table 3.2: Descriptive Statistics for the Economic Growth

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Mean	1.44	1.58	0.99	1.37	1.51	4.63	2.69	9.15	2.68	3.08
Median	2.06	1.70	1.35	1.99	2.01	4.61	4.63	8.96	2.47	1.81
Maximum	3.76	4.35	5.21	4.41	5.09	8.75	7.87	13.60	10.69	30.34
Minimum	-4.91	-5.38	-5.52	-4.42	-5.99	-0.98	-7.08	2.42	-7.80	-3.12
Std.Dev.	1.95	2.20	2.21	2.25	2.81	2.36	4.83	2.48	4.34	6.78
Skewness	-1.70	-1.31	-0.78	-0.91	-0.92	-0.31	-0.88	-0.37	-0.13	2.95
Kurtosis	6.02	5.49	4.89	3.25	3.27	2.60	2.42	3.81	3.13	12.60

The descriptive statistics for trade openness is reported in Table 3.3. The mean of openness (over the 23 years) ranges from a low value of 23.22 percent of

GDP for Japan to a high value of 75.93 percent of GDP for Sweden. The minimum openness value for any one year belongs to Japan, which is around 15.92 percent while the maximum is 85.89 percent for Germany. The standard deviation of trade openness ranges from a low of 4.87 for the UK to a high of 15.04 for Germany.

Table 3.3: Descriptive Statistics for the OPEN

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Mean	53.55	61.08	23.22	50.28	75.93	32.17	45.47	43.98	43.43	60.10
Median	53.91	61.39	20.39	53.11	77.86	26.44	47.74	42.06	43.09	61.03
Maximum	63.01	85.89	35.23	60.24	93.36	55.75	57.75	64.77	56.05	81.81
Minimum	44.03	40.64	15.92	35.51	51.72	15.24	30.48	29.62	29.23	42.31
Std.Dev.	4.87	15.04	6.05	8.15	12.29	13.43	7.99	11.36	7.30	10.59
Skewness	0.01	0.23	0.64	-0.74	-0.63	0.49	-0.59	0.50	-0.06	0.09
Kurtosis	2.78	1.67	2.04	2.20	2.34	1.77	2.43	1.91	2.18	2.49

As for the GOVT, the descriptive statistics is reported in Table 3.4. The least of mean of government size is in Nigeria with a value of 10.22 percent of GDP. The highest mean of GOVT is in UK with 19.2 percent of its GDP (over the 23 years). The minimum GOVT value for any one year belongs to Nigeria that is around 5.47 percent while the maximum is 27.49 percent for Sweden. The standard deviation of government size ranges from a low of 0.48 for the Germany to a high of 3.45 for Nigeria.

Table 3.4: Descriptive Statistics for the GOVT

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Mean	19.2	18.76	16.99	17.80	25.48	11.435	12.45	14.36	11.89	10.22
Median	18.9	18.91	17.73	17.35	25.42	11.439	12.33	14.12	11.81	8.62
Maximum	22.3	19.55	20.43	20.52	27.49	12.79	14.84	15.86	14.35	17.29
Minimum	16.8	17.49	13.29	16.28	24.06	10.28	10.25	13.04	9.71	5.47
Std.Dev.	1.55	0.48	2.24	1.315	0.82	0.693	1.164	0.947	1.42	3.45
Skewness	0.24	-0.94	-0.15	1.113	0.68	0.295	0.42	0.134	0.143	0.60
Kurtosis	2.08	3.65	1.85	2.985	3.270	2.305	2.91	1.60	2.068	2.04

As for the EDUC, the descriptive statistics is reported in Table 3.5. The least of mean of EDUC (over the 23 years) is in India with a value of 4.33 years and the highest mean of EDUC (over the 23 years) is in UK with 11.54 years. The minimum EDUC value for any one year belongs to Nigeria and India, which is around 3 years while the maximum is 13.1 years for Germany and UK. The standard deviation for the EDUC data ranges from a low value of 0.67 for the Japan to a high value of 1.49 for Germany.

Table 3.5: Descriptive Statistics for the EDUC

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Mean	11.54	11.01	10.76	8.291	11.40	4.33	5.69	6.33	10.94	4.73
Median	11.8	10.8	10.8	8.5	11.5	4.5	5.6	6.5	11.4	5.2
Maximum	13.1	13.1	12	9.6	12.1	5.6	7.6	7.4	12	5.7
Minimum	7.9	8.8	9.6	6	10.5	3	4.5	4.8	9.2	3
Std.Dev.	1.40	1.49	0.67	1.029	0.506	0.81	0.91	0.77	0.91	0.84
Skewness	-1.1	0.038	-0.058	-0.67	-0.34	-0.19	0.583	-0.52	-0.63	-0.87
Kurtosis	3.72	1.51	2.054	2.568	1.700	1.745	2.427	2.055	1.952	2.299

As for the gross fixed capital formation, the descriptive statistics is reported in Table 3.6. The least of mean of investment is in Nigeria with a value of 9.33% of GDP. The highest mean investment is in China with 40.88% of its GDP. China has the highest level of investment for any one year with an investment rate of 47.58 percent of GDP. The standard deviation for the GFCF data ranges from a low value of 1.56 for the UK to a high value of 23.70 for Nigeria.

Table 3.6: Descriptive Statistics for the GFCF

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Mean	18.65	21.82	25.06	25.41	22.42	28.99	21.24	40.88	36.23	9.33
Median	18.78	22.31	24.30	25.42	22.00	26.05	21.32	41.39	36.27	3.36
Maximum	22.03	25.59	32.49	31.34	29.46	39.58	26.62	47.58	43.74	59.30
Minimum	15.29	18.07	19.67	20.23	18.95	21.29	14.94	34.92	23.29	-26.23
Std.Dev.	1.56	2.25	3.90	3.31	2.17	6.27	3.13	4.12	4.58	23.70
Skewness	-0.33	-0.03	0.42	0.29	1.40	0.45	-0.32	0.23	-0.76	0.47
Kurtosis	3.04	1.65	2.05	2.00	2.99	2.45	2.71	6.30	4.10	2.37

For more information about other descriptive statistics of variable, see Appendix A.1 to A.4.

Chapter 4

RESEARCH METHODOLOGY

In this chapter, we discuss research methodology adopted for our empirical analysis, this among other includes, a brief insight into panel data methods, the panel unit root tests, panel fixed-effect model, panel random-effect model, the Hausman test and the general least square estimation technique.

4.1 Panel Data Techniques

Panel data techniques, is a combination of cross sectional analysis and time series analysis. This implies that, panel data methods are made up of both cross section dimension and time dimension. The cross-sectional dimension of panel data methods is related to the use of countries, firms, and markets among others, while the time series dimension aspect of panel data methods is related to time span of these individuals, daily, monthly or annual frequency data. As mentioned in chapter 3, the panel data is the subject of the most innovative activities of econometrics literatures. The advantages of using panel data estimation techniques are summarized below:

1-Panel data estimation methods can identify and estimate the effects that are simply not detectable in cross-sectional or time series analysis. It goes a long way to help in interpreting complicated issues as related to the dynamic behavior of the variables (Baltagi, 2005).

2- Longitudinal panel data account for more variability of the variables, it produces less collinearity results, they are more informative, deals with complicated dynamic behavior of the models and they are more robust and efficient, through combination of countries with either heterogeneous or homogeneous features over time periods.

3- It provides more information and measure effects which cannot be observed by using time series and/or cross-sectional analysis.

4- Panel data approach minimizes estimation bias that one might encounter as a result of accumulated data from different countries into broad aggregate (Gujarati and Porter, 2012). The panel data longitudinal regression model is shown in Equation 2.

$$Y_{it} = \alpha Z_i + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \qquad \text{(Equation 2)}$$

If we assume that, $C_i = \alpha Z_i$, then we rewrite Eq. (2) as

$$Y_{it} = C_i + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \qquad \text{(Equation 3)}$$

Where, X_{it} represents the explanatory variables and k is the number of explanatory variables. i implies cross sections (i= 1,2,,…,N) that is the twenty countries, while t implies time periods (t = 1,2,…,T.) that is the 23 years. α is scalar, while β is the coefficient estimates; the subjects effect is αZ_i, where, Z_i is a constant term and set of the country variables, which may be observed (sex, location, and so on) or unobserved (behaviors, skill or preferences, policy, environmental factors and so on), all of which can be constant over a given period. If Z_i is observable for all countries, then the model can be run as an Ordinary Least Squares (OLS) model. However, problem arises when C_i is unobservable, which will be the case in this study.

There are 3 types of panel data estimation methods, they are discussed below:

1- The Pooled method: This method or type of panel data model considers linear equation for all explanatory variables, the intercept term (α) are constant for all the individual unit. while Z_i contains only a constant terms and slope vector is "β". Therefore, there is no difference between countries, that is, all countries assume to be homogeneous. On the other hand, explanatory variables are assume to be exogenous and do not depend on the value of ε_{it}, which is assume to be independently identically distributed with zero mean and constant variance. If the assumption holds, then, the OLS become appropriate estimation model. The OLS regression model is given in Equation (4).

$$Y_{it} = \alpha + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \qquad \text{(Equation 4)}$$

The pooled model of the panel data takes into consideration certain assumptions. These assumptions are stated as follow:

A. $E[\varepsilon it \mid Xi1, Xi2,..., XiTi] = 0$,

B. $Var[\varepsilon it \mid Xi1, Xi2,..., XiTi] = \sigma^2 \varepsilon$,

C. $Cov [\varepsilon it, \varepsilon js \mid Xi1, Xi2, XiTi] = 0$ if $i \neq j$ or $t \neq s$.

D. i=1,…,N and t= 1,2,…,Ti

The first assumption A indicate that condition means of error term is zero and constant. Assumption B implies that the variance of each disturbance term is constant on the chosen values of independent variables (homoskedasticity). Assumption C on the other hand, indicate that explanatory variables should be independent of the error term, such that, they are not correlated (no serial correlation), while in assumption D ‘i’ denotes individuals and Ti shows that each individuals may represent a different number of times, which is refer to as unbalance panel data. For this study, we used balanced panel dataset. In pooled panel data model analysis, it is observed that, the coefficient of determination (R^2) are very high, while the Durbin-Watson statistic are somewhat low (Gujarati and Porter 2012, page 594). This implies that, the pooled panel data model is prone to autocorrelation problem.

Besides, one can also encounter heterogeneity problem which is caused by pooling together different countries over a period time. Thus, the error term (ε_{it})may be correlated with some regressors, which will make the estimated coefficients biased and inconsistent.

Generally, in the panel data estimation method; fixed-effects model and random-effects model are commonly and widely applied. If unit-specific or time-specific effects are assumed to be fixed, the model is called fixed-effects model. The term "fixed effects" denotes nonrandom quantities are accounted for the heterogeneity characteristics across cross-sections. On the other hand, if these specific-effects are assumed random and not correlated with the independent variables, the model is "random-effects model". In fact, random effects models include the individual effects as a component of the error term (Baltagi, 2013). Fixed effects model and random effects model are represented in Equation (5) and (6) below:

2- Fixed effects model (FEM):

$$y_{it} = \alpha_i + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{i,t} \qquad \text{(Equation 5)}$$

Where i=1,2,…..,k and T=1,2,…..,T and X represent vector of independent variables with K variables, while ε_{it} contains two parts, the first part indicate that all unobserved factors varies across cross-sections but are constant over time (Fixed-effect model), while the other one indicate that all unobserved factors varies across cross-sections and time (Random-effect model).

3- random effects model (REM):

$$y_{it} = \sum_{k=1}^{K} \beta_k X_{kit} + (\alpha_i + u_i) + \varepsilon_{i,t} \qquad \text{(Equation 6)}$$

Where α_i is a random variable andui is a random error term. In many situations, there is an uncertainty that whether the unit dependent unobserved effects are correlated with one or more of the explanatory variables, thus, specifying the optimal model become necessary. However, in order to confirm the appropriate panel model, that is, whether the fixed-effects model or REM is most suitable empirical model, the Hausman test is applicable. In the following section, we discuss the panel data unit root test for stationary of the variables of interest, after which we conduct Hausman test to ascertain the appropriate model for the study.

4.1.1 Panel Data Unit-Root Tests

Stationarity properties in a panel data studies is a crucial empirical analysis that should be examined before empirical estimations. Stationary properties of a variable indicate that the mean, variance, covariance properties of such variable are constant over time. On the other hand, non-stationary properties of a variable indicate that, the mean or variance or both are not constant overtime. Here, we briefly describe the five panel unit root tests such as; Levin, Lin, and Chu (LLC, 2002), Breitung (B-tstat, 2000), Im, Pesaran and Shin (IPS, 2003), Fisher-Augmented Dickey Fuller (ADF) and Fisher- Phillips-Perron (PP) unit root tests (2003). We consider basic AR (1) process for the longitudinal panel data method in Equation (7) as follow:

$$\mathrm{Y_{it}} = \mathrm{\rho_i Y_{it-1}} + \mathrm{\delta_i X_{it}} + \mathrm{\varepsilon_{it}} \qquad \text{(Equation 7)}$$

In Equation (7) subscript i is individual units, which can be observed during the period, t is time $\mathrm{X_{it}}$ indicates exogenous variables in the model, $\mathrm{\rho_i}$ is the autoregressive coefficient and $\mathrm{\varepsilon_{it}}$ shows mutually independent disturbance. If $|\rho_i| < 1$, Yi is said to be stationary, and when,$|\mathrm{\rho_i}| = 1$, Yi is non-stationary. The unit root test can be observed on the level (raw data), First difference, or second difference basis[3], by estimating either with individual constant terms or individual constants and trends. This is shown in ADF unit root regression in Equation (8) under the assumption of individual without trend and in Equation (9) with trend and individual constant term respectively:

$$\Delta Y_{it} = \varphi_i + \alpha Y_{it-1} + \sum_{j=1}^{p_i} \beta_{ij} \Delta Y_{it-j} + \delta_i X'_{it} + \epsilon_{it} \qquad \text{(Equation 8)}$$

$$\Delta Y_{it} = \varphi_i + \alpha Y_{it-1} + \sum_{j=1}^{p_i} \beta_{ij} \Delta Y_{it-j} + \delta_i X'_{it} + \varphi_i t + \epsilon_{it} \qquad \text{(Equation 9)}$$

where, $\alpha = 1 - \rho$, φ_i shows individual fixed effects and $\varphi_i t$ indicates individual intercepts and individual trends. There are two assumptions about the ρ_i in the Equation (7). First, we assume that $\rho_i = \rho$ for all i. The LLC and B-stat, tests all use for this assumption. Second, ρ_i vary across cross-sections. The IPS, and Fisher-ADF and Fisher-PP tests take this form.

4.1.1.1 LLC and B-tstat Tests

The method of LLC derives estimates $(\alpha = 1 - \rho)$ from ΔY_{it} and Y_{it} that are standardized and free of autocorrelations and deterministic components. The B-tstat method differs from LLC.

In the B-tstat only the autoregressive portion (ρ_i) is removed and it requires specification of the lag length used in each cross-section, ADF regression, and the exogenous regressors. If consider basic ADF in Equation (10):

$$\Delta Y_{it} = \alpha Y_{it-1} + \sum_{j=1}^{p_i} \beta_{ij} \Delta Y_{it-j} + \delta X'_{it} + \epsilon_{it} \qquad \text{(Equation 10)}$$

Where, $\Delta Y_{it} = Y_{it} - Y_{it-1}$, ρ_i is the same across cross-sections $(\rho_i = \rho)$, $\alpha = 1 - \rho$, but allow the lag order from first difference term. Therefore, hypothesis for the panel unit root methods are written as below:

[3]Eviews note, (2017), Advanced Univariate Analysis, Univariate Time Series Analysis,Panel Unit Root Testing. http://www.eviews.com/help/helpintro.html#page/content/advtimeser-Panel_Unit_Root_Testing.html

$\mathrm{H0}: \alpha = 0$ unit root

$\mathrm{H1}: \alpha < 0$ no unit root

4.1.1.2 IPS, and Fisher-ADF and Fisher- PP unit root tests

The Im, Pesaran, and Shin, Fisher-ADF and Fisher- PP panel unit root tests are used for individual unit root processes, in such a way that $\rho_i$varies across cross-sections. If π_i denotes the p-value from the individual unit root test for cross-section i, then across time dimension we specified Equation (11) as follow:

$$-2\sum_{i=1}^{N} \log(\pi_i) \rightarrow X_{2n}^{2} \qquad \text{(Equation 11)}$$

For both the Fisher- ADF and Fisher- PP panel unit root tests, one can conduct the unit root tests, for the exogenous regressors under the assumption of; individual constants or individual constant and trend terms. Moreover, one needs to specify the lag length.

In this study, we estimated all the panel unit root tests methodology mentioned above According to Equation (10), we specify the hypothesis as shown below:

$\mathrm{H0}: \alpha = 0$ unit root

$\mathrm{H1}: \begin{cases} \alpha & = 0 \\ \alpha & < 0 \end{cases}$ no unit root.

4.1.2 Hausman Test

In the panel regression analysis, there is an assumption of error term being independent of the explanatary variables in an empirical model that is, $E(\epsilon_{it}|X_{it}) = 0$. This indicate that, the error term contain unobserved individual effects (C_i in Equation (3)) which must not be correleted with the explanatory variables (X_{it}). If there is a correlation between C_i and X_{it},, that is, $E(\epsilon_{it}|X_{it}) \neq 0$, then the assumption of no serial correlation is violated. Thus,, $\hat{b}_{REM}$ is baised and inconsistence.

The Hausman test is used to determine whether the difference in intercepts from the origin of sectional units are fixed or random. It is used to test, which of the panel data estimation techniques is suitable for empirical analysis, i.e. whether the FEM or REM is an appropriate and more efficient estimation model. It has only a little justification for treating the individual effects as uncorrelated with other regressors. Hausman (1978) tested that the covariance of an efficient estimator b_{FE} with a difference of an inefficient estimator b_{REM} is zero. We estimated covariance matrices of the slope coefficient of the FEM and REM excluding the constant term. The chi-squared test is based on the Wald statistic criterion for Hausman test. We depict this in Equation (12):

$$W = \chi^2 \sim (df = k) \qquad \text{(Equation 12)}$$

Where, K is number of independent variables, and W is equal to the χ^2 by K degrees of freedom. The Hausman's hypothesis is written as below:

H0: $\begin{matrix} E(\epsilon_{it}|Xit) = 0 \\ \mathrm{corr}(Ci, Xit) = 0 \\ bFE = bREM \end{matrix}$ REM i*s* acceptable

H1: $\begin{matrix} E(\epsilon_{it}|Xit) \neq 0 \\ \mathrm{corr}(Ci, Xit) \neq 0 \\ bFE \neq bREM \end{matrix}$ FE is acceptable

This study adapts to different methods in the context of growth over the sample period.

4.2 Generalized Least Square Method (GLS)

In the panel analyses framework, the sampling design have the potential sources of correlation between observations which is usually refer to as clustering problem. This problem is caused by the presence of a common unobserved random shock at the group level that leads to correlation relationships between all observations within each group, hence, biased standard error and misleading inference. To solve this problem and obtain robust coefficient estimates, the standard error obtained through Feasible Generalized Least Square (FGLS) estimation is suggested by many scholars (Baltagi, 2013).

In the presence of heteroskedasticity which is as a result of the clustering problem, the cross-section heteroskedasticity is applied for a different residual variance for each cross-section while the residuals between different cross-sections and different periods are assumed to be 0. To eliminate this problem, FGLS estimation is required. It is also refer to as Generalized Least Square (GLS), the GLS estimation account for heteroskedasticity in the unit dimension (Wooldridge, 2010). In addition, clustering also lead to autocorrelation problem. In such situation, robust estimators are required. The popular form of the robust covariance which is applied is the White cross-section method. This method assumes that the errors are contemporaneously correlated.

The method pool regression as a multivariate regression (with an equation for each cross-section), and estimates robust standard errors for the system of equations. Hence, this technique provides robust estimators to cross-equation (contemporaneous) correlation and heteroskedasticity (Wooldridge, 2002).

4.3 Results

This section forecasts the growth by using standard fixed-effect model discussed in previous part. We used the GLS method because it provides robust results for the model's key predictions of econometric problems such as autocorrelation and heteroskedasticity. In this section, we explained the unit root tests, Hausman test, and the GLS fixed effect’s results.

4.3.1 Unit Root Test Results

Before proceeding with the panel data estimations, it is crucial to estimate the stationarity properties of the variables of interest in order to avoid spurious regressions. Table 4.1 presents the panel unit root tests results conducted for the study. Empirically, the null hypothesis can be rejected for all variables at the first time. Therefore, all variables are stationary in the level with intercept. If the unit root tests t-statistic values obtained is less than the critical values at 1%, 5%, and 10%, then, we fail to reject the null hypothesis of non-stationary at levels, otherwise, we reject the alternative hypothesis and conclude that, the variables are stationarity at levels.

As the first step in empirical estimation, we checked for stationarity properties of the variables of interest. As presented in Table 4.1, all the variables are stationary at level data at 1%, 5% or 10% significant levels, under the unit root tests specification of individual intercept.

For more information on the unit root tests results, please see the Appendix B.

Table 4.1: Unit Root Test Results

Variables	At level-individual intercept				At level-individual intercept and trend					H0: non-stationary
	L.LC[2]	**I.P.S**	**ADF**	**PP**	**L.LC**	**B. t-stat**	**I.P.S**	**ADF**	**PP**	
Growth	-6.947 (0.0000)[1]***	-6.736 (0.0000)***	126.787 (0.0000)***	162.732 (0.0000)***	-6.804 (0.0000)***	-2.300 (0.010)***	-4.581 (0.0000)***	93.012 (0.0000)***	124.750 (0.0000)***	Reject- It is stationary at the level
GFCF	-1.836 (0.0033)***	-2.547 (0.0054)***	68.165 (0.0036)***	70.867 (0.0019)***	-1.605 (0.0542)**	-2.631 (0.0042)***	-3.125 (0.0009)***	76.5562 (0.0004)***	164.150 (0.0000)***	Reject- It is stationary at the level
INF	-7.2641 (0.0000)***	-6.972 (0.0000)***	129.745 (0.0000)***	134.537 (0.0000)***	-4.854 (0.0000)***	-0.340 (0.3667)	-4.264 (0.0000)***	85.568 (0.0000)***	105.053 (0.0000)***	Reject- It is stationary at the level
TOT	-1.31 (0.0949)*	-2.096 (0.0180)**	59.309 (0.0252)**	62.174 (0.0139)**	-1.131 (0.1010)***	-2.0855 (0.0185)**	-1.398 (0.0810)*	51.015 (0.100)*	60.295 (0.0206)**	Reject- It is stationary at the level
POPUL	1.522 (0.936)	-0.224 (0.412)	57.46 (0.0411)**	61.367 (0.0165)**	-5.918 (0.0000)***	0.196 (0.577)	-5.350 (0.0000)***	122.906 (0.0000)***	55.920 (0.048)**	Reject - It is stationary at the level
GOVT	-2.525 (0.0058)***	-2.876 (0.0020)***	67.37 (0.0043)***	59.43 (0.0245)**	-3.084 (0.0010)***	-3.786 (0.0001)***	-2.360 (0.0091)***	57.63 (0.0350)**	49.538 (0.143)	Reject- It is stationary at the level
INIGDPPC	-0.878 (0.190)	3.531 (0.999)	21.605 (0.9923)	22.101 (0.9903)	-1.039 (0.1492)	1.643 (0.949)	-0.0964 (0.461)	45.93 (0.239)	35.559 (0.670)	Not reject- It is stationary at the 1st difference(Apendix B)
OPEN	-1.5958 (0.0595)*	-1.382 (0.0833)*	58.314 (0.0301)**	55.582 (0.0517)*	-2.605 (0.0046)***	-2.640 (0.0041)***	-2.184 (0.0145)**	62.052 (0.0143)**	55.807 (0.0495)**	Reject- It is stationary at the level
EDUC	-1.92978 (0.0268)**	-2.2673 (0.0117)**	78.1821 (0.0003)***	75.5328 (0.0006)***	-2.3922 (0.0084)***	-3.8542 (0.0001)***	-7.1865 (0.0000)***	51.3171 (0.105)*	48.1097 (0.1987)	Reject- It is stationary at the level
Obs.	420	420	420	440	420	400	420	420	440	

*** Indicates significant at 1%. ** Indicates significant at 5%. * Indicates significant at 10%. 1) Amount of p-value is reported in parenthesis. 2) Levin, Lin, and Chu (LLC, 2002), Breitung (B-tstat, 2000), Im, Pesaran and Shin (IPS, 2003), Fisher- Augmented Dickey Fuller (ADF) and Fisher- Phillips-Perron (PP).

4.3.2 Hausman Test Results

In order to confirm the appropriate model for our study, that is, evaluate whether the FEM or REM is suitable for model estimation we carried out the Hausman specification test. The FEM removes time-invariant characteristics which make it possible to assess the net effect of the predictors on the estimation outcome. First, we test regression with random effects in cross section in three scenarios (Table 4.2).

Table 4.2: Hausman Test Results

Data Set	**Test Statistic**
Whole 20 countries, yearly data	41.74 (0.0000)***
Whole 20 countries, periodic non overlapping	21.93 (0.0050)***
Whole 20 countries, periodic overlapping	124.44 (0.0000)***
Developing countries, yearly data	14.12 (0.0785)*
Developing countries, periodic non overlapping	15.47 (0.0506)*
Developing countries, periodic overlapping	105.73 (0.0000)***
Developed countries, yearly data	70.90 (0.0000)***
Developed countries, periodic non overlapping	18.40 (0.0184)**
Developed countries, periodic overlapping	204.50 (0.0000)***

***$p < .01$, **$p < .05$, *$p < .1$
The number above the parenthesis is chi-square value.

The null hypothesis of the test is specified under the assumption that, there is similarity between the coefficients of the models i.e. the FEM and REM, while the alternative hypothesis is specified under the assumption that, the fixed-effects estimation is appropriate for model estimation than the random-effects model.

According to results reported in Table 4.2 for Hausman specification test, the p-value for whole annual panel, mean 4-year overlapping and mean 4-year non-overlapping data is 0.0000, this indicate statistical significance at 1% level. The annual panel for mean 4-year non-overlapping data for developing countries is statistically significant at 10% level, while mean 4-year overlapping for developing countries is statistically significant at 1% level. Moreover, the p-value of an annual panel, mean 4-year overlapping and mean 4-year non-overlapping data for developed countries is statistically significant at 1% and 5% levels.

The statistical summary of the Hausman test yields statistically significant coefficients for the three scenarios estimated. We reject the null hypothesis that the two methods are similar in favor of the alternative hypothesis that FEM is an appropriate model for the study. Thus, we found that the FEM is suitable for our empirical analysis rather than the random effect model.

4.4 GLS Fixed Effect Estimation Results

We estimate the GLS fixed-effect model for three different samples of data: (i) whole 20 countries, (ii) 10 developing countries only and (iii) 10 developed countries only. These results are presented in Tables 4.3, 4.4 and 4.5 respectively. Furthermore, for each sample we use three different data formatting: (i) annual data, (ii) periodic data of 4 years with overlapping years, and (iii) periodic data of 4 years with non-overlapping years. These are taking place in Column 1, Column 2 and Column 3 in each table respectively.

We do so in order to see, how the panel estimation results differ from each other, we divided the period into shorter spans, re-estimate the growth equation over shorter consecutive intervals, and conclude that, the GLS fixed-effect estimation results are not different from the conventional FEM results.

Now, let us present our estimation results with whole 20 countries in Table 4.3 below.

Table 4.3: Panel Data Fixed Effect Model for Whole 20 Countries

Variables	Yearly data (1)	Periodic of 4 years overlapping data (2)	Periodic of 4 years non-overlapping data (3)
GFCF	0.165853 (0.0000)[1]***	0.205222 (0.0000) ***	0.09363 (0.0303) **
GOVT	-0.66608 (0.0000)***	-0.525157 (0.0000) ***	-0.43320 (0.0000) ***
INF	-0.001617 (0.0000) ***	-0.001901 (0.0000)***	-0.001881 (0.0363)**
INIGDPPC	-4.58985 (0.0001) ***	-8.045096 (0.0000) ***	-1.90809 (0.0086) ***
OPEN	0.05860 (0.0200) ***	0.070226 (0.0000) ***	0.036748 (0.0520) *
EDUC	0.418142 (0.0368)**	0.876891 (0.0000) ***	-0.00056 (0.5284)
POPUL	-1.29227 (0.0011) ***	-1.87111 (0.0000) ***	-0.75242 (0.0593) *
TOT	-0.033266 (0.0000) ***	-0.019380 (0.0010) ***	-0.02620 (0.0302) **
C	49.10529 (0.0000)	72.1248 (0.0000)	25.84277 (0.0000)
R^2	0.500218	0.780012	0.763674
D.W	1.31839	0.441127	2.192149
Obs.	460	400	120

*** Indicates significant at 1%.
** Indicates significant at 5%.
* Indicates significant at 10%.
1) Amount of p-value is reported in parenthesis.

In Column 1 in Table 4.3, we report estimation results using annual frequency data. The result shows that GFCF, INF, TOT, GOVT, INIGDPPC, OPEN, and POPUL are statistically significant at 1% level, while, EDUC is so at 5% level.

We found that, INF, TOT, GOVT, POPUL, and INIGDPPC exhibit a negative impact on growth, while GFCF, OPEN, and EDUC have a positive impact on growth. Based on these results we can see that all of our explanatory variables have the correct sign as indicated in our theoretical expectations.

The Central Bank and policymaker's decisions are based on attaining sustainable economic growth and price stability. Thus, it is a worthy macroeconomic objective to control the level of inflation through price stability in some developing countries, such as Iran, India and Nigeria (Appendix A.2). In general, low and high inflation rate can be problematic for saving, investment, household consumptions and production decisions. In our study, INF is statistically significant, and has a negative impact on growth, such that 1% increase in the level of INF rate is associated with a decline of 0.16 % in the annual growth rate. This is line with the findings of Aydin et.al (2016).

The estimated coefficient of GFCF is positive and statistically significant at 1% level. As reported in literature, that physical capital or investment has a crucial role to play in enhancing economic growth. High level of GFCF would create more opportunities for investors, hence, high production capacity and more return on domestic investment (Swamy, 2015, Adhikary, 2011; Ilegbinosa et. al, 2015).

On the other hand, the estimated coefficients for INIGDPPC are -4.58, -8.04 and -1.90, respectively in Column 1, 2 and 3. This indicates that INIGDPPC has a negative impact on growth. The INIGDPPC coefficient in the mean 4-year overlapping is less than two other scenarios because the accumulated rate of convergence is slightly less than its annual value. The modern growth theory shows that lower INIGDPPC speeds up convergence. The lower level of INIGDPPC, relative to the long-run steady state, the faster is the growth rate. It has been argued that, countries with lower capital-labor ratio tend to have higher rate of return and grow faster vice versa. The convergence is conditional because the capital and output per person depend on saving, investment, OPEN, POPUL, GOVT, EDUC and some different characteristic across countries (Barro and Sala-i-Martin, 2004).

The estimated coefficients for EDUC are 0.41 and 0.87 in Columns 1 and 2 respectively, with a p-value of 0.03 and 0.00. Hence, the variables are statistically significant at 1% level. The estimated coefficients means that 1 year of increase in schooling increases growth rate by 0.41% and 0.87% in Column 1 and 2 respectively. In the Column (1) and (2), the EDUC would enhance the growth rate through technological advancement. Increases in the EDUC imply that fraction of human capital in the growth model increases, since labor force is one of the important factors that contribute to growth. Increase in the EDUC would lead to high efficiency of any educated person, which would lead to increase in labor marginal productivity, while in the last Column; EDUC is statistically insignificant with the negative sign.

The coefficient of OPEN is positive and statistically significant in the three Columns in the Table 4.3. The OPEN coefficient determines the trade volume of the origin country versus other competing countries. The greater the OPEN, the greater the amount capital inflow into the country through exports. As a result, it has a statistically significant positive effect on the growth rate and investment in the country of origin (Gries Redlin, 2012, Sokolov et al., 2016, and Barro, Sala-i-Martin, 2004).

The estimated coefficient of TOT is negative and statistically significant. TOT is an exogenous variable to a country's growth, because it depends on world prices. If the price of oil increases, it will be beneficial for crude oil exporters and leads to increase in the level of investment, employment and output in an oil exporter countries. On the other hand, the resources of those countries will shift to an oil importing countries, while this shift in resources is harmful to development and growth of such economy (Kalumbu and Sheefeni, 2014 and Won, 2010). Therefore, the terms of trade on economic growth is known to be ambiguous. As you can see in Table 4.3, the estimated coefficients of TOT are -0.033, -0.019 and -0.026 for yearly observation, periodic of 4 years in an overlapping way, and for periodic of 4 years in a non- overlapping way, respectively.

According to the growth theory, GOVT has a direct impact on productivity or the security of property rights. The estimated coefficients of government size are -0.666, -0.525 and -0.433 and they are statistically significant for yearly observation, periodic of 4 years in an overlapping way, and for periodic of 4 years in non-overlapping way respectively. The results mean that 1 % increase in government size lead to a fall in the growth rate by 0.66%, 0.52% and 0.43% in Column 1, 2, and 3 respectively.

In addition, the estimated coefficients of POPUL is statistically significant and of negative sign. The estimated coefficients are -1.29, -1.87, and -0.75 for annual data, periodic of 4- year overlapping and periodic of 4- year non-overlapping data respectively. For annual data, the coefficient estimate of -1.29 indicates that 1% increase in population decreases the GDP per capita growth rate by 1.29%. The POPUL decreases the household saving by increasing cost of raising children. The children need more educational attainment, health insurance, food and so on. Therefore, the most of the resources should be devoted to the cost of raising children, rather than to production of goods and services. Thus, growth rate will drop due to a decline in saving (Klasen and Lawson, 2007).

In the Table 4.3 the results in Column 2 show that all explanatory variables are statistically significant at 1% level. Furthermore, the signs of all variables are the similar to the one reported in Column 1. In the 4-year overlapping data, all the estimated coefficients are bigger than annual data except TOT coefficient, which shows that average 4- year overlapping data explained changes in growth rate better than the annual data. In addition, we also found that the growth rate of GDPPC depends on the high contribution of GFCF, OPEN, and EDUC respectively.

Furthermore, in the Table 4.3 Column (3), the result shows that GOVT, INIGDPPC are highly statistically significant with the negative sign. The GFCF, INF, and TOT are statistically significant at the 5% level, with the similar signs in Column (1). The POPUL is negatively and OPEN positively statistically significant at 10% level. However, when data is considered as a mean 4-year non-overlapping, EDUC become statistically insignificant.

In Table 4.4 and Table 4.5 we present the estimation results of the regressions based on samples of developing countries only and developed countries only respectively. As shown in Table 4.4, we have 3 Columns similar to Table 4.3 for developing countries. In the Table 4.4, Column (1), the same variables are included in the model; the INF, TOT GOVT and INIGDPPC have the statistically significant impact on growth at 1% level.

Table 4.4: Panel Data Fixed Effect Model for Developing Countries

Variables	Yearly data (1)	Periodic of 4 years overlapping data (2)	Periodic of 4 years non-overlapping data (3)
GFCF	0.086554 (0.0298)[1] **	0.12035 (0.0000) ***	0.086864 (0.0416) **
GOVT	-0.476996 (0.0000) ***	-0.62771 (0.0000) ***	-0.48240 (0.0366)**
INF	-0.001288 (0.0000) ***	-0.001602 (0.0000) ***	-0.002985 (0.0000) ***
INIGDPPC	-2.8397 (0.0001) ***	-8.57243 (0.0000) ***	-0.969051 (0.0104) **
OPEN	0.065724 (0.0113) **	0.1279 (0.0000) ***	-0.01854 (0.4737)
EDUC	1.186508 (0.0016)***	2.30116 (0.0000) ***	0.002003 (0.0195) **
POPUL	1.10092 (0.2617)	0.4357 (0.3572)	-1.64982 (0.0295)**
TOT	-0.01978 (0.0072)***	0.0051 (0.0695) *	0.036382 (0.1000)*
C	19.980 (0.0006)	55.5679 (0.0000)	15.02386 (0.0139)
R^2	0.70202	0.97222	0.561767
D.W	1.63355	1.38812	1.655655
Obs.	230	200	60

*** Indicates significant at 1%.
** Indicates significant at 5%.
* Indicates significant at 10%.
1) Amount of p-value is reported in parenthesis.

The coefficient of the INIGDPPC is negative and statistically significant, (-2.83, -8.57 and -0.96), at 1% level for yearly observation, periodic of 4 years in an overlapping way and periodic of 4 years in a non-overlapping way. This shows conditional convergence in developing countries. The developing countries with low initial GDP per capita grow faster than the countries with relatively higher initial GDP per capita.

In the Table 4.4 Column (2), the TOT estimated coefficient is positive and statistically significant at 10% level. This indicates that the TOT of the developing countries has positive impact on growth. According to Barro (1996, p 20) "movements in real GDP occur only if the shift in terms of trade stimulates a change in domestic employment and output." The developing countries with high export prices obtain benefit from their competitors. This gain pushes up the investment, productivity capacity level, employment rate, and living standards in developing countries. The other explanatory variables are statistically significant at 1% level similar to results reported in Table 4.4.

In the Table 4.4 Column (3), the GFCF, GOVT, INIGDPPC, EDUC, and POPUL are also all statistically significant at 5% level. The inflation rate has negative coefficient at 1% significance level, while TOT exhibit positive statistically significant impact on growth at 10% significance level. The estimated OPEN is statistically insignificant for the 4-year non-overlapping scenario. On the other hand, INF, POPUL, INIGDPPC, and GOVT have the negative impact on growth but EDUC, TOT, and GFCF have the positive impact on the growth rate.

As shown in Table 4.5, the result of fixed-effect model was extracted for developed countries. Generally, INF and OPEN are statistically significance at 5% for yearly observation, periodic of 4 years in an overlapping way while the other explanatory variables are statistically significant at 1% level. The estimated coefficients for POPUL, INIGDPPC, INF, GOVT and TOT have negative signs, while that of OPEN, EDUC and GFCF have positive impact on growth.

Table 4.5: Panel Data Fixed Effect Model for Developed Countries

Variables	Yearly data (1)	Periodic of 4 years overlapping data (2)	Periodic of 4 years non-overlapping data (3)
GFCF	0.227605 (0.0024)[1] ***	0.22456 (0.0000) ***	0.258695 (0.0021) ***
GOVT	-0.868544 (0.0000) ***	-0.47811 (0.0002) ***	-0.551329 (0.0000) ***
INF	-0.007474 (0.0414)**	-0.010532 (0.0383) **	-0.011134 (0.0647)*
INIGDPPC	-8.37844 (0.0010) ***	-9.90359 (0.0000) ***	-19.7290 (0.0000) ***
OPEN	0.07219 (0.0460) **	0.050415 (0.0184) **	0.065388 (0.0017)***
EDUC	0.64081 (0.0001) ***	0.70894 (0.0000) ***	-0.000199 (0.5937)
POPUL	-1.406256 (0.0000) ***	-1.12241 (0.0000) ***	-0.883763 (0.0123) **
TOT	-0.049034 (0.0000) ***	-0.03825 (0.0000) **	-0.029025 (0.0059) ***
C	93.83796 (0.0000)	101.4432 (0.0000)	208.7986 (0.0000)
R^2	0.464580	0.62833	0.803644
D.W	1.12489	0.37271	1.61854
Obs.	230	200	60

*** Indicates significant at 1%.
** Indicates significant at 5%.
* Indicates significant at 10%.
1) Amount of p-value is reported in parenthesis.

In the periodic of 4 years in an overlapping scenario, TOT, INF, and OPEN are statistically significant at 5% and the rest of the explanatory variables are highly statistically significant with the same signs in the first scenario. In the periodic of 4 years in a non-overlapping scenario, INF is statistically significant at 10% level, the GFCF, INIGDPPC, GOVT, and TOT are negative and statistically significant at 1% level, but POPUL is statistically significant at 5% level. These explanatory variables improve the goodness-of-fit and decrease the error value. EDUC does not have statistically significant impact on growth in the 4- year non-overlapping data. Thus, we can conclude that, the GFCF, INF, INIGDPPC, GOVT, POPUL, TOT, and

OPEN have influence the prediction of changes in growth, within the developed countries sampled.

In Table 4.5 Column (3), the estimated coefficient shows that EDUC is insignificant in developed countries. The level of EDUC tends to negatively impact on growth. The estimated coefficient is -0.00019, since, increase in INIGDPPC and EDUC would reduce growth in the equation 1 (refer to chapter 3) by diminishing return to input factors. So a country with high level of INIGDPPC and EDUC tends to grow at a slower rate (Barro and Sala-i-Martin, 2004).

If we compare Table 4.4 and Table 4.5, the estimated coefficient of INIGDPPC for developed countries is less than that of the developing countries in 3 scenarios. This indicates that the developing countries grow faster than the developed countries. We also found that, the developing countries EDUC estimated coefficient is higher than that of the developed countries for 3 scenarios. This implies that the higher investment in human capital, speed up the convergence process. Thus, education attainment is a crucial yardstick for measuring technological progress and increase in economic growth, since improving skills and knowledge of the labor force are crucial factors when considering efficiency gain.

Chapter 5

ARTIFICIAL NEURAL NETWORK/GENETIC ALGORITHM METHOD

The second part of our methodology entails the application of neural network as nonparametric methods to examine the overall effects of growth determinants on the economic growth in a panel data context. In fact, this section is the novelty of this thesis in the sense that, -to the best of our knowledge – no other work has used ANN methodology in a panel data framework in the growth studies. Our study follows the artificial neural network technique developed by Giovanis (2008) to estimate and compare in-sample forecasting performance of the panel-based fixed-effect models with that of the neural networks (NNs). Furthermore, in order to define the efficiency of artificial neural network approach in predicting growth, a hybrid version of the ANN approach known as genetic algorithms (GAs) is employed. In this regard, root-mean-square error (RMSE) statistics are used to determine the predictive power of the estimation model against other standard panel-based regression estimations.

5.1 Artificial Neural Network (ANN)

In this section, we discuss the application of ANN approach and some other special properties of the NNs. The ANN approach has two types of machine learning; the supervised machine learning and the unsupervised machine learning. The unsupervised machine learning executes a task of inferring a function to express hidden framework from "untagged" data. This approach does not require categorization or classification of the variables under observations. This approach makes use of the input variables without considering the output variables. Just as the name implies, it does not need supervision during training process. The supervised machine learning on the other hand indicates that, both the input variables and its relevant output variables are present. The most famous type of supervised neural network machine learning in data analysis is the Multilayer Perceptron (MLP). The MLP network has high capacity of management in terms of its application to non-linear functions. The MLP network is a feed-forward[4] network compose of an input layer, hidden layers (make up of sigmoid neurons), an output layer, and the numbers of activity function perceptron nodes of each layer (see Figure 5.1). These individual perceptron nodes through their weight coefficients obtain past layers outputs and forward it to the succeeding layer.

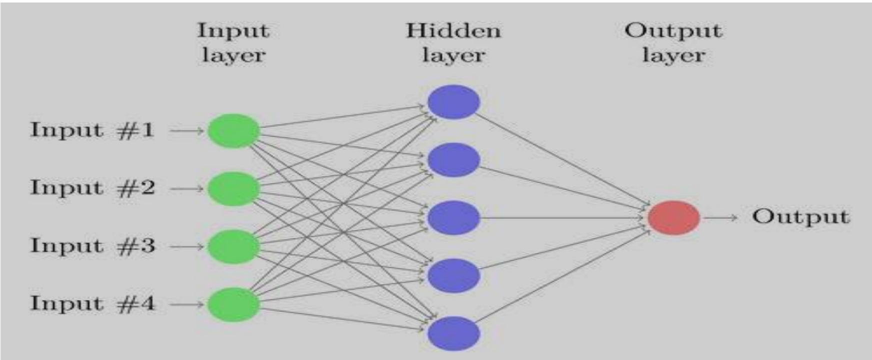

Figure 5.1: The architecture of the feed-forward Neural Network with one hidden layer

The network operation also depends on relationships between those layers and neurons. Thus, it is better to first define those layers and then, the interconnection between the layers should be determined.

[4] Most links are of this type, in which signals travel in only one direction. There is no feedback from output to input, and the output of each layer has no effect on that layer (in this study author used forward connection).

The Input layers are receiving raw data that was fed into the network. The hidden layers on the other hand, are layers that their operation is determined by inputs and weights of the relationships between the input layers and the hidden layers. The weights between the input nodes and hidden nodes determine when a hidden node must be activated. Meanwhile, the output layers, operation depends on the activities of the hidden unit and weight of the relationship between the hidden nodes and the output nodes. There are also single-layer and multi-layer networks for output units. The single-layer shows all units are linked to one layer. The single-layer is most widely used and has higher computational potential than the multi-layered network. The units in this network are numbered by layer in multi-layers networks, rather than the conventional global numbering.

The hidden layers are composed of hidden neurons or nodes. In this dissertation, we considered hidden layer with 10 nodes. Each node creates a linear regression by using their weight coefficients. Therefore, in applying linear regression analysis, the study estimates a linear integration of the inputs and weights known as net input function. Afterwards, we examined hyperbolic tangent function (Tansig) via non-linear activation function. We considered the Tansig as a non-linear activation function, squeezed via a unit step function to generate the projected output.

This is done in order to allow a nonlinear realtionship between the weighted inputs and output. The combination of nonlinear function gives MLPs their modeling flexibility. A regression (multilayer perception) model such as MLP depends on unknown parameters that should be estimated from the data. The MLP method chooses the variables data to calculate error function. The error functions can be thought of as a measure of the distance between forecast data and the real data. The objective of this function is to find set of parameter estimates that minimize the error or RMSE (Masters, 1993).

In this study, in order to determine the output neuron, the input neurons are weighed by training process weight coefficients, known as Levenberg–Marquardt approach. This method gives the real weight to each neuron. The weight coefficient is as below:

$$w_{k+1} = w_k - a_k g_k \qquad \text{(Equation 13)}$$

Where, w_k is the weight coefficient of network, g_k represents gradient of network's output error, while a_k is learning coefficient of network. This approach can also be referred to as gradient descent algorithm which is a kind of numerical optimization techniques. The numerical optimization technique accelerates the rate of computation to reach gradient error, thereby lowering the mass of computation.

Besides, one of the major reasons for employing the tools of neural network in-sample forecasting is due to the non-linearity properties of time series and cross sectional data. The ANN on the other hand, is suitable for modeling non-linear relationship inherent in data, without preexisting knowledge regarding the nexus between the input variables and the output variables. In applying ANN approach, optimization techniques should be used to examine best value of the predictions. First, one has to choose the values to start with. The starting values are equivalent to an initial guess at the parameter values. These values are updated to improve the estimates and decrease the error. These processes continue until there is no any further progress in the result. The neural network is suitable for most economic applications and predictions but it can be generalized further. However, since the feed-forward MLP suffered from the low convergence in real data, several researchers have made use of different methods for enforce optimization (Sexton et al., 1998). Thus, the current study also optimized the weights of the ANN via the Genetic algorithms (GA).

5.2 Genetic Algorithm Method (GA)

Genetic algorithm is described as a feature of natural evolution and has been used in-sample forecasting processes as a tool of numerical optimization. The chromosomes as a set of properties can be mutated for each candidate solution. The use of GA to learn the weights of a neural network can be much faster than other methods. GA operators include crossover, selection and mutation, so, in GA we have 3 components or stages: the selection stage, the crossover stage, and the mutation stage. The selection stage is a stage that one comes by after ranking variables (chromosomes in any population) among all chromosomes. Some of these variables can be selected to produce new generation. Meanwhile in crossover stage, two chromosomes are randomly chosen, and recombined from good individuals until creating the better ones. Mutation stage on the other hand, is a stage after the crossover stage.

In mutation stage, the values of selected chromosomes are changed up to the new chromosome, while mutation creates new chromosomes by changing gene content inside of each chromosome.

As presented in Fig. 5.2, among the chromosomes in the population, the most elegant are selected and two chromosomes among them have crossover, randomly. Crossover for each pair of chromosomes is considered between 0.6 and 0.95 such that, this number is called a crossover rate or probability of crossover (PC). If the crossover operation is not done on a pair of chromosomes, then children are produced as repetitions of the parents.

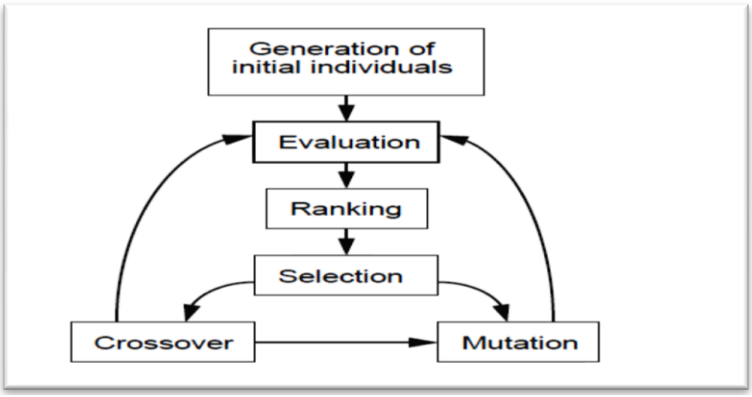

Figure 5.2: General Structure of Genetic Algorithm
Source: by author

The mutation operator has selected a gene from a chromosome randomly and then it alters its content. The probability of action on each chromosome mutation is random and it is called the rate of mutation or mutation probability, indicated by PM. Usually, this number is considered very small (for example, 0.001). To get maximum coverage of GA the mutation process is used which helps to achieve new different arguments which are substantially different from the original ones.

In general, GA is developed for solving problem on a set of assumptions called “population” which is replaced with new hypotheses periodically and produces a very large set of possible solutions. In each repetition, each solution is evaluated using a "fitness function". However, some of the best solutions (hypotheses) will be used to generate new solutions. Thus, this will result in development of solutions and among the solutions obtained; those that do not pass the upper threshold of the mean square error are selected. Using an iterative process, the weight of each of the prediction approaches that forms chromosomes for the proposed algorithm is updated and this process then continues using gene operators until the amount of error or loss function decreases appropriately. Finally, the best chromosomes or prediction weights are obtained, in such a way that the search space has evolved to reach its optimal solution. This method can be very effective in cases where a proper selection of parameters has been applied.

Genetic algorithms have numerous applications: economics, optimization, automatic programming, machine learning, operations research, studies of evolution and learning, ecology, and social systems (Siedlecki, and Sklansky, 1989).

Advantages of GA compared to other models of forecasting are as follows: (i) genetic algorithms search in a population of solutions, not in a single answer. The genetic algorithm has no special mathematical needs, and regardless of internal problem performance, it solves optimization problems. This algorithm is able to solve any linear or nonlinear restrictions defined on continuous, discontinuous or mixed search space. (ii) The structure of genetic algorithm operators enables this algorithm to find optimal general solutions successfully. In traditional methods the search is carried out by making comparison with neighbor points, while movement is conducted toward relative optimum points.

5.3 The ANN/GA Method Based for Panel Data

Among all growth theory prediction methods, application and combination of ANN/GA machine learning methods was selected to examine the topic, since recent studies have demonstrated that optimal results are achieved using the combined prediction method as an alternative to individual models. This is because, the prediction combination has several advantages that can cover weaknesses in individual models and transfer the best method for prediction. Even if the best model at any point in time can be identified, combining the two is still an attractive strategy because it has the benefit of diversity. However, its success will depend on the method used to extract the combining weights. Given that GA provides a random global search method to solve complex optimization problems, the ANN/GA method has been used to solve the problem and achieve optimal weights.

Meanwhile due to limitations associated with individual growth prediction models, different models has been used to predict growth rate such that, each has strengths and weaknesses over the other models. Combining two different predictions methods improve accuracy. In many cases, performance can be improved significantly by simple averaging of a prediction (Vafaie and De Jong, 1992, 1993).

Various studies have used different classes of machine learning in time series analysis, however, to the best of our knowledge, up till date, only a few studies have adopted NN approach in panel data context, to conduct in-sample forecasting comparison with other methodologies like fixed-effect models. We follow the methodology of Giovanis (2008) in order to feed our inputs to the network in a panel framework. In addition to our independent variables in Equation (1), two more class of inputs are introduced in the model. The first one is time variable and the cross-sectional county-specific dummy variable. Using this framework helps us to compare the neural network estimators with the fixed-effect model estimators. This study after the pre-processing the dataset for ANN, we used MLP_GA with 10 sigmoid neurons, one hidden layer, activation function of Tansig[5],and linear activation function for output layer with Levenberg–Marquardt method is used for training.

5.4 ANN/GA Result

As reported in Table 5.1, in univariate time series application, each input neuron represents the explanatory variables, while the output neuron represents the dependent variable or MLP network forecasts. A weight (connection strength-Equation (13)), w_k, is associated with each link, and a network is trained (learned) by

[5] *Tansig* is a transfer function. Transfer functions calculate a layer's output from its net input.

modifying these weights, thereby modifying the network function that maps inputs to outputs. Figure 5.3 reports the results of MLP_GA neural network. In each scenario, based on the different dataset; all the independent variables (same as what we used in fixed-effect model) are introduced into the model as an input, to capture the growth rate. We used 10[6] sigmoid neurons for one hidden layer and the network is iterated 100 times to get the optimized layer weights using the GA. In the GA, selection mode is assigned randomly and the population size is set up to 50[7]. The activation function is Tansig and the linear activation function is used for the output layer. In addition, the Levenberg-Marquardt method, as shown in Equation 13, is applied to define the weight coefficients for training (Giovanis, 2008). The structure of the used network is demonstrated in Figure 5.3.

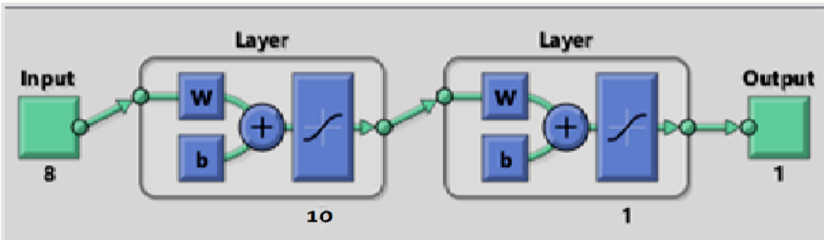

Figure 5.3: The Structure of MLP Network

As it shown in the Table 5.1, the other models (fixed effect method for three scenarios) outperform the accuracy of the whole data set. In the case of using the datasets with whole countries and with developing countries only, the periodic data with overlapping format has lower RMSE compare to the non-overlapping point of view, while for the developed countries only dataset, the non-overlapping format outdo by the overlapping scenario. The results are comparable with the RMSEs,

[6] Three different number of neurons [5, 10, 15] are tested to obtain neurons number with minimum error in hidden layer. Since 10 neurons had lowest Akaike Information Criterion (AIC) comparing to the other alternatives, therefore we adopt 10 neurons as an optimal number.

[7] Population size is considered randomly from 20 to 100, according to the length of chromosome. For our study, size 50 is an optimal number.

which are extracted from the panel-based FEM analyses (for RMSE formula see Appendix C). As reported, in all cases, the performance of neural network as a non-parametric regression tool is robust and consistent when compared with that of the fixed-effect model.

It is noteworthy to mention that, if we had used the pooled dataset for the neural network methodology, then, the results would not be comparable with the panel-based fixed-effect analyses due to the impact of cross sectional dependency and heterogeneity. Therefore, to correct the observed shortcomings, we developed the hybrid model of ANN, which is suitable for comparison analysis.

Table 5.1: Value of Root Mean Square Error

Data Set	RMSE-ANN	RMSE_Fixed
Whole 20 countries, yearly data	0.014	2.84
Whole 20 countries, periodic non overlapping	0.24	3.02
Whole 20 countries, periodic overlapping	0.06	2.42
Developing countries, yearly data	0.12	3.69
Developing countries, periodic non overlapping	0.85	3.28
Developing countries, periodic overlapping	0.43	2.06
Developed countries, yearly data	0.25	1.35
Developed countries, periodic non overlapping	0.65	3.35
Developed countries, periodic overlapping	0.72	2.93

Moreover, when we combine the ANN and GA methods, the predicting power of growth model become more stable than traditional panel method for the 3 scenarios.

Certainly, upon the introduction of the ANN/GA, the accuracy of the model for nonlinear framework changed. This is acceptable since using hybrid models need to introduce two different categories of new independent variables, namely; the time variable which represent horizon of interest and cross section variable, which cover somewhat the heterogeneity of the model. Table 5.1 revealed the nonlinear characteristic of the panel dataset and the efficient performance of the neural network in case of capturing the nonlinear phenomena (Giovani, 2008; Tohidi et al, 2015).

Chapter 6

CONCLUSION

This dissertation attempts to identify the determinants of the economic growth by using panel fixed effect model. Moreover, it also used ANN/GA in a panel data setting for growth prediction. As such, it is one of the rare studies using ANN/GA model for growth study. The results are also compared based on RMSE values in order to evaluate comparative performance of the two models. However, this dissertation's real contribution to economics literature lays in its pioneering attempt to use ANN/GA method in a panel data framework in economic growth studies.

To this end, this study uses a panel data of 20 countries and 23 years from 1990 to 2012. Of the 20 countries, ten are developed countries and ten are developing countries. The developed countries in the sample are United Kingdom, Germany, Japan, Spain, Norway, New-Zealand, Sweden, France, Australia, Greece, while the developing countries are India, Venezuela, Turkey, China, Nigeria, Iran, Russia, Ukraine, Pakistan, and Brazil.

We believe that the topic is particularly important. In fact, issues pertaining to imbalance in the economic growth of nations of the world have been one of the most controversial issues in last two decade. Almost all economists and policy-makers question why some countries stagnate, experience persistent/slower growth overtime while others grow faster. There are numerous studies on this field, however new theories, better data and new estimation methodologies continue to contribute to the field. In line with this, we introduce a new economic growth in-sample forecasting method by using nonlinear hybrid ANN/GA technique in a panel data framework. This model is built on two main studies. The growth model is based on Barro (1996), and the new estimation methodology adopted from Giovani's paper (2008).

The explanatory variables employed for our empirical analysis are initial GDP per capita (INIGDPPC) -capturing the effects of convergence-, gross fixed capital formation (GFCF), human capital (EDUC), trade openness (OPEN), terms of trade (TOT), government size (GOVT) -that is government final consumption expenditure as percentage of GDP, inflation rate (INF) and population growth rate (POPUL).

In sum, the aim of this study is to evaluate and forecast the impact of the specified macroeconomic variables on the economic growth by using a panel data of 20 sampled countries. The work uses panel-data fixed-effect model and machine learning approach of artificial neural network and genetic algorithm, and attempts to compare and choose between the two models the more robust and efficient methodology in evaluating growth of the countries.

Thus, the main objective of this dissertation is to compare the power of various methods in predicting the economic growth in selected countries (developing and developed). One contribution of this dissertation is to show that the combined Artificial Neural Networks with Genetic Algorithms has greater predictive power than traditional panel method with regard to changes in growth rate. In fact, ANN/GA in a panel framework has not yet been used by previous studies investigating the determinants of economic growth. Furthermore, the study intends to evaluate whether INF, GFCF, INIGDPPC, OPEN, TOT, EDUC, GOVT and POPUL are sound macroeconomics variables and suitable in evaluating growth theory in economic literature. This was done using panel data from the period 1990 to 2012 and concluding with proposed adjustments to the neural network model in order to improve the accuracy of its predictions.

In this dissertation, the real GDP per capita growth rate was predicted by using three different data formatting in each of three different samples of countries for two methods. The two methods used are the hybrid model of ANN/GA and fixed effect panel regression. The three different samples are based on whole of 20 countries, 10 developed countries only, and 10 developing countries only. Data formatting includes yearly data, periodic data of 4-year overlapping, and periodic data of 4-year non-overlapping data.

The fixed effect method results have shown that the INF, INIGDPPC, GOVT, POPUL, TOT, OPEN, EDUC and GFCF are all statistically significant and suitable to predict the growth rate for "***panel data with whole 20 countries***" in all 3 data-formatting, except the EDUC in periodic 4-year non-overlapping data.

For example, the estimated coefficient for inflation rate is negative and has statistically significant impact on economic growth. Similarly, the INIGDPPC has the negative effect on the growth rate because it captures the effects of convergence, that is the higher the initial GDP per capita, the lower is the growth rate. On the other hand, gross fixed capital formation in a country facilitates an increase in production and thus higher per capita real GDP or growth rate. In line with this, our results indicate the gross capital formation has a positive effect on the real GDP per capita growth rate.

Most trade theories state that trade openness in a country has a positive impact on real GDP per capita growth rate. In line with these theories, we find that trade openness has a positive sign on economic growth. The estimated coefficients for EDUC turns out to be statistically significant and of positive sign for both yearly data and periodic 4-year overlapping data. This is considered to have a positive impact on growth. The higher mean years of schooling (human capital) raises the ability to absorb new technological advancement, which turns out to have positive impact on economic growth in the long-run.

Meanwhile, government size has negative and statistically significant impact on growth. We can explain this in two ways: First, if government consumption is financed by the tax system, it will distort competition within the private sectors.

Second, if government expenditures is financed by borrowing, it may lead to higher interest rate and higher debt which may destroys macroeconomic stability and raises private sector investment costs.

The population growth rate has also negative and statistically significant impact on the economic growth. High population growth rate decreases private saving which in turn retard such country's economic growth. Finally, the estimated coefficient for terms of trade is found to be negative and statistically significant. This is in line with several previous papers which were already quoted in our theoretical expectations section.

When we are using ***a "sub-sample of 10 developing countries only"***, the fixed-effect model results indicate that, the explanatory variables are all statistically significant to predict the growth rate for developing countries in 3 data formatting, except for the population growth rate in yearly data and periodic of 4-year overlapping data as well as except for trade openness in periodic 4-year non-overlapping data.

The human capital, gross fixed capital formation, and trade openness have a positive impact on growth rate, while government size, inflation rate, initial real GDP per capita and population growth rate exhibit negative impact on growth. The estimated coefficient for terms of trade is negatively statistically significant for annual panel data and it is positively statistically significant for other periodic data formatting. This might be in line with the previous mixed results for terms of trade.

When we are using ***a "sub-sample of 10 developed countries only"***, the fixed-effect model results also indicate that all explanatory variables are statistically significant to predict the growth rate for developed countries in all 3 data formatting, with exception of the human capital in periodic 4-year non-overlapping data. Meanwhile increase in government size, inflation rate, initial GDP per capita, population growth rate and terms of trade would decrease the growth rate, while increase in human capital, gross fixed capital formation and trade openness would increase the growth in developed nations.

We used the artificial neural network and genetic algorithm to predict the growth rate by using the same independent variables for 3 different data-formatting. We introduced MLP-GA feed-forward neural network, which can give the proper estimation, against traditional statistics and econometric estimations. Moreover, the root of mean square error was used to compare the predictive power of two different methods.

Empirical results indicated that, in fixed-effect model the explanatory variables have statistically significant impacts on economic growth. Moreover, the predictive power of the ANN/GA model is greater than that of the fixed panel model. This is because, the RMSE value for ANN/GA is lesser than that of the FEMs. Thus, ANN/GA has greater growth- in sample forecasting power than the traditional panel-based model in 3 data formatting - annual panel, periodic of 4-year overlapping and periodic of 4-year non-overlapping data-.

Finally, we conclude that, the genetic algorithm combined with the multiplier neural network method has the greatest impact in predicting country's economic growth.

REFERENCES

[1] AbuDalu, A., Ahmed, E. M., Almasaied, S. W., Elgazoli, A. I. (2014). The Real Effective Exchange Rate Impact on ASEAN-5 Economic Growth. *International Journal Econ Management*. Sci 3:174. doi: 10.4172/2162-6359.1000174.

[2] Adhikary, B. K. (2011). FDI, Trade Openness, Capital Formation, and Economic Growth in Bangladesh. *A Linkage Analysis International Journal of Business and Management,* 6 (1).

[3] Afonso, Ó. (2001). The Impact of International Trade on Economic Growth. Working paper. Faculty of Economics. University Do Porto. Portugal. (pp. 1-23).

[4] Andres J. and I. Hernando (1997). Does inflation harm Economic Growth? Evidence for the OECD, Banco de Espana. *Working Paper* 9706.

[5] Aydın, C., Eseb, Ö., and Bayrak, M. (2016). Inflation and Economic Growth: A Dynamic Panel Threshold Analysis for Turkish Republics in Transition Process. ScienceDirect: Social Behavioural Sciences, *5th International Conference on Leadership, Technology, Innovation and Business Management.* 229. (pp. 196 – 205).

[6] Baltagi, B. H. (2013*).* *Econometric Analysis of Panel Data*. 5rd Ed. John Wiley & Sons.

[7] Barro, R. J. (1996). Determinants of Economic Growth – A Cross-Country Empirical Study. *National Bureau of Economic Research* (NBER). Working Paper No. 5698.

[8] Barro, R. J. (1995). Inflation and Economic Growth. *National Bureau of Economic Research* (NBER). Working Paper No. 5326.

[9] Barro, R. J., and Sala-i-Martin, X. (2004). *Economic Growth*. MIT Press, Cambridge Mass, 2004.

[10] Becker, G. S., Glaeser, E. L, and Murphy, K. M.(1999). Population and Economic Growth. *American Economic Review Papers and Proceedings* 89, no. 2. (PP. 145-154).

[11] Becker, R.A., Chambers, J.M., and Wilks, A.R. (1988). *The new S language*. Chapman and Hall: New York.

[12] Binner, J. Bissoondeeal, R., Elger, T., Gazely, A., and Mullineaux, A. (2005). A comparison of linear forecasting models and neural networks: an application to euro inflation and euro division. *Applied Economics*, no 37(6). (pp. 665-680).

[13] Blake, A.P. (1999). An Artificial Neural Network System of Leading Indicators. *National Institute of Economic and Social Research*. Unpublished paper.

[14] Blattman, C., Hwang, J., and Williamson, J. G. (1997). The Impact of the Terms of Trade on Economic Development in the Periphery, 1870-1939: Volatility and Secular Change. *NBER Working Paper* No. 10600-NBER Program(s): DAE ITI.

[15] Buigut, K., Soi, N., Koskei, I. Kibet, J. (2013). The effect of capital structure on the share price of listed firms in Kenya. A case of energy listed firms. *European Journal of Business and Management* 5 (9). (pp. 29-35).

[16] Chaudhuri, T. D., and Ghosh, I. (2016). Artificial Neural Network and Time Series Modeling Based Approach to Forecasting the Exchange Rate in a Multivariate Framework. *Journal of Insurance and Financial Management*. Vol. 1, Issue 5. (pp. 92-123).

[17] Checherita, C., and Rother, P. (2010). The impact of high and growing government debt on economic growth: an empirical investigation for the Euro area. *European Central Bank*, *Working paper series*,No1237. (pp. 1-40)

[18] Cohen, D., and Soto, M. (2007). Growth and human capital: Good data, good results. *Journal of Economic Growth*. 12(1).(pp. 51–76).

[19] Dawson, P. J., and Tiffin, R. (1998). Is there a Long-run Relationship between Population Growth and Living Standards? The case of India. *Journal of Development Studies* 34, no. 5. (pp149-56).

[20] Demir, A., Shadmanov, A., Aydinli, C., and Eray, O. (2015). Designing a forecast model for economic growth of Japan using competitive (hybrid ANN vs multiple regression) models. *ECOFORUM*, Volume 4, Issue 2 (7). (pp. 49).

[21] Devarajan, S., Swaroop, V. and Zou, H. (1996). The Composition of Public Expenditure and Economic Growth. *Journal of Monetary Economics*, 37(2), (pp. 313-344).

[22] Durlauf, S. N., Paul, A., and Poughkeepsie, NY. (2004). Growth Econometrics. Vassar College *Economics Working Paper* No. 61

[23] Dutt, A. K., and Ros, J. (2008).International Handbook of Development Economics, Volumes 1 & 2. Handbook. Elgar original reference. ISBN:9781845423278.

[24] Easterly, W., Kremer, M., Pritchett, L., and Summers, L. (1993). Good Policy or Good Luck? Country Growth Performance and Temporary Shocks. *Journal of Monetary Economics.* 32 (3). (pp. 459-83).

[25] Eviews note. (2017). Advanced Univariate Analysis, Univariate Time Series Analysis, Panel Unit Root Testing. Retrieved from http://www.eviews.com/help/helpintro.html#page/content/advtimeser-Panel_Unit_Root_Testing.html.

[26] Feng, L., & Zhang, J. (2014). Application of artificial neural networks in tendency forecasting of economic growth. *Economic Modeling,* 40. (pp. 76–80).

[27] Ferrara, L., Guégan, D., and Rakotomarolahy, P. (2010). GDP now casting with ragged-edge data: a semi-parametric modeling. *Journal of Forecasting*. 29, (pp. 186-199).

[28] Fischer, S. (1993). The Role of Macroeconomic Factors in Growth. NBER woeking aper.no.4565. *Journal of Monetary Economics,* 32. (PP.36)

[29] Geisendorf, S. (1999). Genetic Algorithms in Resource Economic Models. Retrieved from HTTP:// www.santafe.edu/media/workingpapers/99-08-058.pdf.

[30] Ghosh, A., and Philips, S. (1998). Inflation, Disinflation, and Growth, *IMF Working Paper*.

[31] Gillman, M.(2009). Inflation Theory in Economics: Welfare, Velocity, Growth and Business Cycles. *Routledge international studies in money and banking*. Published by Routledge.

[32] Gjylapi, D., Proko, E., and Hyso, A. (2016). Genetic Algorithm Neural Network model vs Back propagation Neural Network model for GDP Forecasting. ResearchGate. Conference: Recent Trends and Applications in Computer Science and Information Technology (RTA-CSIT).

[33] Giovanis, E. (2008). A Panel Data Analysis for the Greenhouse Effects in Fifteen Countries of European Union. *MPRA Paper*. No. 10321, posted 9. UTC.PP 1-38.

[34] Gries, T., and Redlin, M. (2012). Trade Openness and Economic Growth: A Panel Causality Analysis, Center for International Economics, University of Paderborn, Germany.

[35] Gujarati, D.M., Porter, D.C., (2012). *Basic Econometrics*. 5th Edition.

[36] Gylfason, T. (1991). Inflation, Growth, and External Debt: A View of the Landscape. *The World Economy*, 14, 3 (Sept). (pp. 279-298).

[37] Harberger, A.C. (1950). Currency depreciation, income and the balance of trade. *Journal of Political Economy,* no 58. (pp. 47-60).

[38] Headey, D. D., and Hodge, A. (2009). The effect of *Population growth on economic growth: history and sect oral considerations. Unpublished paper. London school of economics. Dev. Rev.* 35(2), (pp. 222).

[39] Heravi, S., Osborn, D. R., & Birchenhall, C. (2004). Linear versus neural network forecasts for European industrial production series. *International Journal of Forecasting*, 20 (3). (pp. 435–446).

[40] Hernandez, F. J. A., Ortiz, J. S., Alejandre, M. P. A., and Cruz, L. X. V. (2017). Bayesian Analysis for a Fractional Population Growth Model. *Journal of Applied Mathematics*. Vol 2017. Article ID 9654506. 9 pages.

[41] Hoptroff, R. C., Bramson, J., and Hall, T. J. (1991). Forecasting economic turning points with neural networks to combine financial forecasts. *IEEE transactions on evolutionary computation.* Vol.1, No. 1. (pp. 347-352).

[42] Ilegbinosa, I. A., Micheal A., and Watson, S. I. (2015). Domestic Investment and Economic Growth in Nigeria from 1970-2013. *An Econometric Analysis - Canadian Social Science,* Vol. 11, No. 6, 2015, (pp. 70-79). DOI:10.3968/7009.

[43] Imoisi, A. I., Abuo, M., and Sogules, I. W. (2015). Domestic Investment and Economic Growth in Nigeria from 1970-2013: An Econometric Analysis. Canadian Academy of Oriental and Occidental Culture.ISSN 1712-8056[Print] ISSN 1923-6697[Online] *Canadian Social Science,* Vol. 11, No. 6, 2015, (pp. 70-79). DOI:10.3968/7009.

[44] Javed, Z. H., Qaiser, I., Mushtaq, A., Sai-ullaha, I. A. (2012). Effects of International Trade on Economic Growth: The case study of Pakistan. *International Journal of Academic Research in Progressive Education and Development.* 1(2). (pp. 103-113).

[45] Joint Report by the Economic Policy Committee (Quality of Public Finances) and the Directorate-General for Economic and Financial Affairs.

(2010). Efficiency and effectiveness of public expenditure on tertiary education in the EU. Working Document. *European Economy Occasional Papers*. No.70

[46] Kalumbu, S. A., and Sheefeni, J. P. S. (2014). Terms of Trade and Economic Growth in Namibia. International Review of Research in Emerging Markets and the Global Economy (IRREM), *an Online International Research Journal.* ISSN: 2311-3200, Vol: 1 Issue 3. (pp. 90).

[47] Klasen, S., and Lawson, D. (2007, June 12). The Impact of Population Growth on Economic Growth and Poverty Reduction in Uganda. *Researchgate.* Retrieved from https:// www.researchgate.net/publication/ 5081419.Source: RePEc.

[48] Koutmos, G., and Boot, G. G. (1995). Asymmetric volatility transmission in international stock markets. *Journal of International Money and Finance*. vol. 14, issue 6. (pp. 747-762).

[49] Li, S., Xu, M. and Sun, B. (2014). Long-term hydrological response to reforestation in a large watershed in southeastern China Hydrol. *Process.* 28 5573–82.

[50] Lucas, R.E. (1988). On the mechanics of economic development. *Journal of Monetary Economics,* 22. (pp 3 –42).

[51] Mankiw, N.G., Romer, D., and Weil, D.N. (1992). A contribution to empirics economic and growth. *The Quarterly Journal of Economics*. (pp. 407-437).

[52] Martin, M., & Herranz, A. (2004). Human Capital and Economic Growth in Spanish Regions. *IAER*. 10(4).(pp. 249-256).

[53] Masters, T. (1993). *Practical Neural Network Recipes in C++*.Academic Press Sandiego,CA.

[54] Medee, P.N and Nembee, S.G (2011). Econometric Analysis of the impact of Fiscal Policy Variables on Nigeria's Economic Growth (1970 - 2009). *International Journal of Economic Development Research and Investment*, 2(1).(pp. 171-183).

[55] Mendoza, E. G. (1997). Terms-of-trade uncertainty and economic growth.Journal of Development Economics. Elsevier Science B.V, Journal of Development Economics. *Journal of Development Economics*. Vol. 54. (pp. 23-356).

[56] Pena, D., Tiao, G.C., and Tsay, R.S. (2003). A course in time series analysis. *Wiley Series in Probability and Statistics: New York*.

[57] Phillips, B. C.P., and Perron, P. (1988). Testing for a Unit Root in Time Series Regression". *Biometrika* 75, no. 2, (pp 46-335).

[58] Potančoková, M., K.C., S. and Goujon, A. (2014). Global Estimates of Mean Years of Schooling: A New Methodology. IIASA Interim Report. *IIASA*, Luxemburg, Austria, IR14005.

[59] Ramsey, F. P. (1928). A Mathematical Theory of Saving. *Economic Journal*. 38 (152). PP 543–559. JSTOR 2224098.

[60] Romer, P. M. (1989a). Endogenous technological change. *National Bureau of Economic Research*, Cambridge, MA.

[61] Romer, P. (1989b), "Human Capital And Growth: Theory and Evidence," *NBER Working Papers* 3173, National Bureau of Economic Research, Inc.

[62] Sameti, M., Farahmand Sh., Koleyni K., and Aghaeifar R. (2013). Comparing the ANN and Linear Regression in Estimation of the Growth Model (The Case of MENA*). Iranian Economic Review*, Vol.17, No. 1.

[63] Samimi, A. J., Sadeghi, S., and Sadeghi, S. (2011). The impact of the terms of trade volatility on Economic Growth: Evidence from Oil Exporting Countries. *International Journal of Economics and Management Engineering*, 1(1). (pp. 50-53).

[64] Finnegan, J., and Sexton, K. (1998). Community-based environmental decisions: analyzing power and leadership. In: Sexton K, Marcus AA, Easter KW, et al., editors. Better Environmental Decisions. Washington, DC: Island Press, 1999, 331–51.

[65] Shi, P., Chen, Z., and Xie, G. (2006). Using artificial neural network trained with the genetic algorithm to model GDP prediction. *Center for Systems and Control, LTCS, and Dept. of Mech. & Eng. Sci.* Peking University.

[66] Siedlecki, W., and Sklansky, J. (1989). A note on genetic algorithms for large scale pattern selection. *Pattern Recognition Letters*, vol.10, (pp. 335-347).

[67] Sokolov, S. M., Milovancevic, M., & Mladenovic, I. (2016). Evaluation of trade influence on economic growth rate by computational intelligence approach. *Physica A.*: PHYSA 17448. PII: S0378-4371(16)30548-9.

[68] Solow, R. M. (1956). A contribution to the theory of economic growth". Quarterly Journal of Economics. *Oxford Journals.* 70 (1). (pp. 65–94). doi:10.2307/1884513. JSTOR 1884513.

[69] Ssewamala, F. M., Nabunya, P., Ilic, V., Mukasa, M. N., & Damulira, C. (2015). The relationship between family economic resources, psychosocial well-being, and educational preferences of AIDS-orphaned children in southern Uganda: Baseline findings. *Global Social Welfare*. DOI: 10.1007/s40609-015-0027-z.

[70] Stewart, M. B. (2004).The Employment Effects of the National Minimum Wage". *Economic Journal,* 114. (pp. 10-16).

[71] Swamy, V. (2015). Government Debt and Economic Growth: Decomposing the Cause–Effect Relationship. *IEG Working Paper*, No. 360.

[72] Swan, T. W. (1956). Economic growth and capital accumulation". Economic Record. *Wiley*. 32 (2). (pp. 334–361). Doi:10.1111/j.1475-4932.1956.tb00434.

[73] Tan, C. J. K. (2012). International Trade and Economic Growth: Evidence from Singapore. A paper Submitted in partial Fulfillment of the requirement for the degree of master's of Arts in the Graduate School of Arts and Science, Columbia University.

[74] Thornton, J. (2001). Population Growth and Economic Growth: Long-run Evidence from Latin America. *Southern Economic Journal*, 68, no. 2. (pp. 464-68).

[75] Thuku, K, Paul, G., & Almadi, O. (2013). The Impact Of Population Change On Economic Growth In Kenya. *International Journal of Economics and Management Sciences*, Vol. 2, No. 6, 2013, (pp. 43).

[76] Tkacz, G., and Hu, S. (1999). Forecasting GDP Growth Using Artificial Neural Networks. *Bank of Canada Working Paper*. (pp. 3-99).

[77] Todaro, M., Smith, S. C. (2012). *Economic Development*. Addison Wesley Longman, Inc., New York.

[78] Tohidi, A., Zare Mehrjerdi, M. R., Mehrabi Boshrabadi H., and Nezamabadi, P. H. 2015)"Evaluation of Artificial Neural Network-Panel Data Hybrid Model In Predicting Iran's Dried Fruits Export Prices." 95-116.

[79] Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach*, New York, NY: Oxford University Press.

[80] United Nations. (2011). Millennium goals. Retrieved August 15th, 2011, Retrieved from http://www.un.org/millenniumgoals/bkgd.shtml.)

[81] Vafaie, H., and De Jong, K. (1993). Robust feature selection algorithms. Proc. of the fifth conference on tools for artificial intelligence, Boston, MA: IEEE Computer Society Press. (pp. 356-363).

[82] Vafaie, H., and De Jong, K. (1992). Genetic algorithms as a tool for feature selection in machine learning. In Proceedings of the 4th International conference on tools with artificial intelligence. Arlington, VA. (PP.200-204).

[83] Wacziarg, R., and Welch K. H. (2008). Trade Liberalization and Growth: New Evidence. *World Bank Economic Review*, 22(2). (pp. 187-231).

[84] Won H., T.(2010). Terms of trade and economic growth in Japan and Korea: an empirical analysis. *Empirical Economics Journal*. Volume 38, Issue 1, pp 139–158.

[85] Wooldridge, J. M. (2002, and 2010), "*Econometric Analysis of Cross Section and Panel Data*". MIT Press.

[86] Yabu, N., & Kessy, N. J. (2015). Appropriate Threshold Level of Inflation for Economic Growth: Evidence from the Three Founding EAC Countries. *Applied Economics and Finance* Vol. 2, No. 3; August 2015 ISSN 2332-7294 E-ISSN 2332-7308. Published by Red-Fame.

[87] Yanikkaya, H. (2003). Trade Openness and Economic Growth: A Cross-Country Empirical Investigation. *Journal of Development Economics*, 72, (pp. 57 – 89).

[88] Yatchew, A. J. (1998). Nonparametric regression techniques in economics. *Journal of Economic Literature.* 36, (pp. 669-721).

[89] Young, A. (1991). Learning by doing and the dynamic effects of international trade. *Quarterly Journal of Economics.* 106, (pp. 369– 405).

APPENDICES

Appendix A: Descriptive Statistics

Appendix A.1: Descriptive Statistics for the INIGDPPC

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Me.	34641	37898.1	40219.6	27460.5	44054.1	862.022	8076.84	2143.51	4694.75	1572.4
Med	35658.1	38220.6	40199.2	28551.1	44693.8	794.481	7845.69	1761.14	4281.36	1320.2
Max	40477.8	44223.6	43882.7	32461.8	53421.2	1460.48	10818.0	4919.53	6455.21	2363.6
Min	27744.7	32339.0 3	35030.2	21689.8	35238.9	540.511	6039.91	708.825	3305.32	1238
S.D	4465.83	3405.8	2251.14	3698.08	6464.63	280.916	1394.24	1271.11	921.101	404.66
Sk.	-0.304	0.1796	-0.254	-0.187	0.1025	0.70623	0.44350	0.78621	0.5773	0.835
Kr.	1.575	1.966	2.5866	1.4538	1.4681	2.34109	1.98890	2.46205	2.026	2.010

Appendix A.2: Descriptive Statistics for the INF

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Mean	2.70	2.03	0.37	3.45	2.29	7.88	44.86	4.67	19.49	20.24
Median	2.32	1.72	0.06	3.37	1.93	8.35	54.40	3.06	17.21	12.88
Maximum	7.53	5.08	3.30	6.72	10.47	13.87	106.26	24.24	49.66	72.84
Minimum	0.79	0.31	-1.35	-0.29	-0.49	3.68	6.25	-1.41	7.63	5.38
Std.Dev.	1.73	1.17	1.22	1.55	2.70	3.25	33.43	6.23	9.07	18.85
Skewness	1.60	1.12	0.91	0.05	1.98	0.20	0.13	1.80	1.66	1.68
Kurtosis	4.98	3.84	3.29	3.44	6.39	1.84	1.56	5.79	6.33	4.49

Appendix A.3: Descriptive Statistics for the TOT

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Mean	96.13	106.55	108.88	102.76	96.63	94.09	123.36	102.63	71.75	80.87
Median	96.55	105.84	97.85	102.35	96.96	94.58	123.78	101.73	67.94	79.51
Maximum	111.19	114.54	180.26	119.86	106.13	115.08	152.68	118.23	103.36	113.40
Minimum	83.37	99.48	71.37	88.09	85.76	73.69	100.00	79.71	39.74	44.48
Std.Dev.	8.40	4.61	32.87	9.58	6.43	12.91	13.94	8.64	18.12	18.21
Skewness	0.21	0.16	1.13	0.28	-0.11	-0.16	0.24	-0.25	0.06	-0.09
Kurtosis	1.98	1.85	2.798	2.22	1.60	1.75	2.61	3.65	2.15	2.02

Appendix A.4: Descriptive Statistics for the POPUL

	UK	Germany	Japan	Spain	Sweden	India	Turkey	China	Iran	Nigeria
Mean	0.48	0.09	0.15	0.81	0.50	1.71	1.50	0.82	1.44	2.57
Median	0.38	0.06	0.21	0.46	0.56	1.73	1.54	0.73	1.29	2.55
Maximum	0.79	0.86	0.38	1.85	0.85	2.07	1.80	1.47	2.59	2.69
Minimum	0.24	-1.69	-0.20	0.06	0.06	1.29	1.17	0.48	1.12	2.50
Std.Dev.	0.22	0.50	0.16	0.66	0.27	0.24	0.17	0.31	0.37	0.07
Skewness	0.38	-1.69	-0.76	0.51	-0.33	-0.22	-0.39	0.54	1.50	0.57
Kurtosis	1.41	8.32	2.84	1.52	1.74	1.83	2.41	2.01	4.95	1.80

Appendix B: INIGDPPC Unit-Root Test without Trend and with Trend

Variable / First Difference		INIGDPPC	Obs.
Without Trend	L.LC[2]	-6.18372 (0.0000)[1]***	400
	I.P.S	-6.03619 (0.0000)***	400
	ADF	116.970 (0.0000)***	400
	PP	166.412 (0.0000)***	420
With Trend	L.LC	-5.9827 (0.0000)***	400
	B. t-stat	-1.1754 (0.1199)	380
	I.P.S	-4.2274 (0.0000)***	400
	ADF	91.4742 (0.0000)***	400
	PP	143.833 (0.0000)***	420

***Indicates significant at 1%.
** Indicates significant at 5%.
* Indicates significant at 10%.
1) Amount of p-value is reported in parenthesis.
2) Levin, Lin, and Chu (LLC, 2002), Breitung (B-tstat, 2000),
Im, Pesaran and Shin (IPS, 2003),
Fisher- Augmented Dickey Fuller (ADF) and
Fisher- Phillips-Perron (PP).

Appendix C: Root Mean Square Error

RMSE $= \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}$, n= number of observation=460,i= number of cross sections, $\hat{y}$is a dependent variables after regressing two ways fixed panel and ANN/GA methods or estimation value of Growth, and y_i is actual value of dependent variable.

$$RMSE_{\text{Whole FIXED PANEL DATA}} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} = \sqrt{\sum_{1}^{230} \frac{(y_i - \hat{y}_i)^2}{230}} = 2.84$$

$$RMSE_{\text{FIXED PANEL DATA}-\text{developing}} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} = \sqrt{\sum_{1}^{230} \frac{(y_i - \hat{y}_i)^2}{230}} = 3.69$$

$$RMSE_{\text{FIXED PANEL DATA}-\text{developed}} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} = \sqrt{\sum_{1}^{460} \frac{(y_i - \hat{y}_i)^2}{460}} = 1.35$$

$$RMSE_{\left(\frac{\text{ANN}}{\text{GA}}\right)\text{whole data}} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} = \sqrt{\sum_{1}^{460} \frac{(y_i - \hat{y}_i)^2}{460}} = 0.014$$

$$RMSE_{\left(\frac{\text{ANN}}{\text{GA}}\right)\text{developing countries data}} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} = \sqrt{\sum_{1}^{230} \frac{(y_i - \hat{y}_i)^2}{230}} = 0.12$$

$$RMSE_{\left(\frac{\text{ANN}}{\text{GA}}\right)\text{developed countries data}} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} = \sqrt{\sum_{1}^{230} \frac{(y_i - \hat{y}_i)^2}{230}} = 0.25$$

Appendix D: The Panel Unit Root Tests

Test	LLC	B- t stat	IPS	Fisher-ADF	Fisher-PP
H_0	Unit root	Unit root	Unit root	Unit root	Unit root
H_1	No Unit root	No Unit root	Some cross sections without Unit root	Some cross sections without Unit root	Some cross sections without Unit root
Component of each method	No exogenous variables-Fixed effect- and Individual effect and individual trend.	No exogenous variables-Fixed effect- and Individual effect and individual trend.	Fixed effect- and Individual effect and individual trend.	No exogenous variables-Fixed effect- and Individual effect and individual trend.	No exogenous variables-Fixed effect- and Individual effect and individual trend.

Source: Univariate time series analysis -Eviews tutorial file by Eviews official website (http://www.eviews.com)