EFFECT OF DIFFERENT LEVELS OF EDUCATION ON ECONOMIC DEVELOPMENT IN TURKEY: A PANEL ANALYSIS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MÜŞERREF HÜSAMOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF ECONOMICS

DECEMBER 2008

Approval of the Graduate School of Social Sciences

Prof. Dr. Sencer Ayata
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science

Prof. Dr. Erol Taymaz
Head of Department

This is to 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 Master of Science

Prof. Dr. Aysıt Tansel
Supervisor

Examining Committee Members

Assist. Prof. Dr. Dürdane Şirin Saraçoğlu (METU, Econ)

Assoc. Prof. Dr. Sibel Kalaycıoğlu (METU, Sociology)

Prof. Dr. Aysıt Tansel (METU, Econ)

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Müşerref Hüsamoğlu

Signature :

ABSTRACT

EFFECT OF DIFFERENT LEVELS OF EDUCATION ON ECONOMIC DEVELOPMENT IN TURKEY: A PANEL ANALYSIS

HÜSAMOĞLU, Müşerref
M.S., Department of Economics
Supervisor : Prof. Dr. Aysıt Tansel

December 2008, 153 pages

In this study, I aimed to examine the impact of different levels of education on real GDP (and real GDP per workforce) in Turkey, and hence the relationship between different levels of education and the standard of living is estimated by panel data techniques. The panel data set in the study is constructed by pooling 67 provinces of Turkey over the period of 1975-2000. Furthermore, in the empirical work, two models are employed: the model introduced by Knowles (1997) and the augmented Solow model with different levels of education. The panel data estimation of the Knowles's model implies that the secondary level of schooling has the greatest contribution to real GDP, while the augmented Solow model implies that the higher level of schooling has the largest impact on real GDP per workforce.

Keywords: Different Levels of Education, Development, Panel Data, Turkey

ÖZ

FARKLI EĞİTİM DÜZEYLERİNİN TÜRKİYE'DE EKONOMİK KALKINMAYA ETKİSİ: PANEL ANALİZİ

HÜSAMOĞLU, Müşerref
Yüksek Lisans, İktisat Bölümü
Tez Yöneticisi : Prof. Dr. Aysıt Tansel

Aralık 2008, 153 sayfa

Bu çalışmada, farklı eğitim düzeylerinin Türkiye'deki reel GSYİH (ve işgücü başına düşen reel GSYİH) üzerindeki etkisini ölçmek amaçlanmıştır ve bu nedenle farklı eğitim düzeyleri ile yaşam standartları arasındaki ilişki panel veri yöntemleri ile tahmin edilmiştir. Çalışmada kullanılan panel veri seti Türkiye'deki 67 ilin 1975-2000 dönemi için bir araya getirilmesi ile elde edilmiştir. Ayrıca, ampirik çalışmada iki model kullanılmıştır: Knowles (1997) tarafından ortaya konulan model ve farklı eğitim düzeyleri ile genişletilmiş Solow modeli. Knowles modelinin panel veri tahmini ortaöğretim düzeyindeki eğitimin reel GSYİH'a en fazla katkı sağladığını ortaya koyarken, genişletilmiş Solow modeli yükseköğretimin işgücü başına reel GSYİH üzerinde en fazla etkiye sahip olduğunu ortaya koymaktadır.

Anahtar Kelimeler: Farklı Eğitim Düzeyleri, Kalkınma, Panel Veri Seti, Türkiye

To My Mum

ACKNOWLEDGMENTS

I would like to express my thanks to my supervisor Prof. Dr. Aysıt Tansel for suggesting the topic of this thesis and for her encouragement and support through the preparation of the thesis. I also wish to thank the members of my examining committee, Assist. Prof. Dr. Dürdane Şirin Saraçoğlu and Assoc. Prof. Dr. Sibel Kalaycıoğlu for their valuable suggestions and comments.

I wish to express my appreciation to my head of the Department of Social Policy in the State Planning Organization İlyas Çelikoğlu, and my colleges especially Sırma Demir Şeker and Sinem Çapar for their great support and encouragement. Additionally, I desire to thank the Scientific and Technological Research Council of Turkey for their financial support for my M.S. degree.

I should also mention the encouragement, understanding and support of my family, Murat Küçükbayrak, Ruken Küçükbayrak and Murat Özer during the preparation of my thesis. It would have been very difficult to complete this thesis without them.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The determination of the determinants of output level and the sources of economic growth is one of the significant issues constituting the core of the economic growth models. The supporters of neoclassical growth models assuming decreasing returns to factors of production claim that the economic growth of a country in the long run is solely determined by the exogenous technological progress, and that a country with lower per capita output grows faster than the ones with relatively higher per capita output so that the faster growing country is able to catch up with the others after some time. However, the empirical findings generally do not support the neoclassical claims about the long run behavior of the economies. Therefore, the inadequacy of the neoclassical models in explaining the economic growth has led the economists to seek for more convincing models in determining the sources of aggregate output and growth. A welcome extension of the neoclassical growth theory is the study of Mankiw et al. (1992) in which the theory is modified with the inclusion of human capital. The augmented model introduced by Mankiw et al. concludes that the output level of a country could be explained better with the inclusion of human capital.

The recent growth theory suggests that the human capital, which could be viewed as the stock of knowledge used in the production process is an important determinant of the aggregate output and growth in an economy. In fact, the concept of human capital plays a central role in endogenous growth theory. In endogenous growth theory, human capital, when used with the other classical factors of production such as physical capital and labor, leads to economic growth by means of innovation of new technology, and imitation and adoption of technologies abroad. In addition, accumulation of human capital produces spillover effects. That is, not only the individual benefits from the knowledge gained, but also the whole society benefits from this knowledge through diffusion of knowledge across sectors and industries (Cheng and Hsu, 1997). The knowledge spillover causes increasing returns and hence it could provide a positive growth for an economy in the long run without technological improvement.

Human capital is defined by Organization for Economic Cooperation and Development (OECD) as “the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being” (OECD, 2001:18). Those skills and knowledge could be improved through education, learning by doing or on the job training. Nevertheless, human capital is a more complex input that comprises more than knowledge capital (McDonald and Roberts, 2002). Indeed, the attributes and competencies of individuals producing economic value could be affected by health conditions of those individuals; and hence accumulation of human capital could also be provided by medical care or nutrition support for the ones involved in the production process. Therefore, the human capital in growth theory is viewed as being composed of two main components: education and health. Yet, many studies on output level and economic growth, like the present study, are mainly focused on the education capital as a proxy for human capital due to lack of data. In this study, I provide evidence on the impact of education on real gross domestic product (GDP) and real GDP per workforce in Turkey.

The impact of education on output level and economic growth differs in various empirical studies. In fact, there are contradictory empirical results about the effect of education on growth or development. Some studies find insignificant or negative effect of education capital (e.g. Islam, 1995; Temple, 2001), while many studies emphasize the importance of education in aggregate production and growth (e.g. Oketch 2006; McMahon, 1998). The letter studies provide evidence on the positive impact of education on level of output and growth, however it is not clear which level of education positively affects output level (Gyimah-Brempong et al., 2006). Does educational attainment at different levels affect output positively? If so, which level has the greatest impact on output level? Providing answers to those questions in the Turkish context is the objective of this study. Hence, the present study aims to examine the impacts of different levels of education, which is proxied by the educational attainment of the workforce, on real GDP and real GDP per workforce in Turkey. In this regard, the educational attainment of the labor force is differentiated into four levels: labor force without formal schooling, and with schooling at the basic, secondary and higher levels.

To the best of our knowledge, an academic study regarding the relationship between different levels of education and output level for Turkey has not been conducted before. Furthermore, panel data is utilized in investigating the relationship between different levels of education and output level in Turkey. Hence, the main contributions to the empirical literature of the present study will be the usage of panel data and the differentiation of education capital in different levels. In that respect, based on the assumption of Cobb-Douglas production function, two models regarding the effects of differentiated education capital on real GDP and real GDP per workforce in Turkey are estimated with panel data in this thesis. One is the model introduced by Knowles (1997) and the other is the augmented Solow model with differentiated education capital produced.[1]

In the second chapter of this thesis, the literature regarding the relationship between human capital, especially in the form of education, and the output level and economic growth are reviewed. First, the neoclassical and endogenous growth theories are presented; and the role of human capital in those models are outlined. Moreover, in this chapter the models in literature examining the effect of different

[1] The augmented Solow model used in this study is an extension of the model introduced by Islam (1995).

education levels on output level and growth, and the models using panel data within growth theories are presented.

In Chapter 3, proxies for the variables commonly used in empirical studies, which investigate the relationship between human capital and output or economic growth, are mentioned. As most of the models in growth literature are generally derived from production functions, the variables used in the growth models are labor, human capital and physical capital variables.

In the fourth chapter, the recent developments in the education sector in Turkey are presented. The structure of education in Turkey are discussed by using indicators such as the enrollment ratios, average years of education of the workforce, the number of schools, students and teachers, the resources allocated to education and the literacy rates.

In the fifth chapter, the estimation methods used in this thesis are presented. Since the models in the present study are estimated by ordinary least squares (OLS) and panel data estimation techniques; these techniques and their advantages and disadvantages are discussed in Chapter 5. Moreover, in order to determine which of the methods is more appropriate for the models estimated test statistics are presented.

In Chapter 6, data source and the variables used in the models estimated in this study are discussed. In the present study, the panel data is available at the province level for the years 1975, 1980, 1985, 1990 and 2000. The estimation results that are obtained from the OLS and fixed and random effect estimations are presented in Chapter 7. In addition, the results of the hypothesis tests for choosing the best method of estimation for each model are provided. Additional tests such as for heteroscedasticity, autocorrelation and omitted variables are performed for the most appropriate method of estimation of each model.

The main findings of the thesis are given in Chapter 7. In addition, the suggestions about the educational system of Turkey based on the main findings are also included in this chapter.

CHAPTER 2

REVIEW OF LITERATURE

In this chapter, the literature regarding the relationship between human capital and output level (and economic growth) will be discussed. In this respect, the theories of neoclassical growth models, namely the Solow, Augmented Solow and Ramsey-Cass-Koopmans models and endogenous growth models will be presented. Then, since the models in this study aim to examine the impacts of different levels of education on real GDP (and real GDP per capita) within the panel data framework, the models in literature that use panel data techniques and those that differentiate between different education capitals will be discussed.

2.1. Theories of Economic Growth and the Role of Human Capital

One of the main objectives of empirical studies in growth models is to determine the determinants of output level and the sources of economic growth. The earlier studies focus on the classical factors of production such as physical capital and labor. On the other hand, the later studies explore the additional factors such as human capital in determining the determinants of output level and the sources of economic growth. For instance, Lau, Jamison and Louat (1991) investigate the effect of human capital, in the form of education, on aggregate real output and productivity; and they conclude that education capital is a significant determinant of output level for developing countries. In their empirical study, they use three different proxies for the education capital: primary education, secondary education and total education. They find that primary education does not have a significant effect on output level, while the secondary education's effect on output is significantly positive. The empirical findings of Barro (2001) is similar to those obtained by Lau et al. in the sense that the educational attainment of adult males at the primary level does not have statistically significant impact on the growth rate of output per capita. However, the secondary and higher schooling of males is found to affect the growth of per capita output significantly. In his study, Barro examines the role of qualitative factors of education in the determination of economic growth as well as the quantitative factors. He concludes that the quality of education, measured by test scores of international examinations, is more important than the quantity of schooling, represented by average years of attainment of adult males, in determining the output growth.

Moreover, there are many other economists emphasizing the importance of human capital formation in determining the determinants of output level and the sources of economic growth such as Schultz (1960; 1961), Temple (1999), Barro and Lee (1994), Oketch (2006), Hanushek and Kimko (2000), Cabelle and Santos (1993), McMahon (1998), Bassanini and Scarpetta (2002). Some studies suggest that the impact of human capital, in the form of education, is negative or insignificant. For instance, Temple (2001) finds that the increases in educational attainment have done little to raise the growth of output in less developed countries. However, the indirect effects of education capital through total factor productivity are excluded in his model. This could be the reason as to why education has little impact on output growth. Studies including human capital as a factor of production do not produce a certain result regarding the relationship between human capital and output level and economic growth.

In this section, the literature regarding the impact of human capital is presented. Firstly neoclassical growth theories, ignoring the role of human capital, will be discussed. Then, endogenous growth theories, emphasizing the importance of human capital, will be presented.

2.1.1. Neoclassical Growth Models

Neoclassical growth models are based on the assumption of neoclassical production function having the properties of constant returns to scale, positive and diminishing returns to physical capital and satisfying the Inada conditions.[2] In addition to the neoclassical production function, an exogenous technological progress is assumed in those models, that is, the growth rate of technology is assumed to be given as a constant. Under those assumptions, the supporters of neoclassical theory contend that the only way for economies to grow at a positive rate in the long run is the technological progress. Further, a country having initially low output per capita grows faster than the ones with initially high per capita output. That is, the neoclassical economists support the absolute convergence when explaining the cross country relationship on growth.[3] However, the main inadequacy of those models, except for the augmented Solow model, is the absence of "human capital" which is the primary focus of the endogenous growth theory. In order to have some idea about the neoclassical growth models, in the following sub sections, the general structure of such models will be examined under the Solow growth model, which takes the saving rate as exogenous; the augmented Solow model which is an extension of the Solow's model; and the Ramsey-Cass-Koopmans model assuming an endogenous saving rate.

2.1.1.1. Solow and Augmented Solow Growth Models

Most of the economic models developed under the neoclassical theory are based on the model constructed by Solow (1956). In his article, Solow aims to see the long run behavior of an economy and to do so he examines the behavior of both of the physical capital and the labor by using a standard neoclassical production function.

---

[2] The property of the marginal product of capital (labor) approaching infinity as capital (labor) goes to zero and the marginal product of capital (labor) approaching zero as capital (labor) goes to infinity are known as the Inada conditions.

[3] For a detailed discussion for the concept of convergence see Appendix C.

Solow finds that the behavior of the economy in the long run is determined by the saving rate, the rate of population growth and the exact shape of the production function. The theory of Solow depends on some basic assumptions. He assumes that the rate of saving is exogenous, the amount of labor is being supplied inelastically and the growth rate of labor force is constant through time. Moreover, he ignores the technological progress and the depreciation of physical capital by taking them equal to zero.

In the article of Solow (1956), unfortunately, he does not include human capital as a factor determining the real output of an economy. Subsequent authors write about the extended Solow model including physical and human capital together such as Barro and Sala-i Martin (2004). Barro and Sala-i Martin examine the Solow-Swan model with respect to the dynamics of economic growth. In the light of the empirical results, they find that a reasonable, in fact an observed, speed of convergence requires a relatively higher share of total capital in output than expected when only physical capital is considered.[4] For instance, for a speed of convergence of 2 percent per year, the neoclassical model requires a share of 0.75 for the capital input. A capital share of 0.75, however, is too narrow to include only the physical capital. They conclude that it could be reasonable to expand the concept of capital to include the human component as well. Hence, in addition to investigating the Solow-Swan model; Barro and Sala-i Martin examine the growth dynamics of economies using the Solow model augmented to include the human capital. They find that the inclusion of human capital into the production function produces a convergence rate which is empirically meaningful.

Mankiw, Romer and Weil (1992) also discuss the importance of human capital within the Solow model framework.[5] In the empirical study, Mankiw et al. estimate

---

[4] The speed of convergence is the rate at which an economy converges to its steady state.

[5] One of the models estimated in this study is the extension of the augmented Solow model introduced by Islam (1995). The augmented model of Islam is derived from the model constructed by Mankiw,

first the textbook Solow model and then they extend the model to include human capital as well as physical capital. In the textbook Solow model, they start with the Cobb-Douglas production function of the form $Y(t) = K(t)^{\alpha}(A(t)L(t))^{1-\alpha}$, where Y is the output level, K is the stock of physical capital, L is the amount of labor and A denotes the level of technology. They assume that the level of technology and the amount of labor grow at the constant rates g and n respectively. They estimate the following equation by OLS within the Solow model,

$$\text{In}(Y/L) = a + \frac{\alpha}{1-\alpha}\text{In}(s) - \frac{\alpha}{1-\alpha}\text{In}(n+g+\delta) + \varepsilon \qquad (2.1.1)$$

where s is the saving rate and δ is the rate of depreciation of physical capital.[6] In addition, they measure n as the average rate of growth of the working age population (15-64 aged). s is measured by the average share of real investment in real GDP and Y/L is measured by the real GDP divided by working age population. They conclude from the OLS estimation of the equation (2.1.1) that the differences in saving and population growth explain a large fraction of the cross country variation in output per capita within the Solow model framework. Moreover, the estimates of the effects of the saving rate and the population growth on the real GDP per capita is found to have the predicted signs with high significance; a higher saving rate and a lower population growth rate yield a higher output level. However, the coefficients are much larger than what Solow predicts when human capital is excluded.

The Cobb-Douglas production function used in the augmented Solow model is $Y(t) = K(t)^{\alpha}H(t)^{\beta}(A(t)L(t))^{1-\alpha-\beta}$, where all the variables in this function is the same as above and H is defined as the stock of human capital. In that case, Mankiw et al. (1992) assume a common depreciation rate of both physical and human capital. They

---

Romer and Weil (1992). Hence, the augmented Solow model of Mankiw et al. is discussed in this section in detail.

[6] They assume that $g+\delta$ is 0.05.

employ OLS to estimate the equation, which is directly obtained from the Cobb-Douglas production function, of the form

$$\text{In}(Y/L) = \text{In}A(0) + gt + \frac{\alpha}{1-\alpha}\text{In}(s_k) - \frac{\alpha}{1-\alpha}\text{In}(n+g+\delta) + \frac{\beta}{1-\alpha}\text{In}(h^*) \tag{2.1.2}$$

where A(0) is the initial level of technology, s_k represents the share of real physical capital investment in real GDP, and h* is the steady state level of human capital per effective worker. In the augmented Solow model, they use the percentage of the population in the secondary school to measure the effect of human capital. They find that in the augmented Solow model, the differences in saving rate, population growth and human capital explain a larger fraction of the cross country variation in output per capita than the Solow model. Including human capital in the Solow model yields more plausible estimates for the impacts of the saving and population growth rates.

In their study, the convergence picture of the countries used in the empirical model is examined with and without controlling for investment, growth of working age population and school enrollment. In the case of the absence of any condition, the countries with initially lower levels of output per capita are not found to have a tendency to grow faster than the ones with higher levels of per capita output. On the other hand, there is a strong evidence for the conditional convergence. That is, a country would have a tendency to grow faster if it has initially lower output per capita, when the savings, population growth and human capital are each assumed to be equal among the countries of interest.

Grammy and Assane (1996) improve the results obtained from the augmented Solow model introduced by Mankiw et al. (1992) by using broader measures of human capital.[7] They find that the estimated coefficients of saving and population growth

[7] In the empirical study, they employ two measures of human capital: the United Nation's Human Development Index (HDI) and Economic Liberty Index (EDI). The HDI is constructed as an unweighted average of relative distances measured in longevity, educational attainment and access to

rates become smaller when broader measures of human capital are employed; and obtain an evidence of conditional convergence at a faster rate than predicted by Mankiw et al.

Contrary to the economists supporting the importance of human capital within the augmented Solow model, Hamilton and Monteagudo (1998) emphasize that investment in physical capital seems to be more important than investment in human capital for economic growth. Their empirical study is based on the model constructed by Mankiw et al. (1992); however, their model takes the change in the average annual growth rate of output per worker between the periods 1960-1970 and 1975-1985 as the dependent variable in the regression analyses. Hamilton and Monteagudo find that the change in the share of physical investment in real GDP between the periods 1960-1970 and 1975-1985 has significantly positive effect on the change in the output growth over the same periods. However, the change in the fraction of resources devoted to education is related negatively to the output growth rate.[8]

2.1.1.2. Ramsey-Cass-Koopmans Model

In Solow and augmented Solow models, it is not allowed for households and firms to behave optimally. However, as Barro and Sala-i Martin (2004) state that by not allowing consumers and producers to behave in an optimal way, the growth analysis does not permit us to examine how the incentives affect economies. Moreover, it is difficult to analyze how economy reacts to the changes in variables affecting the level of income such as tax rates without allowing consumers to behave optimally. So, the models of economic growth, depending on optimal behavior of individuals

---

resources. The ELI is constructed as summary indexes which are based on fifteen features of economic liberty such as freedom of information.

[8] The conclusion about the effect of human capital investment on the change of the rate of output growth in their study based on the fact that they use the percentage of the population in the secondary school to proxy for human capital, so the result in fact should reflects the effect of the secondary schooling. Thus, this deduction may be misleading in the sense that the proxy for human capital used is not perfect.

units, constitute an important part of the neoclassical growth theory. The logic behind the "optimality" is to allow consumers to maximize their lifetime utility subject to an intertemporal budget constraint together with the optimal behavior of the firms. This specification of consumer behavior is developed by Ramsey (1928), Cass (1965) and Koopmans (1965). Nevertheless, as in the case of Solow model, unfortunately, the optimal theory introduced by Ramsey, Cass and Koopmans does not include the human capital as a factor affecting output level and growth. On the other hand, in this section, the optimal theory will be presented in order to show the complete picture of neoclassical growth models.

Ramsey (1928) is the first economist who discusses the idea of optimal saving. Cass (1965) and Koopmans (1965) develop the idea proposed by Ramsey. Ramsey is mainly concerned, under some appropriate assumptions, how much a nation should save in order to reach or approach bliss, the maximum obtainable rate of utility, after some time. In this study, he assumes that the number of people in an economy is constant and there are diminishing returns for physical capital and labor. With those simplifying assumptions, he finds that the optimal saving rate should satisfy the condition that when it is multiplied by the marginal utility of consumption, it should always equal the maximum possible rate of utility minus the actual rate of utility.

The problem of optimal saving in a simplified economy, closed and centralized, is also discussed by Cass (1965). In his paper, the social welfare is represented by the total discounted utility of consumption per capita, and it is maximized in the absence of technological progress. The maximization process yields a unique optimum growth path. He concludes that if an economy is initially on the optimal growth path, it would finally reach the point where the maximum possible consumption level is attained. Moreover, both the consumption per capita and physical capital per capita is increasing (decreasing) on this unique growth path whenever initial physical capital per capita is below (above) the optimal one. However, with regard to the behavior of the saving rate, he find that its behavior on the optimal growth path is ambiguous even with the simplified economy described, and the behavior of saving rate only depends on the particular parameters chosen on the model of interest. In addition, Cass concludes that when an economy reaches the level where optimal consumption is attained, it stays there forever.

Like Cass (1965), the optimal behavior of economies is also discussed by Koopmans (1965). He examines the conditions which are required for the existence and the uniqueness of an optimal growth path with a one sector model having constant level of technology and steadily increasing labor force. By doing so he makes some significant assumptions, having the power to change the main conclusions he makes, such as the absence of depreciation of physical capital and diminishing returns to physical capital and labor. Even though the existence of those difficulties arising from his assumptions; Koopmans is able to construct the optimal growth path for consumption per worker and physical capital per worker.

2.1.2. Endogenous Growth Models

In neoclassical models, the only way to obtain a positive growth in the long run is the technological progress, and without such an improvement an economy converges to its steady state with zero per capita growth. The reason behind the convergence in those models is the assumption of diminishing returns to factors of production. Contrary to the neoclassical models, in endogenous growth models, it is possible to obtain a positive growth without technological development, even in the long run, due to the absence of diminishing returns. One way to eliminate diminishing returns is to insert the "human" component into the concept of capital (Barro and Sala-i Martin, 2004). The endogenous growth theory, generated from within a system as a result of internal processes, emphasizes the improvement of human capital leading to economic growth through the development of new forms of technology and efficient and effective means of production. On the other hand, these kinds of models are not consistent with the empirical evidence on convergence as opposed to the neoclassical growth models, that is, the empirical results in endogenous growth models do not support for absolute convergence.

Romer (1989) is one of the economists investigating the role of human capital within the framework of endogenous growth theory. In the study, he explains both the theoretical and the empirical dimensions of endogenous growth by considering how do knowledge and science affect production. In the empirical part, the literacy rate of the population is taken as a proxy for human capital. For a cross section of countries during the period of 1960-1985, Romer uses two methods of estimation: OLS and the instrumental variable estimations. The OLS estimation implies that the initial level of output affects the growth rate negatively while the initial literacy rate affects positively. However, the literacy rate is not significant. Moreover, the instrumental variable estimation does not yield a significant estimate for the effect of initial literacy rate when the share of GDP devoted to investment is taken as an explanatory variable in the regression equation. On the other hand, the exclusion of physical investment from the regression equation causes the literacy rate to become significant. Regarding this situation, he concludes that the literacy rate has no additional explanatory power in the growth regression; nevertheless, it helps to explain the rate of investment and hence it impacts the rate of output growth indirectly.

Benhabib and Spiegel's (1994) study examines the effect of human capital on output growth by using cross country estimates of physical and human capital stocks. Contrary to the results emphasizing the role of human capital on the determination of output growth, they find that human capital has an insignificant effect on growth. However, as an alternative they construct a model where the total factor productivity growth rate depends on human capital. This specification yields a positive effect for human capital. The latter model is based on the idea that human capital could affect the growth of technological progress through enhancing the ability of a country to develop its own technological innovations and the ability to adapt and implement technologies developed abroad; and hence it could affect output growth indirectly through total factor productivity growth.

Papageorgiou (2003) takes the model of Benhabib and Spiegel (1994) as the starting point in order to examine the effect on economic growth of human capital accumulation. He improves their study in two respects. Papageorgiou firstly assumes that human capital affects growth not only through the improvement of technological innovations but also through output production. He differentiates the human capital as being proxied by primary and post-primary (secondary and tertiary) education. Papageorgiou concludes that the structural specifications allowing human capital to operate as a facilitator of technological progress are more successful in explaining the growth rather than the standard growth accounting specification. He finds the primary education contributes mainly to the production of final output, whereas the post-primary education contributes mainly to adoption and innovation of technology.

McMahon (1998) provides empirical evidence for the importance of the human capital as a central determinant of the growth process within the endogenous growth theory. In this study, the effects of primary, secondary and higher education, in the forms of enrollment rates and of public investment, on per capita GDP growth in East Asia are examined; and it is concluded that the per capita growth in this region could be determined partially by investment in human capital.

Oketch (2006) also examines the determinants of economic growth in African countries within the endogenous theory framework. In the empirical study, the effect of human capital development produced by formal schooling, on the economic growth is estimated. It is concluded that investment in human capital as well as physical capital could be interpreted as causal factors contributing per capita growth. Hanushek and Kimko (2000) find similar results regarding the impact of human capital in the form of education. They find that the quality of schooling has a central role in determining economic growth.[9] They conclude that the quality of schooling has a causal impact on output growth. Cheng and Hsu (1997) also investigate the causal effect of human capital on growth in Japan by using a time series data over the period 1952-1993; and they find a strong evidence for the bidirectional causality between human capital and economic growth.

Barro (1991), in his study, examines the relationship between the growth of real GDP per capita and the initial level of it in order to look for convergence across countries. He finds that given the initial level of human capital, there is a negative correlation between initial level of output per capita and growth rate of output implying that the poor countries are able to catch up with the rich ones. In addition, OLS estimation of this model implies that the correlation between initial level of human capital and growth rate of real GDP per capita is positive so that the growth rate is positively related to initial level of human capital given the starting amount of real GDP per capita.[10] In a later study, Barro and Lee (1994) find similar impacts of initial values of real GDP per capita and human capital on growth. In this study, they try to determine the sources of economic growth which systematically differs across countries of different development levels. One of the determinants described in the study is the conditional convergence effect, that is, a country will have a tendency to grow faster if the initial per capita output is lower relative to its initial level of human capital.

Cabelle and Santos (1993) examine endogenous growth, including both physical and human capital as factors of production, in an optimal growth model framework. They deal with an economy in which agents may devote a part of their nonleisure time to going to school in order to increase the productivity of labor being supplied. Cabelle and Santos investigate the dynamics of such an economy within the standard optimal

[9] The quality of schooling used in the empirical study is proxied by the international test scores of mathematics and science.

[10] In this study, the convergence is obtained when it is conditioned on human capital.

growth model and find the necessary and sufficient conditions for an optimal balanced growth path to exist. In their study, Cabelle and Santos conclude that if the ratio of physical capital to human capital is initially low, the economy would accumulate physical capital, otherwise; it would decumulate it. This conclusion places human capital as a key factor since the relative availability of human capital determines the accumulation of physical capital.

2.2. The Panel Data Approach to Growth Models Including Human Capital

In this thesis, in order to examine the relationship between human capital in the form of education and real GDP (and real GDP per workforce) in Turkey, the empirical models being constructed will be estimated mainly with the panel data techniques. Therefore, in this section the studies investigating this relation within the panel data framework are presented. On the other hand, the panel data approach regarding the relation between output level or growth and human capital is not so common in the literature. Islam (1995) is one of the first researchers examining the usual relation with a dynamic panel data model. The extension of Islam's model including different levels of education is estimated in this study. Hence, the study of Islam will be discussed in a more detailed way than the other studies mentioned in this section.

Islam (1995) looks for the differences of the results obtained from the estimation of single cross section regressions, specifically the model formed by Mankiw et al. (1992), and from panel data regressions regarding the relationship between factors of production, especially human capital, and per capita output. He finds that the estimated conditional convergence rate becomes higher and the estimated elasticity of output with respect to physical capital becomes lower in a panel data framework due to the method of controlling the individual country effects in such a model. Contrary to the studies using either OLS estimation or the pooled regression, the model including human capital variable as a determinant of the income per capita does not yield different results compared to the model without human capital in his panel data model. The reason behind such similarities is attributed to incorporating the temporal dimension of human capital variables into growth regressions.

In the empirical part of his article, Islam (1995) assumes a Cobb-Douglas production function with labor-augmenting technological progress. The production function is of the form $Y(t) = K(t)^{\alpha}(A(t)L(t))^{1-\alpha}$. He uses single and pooled OLS estimations in addition to the estimation of panel data in order to be able to compare the results previously obtained, especially those of Mankiw et al. (1992). Moreover, the data he employs is the same as that used by Mankiw et al. Islam estimates the following equation by OLS

$$\mathrm{Iny}(t_2)=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\mathrm{In}(s)-(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\mathrm{In}(n+g+\delta)+e^{-\lambda\tau}\,\mathrm{Iny}(t_1)+(1-e^{-\lambda\tau})\mathrm{InA}(0)$$
$$+g(t_2-e^{-\lambda\tau}t_1) \qquad (2.2.1)$$

where y represents the per capita output, s, n, g and δ are the saving rate, the population growth rate, the rate of technological progress and the depreciation rate of physical capital respectively. α is the share of physical capital in total output, λ equals to $(n + g + \delta)(1-\alpha)$, and τ is $t_2 - t_1$, where t_1 and t_2 are 1960 and 1985 respectively. In the equation (2.2.1), the term A(0) reflects not only the level of technology, but also represents resource endowments, intuitions, climate and so on; and hence, it could differ across countries. This fact is reflected with the equality $\mathrm{InA}(0) = a + \varepsilon$, where ε is the country specific shift term, and a is a constant. However, since the OLS regression is not able to differentiate the effect of the country specific variables contained in InA(0); the term $(1- e^{-\lambda\tau})\mathrm{InA}(0)$ is included in the disturbance term in the OLS estimation. Moreover, the term $g(t_2 - e^{-\lambda\tau}t_1)$ reflects the time specific effects and due to the same reasoning that the OLS estimation could not differentiate the time specific effects, it is also included in the error term of the OLS regression.

Islam (1995) compares the results of OLS estimation of the equation (2.2.1) with the results obtained by Mankiw et al. (1992) from the OLS estimation of

$$In(y(t))\text{-}In(y(0))=(1\text{-}e^{-\lambda t})\frac{\alpha}{1-\alpha}In(s)\text{-}(1\text{-}e^{-\lambda t})\frac{\alpha}{1-\alpha}In(n+g+\delta)\text{–}(1\text{-}e^{-\lambda t})In(y(0)) \qquad (2.2.2)$$

where y, s, λ, α, n, g and δ are the same as above and t represents the year 1985. Moreover, the y(0) represents the GDP per working age population in 1960. The OLS estimation, based on (2.2.1), yields estimators for the initial output level variable that are very close to those obtained by Mankiw et al. from the equation (2.2.2).[11]

After estimating the OLS regression, the total time period (1960-1985) is divided into five equal time intervals so as to make the pooled OLS and panel estimations possible. The equation estimated with the pooled OLS is given in the form of

$$y_{i,t}=\gamma y_{i,t-1}+\sum_{j=1}^{2}\beta_j x_{it}^{j}+v_{it} \qquad (2.2.3)$$

where x_{it}^{1} and x_{it}^{2} are In(s) and In(n+g+δ), respectively. Islam finds that the estimated coefficients from (2.2.3) do not produce very different results than single cross section estimation of the equation (2.2.1). The main difference is that the estimated coefficient for the term In(s) obtained by single OLS is much higher than that obtained by pooled estimation.

[11]The difference of the regression equations estimated with OLS by Islam and Mankiw at al. is the dependent variable. In fact, Islam uses Iny(t) as the dependent variable, while Mankiw et al. use Iny(t)-Iny(0) as the dependent variable. However, to make a plausible comparison Islam also estimates the equation (2.2.1) by taking the dependent variable as the log difference in GDP per capita for between the years 1960 and 1985.

In the panel data estimation, Islam uses the methods of least squares with dummy variables (LSDV) and Minimum Distance (MD) estimators. The regression equation estimated within the panel data framework is obtained from the following equation

$$\mathrm{Iny}(t_2)=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\mathrm{In}(s)-(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\mathrm{In}(n+g+\delta)+e^{-\lambda\tau}\mathrm{Iny}(t_1)+(1-e^{-\lambda\tau})\mathrm{InA}(0)$$
$$+g(t_2-e^{-\lambda\tau}t_1) \qquad (2.2.4)$$

Thus, the equation (2.2.4) represents a dynamic panel data model in the form $y_{i,t} = \gamma y_{i,t-1}+\sum_{j=1}^{2}\beta_j x_{it}^{j}+\eta_t+\mu_i+v_{it}$, where $y_{i,t}$=lny(t_2), $y_{i,t-1}$=lny(t_1), $\gamma=e^{-\lambda\tau}$, $\beta_1=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}$, $\beta_2=-(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}$, x_{it}^{1}=In(s), x_{it}^{2}=In(n+g+δ), $\mu_i=(1-e^{-\lambda\tau})$InA(0) and $\eta_t=g(t_2-e^{-\lambda\tau}t_1)$. The resulting estimated coefficients from the two methods of LSDV and MD are very similar to each other, but are different from those obtained by either single or pooled OLS. The panel data estimation results in higher rates of convergence, which is represented by λ, and more plausible estimates of the elasticity of output with respect to capital, α, when compared to the single cross section and pooled OLS estimation even without including human capital.

After the inclusion of human capital as an explanatory variable to the model, the same estimations presented above are performed again. The inclusion of human capital in the OLS regression results in higher λ and lower α; however, it yields lower λ and higher α for the pooled OLS estimation. On the other hand, regarding the estimated share of physical capital and the rate of convergence, the panel data estimation with human capital produces similar results compared to those obtained without human capital variable. In addition, the estimation of the same model including human capital, which are proxied by average years of schooling, results in negative impact of schooling when panel data estimation methods are used. Yet, this consequence is attributed to the indirect effect of human capital through technology.[12]

Dessus (2001) also analyzes the relationship between human capital, in the form of education, and growth within the panel data framework; and finds that, even using quality of schooling variables as proxies for education, human capital contributes negatively to growth as Islam (1995) finds. On the other hand, this result is attributed to the heterogeneity of the slope coefficients for human capital. A significant positive impact of human capital is obtained by relaxing the assumption of homogeneity.

McDonald and Roberts (2002) examine the effects on the output level of human capital within the augmented Solow model framework. They develop a model incorporating both health and education capital, as a proxy for human capital, in a dynamic panel data framework. So, the constructed model is different from the Islam's (1995) in the sense that they use a more comprehensive human capital variable. In addition, McDonald and Roberts emphasize the importance of human capital in the form of both education and health. They conclude that the role of different forms of human capital changes as the level of income changes. Actually, in determining output level of a country, health capital seems to be more important at low income levels whereas the education capital seems to be more important at high income levels.[13]

Gyimah-Brempong et al. (2006) investigate the impact of human capital, in the form of higher education, on the per capita output growth in African countries within the

[12] Evidence regarding the indirect effect of human capital through technology is found in his study. It is the significantly positive correlation between the level of technology and human capital variable that generates this indirect effect.

[13] In an earlier study McDonald and Roberts (1999) test the restriction imposed in cross section studies such as the assumption of common initial technologies across countries. They find that unlike this assumption, there are systematic differences across countries for the data and samples they use. So, they conclude that the assumption of common initial technologies across countries, which is required to implement cross section estimation methods, is unlikely to be valid and hence they propose the panel data approaches in model estimations.

augmented Solow model introduced by Mankiw et al. (1992). To do so, they use a panel over the period of 1960-2000 for 34 African countries; and it is concluded that the effect of the higher education on output growth is significantly positive and it is twice as large as the impact of physical capital investment.

Bassanini and Scarpetta (2001) examine the long run relationship between the human capital, being proxied by the average years of schooling of the population aged 25-64, and output by using pooled mean-group estimation method for a panel of 21 OECD countries in the period 1971-1998. They find that the long run estimated elasticity of output per working age population to human capital is significantly positive. In addition, Middendorf (2005) also investigates the impact of human capital on the growth rate of per capita GDP in 29 OECD countries over the period 1965-2000. He uses the average years of schooling and secondary school attainment of the population aged 25 and over to proxy for human capital. He concludes that the human capital has significant impact, which is positive, when the fixed effect estimation is employed.

Güngör (1997) estimates the relationship between human capital in the form of education and industrial output for the period of 1980-1990 in Turkey within the panel data framework. As in the present study, the panel data used in her study is constructed by pooling the data for 67 provinces of Turkey. On the other hand, contrary to the present study, the output as a dependent variable is restricted to those in industrial sectors. The average years of schooling of the employed workforce in those sectors is used to proxy for education capital in Turkey. As a result of the panel estimation, she finds that the average educational attainment of the employed workers in industry has a positive and significant impact on the industrial output for Turkey.

2.3. Models with Different Education Levels

In this thesis, the aim is to distinguish the impact of different levels of education on real GDP (and real GDP per workforce) of Turkey. Hence, in this section, the models in the literature constructed to examine the effects of different education levels on growth rate or level of output (and per capita output) will be discussed. In addition, the model introduced by Knowles (1997) and the augmented Solow model with different levels of education will be presented extensively in this section, since those models will be estimated for the case of Turkey within the panel data framework in this study.

Most of the studies regarding the relationship between education at different levels and output rely on the usage of time series data. For instance, Kar and Ağır (2006) examine the relationship between human capital, in the forms of education and health, and per capita GDP in terms of causality for Turkey in the period of 1926-1994. In their study, the shares of health and education expenditures in per capita GDP are used as proxies for human capital. Kar and Ağır conclude that the education expenditures has causal a impact on output per capita. However, the reverse causality exists for the relationship between per capita output and health expenditures in Turkey.

Sarı and Soytaş (2006) examine the relationship between real GDP and enrollment rates in primary, middle and high schools and in the universities in Turkey for the period 1937-1996. In fact, they investigate the causal impact of different levels of education, proxied by enrollment rates, on real GDP. Firstly, the cointegration between the different levels of school enrollments and real GDP is tested. Cointegrated relationships are found which imply the existence of a long run equilibrium relationship between education and output in Turkey. Then, the vector error correction modeling is employed in order to test for causality. Sarı and Soytaş find that primary and secondary school enrollment Granger cause output. There is bidirectional causality between university enrollments and real GDP in Turkey.

Moreover, Doğan and Bozkurt (2003) also examine the relationship between the enrollment rates in primary and high schools and in the universities on the one hand and the per capita GDP on the other hand in Turkey for the period of 1983-2001. As in the study of Sarı and Soytaş, the cointegration is tested for the education variables and real GDP. The evidence is found implies a long term relationship between output per capita and enrollment rates in high schools and universities in Turkey. In addition, the error correction model used in the study implies that there is bidirectional causality between the high school and university enrollments and per capita GDP in Turkey.

In order to determine which level of education affects economic growth in Turkey, Deniz and Doğruel (2008) estimate a vector autoregressive regression model using annual data covering the period 1930-2004. They use the number of students per teacher at primary, secondary and higher levels to proxy for the quality of education at different levels. The estimation results of their study indicates that the quality of education for primary and secondary schools in Turkey have long run growth effects.

Lin (2006) investigates the impact on real output of the educational attainment for the Taiwanese economy in the period of 1964-2000. In the empirical study, the effects of different education levels, namely primary, junior-high, senior-high and college, are obtained with the creation of different indices of educational achievement.[14] Lin finds that the effect of education without weighting by any level has less impact on real output than that of each indexed variable obtained by weighting different educational attainments. Moreover, the separate estimations obtained by using different indices as independent variables imply that the primary education has the greatest impact on output for the Taiwan economy.

[14] Each index is obtained by giving a higher weight to the relevant educational level when calculating the average number of years of formal education per person.

Liu and Armer (1993) also examine the effects on output of the different levels of education in Taiwan over the period 1953-1985. In the empirical study, the percentage of adult population who completes primary, junior-high, senior-high school and college education are taken as proxies for education variables. Liu and Armer find a similar result to those obtained by Lin (2006) in the sense that the effects of primary and junior-high school attainment have the greatest impact on output growth in Taiwan over the period 1953-1985. However, contrary to Lin, the empirical findings of their study imply that the senior-high and college education have no significant impact.

Slef and Grabowski (2004) investigate the impact of education on growth of output and causality in India for the time period of 1966-1996. In their study, education capital is broken down into categories of primary, secondary and tertiary levels. The relationship between the growth of GDP per capita and each education level is developed. The education capital, in the empirical study, is measured by two commonly used proxies: enrollment rates and mean years of education at each level. The empirical evidence implies a strong correlation between various levels of education and per capita output growth. Furthermore, it is concluded that while the tertiary education has no causal impact on growth; the secondary education has a weak causal impact but the primary education has a strong one.[15]

Petrakis and Stamakis (2002) examine the relationship between growth effect of education and level of development. They divide the sample into three subgroups: less developed, developed and advanced countries, and, they also break down the education into three different levels: primary, secondary and higher education.[16] Petrakis and Stamakis exploit the Weighted Least Squares (WLS) estimation in the empirical analysis; and they conclude that higher education has a crucial role on

[15] The conclusions about causality do not differ much whichever two proxies is used.

[16] Petrakis and Stamakis form the sample sub groups including countries from different geographical regions in order to avoid the problem of possible multicollinearity.

affecting the growth rate of GDP per capita in advanced economies. Primary and secondary levels of education are more important for the groups of less developed and developed countries. In addition, the test of whether the educational contribution of each level differs significantly within each sub sample (for less developed, developed and advanced countries) is conducted. They conclude that there is a significant difference in terms of the effect of each education level on output growth within each group.

According to Petrakis and Stamakis (2002) primary and secondary education are more important for growth in less developed countries, while higher education is more important in developed countries. Similarly, Slef and Grabowski (2004) find that the primary education has a stronger effect on growth than the secondary education.

In the rest of this section, the model of Knowles (1997) and the augmented Solow model regarding the relationship between differentiated education capital and output level will be presented in a more detailed way because those are the models that will be estimated in this thesis.

2.3.1. The Model of Knowles

In this thesis, one of the models that will be used to examine the impacts of different levels of education on real GDP (and real GDP per workforce) in Turkey is the one which is introduced by Knowles (1997). So, in this sub section the model constructed by Knowles is presented.

Knowles assesses how different education levels affect aggregate output. To do so he disaggregates the labor force on the basis of highest level of schooling regardless of completing the relevant level of schooling, that is, a labor should enter a given level of schooling, not necessarily complete, to characterize that level. He estimates a model for a cross section of 77 countries for the whole sample and for the sub samples of high income and low income countries.

Knowles uses the Cobb-Douglas production function defined by $Y = AK^{\alpha}L_1^{\beta 1} L_2^{\beta 2} L_3^{\beta 3} L_4^{\beta 4}$, where Y is the real output, A is the level of technology and, L_1, L_2, L_3 and L_4 are the labor with no formal schooling (or illiterate), with primary, secondary and tertiary schooling respectively. In the empirical model, he estimates the equation, which is directly obtained from the Cobb-Douglas production function,

$$\text{In}Y = a + \alpha \text{In}K + \beta_1 \text{In } L_1 + \beta_2 \text{In } L_2 + \beta_3 \text{In } L_3 + \beta_4 \text{In } L_4 + \varepsilon \quad (2.3.1)$$

The data on real output and population used are from Summers and Heston (1991) and the physical capital stock are from Benhabib and Spiegel (1994). Knowles estimates the equation (2.3.1) with single OLS (for the year 1985) for the whole sample and two sub samples of low and high income countries with those data.

He tests whether pooling the data of high income and low income countries is appropriate or not by using the Chow test; and he uses the Jarque-Bera test for normality of residuals. Moreover, the RESET test for model misspecification is performed. The normality of residuals and correct model specification are both rejected for the full sample but not for the two sub samples. Furthermore, the null hypothesis of pooling the data is rejected implying that it is not appropriate to pool the sub samples, which supports the results of Jarque-Bera and RESET tests.

The main conclusion of his OLS estimation is that the tertiary education has the greatest impact on output level because the marginal productivity of labor increased with the level of education in both low and high income countries sub samples.[17] Moreover, the unskilled labor force, proxied by low level of educational attainment

[17] It should be noted that the coefficients in the equation (2.3.1) reflect the elasticities, not the marginal products. However, the marginal products are calculated according to the rule $MP_{Li} = Y/L_i * \beta_i$ for i = 1,2,3,4.

of the workforce with no formal schooling, is productive in less developed countries but not in highly developed countries due to the zero marginal productivity of labor with no formal schooling (L_1) in high income countries.

2.3.2. Augmented Solow Model with Different Levels of Education

As well as the model of Knowles (1997), the augmented Solow model with different levels of education is also estimated in this thesis. In order to construct the augmented Solow model being employed in the empirical analysis of the study, the augmented model introduced by Islam (1995) is extended to include differentiated education capital. Hence, firstly, the augmented Solow model used by Islam is presented and then, this model is extended to include educational capital at different levels.

In his empirical study, Islam uses the augmented model introduced by Mankiw et al. (1992) within the panel data framework. So, consider the augmented equation being estimated by Mankiw et al. with a cross sectional data, $\ln(Y/L)=\ln A(0)+gt+\frac{\alpha}{1-\alpha}\ln(s_k)-\frac{\alpha}{1-\alpha}\ln(n+g+\delta)+\frac{\beta}{1-\alpha}\ln(h^*)$. Yet, this equation is valid only at the steady state; on the other hand, the out-of-steady-state equation, which is employed by Islam, is as follows[18]

$$\ln y(t_2)=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(s_k)-(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(n+g+\delta)+e^{-\lambda\tau}\ln y(t_1)+(1-e^{-\lambda\tau})\frac{\beta}{1-\alpha}\ln(h^*)+(1-e^{-\lambda\tau})\ln A_i(0)+g(t_2-e^{-\lambda\tau}t_1) \qquad (2.3.2)$$

In order to see the out-of-steady-state dynamics for economies it is more plausible to estimate the equation (2.3.2), hence, in this study, the extension of this equation including different education levels as proxies for human capital will be employed.

[18] For the construction of the equation (2.3.2) in a detailed way see Mankiw et al. (1992).

In the equation (2.3.2), h* represents the steady-state level of human capital per effective worker and it could be proxied by different educational attainments of the labor force in separate equations or in a single equation including all levels simultaneously. In this study, the stock of human capital, H, is proxied by the educational attainment of the workforce at four different levels (L_1: the amount of the labor force having no formal schooling, L_2: the amount of the labor force having basic level of schooling, L_3: the amount of the labor force having secondary level of schooling and L_4: the amount of the labor force having higher level of schooling), and those education levels are taken in a single equation.[19] Then, the resulting model is given by

$$\mathrm{Iny}(t_2)=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}[\mathrm{In}(s_k)-\mathrm{In}(n+g+\delta)]+e^{-\lambda\tau}\mathrm{Iny}(t_1)+\beta_1\mathrm{In}(L_1/L)+\beta_2\mathrm{In}(L_2/L)+\beta_3\mathrm{In}(L_3/L)$$
$$+\beta_4\mathrm{In}(L_4/L)+(1-e^{-\lambda\tau})\mathrm{InA}_i(0)+g(t_2 - e^{-\lambda\tau}t_1) \qquad (2.3.3)$$

In this study, the natural logarithm of the level of human capital per effective worker is approximated by the sum of the natural logarithms of different levels of education divided by total labor force, in the equation (2.3.3). It should be emphasized that all levels of education as proxies of human capital are taken in a single equation simultaneously; nevertheless, a person having secondary level of education has also basic level of education, so there may be correlation between the education variables in this equation. On the other hand, this model could be helpful to see the impact of different levels of education on the per capita output at the same time within the augmented Solow model framework.

Moreover, in the panel data framework, the equation (2.3.3) could be written in the form $y_{i,t}=\gamma y_{i,t-1}+\sum_{j=1}^{5}\beta_j x_{j,it}+\eta_t+\mu_i+v_{it}$, where $x_{j,it}$ represents $\mathrm{In}L_j/L$ for j=1,2,3,4 and

[19] The variables L_1, L_2, L_3 and L_4, which are used in the empirical studies of this thesis, will be explained in Chapter 6.

In(s_k)-In($n+g+\delta$) for j=5 and the other variables are the same as defined in the section (2.2). In addition, the estimation results of this equation will be given in Chapter 7.

CHAPTER 3

REVIEW OF THE VARIABLES USED IN LITERATURE

In studies examining the determinants of output level and sources of economic growth, empirical models, generally based on aggregate production function, are used. The factors used in production functions in those models are mainly labor, physical and human capital. In this regard, the variables, which are used to proxy for these factors of production similar to the ones being employed in this thesis, will be discussed in this chapter. Firstly, the variables desired to be explained within the growth literature are discussed, then the proxies for the labor, human and physical capital variables used in the empirical literature will be explained.[20]

3.1. Explained Variable

In the empirical growth literature, the determinants of output level and the sources of economic growth are main concerns which are desired to be explained. This is due to the fact that if those sources are determined, then the limited resources of the economies could be allocated more efficiently to the factors that produce more economic value. Hence it could be possible to increase the standard of living of these economies. To do so, *the level of real GDP and its growth rate; or the growth rate of real GDP per capita (per worker or per working age population) and* its *level* are commonly used in literature. In addition, *the average annual growth rate of real GDP per capita (per worker or per working age population)* is also employed in empirical models. Moreover, the growth models generally use output level variables

[20] In this chapter, not only the variables which are included directly in the models estimated in the present study such as education are presented, but also the variables which are not included in these models are presented.

in the logarithmic form. In order to examine the growth rate of those variables, economists employ the log difference of the real GDP per capita (per worker or per working age population).[21]

3.2. Labor Variable

One of the factors of production that produce output and hence cause economic growth is the labor input. The aggregate level of labor input, used in the production processes, is generally measured as the labor force in the empirical literature. In some cases, however, the adult population aged 15 and over; or aged 25 and over; or the working age population are used. In addition, since the employed people, in fact not the whole labor force, are the ones entering into the production processes and hence affecting the output level and growth; some studies preferred to use the numbers of the employed population as the labor input in the production function. In this study, on the other hand, the labor input is not included directly in models estimated; instead the educational attainment of the labor force at different levels are employed as education capital input. That is, in this study, the labor input is included indirectly in the production.

3.3. Human Capital Variable

Human capital is another factor of production used in empirical literature, especially in endogenous growth models, to explain the determinants of output level and growth. Human capital has both quantitative dimension such as number of people, the proportion entering upon useful work and hours worked; and a qualitative dimension such as knowledge, skills and similar attributes affecting human capabilities to do productive work (Schultz, 1961: 8). The quantitative aspect of

[21] Since taking the logarithm of a standard Cobb-Douglas production function simplifies the calculations in regression estimates, the widespread usage of the Cobb-Douglas production function means researchers often take the logarithm of output variable in regressions. Moreover, as the log difference of output variable is equivalent to the growth, the log difference of real GDP per capita (per worker or per working age population) also used in regression equations.

human capital could be measured easily and its effect on output would become more apparent in the empirical framework. However, the capabilities to increase the value of productivity of labor representing qualitative dimension of human capital are difficult to measure, and they are more valuable with respect to output production.

In literature, human capital proxies generally take the form of either educational attainment or health. The educational attainment is commonly used in the literature and has relatively more motivation to be employed in the empirical models. In this study only the education capital is used in order to proxy for human capital. In this section, however, in order to see the whole picture regarding human capital, the variables employed for proxying both education and health capital are discussed.

3.3.1. Education Capital

In empirical literature, in order to proxy for human capital, both qualitative and quantitative measurements for education capital are employed. However, due to lack of data, quantitative measurements such as *enrollment rates and average years of schooling* are more popular than the qualitative ones such as *pupil-teacher ratio, dropout rates, test scores for various disciplines, spending per pupil as a fraction of GDP per capita, the ratio of total expenditure on education to real GDP and the ratio of estimated average salaries of teachers in per capita income.* Moreover, the variables proxying education quality are difficult to measure and, this difficulty hinders the usage of qualitative measurements of education in empirical studies as in the present study. However, in order to provide a review of the possible proxies of educational capital, the commonly used proxies, both qualitative and quantitative, will be presented in this section.

3.3.1.1. Enrollment Rates

Enrollment rates, which are commonly used in empirical work, are flow variables measuring quantitative additions to human capital without regarding their quality. The commonly employed enrollment rates have, in fact, two versions: net and gross enrollment rates. The former for a given level is computed by dividing the number of enrollments at that level of education and at the corresponding age range by the whole population in the same age range. On the other hand, the latter for a specified level is calculated as the number of children enrolled at that level divided by the population of the persons of the designated school age. Although, the gross enrollment rates have relatively widespread usage, there would be a propensity for gross ratio to overstate the accumulation of education capital due to its definition.[22]

One of the variables proxying education capital used by McMahon (1998) is the *gross enrollment rates* emphasizing access and quantitative dimension of education. Furthermore, in order to examine the effect of education on growth of output, Barro (1991) uses the *primary and secondary school enrollment rates* to proxy for educational attainment. Nevertheless, Hanushek and Kimko (2000) mention that since schooling flow variables do not represent either the relevant stock of human capital or even changes in the stock during the periods of educational and demographic transition; the usage of such variables are inappropriate. Hence, even though the data is available, the enrollment rates are not employed as proxies for educational attainment in this study.

3.3.1.2. Adult Literacy

Some studies such as of Romer's (1989) uses the *adult literacy* rate being simply defined as the percentage of population aged 15 and over who can read and write to proxy for education capital. In spite of the fact that the adult literacy rate contrary to

[22] Due to the repeats or the drop outs, the gross enrollment rates sometimes exceed 1. On the other hand, net enrollment ratios are always less than 1.

the enrollment rate, which is a flow variable, reflects a measure for the stock of education capital; it is probably underestimates the education capital stock. In fact, for productivity, other aspects of labor such as numeracy, logical and analytical reasoning would be required, so the usage of adult literacy rate imply that the education beyond the most elementary level does not contribute significantly to the productivity of labor (Barro and Lee, 1993). On the other hand, literacy rates have the advantage of being available more often and they are easy to measure than other variables like enrollment rates.

3.3.1.3. Average Years of Schooling

One of the commonly used proxies for human capital in the form of education is the *average years of schooling* of labor force or working age population. It could be simply defined as the weighted average of the share of people in labor force or working age population in different levels of schooling. The weights are generally taken as the duration of the corresponding levels of education, and different schooling levels are mostly taken as primary, secondary and higher education. For instance, Psacharopoulos and Arriagada (1986) compute *mean years of schooling of the labor force* variable as the weighted average of the percentage of persons in the labor force in each level of schooling (no education, complete and incomplete primary and secondary, and higher education) and the weights are the duration in the years of schooling of each level. Moreover, Lau et al. (1991) use *average years of schooling of the working age population* to proxy for education capital. The averaged number of years of schooling completed per person of working age population, in their empirical study, is calculated as the total number of years of schooling completed by all individuals in that population divided by the number of people in the working age population.

In the panel data study of Islam (1995), the *average schooling years in total population over age 25* is employed to proxy for education. This variable is borrowed from the study of Barro and Lee (1993) describing a data set on educational attainment for cross section of countries over five year periods from 1960-1985. Klenow (1997) also uses the Barro and Lee's (1996) measure of average years of schooling, as well as enrollment rates, to measure the effect of education on growth with an updated data set.[23]

3.3.1.4. School Attainment at Specific Levels

Sometimes, in order to differentiate the effect of education at different levels on output level or growth as in the present study, the labor force or working age population is divided into categories with respect to highest level of schooling attained. For instance, Knowles (1997) *differentiates* the population aged 15-64 with respect to particular level of schooling namely primary, secondary and tertiary under the assumption of full employment of those aged 15-64. Such an assumption obviously leads to an imperfect proxy for the employed workforce since the whole population of aged 15-64 is unlikely to be fully employed. Furthermore, in his work, to characterize a given level of schooling, it is enough for a worker to enter that level of schooling, not necessarily completed; so the variables being used for schooling levels do not represent the success of schooling and hence those variables could be inappropriate.

Petrakis and Stamatakis (2002) also use *differentiation* of education capital. In their empirical study, the *completion rates at different educational levels* are employed. The completion rate for a specified education level is defined as the percentage of those who completes the corresponding level in the population enrolled at that level. Those rates used in the empirical study, are broken into three categories: lower, secondary and higher education. However, the completion rates at each level differ

---

[23] The difference in the calculation of mean years of schooling in the two studies of Barro and Lee is the adult population being concentrated on. Whereas the population aged 25 and over are used in the earlier study, they extend the range to cover the population aged 15 and over as well. On the other hand, it is more plausible to use the wider range to estimate the school attainments since in many developing countries a large proportion of labor force is younger than 25 (Barro and Lee, 1996).

from the variables used in the model developed by Knowles in the sense that completion of an education level is not a requirement for Knowles's (1997) definition. In addition, the completion rates they use give measures of the successes of the school system in producing graduates and reflect the quality of education.

3.3.1.5. International Test Scores

One of the qualitative measurements for the education capital, which has a common usage in empirical literature, is *students' test scores* from international examinations. For instance, Barro (2001) uses *international test scores of students* from mathematics, science and reading examinations measured as percentage of correct answers of students. One advantage of using test scores in empirical models is that those comprise the effects of the factors not only resulting from education system but also from outside it. However, two points in using test scores of students needs to be handled: firstly the factor affecting output level and economic growth is, in fact, the quality of labor force, not the quality of students; secondly the current performance of students may lead a future growth, and may not probably cause an immediate result due to the time lag between the school enrollment and the entry into labor force (Hanushek and Kimko, 2000).

3.3.1.6. Investment Devoted to Education

Another commonly employed qualitative measurement for education capital is the *amount of investment devoted to education*. For example, one of the variables McMahon (1998) uses to proxy for education capital is *ratio of gross public investment devoted to education in GDP* representing a significant aspect of the qualitative dimension of education. In order to separate the effects of various levels of education on per capita growth, the public investment allocated to education is differentiated into three levels in his empirical study (investment in primary, secondary and higher education). Although public investments in education are taken as a measure of quality, clearly it is not the best choice since the private investments are excluded. On the other hand, McMahon mentions that the investment in education is important for quality in poor countries since it represents investment in textbooks, libraries and other input, which contribute to educational effectiveness.

3.3.2. Health Capital

Despite the fact that human capital is usually considered in the form of only education, the indicators representing health conditions of countries are also employed to proxy for human capital. In addition, even though the human capital is not taken as in the form of health in this study, the commonly used health indicators in literature will be mentioned in the present section since those variables, as well as indicators of education, are important due to reflecting the part of total factor productivity in countries. In fact, the more healthy workers are in production process, the more output will be available due to high productivity.

3.3.2.1. Life Expectancy at Birth

One of the commonly used proxies for human capital in the form of health is the *life expectancy at birth.* The life expectancy at birth is defined as the average number of years a person is expected to live from the time of his birth. For instance, Barro and Lee (1994) use life expectancy at birth in order to represent the level of health of the countries included in their empirical study. Moreover, McDonald and Roberts (2002) also employs life expectancy at birth as one of the human capital proxies.

3.3.2.2. Infant Mortality

Infant mortality is another proxy for health capital and this variable has also a widespread usage in empirical literature. Infant mortality could be defined as the number of infant deaths before one year of age per 1000 live births. McDonald and

Roberts (2002) use, as well as life expectancy at birth, infant mortality as a health indicator. They state that although health capital is not enough alone to proxy for human capital, it is clear that human capital is a more complex input consisting of more than knowledge.

3.4. Physical Capital Variable

Physical capital, which could be defined as the machinery, equipment, plant and buildings used in the production process, is an important factor employed to determine the determinants of output level and the sources of economic growth. Nevertheless, in empirical literature, it is difficult to measure the variables that could be viewed as physical capital. On the other hand, a common approach to estimate the stock of physical capital in empirical literature is the *perpetual inventory method* (PIM). For example, Hulten and Isaksson (2007), Lau et al. (1991) and Papageorgiou (2003) use this method in their empirical studies. Under the PIM, the capital stock at the end of year t, available for the production of the following year, K_{t+1}, could be estimated as the sum of the depreciated amount of capital left from the preceding year, $(1-\delta)K_t$ and the new capital added through new investment during the year, I_t. That is, the stock of physical capital could be obtained from $K_{t+1}=(1-\delta)K_t + I_t$, where δ is the depreciation rate of physical capital.[24]

By a simple recursive process, the stock of physical capital could be represented as the weighted sum of the initial stock of capital and the series of investment levels. In other words, K_t could be obtained from $K_t = (1-\delta)^t K_0 + \sum_{i=1}^{t}(1-\delta)^{t-i} I_i$. This method requires a time series data for investment, an initial stock of physical capital and the rate of depreciation. The depreciation rate used in empirical studies are generally

[24] It should be noted that, in the PIM, the new investment made within the year t is assumed to enter the production process in the following year, t+1. However, it is also possible for the new investment to enter in the production process in the same year. For instance, Lau, Jamison and Louat (1991) use the assumption that the investment in year t enters the production process in that year.

taken as four, five, seven or ten percent. Moreover, the time series data on investment are available for almost all countries. (Hulten and Isaksson, 2007: 42).[25] On the other, the difficulty of this method is to obtain an initial value for the physical capital stock. Nevertheless, there are several ways to get an initial stock and one way, for example, could be obtained through the relationship $K_0 = I_0/(\dot{g} + \delta)$, where $\dot{g}$ is the average annual growth rate of investment during the period being under consideration (Yaşar, 2008).[26]

Lau, Jamison and Louat (1991) use the time series of *utilized capital stocks* which is calculated by the multiplication of the rate of utilization and the estimated time series of capital stocks obtained by PIM in order to proxy for physical capital.[27] In the estimation of physical capital stock with PIM, the initial stock of physical capital, in 1945, is taken as zero and the depreciation rate is taken as 5 percent per annum.

In some studies such as Barro (1991; 2001), Barro and Lee (1994; 1996) and McMahon (1998), instead of estimating the stock of physical capital directly from the investment figures, the ratio of real gross investment to real GDP is used to proxy for physical capital.[28] McDonald and Robert (2002) also use the rate of investment in real GDP per worker to measure the physical capital.

---

[25] In empirical studies, the series of real investment, rather than gross one, is generally used to get more plausible results.

[26] There are other methods for obtaining initial stock of physical capital, for instance, Benhabib and Spiegel (1994) use the coefficient estimates for the physical capital variable, K, from the regression equation $\log Y = A + \alpha \log K + \beta \log L + \gamma \log H + \varepsilon$, with the data in which the physical capital stock is available.

[27] The annual rates of utilization are estimated as the ratio of actual real GDP to potential real GDP in the empirical study of Lau et al. (1991).

[28] Barro (1991) uses real private and public investments which are taken both separately and together.

CHAPTER 4

BACKGROUND OF EDUCATION IN TURKEY

In Chapter 2, the literature regarding the role of human capital, especially in the form of education, in the growth theories are presented. The development of education, which is an important component of human capital, over time in Turkey, will be discussed in this chapter. The present situation and developments over time of the educational sector will be presented by using indicators such as enrollment ratios, average years of education of workforce, the number of schools, students and teachers, the resources allocated to education and the literacy rates in Turkey.

The literacy rate is an important indicator regarding the educational attainment of the population at the basic level. It is commonly employed in the empirical literature due to easier accession of the data. Furthermore, the literacy rate increases with the level of development of countries, and it reaches about 100 percent in advanced economies. The literacy rates of the population aged 6 years and over for the period 1935-2000 in Turkey are given in Table 1. The literacy rates for both males and females in Turkey show an upward trend for the period 1935-2000. Indeed, it rises from 9.8 percent for women and 29.3 percent for men in 1935 to 80.6 percent for women and 93.9 percent for men in 2000. However, in Turkey, the female literacy rate is still too low relative to that of male. This implies a disadvantage for the girls at the beginning of the education process. In addition, the girls, who are able to participate in the education process, withdraw from the educational system as the level of education increases in our country. The values of gender enrollment ratio, indicating the relative size of female gross enrollment ratio as compared to male gross enrollment ratio, at each level of education for Turkey could demonstrate this situation.[29] Those statistics are given in Table 2. As can be seen from Table 2, the gender enrollment ratio in higher education (between 70 and 80 percent) is lower than the gender enrollment ratio in both primary (between 85 and 95 percent) and secondary education (between 75 and 85 percent) in Turkey. That is, ratio of the female enrollment rate to male enrollment rate in higher education is lower than this ratio in primary and secondary education.

Table 1. Literacy Rates (for population age 6 and over)

Census Year	Total (%)	Female (%)	Male (%)
1935	19.2	9.8	29.3
1940(1)	24.5	12.9	36.2
1945(2)	30.2	16.8	43.7
1950(3)	32.5	19.4	45.5
1955	41.0	25.6	55.9
1960	39.5	24.8	53.6
1965	48.8	32.8	64.1
1970	56.2	41.8	70.3
1975	63.7	50.5	76.2
1980	67.5	54.7	80.0
1985	77.4	68.2	86.5
1990	80.5	72.0	88.8
2000	87.3	80.6	93.9

(1)Data of 1940 is estimated by using the data of 1935 and 1945.
(2)Population age 7 and over.
(3)Population age 5 and over.
Source: Devlet Planlama Teşkilatı (2007).

[29] The gender enrollment ratio at a specific level is defined as the female gross enrollment ratio divided by the male gross enrollment ratio at that level multiplied by 100.

Table 2. Gender Enrollment Ratio by Level of Education (1997-2007)[(1)]

	Primary education	Secondary education	Higher education
1997	85.6	74.7	69.6
1998	87.0	75.5	69.4
1999	88.5	74.7	71.0
2000	89.6	74.4	73.6
2001	90.7	75.9	75.2
2002	91.1	72.3	74.3
2003	91.9	77.8	74.1
2004	92.3	80.3	74.7
2005	93.3	78.8	77.2
2006	94.1	79.6	77.6
2007	96.4	85.8	-

(1)The compulsory education was extended to 8 years in 1997 and so the primary education, which is available after 1996, comprises both the primary school and middle school education. In addition, the secondary education and higher education refer to the high school and higher level of school education respectively.
Source: Milli Eğitim Bakanlığı İstatistikleri, Örgün Eğitim 2007-2008.

One of the important indicators of the human capital of countries is the enrollment ratios. Hence, the gross and net enrollment ratios at specific levels of education in Turkey are presented in Table 3. In the last decade, an improvement in the enrollment ratios at each level (primary, secondary and higher education) in Turkey is observed. As can be seen from Table 3, the greatest increase in the enrollment rates (both gross and net) in that period is obtained in the secondary education. Moreover, in Turkey, the primary education enrollment ratios are high and have increasing trends over the last ten years. In fact, in the last decade the primary education gross enrollment rate is increased by 15.0 percentage points to 104.5 percent in 2007. In the same period the net enrollment ratio in primary education is increased by 12.7 percentage points to 97.4 percent in 2007. Although enrollment ratios in the higher education, which has the greatest potential regarding the contribution of the accumulation of human capital, are increasing in the last ten years; they are still too low relative to the enrollment rates at primary and secondary education levels.

Table 3. Enrollment Ratio by Level of Education (1997-2007)(1)

		Primary education	Secondary education	Higher education
1997	gross	89.5	52.8	19.5
	net	84.7	37.9	10.2
1998	gross	94.3	57.1	21.7
	net	89.2	38.9	10.8
1999	gross	97.5	58.8	21.0
	net	93.5	40.4	11.6
2000	gross	100.9	61.0	22.2
	net	95.3	43.9	12.3
2001	gross	99.4	67.9	23.4
	net	92.4	48.1	13.0
2002	gross	96.5	80.7	27.1
	net	91.0	50.6	14.6
2003	gross	96.3	80.9	28.1
	net	90.2	53.4	15.3
2004	gross	95.7	80.9	30.6
	net	89.7	54.9	16.6
2005	gross	95.6	85.2	34.5
	net	89.8	56.6	18.8
2006	gross	96.3	86.6	36.6
	net	90.1	56.5	20.1
2007(2)	gross	104.5	87.5	-
	net	97.4	58.6	-

(1)The compulsory education is extended to 8 years in 1997 and so the primary education, which is available after 1996, comprises both the primary school and middle school education. In addition, the secondary education and higher education refer to the high school and higher level of school education respectively.
(2)Schooling ratios for the 2007-2008 educational year were calculated according to the results of the Address-Based Population Register System 2007 Population Census.
Source: Milli Eğitim Bakanlığı İstatistikleri, Örgün Eğitim 2007-2008.

Table 4. Number of School, Enrollment and Teacher by Level of Education(1)

	Year	School	Student	Teacher(2)
Primary education	1997	45,649	9,102,074	302,982
	1998	44,525	9,512,044	316,991
	1999	43,324	10,053,127	324,924
	2000	36,065	10,460219	345,141
	2001	35,044	10,562,426	375,620
	2002	35,168	10,331,619	390,275
	2003	36,117	10,479,538	384,029
	2004	35,581	10,565,389	399,025
	2005	34,990	10,673,935	389,859
	2006	34,656	10,846,930	402,829
	2007	34,093	10,870,570	445,452
Secondary education	1997	-	-	-
	1998	5,708	2,013,152	139,664
	1999	6,168	2,444,407	143,469
	2000	6,244	2,606,994	141,441
	2001	6,389	2,855,851	145,461
	2002	6,134	3,034,959	148,563
	2003	6,512	3,014,392	160,049
	2004	6,861	3,039,449	167,949
	2005	7,435	3,258,254	185,317
	2006	7,934	3,386,717	187,665
	2007	8,280	3,245,322	191,041
Higher education(3)	1997	71	912,377	56,401
	1998	73	972,180	60,129
	1999	73	1,015,412	65,204
	2000	76	1,091,755	67,880
	2001	76	1,155,686	71,290
	2002	77	1,256,629	76,090
	2003	77	1,320,392	78,804
	2004	77	1,410,760	82,096
	2005	93	1,543,845	84,785
	2006	115	1,608,253	89,329
	2007	130	1,654,650	98,766

(1)The compulsory education was extended to 8 years and so the primary education, which is available after 1997, comprises both the primary school and middle school education. In addition, the secondary education and higher education refer to the high school and higher level of school education respectively.
(2)Total number of teachers includes permanent and contractual teaching staff.
(3) Students enrolled in graduate studies are included. The open universities are excluded.
Source: Milli Eğitim Bakanlığı İstatistikleri, Öğrenci Seçme ve Yerleştirme Merkezi İstatistikleri and Yükseköğretim Kurulu İstatistikleri.

The education system is an important factor in the accumulation of the human capital in an economy; and the number of schools, students and teachers are one of the indicators representing the improvements in an education system. Hence, those numbers by the level of education for the period 1997-2007 in Turkey are given in Table 4. From the Table 4, it can be observed that the number of students and teachers at each education level (primary, secondary and higher education) has been increasing over the last decade in Turkey. In fact, the number of students enrolled in primary, secondary and higher education has increased by 19.4 percent (from 9,102,074 in 1997 to 10,870,570 in 2007), 61.2 percent (from 2,013,152 in 1998 to 3,245,322 in 2007) and 81.4 percent (from 912,377 in 1997 to 1,654,650 in 2007) respectively. In addition, the number of teachers in the primary, secondary and higher education is increased by 47.0 percent (from 302,982 in 1997 to 445,452 in 2007), 36.8 percent (from 139,664 in 1998 to 191,041 in 2007) and 75.1 percent (from 56,401 in 1997 to 98,766 in 2007) respectively. The increase in the number of the students enrolled in the primary education is less than the increase in the number of teachers in that level of education and this could imply an improvement of the quality of primary education. On the other hand, a reverse situation exists for the secondary and higher education because the raise in the number of the students enrolled in those levels is greater than the increase in the number of teachers. Nevertheless, as can be seen from Table 4, the increases in the number of schools in secondary and higher education may demonstrate that there is a development in those levels of education in Turkey.

Table 5. The Share of Educational Investment in Total Investment and the Ratio of the Educational Investment to GDP in Turkey (1970-2004)

	Share in Total Investments (%)	Share in GDP (%)
1970-1974	2.5	0.5
1975-1979	2.0	0.5
1980-1984	2.0	0.4
1985-1989	2.2	0.5
1990-1994	2.8	0.7
1995-1999	3.3	0.8
2000-2004	5.1	1.0

Source: Türkiye Sanayicileri ve İşadamları Derneği (2006)

The resources allocated to education and training is important for accumulation of human capital for countries especially having high youth population such as Turkey. So, the share of educational investment in total investment and in the GDP of Turkey is calculated by Saygılı et al. (2006:62) for the five year periods over 1970-2004. Those shares are provided in Table 5. It can be seen in Table 5 that although the resources devoted to education in Turkey has an increasing trend over the period of 1970-2004, those are still too low. Indeed, the share of the educational investment in the total investment increases from 2.5 percent in the period 1970-1974 to 5.1 percent in the period 2000-2004. Moreover, the ratio of the investment in education to GDP rises just by 0.5 percentage point over the period of 1970-2004. Even though a slight improvement is observed in the resources devoted to education; in order to transform the potential arising from the high youth population of Turkey into an opportunity, investment in education and training should be increased. The allocation of larger resources to education improves the human resources and hence the level of development in Turkey through more qualified and educated population.

One of the significant indicators regarding impact of education in an economy on the output level is the educational attainment of the workforce, which is employed in most of the empirical studies as well as this thesis; and it is the workforce, not the others in society, that makes the real contribution to economy. Hence, in that respect, the amount of labor force by levels of education and the average years of education of the workforce for Turkey are presented in Table 6 and Table 7 respectively.

In Table 6, the number of the workforce in Turkey is divided into six main categories with respect to the highest level of education attained in the period 1988-2006. It could be seen from Table 6 that the number of the labor force in each level increases during this period. In fact, the amount of the labor force having primary level of education (the sum of third and fourth column of the Table), secondary level of education (fifth column of the Table) and higher level of education (last column of the Table) rises by 14.4, 173.9 and 220.5 percent respectively. The good news is that the greatest improvement is realized in the amount of the labor with higher level of schooling. However, although the share of the higher education graduates in the total labor force is increasing over time; it is still too low relative to the labor force having lower levels of education. The improvement in the average years of education of the workforce in Turkey is presented in Table 7. The mean years of education of the labor force is calculated with the amounts of the labor at different levels of education, which is available in Table 6, by assuming the literates without diploma have two years of primary education. As could be seen from Table 7, the average years of schooling of the workforce is increasing continuously except for the year 1994. Indeed, the mean years of education of the labor force rises from 5.22 in 1988 to 7.58 in 2006.

Table 6. The Total Labor Force by Levels of Education (15 and over aged population)[(1)]

	Illiterate	Literate without diploma	Primary school	Junior-high school	High school	Higher education
1988	3,233	1,685	1,0178	1,384	1,954	957
1989	3,198	1,701	1,0634	1,473	1,916	1,008
1990	2,983	1,488	1,0982	1,567	2,070	1,061
1991	2,923	1,370	1,1866	1,604	2,165	1,083
1992	2,602	1,339	1,1910	1,715	2,520	1,176
1993	1,917	1,174	1,1623	1,720	2,647	1,234
1994	2,217	1,234	1,2389	1,830	2,862	1,343
1995	2,151	1,095	1,2217	2,145	3,264	1,413
1996	2,170	934	1,2456	2,106	3,479	1,551
1997	1,990	752	1,2479	2,231	3,609	1,693
1998	1,968	737	1,2599	2,418	3,880	1,783
1999	2,005	797	1,2727	2,514	3,916	1,919
2000	1,985	750	1,2012	2,254	3,990	2,037
2001	1,959	793	1,2093	2289	4,150	2,114
2002	1,728	722	1,1885	2,488	4,449	2,406
2003	1,606	672	1,1440	2,545	4,552	2,624
2004	1,537	890	1,1467	2,619	4,842	2,558
2005	1,324	1,069	1,0764	2,736	5,208	2,827
2006	1,243	1,088	1,0471	2,755	5,352	3,067

(1)Thousands of people.
Source: Türkiye İstatistik Kurumu (Turkish Statistical Agency) (TÜİK).

Table 7. Average Years of Education of Total Labor Force (15 year and over aged population)

1988	5.22
1989	5.25
1990	5.41
1991	5.47
1992	5.71
1993	6.00
1994	5.97
1995	6.17
1996	6.28
1997	6.45
1998	6.55
1999	6.58
2000	6.69
2001	6.75
2002	7.01
2003	7.19
2004	7.19
2005	7.43
2006	7.58

Source: TÜİK and own calculations.

In this chapter, the educational sector background in Turkey and its development over time is presented by using various education indicators. It could be concluded from the above discussion that the education, which is the main component of human capital, in Turkey demonstrates a significant improvement over time. Nonetheless, the education still requires more attention by the policy makers as the indicators regarding the current situation of the education are not still perfect for Turkey.

CHAPTER 5

METHODOLOGY

In this thesis, in order to determine which level of education has significant effect on real GDP (and real GDP per workforce) in Turkey and which level has the greatest impact, the model introduced by Knowles (1997) and the augmented Solow model with different levels of education are estimated with single OLS, pooled OLS and panel data estimation methods. In this chapter, those estimation procedures will be explained rigorously; and at the same time, the hypothesis tests required to find appropriate methods of estimation for those models and the test statistics for those hypothesis, which will be calculated for our models in Chapter 7, are presented.

5.1. Single Cross Section OLS Estimation

The cross sectional data is the one collected for many variables at the same point in time. It is commonly used in econometric analyses of growth models due to easy access to data. For the estimation of such data two methods, which are generally used, could be considered: maximum likelihood and OLS estimations. However, the latter method is used more extensively in empirical analyses since its calculation is much simpler than the former method's. Hence, for simplicity, merely the OLS estimation procedure will be discussed in this section.[30]

The OLS estimation has very attractive properties under certain assumptions and this is why it is commonly used in practice. So, firstly those assumptions to make valid interpretation about the coefficient estimates of a given regression will be mentioned,

[30] In fact, two methods of estimation generally give similar results (Gujarati, 2003:58).

and the powerful statistical properties of the OLS estimators under those assumptions will then be presented. Now, consider the linear regression equation of the form

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + u_i \qquad (5.1.1)$$

where, i corresponds to the i^{th} individual unit and i =1,2, ..., N. Then, for the OLS estimation of the equation (5.1.1) it could be assumed that the random disturbance term, u_i, is independently and identically distributed with zero mean and constant variance. It should be noted that the disturbance term includes the factors which are not explicitly included in the equation but affects the dependent variable; so, this assumption simply implies that the mean effect of the unobservable variables on y is zero. In addition, regarding the disturbance term, it could also be assumed that there is no serial correlation among those disturbances, that is, $cov(u_i , u_j) = 0$ for all $i \neq j$.

Moreover, the assumption that each explanatory variable is uncorrelated with the disturbance term could be desirable. Such an assumption could assist to separate the effects of the disturbance term, being the representative of all the omitted variables influencing the dependent variable, and the explanatory variables (Gujarati, 2003:71).

Another important assumption to interpret the estimates of the coefficients in a regression equation correctly is the absence of multicollinearity. Since each explanatory variable contains different information, this assumption could be useful in order to separate the effect of each explanatory variable.

Finally, the number of observations should be greater than the number of parameters. This is required since one could not estimate k+1 parameters with N observation unless N>k+1.[31]

[31] It should be noted that this requirement is not specific to OLS estimation.

The OLS estimates are best linear unbiased (BLU) and consistent under those assumptions. Moreover, assuming normality of the disturbances, as well as the above assumptions, the estimates of β_0, β_1, …, β_k are themselves normally distributed with means β_0, β_1, …, β_k respectively (Gujarati, 2003:248-249).

Despite the fact that it is easy to estimate a given regression equation with OLS, a single OLS procedure enables only cross sectional information about data. Moreover, the aforementioned assumptions to obtain BLU estimates need not hold in practice. For instance, consider the attempt of estimating an equation of an individual's income on her consumption expenditure with a single cross section OLS. However, in such an equation, the individual's income may not be the only factor affecting the consumption expenditures and there could be other variables such as the number of children, educational attainment and spouse's income. Moreover, those variables being not included in the equation may be related with the individual's income. Then, due to the probability of omitted variables and hence of the violation of the assumption that the explanatory variable and the disturbance term are uncorrelated; the single OLS estimation may lead biased estimates. So, although the OLS estimation provides some powerful properties, this requires some strong assumptions which are not easy to have in empirical work.

5.2. Pooled OLS Estimation

The main reason of using pooled cross sections is to get more precise estimates and test statistics with more power by allowing higher sample size. Pooling of cross sections in different time points may be useful when the relationship between dependent variable and at least some of the explanatory variables remains constant over time (Wooldridge, 2002:409). Unfortunately, structural changes such as economic crisis do not allow for such a smooth relationship over time. Furthermore, populations may have different distributions in different time periods and so, the estimation of pooled cross sections with OLS could lead to little statistical complication. Consider a regression equation of the form $y_{it} = \alpha + \beta x_{it} + u_{it}$, for $i = 1,\ldots,N$ and $t = 1,\ldots,T$, then the estimation of this equation with pooled OLS may lead biased estimates due to the heterogeneity which is likely to occur in the parameters of the equation across individuals or time periods. For instance, assume that the intercept, α, is different for different time periods but same for all individual units with the assumption of homogeneous slope coefficients, β. Then, the pooled OLS estimation of this equation, using all NT observations, yields biases due to ignoring the heterogeneity of intercept across time. On the other hand, in order to reflect this fact and to avoid bias estimates it is allowed for the intercept to change across time by simply introducing time dummies (Wooldridge, 2002:409). In addition, with a similar reasoning the interaction of time dummies with explanatory variables could be included in the regression equation to see the changes over time of the effects of the explanatory variables which interact with the time dummies.

Even though, the dummy variable approach could be helpful in pooled OLS estimation, the calculations could be tedious and the loss of degrees of freedom would be enormous when the number of individual units and time periods is too high. So, the question of under which conditions the pooled regression estimation is appropriate arises. Remember that the usage of least squares estimation with all NT observations requires the assumption of the regression parameters being same over all individuals and time periods. Thus, testing whether the parameters are constant over time and across individual units becomes the first step.

Consider a regression equation of the form $y_{it} = \alpha_{it} + \beta_{it}x_{it} + u_{it}$, for $i = 1,\ldots,N$ and $t = 1,\ldots,T$, where the error term has zero mean and constant variance. For this equation, the heterogeneity of the intercept term and the slope coefficients could be tested. In this context, the homogeneity of the intercept term and slope coefficients could be tested either jointly or separately (Hsiao, 1986:12).

For the sake of simplicity, assume firstly that the parameters are constant over time, but variable across individual units. Then, the following regression equation could be estimated for each individual unit

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}, \text{ for } i = 1,\dots,N \text{ and } t = 1,\dots,T \qquad (5.2.1)$$

In that case, the equality of the slope coefficients and intercept term across individuals could be tested either separately or jointly. Firstly, consider the hypothesis test of H_0^1: $\beta_1 = \beta_2 = \dots = \beta_N$. Then, the restricted model under H_0^1 could be stated as

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}, \text{ for } i = 1,\dots,N \text{ and } t = 1,\dots,T \qquad (5.2.2)$$

The restricted residual sum of squares (RRSS), could be obtained from the within regression of the equation (5.2.2).[32] The unrestricted sum of squares (URSS), is given by $URSS = \sum_{i=1}^{N} RSS_i$, where RSS_i is the residual sum of squares obtained from the regression $y_{it} - \bar{y}_{i.} = \beta_i(x_{it} - \bar{x}_{i.}) + (u_{it} - \bar{u}_{i.})$ for each individual unit i (Hsiao, 1986:14). Then, the F statistics required for testing the hypothesis of H_0^1 is given by $F_1 = \dfrac{(RRSS - URSS)/(N-1)}{URSS/(NT-2N)} \sim F_{(N-1),\,NT-2N}$.[33]

The rejection of the null hypothesis, which occurs in the case that F_1 exceeds $F_{(N-1),\,NT-2N}$ at a proper significance level, implies the heterogeneity of slope coefficients across individual units. In such a situation, the testing of poolability process could be halted. On the other hand, the nonrejection of the hypothesis of the slope coefficients

[32] The method of "within regression" estimation will be introduced and explained rigorously in the section (5.3).

[33] In the given example, there exists only one explanatory variable and the degrees of freedom in the formula of the F statistics is given for one independent variable; however, the degrees of freedom could be generalized to K explanatory variables. For the generalization see Hsiao (1986:15).

being the same could imply an additional test that checks the equality of the intercept term across individuals. This hypothesis is given by H_0^2:$\alpha_1=\alpha_2=\ldots.=\alpha_N$ such that $\beta_1 =\beta_2 =\ldots.= \beta_N$

In that case, the hypothesis of H_0^2 depends on the condition that the slope coefficients of the regression equation are the same among individuals. Then, URSS could be obtained from the within estimation of the equation (5.2.2). The RRSS is obtained from the OLS estimation of the equation

$$y_{it} = \alpha + \beta x_{it} + u_{it}, \text{ for } i = 1,\ldots,N \text{ and } t = 1,\ldots,T \tag{5.2.3}$$

The F statistics for the hypothesis is given by $F_2= \dfrac{(RRSS - URSS)/(N-1)}{URSS/(N(T-1)-1)} \sim F_{N-1,\, N(T-1)-1}$.[34] If F_2 is greater than $F_{N-1,\, N(T-1)-1}$, then the null hypothesis of H_0^2 will be rejected. The rejection of the null hypothesis of H_0^2 means that given the equality of slope coefficients the intercept terms are not same across individual units. On the other hand, the nonrejection of this hypothesis supports the poolability of data in the sense that both the intercept term and slope coefficients are constant across individuals (Hsiao, 1986:12-18).

Moreover, it is also possible to test whether both slope coefficients and intercept term are same across individuals. That is, one could test the following hypothesis H_0^3:$\alpha_1=\alpha_2=\ldots.=\alpha_N$ and $\beta_1=\beta_2=\ldots.=\beta_N$. Under the null hypothesis, the restricted model is given by the equation (5.2.3), as in the previous case. Nevertheless, the

[34] In the given example, there exists only one explanatory variable and the degrees of freedom in the formula of the F statistics is given for one independent variable; however, the degrees of freedom could be generalized to K explanatory variables. For the generalization see Hsiao (1986:16).

unrestricted model is the equation (5.2.2). Then, the F statistics is the following $F_3 = \frac{(RRSS - URSS)/2(N-1)}{URSS/(NT-2N)} \sim F_{2(N-1),\ NT-2N}$.[35]

If $F_3 > F_{2(N-1),\ NT-2N}$, then one rejects the null hypothesis of H_0^3 implying the heterogeneity of intercept term or slope coefficients. On the other hand, the nonrejection of the hypothesis of slope coefficients and intercept being constant among individual units provides a reason to use pooled OLS estimation of the given regression equation (Hsiao, 1986:12-18).

In the above cases, the homogeneity of both slope coefficients and intercept term over time periods are initially assumed. In fact, one could also assume initially that those parameters stay constant across individuals rather than time periods. In such a case, the following regression equation could be estimated for each time periods

$$y_{it} = \alpha_t + \beta_t x_{it} + u_{it}, \text{ for } i = 1,\dots,N \text{ and } t = 1,\dots,T \qquad (5.2.4)$$

Then, the homogeneity of the slope coefficients and intercept term over time periods could be tested either separately or jointly, and similar tests, under the assumption that the intercept and slope coefficients being the same across individuals, could be generated to determine such homogeneity.[36]

The poolability of the data used in this study will be tested with the aforementioned hypotheses by calculating the required F statistics for the models of interest in Chapter 7.

---

[35] In the given example, there exists only one explanatory variable and the degrees of freedom in the formula of the F statistics is given for one independent variable; however, the degrees of freedom could be generalized to K explanatory variables. For the generalization see Hsiao (1986:15).

[36] For a detailed discussion one could look at Hsiao (1986:16).

5.3. Panel Data Estimation

In this thesis, the primary estimation method of the models being constructed depends on panel data techniques. Hence, in this section the estimation methods with panel data will be presented; and simultaneously the advantages and disadvantages of those methods will be given. Then, consider a simple panel data regression with a single explanatory variable of the form

$y_{it} = \alpha + \beta x_{it} + u_{it}$, where $i = 1,2,\ldots,N$ and $t = 1,2,\ldots,T$. (5.3.1)

where, *i* stands for the i[th] cross sectional unit and *t* stands for the t[th] time period.[37]

The estimation of the equation (5.3.1) with panel data depends on the assumption about the intercept and slope coefficients and the disturbance term. Two of the methods commonly used in panel data are the "fixed effect estimation" and the "random effect estimation" approaches. The former based on the assumption of fixed slope coefficients over time and across individuals but, variable intercept over time and/or across individual units (Gujarati, 2003:640). The latter, on the other hand, requires a decomposition of the error term and assumes that the part of the disturbances corresponding to the individual and/or time invariant parts is random. [38]

One of the main advantages of using panel data is allowing for unobservable heterogeneity, which could be time or individual invariant. For example, the geographical position of a province, which could be unobservable, specific to the province and time invariant, could be a factor determining the rate of growth or the level of real GDP. In addition to individual specific effects, time specific effects such as an economic crisis could be allowed in the context of panel data analysis.

---

[37] It must be noted that the equation (5.3.1) is formed under the assumption of poolability, that is, it is assumed to represent a behavioral equation with the same parameters over time and across individuals (Baltagi, 2001:51).

[38] A more comprehensive discussion with regard to those methods will be given in the section (5.3.2).

Then, consider a simple panel data regression equation with one explanatory variable in two-way error component model is given by

$$y_{it} = \alpha + \beta x_{it} + u_{it}, \text{ where } i = 1,2,\ldots,N \text{ and } t = 1,2,\ldots,T \tag{5.3.2}$$

and $u_{it} = \mu_i + \lambda_t + v_{it}$, which is called as the composite error.[39]

The unobserved heterogeneity of individuals is captured in μ_i, which is known as "individual specific effect". It should be noted that μ_i, which is time invariant, accounts for effects of any individual specific factor that is not included in the equation (5.3.2). Furthermore, λ_t represents the unobservable "time specific effects" and the individual invariant λ_t accounts for effects of any time specific factor that is not included in the equation (5.3.2) (Baltagi, 2001:11, 31).

Firstly, assume that the μ_i and λ_t are fixed parameters to be estimated and the remainder disturbances are independent and identically distributed such that $v_{it} \sim IID(0,\sigma_v^2)$. Then, (5.3.2) represents a two way fixed effect error component model. In order to estimate (5.3.2), the assumption that the x_{it}'s are independent from the v_{it} for all i and t is needed. Now, by averaging (5.3.2) over time one could get[40]

$$\overline{y}_{i.} = \alpha + \beta \overline{x}_{i.} + \mu_i + \overline{v}_{i.} \text{ for } i = 1,\ldots,N \tag{5.3.3}$$

---

[39] In a one-way error component model, only the individual specific effect, μ_i, or only the time specific effect, λ_t, is included in the composite error. So, such a model is given by $y_{it} = \alpha + \beta x_{it} + \mu_i + v_{it}$ for only individual specific effect or $y_{it} = \alpha + \beta x_{it} + \lambda_t + v_{it}$ for only time specific effect; and in addition to the two-way error component models, one-way error components models are also used in the estimation of our models in Chapter 7.

[40] We utilize the restriction of $\sum_{t=1}^{T} \lambda_t = 0$. Baltagi (2001) states that this is an arbitrary restriction on the dummy variable coefficients to avoid the dummy variable trap, or perfect multicollinearity. Such a restriction is required; since α, μ_i and λ_t are parameters to be estimated and it is not possible to estimate them separately without imposing an additional restriction.

Similarly by averaging (5.3.2) over individual units one obtains[41]

$$\bar{y}_{t.} = \alpha + \beta \bar{x}_{t.} + \lambda_t + \bar{v}_{t.} \text{ for } t = 1,\ldots,T \tag{5.3.4}$$

Also, averaging (5.3.2) across all observations one can get[42]

$$\bar{y}_{..} = \alpha + \beta \bar{x}_{..} + \bar{v}_{..} \tag{5.3.5}$$

Then, from the equations (5.3.2), (5.3.3), (5.3.4) and (5.3.5) one can deduce that

$$(y_{it} - \bar{y}_{i.} - \bar{y}_{t.} + \bar{y}_{..}) = \beta(x_{it} - \bar{x}_{i.} - \bar{x}_{t.} + \bar{x}_{..}) + (v_{it} - \bar{v}_{i.} - \bar{v}_{t.} + \bar{v}_{..}) \tag{5.3.6}$$

The OLS estimation of (5.3.6) provides the fixed effect estimator of β which is unbiased and consistent. In fact, the estimation of (5.3.6) by OLS results in $\tilde{\beta}_{within}$, the within estimator for the two way error model. Moreover, within estimator of the intercept can be obtained from the equation (5.3.5), that is $\tilde{\alpha}_{within} = \bar{y}_{..} - \tilde{\beta}_{within}\, \bar{x}_{..}$. Since individual and time specific effects are assumed to be parameters, it is possible to estimate them. Indeed, the within estimation of μ_i and λ_t are given by $\tilde{\mu}_i = (\bar{y}_{i.} - \bar{y}_{..}) - \tilde{\beta}_{within}(\bar{x}_{i.} - \bar{x}_{..})$ and $\tilde{\lambda}_t = (\bar{y}_{t.} - \bar{y}_{..}) - \tilde{\beta}_{within}(\bar{x}_{t.} - \bar{x}_{..})$ respectively.

---

[41] We utilize the restriction of $\sum_{i=1}^{N} \mu_i = 0$. Baltagi (2001) states that this is an arbitrary restriction on the dummy variable coefficients to avoid the dummy variable trap, or perfect multicollinearity. Such a restriction is required; since since α, μ_i and $\lambda_{t\ t}$ are parameters to be estimated and it is not possible to estimate them separately without imposing an additional restriction.

[42] We utilize the restriction of $\sum_{i=1}^{N} \mu_i = 0$ and $\sum_{t=1}^{T} \lambda_t = 0$. Baltagi (2001) states that this is an arbitrary restriction on the dummy variable coefficients to avoid the dummy variable trap, or perfect multicollinearity. Such a restriction is required; since α, μ_i and λ_t are parameters to be estimated and it is not possible to estimate them separately without imposing an additional restriction.

One of the main advantages of fixed effect estimation is to allow for the possible correlation between the unobservable effects and the explanatory variables. Actually, this is one of the main reasons of using panel data estimation method in empirical studies. Nevertheless, if variables being observable, but individual or time invariant exist in the regression equation as additional explanatory variables such as time dummies; then the procedure of within regression estimation eliminates the effects of those variables. The random effect estimation, on the other hand, would not produce such a problem; because it assumes the unobservable effects, μ_i and λ_t, being random and hence requires different methods of estimation.

Now, assume that the $\mu_i \sim IID(0,\sigma_\mu^2)$, $\lambda_t \sim IID(0,\sigma_\lambda^2)$ and $v_{it} \sim IID(0,\sigma_v^2)$ independent of each other. Then, (5.3.2) represents a two way random effect error component model. In addition, assume that the x_{it}'s are independent from the λ_t, μ_i and v_{it} for all i and t. Then, the disturbances, u_{it}, are homoskedastic with $var(u_{it}) = \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2$ for all i and t. Since $cov(u_{it}, u_{js})$ equals to σ_μ^2 when $i = j$, $t \neq s$; equals to σ_λ^2 when $i \neq j$, $t = s$ and equals to 0 otherwise; the correlation coefficient is given by $correl(u_{it}, u_{js}) = \sigma_\mu^2 / (\sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2)$ if $i = j$, $t \neq s$; $correl(u_{it}, u_{js}) = \sigma_\lambda^2 / (\sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2)$ if $i \neq j$, $t = s$; $correl(u_{it}, u_{js}) = 1$ if $i =j$, $t = s$ and $correl(u_{it}, u_{js}) = 0$ if $i \neq j$, $t \neq s$.

Since λ_t and μ_i are in the composite error for all i and t, the u_{it} are serially correlated across time and individuals as can be seen above. Such a correlation will lead biased estimators when pooled OLS estimation of (5.3.2) is used. In order to get rid of this correlation one can use feasible generalized least squares (GLS) estimation for the two way error component model. The resulting estimate is the GLS estimate of β, $\hat{\beta}_{GLS}$ (Baltagi, 2001: 11-38).

The within estimation of β is unbiased and consistent whether the individual and time specific effects are treated as fixed or random. Under the assumption of the individual and time specific effects being parameters, the within estimation results in best linear unbiased estimator (BLUE) for β; however, the fixed effect estimation do not produce BLU estimator of β in finite samples when those effects are assumed as random. On the other hand, the GLS estimator of β is, in fact, BLUE when the individual and time specific effects are random (Hsiao, 1986: 34).

5.3.1. Advantages and Disadvantages of Panel Data Estimation

The main advantage of using panel data estimation is that it provides a means of eliminating the possible omitted variable biased in a regression equation due to the individual and/or time specific heterogeneity. Consider the one-way error component model given as $y_{it} = \alpha + \beta x_{it} + \mu_i + v_{it}$, where t = 1,2 and the remainder disturbances are stochastic with $v_{it} \sim IID(0,\sigma_v^2)$. Then, one could obtain $y_{i1} = \alpha + \beta x_{i1} + \mu_i + v_{i1}$ for t =1 and, $y_{i2} = \alpha + \beta x_{i2} + \mu_i + v_{i2}$ for t =2.

The simple OLS regression of one of those two equations results in biased estimators because of the omission of the individual specific variable, μ_i.[43] Moreover, the pooling of two time periods by regarding them as independent could not produce a different result than a biased estimator for β. However, by differencing the latter equation from the former, one could get $y_{i1}- y_{i2} = \beta(x_{i1}- x_{i2}) + (v_{i1}- v_{i2})$ and the estimation of the difference equation yields an unbiased estimator of β due to the elimination of the unobserved individual effect, μ_i. Furthermore, the GLS estimation of $y_{it}=\alpha+\beta x_{it}+\mu_i+v_{it}$, within the random effect estimation framework, also produces unbiased estimates provided that the individual specific effect is random. That is, the use of panel data estimation helps the control for the risk of obtaining biased results.

One of the useful methods of panel data estimation is that the fixed effect estimation allows for unobservable effects to be correlated with the explanatory variables

[43] Since μ_i is unobservable, we are not able to include such a variable in a simple cross section regression and the misspecification yields biased estimators. This is because in the OLS estimation, μ_i is jointly estimated within the intercept term.

(Wooldridge, 2002:421).[44] For instance, in our model the estimation of a regression of real GDP on different educational attainments of the labor force and the physical capital will be performed with different methods of estimation including fixed and random effect estimations. In that case, i's are the provinces of Turkey and t represents time. However, the estimation of such an equation with within estimation could be more useful because, for instance, it is more likely for the schooling of the workforce to be correlated with the capacity of human resources, which are specific to each province and unobservable, affecting the level of output.

Moreover, by providing a large number of data points, panel data sets are able to increase the degrees of freedom and reduce the collinearity among the explanatory variables thereby improving the efficiency of econometric estimates (Hsiao, 1986:1). For instance, time series data sets generally face with the problem of multicollinearity; however this is less likely with a panel data set as the cross section dimension adds a lot of variability and hence reducing the probability of such problem. With the addition of more data points, panel data sets are more informative and thus produce more reliable and precise parameter estimates than pure cross section or pure time series estimation (Baltagi, 2001: 6). Hsiao (1986:3) mentions that a single cross section is likely to produce less precise estimates due to reflecting interindividual differences, and single time series data estimation also usually provide less precise estimates of dynamic coefficients.

Besides those benefits, the panel data estimation provides a better tool for studying dynamics of change. To illustrate, with a labor market survey of TÜİK for a single time period, one could not able see the transitions among the labor market status of an individual unless such a question about the previous status is included. It is just possible to make inference about the status of individuals at that time period such as

[44] It must be noted that the allowance for such a correlation is valid only for the cases of first-differenced and fixed effect estimations. The random effect estimation is based on the assumption of zero correlation between the explanatory variables and the unobserved individual or time specific effects.

employment and unemployment levels. On the other hand, it would be desirable to see the effects of business cycles such as crises and booms on the change of labor market status of individuals to determine which policy will be implemented by the authorities; however such information cloud not be obtained from a single cross section data. In fact, by following individuals over time as they change their status with the help of panel data, it is possible to construct a recursive structure to study the before/after effect (Hsiao, 1986:3).

It should be noted that panel data estimation is not a "miracle" and it may not be useful all the time despite the several aforementioned advantages of it. In fact, some problems of using panel data may arise due to the nature of the data. For example, the collection of such data may include problems of nonresponse, misrecording of responses, memory errors and inappropriate informants and those are likely to lead to presence of measurement error. Misleading results could be obtained as a result of significant measurement errors and one should take into account such possibility when using panel data sets. Furthermore, typical panels contain annual data and cover short span of time periods for each individual relative to the number of individual units, so, in that cases, asymptotic arguments depend on the number of individuals tending to infinity (Baltagi, 2001:9).

One of the disadvantages of the panel data estimation could arise from "selectivity". That is, a panel data may not be randomly drawn from a large population and such a nonrandomness of panel data could produce biased estimates, which is known as selectivity bias. On the other hand, such a situation is not only peculiar to estimation of panel data, but also valid for other types of data such as cross sectional data (Hsiao, 1986:7-8).

Moreover, in order to be specific, the use of panel data estimation is not so widespread in Turkey because of the difficulty of availability of such data, and this makes a study for Turkey about any subject depending on analysis of panel data less possible.

5.3.2. Panel Data Estimation with Fixed and Random Effects

In this study, the model of Knowles (1997) and the augmented Solow model with different levels of education will be estimated with OLS and panel data estimation methods. On the other hand, it is important to determine which method is better. So, in order to find an appropriate estimation method for each model a sequence of hypothesis will be tested. Indeed, OLS estimation of those models versus fixed and random effect estimations will be tested. After testing those hypotheses; if the panel data estimation is the preferable one, then fixed effect versus random effect estimation of the models will be tested by the procedure introduced by Hausman (1978). In that respect, fixed and random effect estimation methods will be discussed rigorously in this section. At the same time, the test statistics for hypothesis used to find proper ways of estimation will be presented.

5.3.2.1. Fixed Effect Estimation

In a two way error component model, a simple panel data regression which is given by the equation (5.3.2) implies the following

$$y_{it} = \alpha + \beta x_{it} + \mu_i + \lambda_t + v_{it}, \qquad (5.3.7)$$

where $i = 1,2,\ldots,N$ and $t = 1,2,\ldots,T$

The estimation of (5.3.7) with the assumption of individual and time specific effects, captured in μ_i and λ_t, being parameters yields fixed effect estimation for this equation. Furthermore, at the beginning of the section (5.3), the calculation of the fixed effect estimator of β is introduced. In addition, as will be discussed below, it is possible to estimate β by using intercept dummies, and the models of interest will be estimated in this study by using the dummy variables method for the two-way error component model.[45]

Then, assume that the equation (5.3.7) has fixed slope coefficients over time and across individuals, but variable intercept over time and across individual units. In that case, the difference of the intercepts could be represented with N+T-2 (N-1 dummies to represent heterogeneity across individuals and T-1 dummies to represent heterogeneity across time periods) dummies.[46] That is, the equation (5.3.7) could be written in the form of

$$y_{it}=\alpha_1+\alpha_2D_{2i}+\alpha_3D_{3i}+\ldots+\alpha_ND_{Ni}+\gamma_1+\gamma_2D_{2t}+\gamma_3D_{3t}+\ldots+\gamma_TD_{Tt}+\beta x_{it}+v_{it}, \tag{5.3.8}$$

for i = 1,…, N and t = 1,…,T, where, the error term ,v_{it}, is assumed to be distributed normally with zero mean and constant variance (Gujarati, 2003:642). In the equation (5.3.8), the dummy variable for individuals D_{ji} takes the value of 1 when i=j for all j =2,…,N and i = 1,…, N; and the dummy variable for time D_{kt} takes the value of 1 when t=k for all k = 2,…,T and t = 1,…, T. In that case, the first individual unit and the first time period are taken as the bases, that is the estimation of α_1 represents the estimated intercept for the first individual unit, and α_2, α_3, …, and α_N represent the estimated differential intercept coefficients (from α_1) of second individual, third individual, …, and the N^{th} individual respectively; and similarly, the estimation of γ_1 represents the estimated intercept for the first time period and γ_2, γ_3, …, and γ_T represent the estimated differential intercept coefficients (from γ_1) of second time period, third time period, …, and the T^{th} time period respectively.

[45] The fixed effect estimation is also called least squares dummy variable estimation in literature.

[46] Since there exists N individual units and T time periods; N-1 and T-1 dummy is taken for individuals and time periods respectively in order to avoid the dummy variable trap.

The within estimation of (5.3.7) and the OLS estimation of (5.3.8) produce same results in the sense that; the estimation of α_1 corresponds to the estimation of μ_1; and the estimation of α_1 plus the estimation of α_2 corresponds to the estimation of μ_2; and so on. Similarly, regarding to the intercepts of the time periods, the estimation of γ_1 corresponds to the estimation of λ_1; and the estimation of γ_1 plus the estimation of γ_2 corresponds to the estimation of λ_2; and so on. Although the two methods generate the same estimators, some differences would occur because of the structure of the regression equations. For example, the estimation of (5.3.8) has the chance of reducing the possibility of multicollinearity among the explanatory variables; however, it could suffer from enormous loss of degrees of freedom due to having too many dummies (Baltagi, 2001:32).

Then, one could test the joint significance of the intercept dummies in order to determine whether the fixed effect estimation is better than OLS by performing an F test. In this regard the hypothesis of H_0^4 will be tested in Chapter 7, where H_0^4:$\mu_1=\mu_2=....=\mu_{N-1}= 0$ and $\lambda_1=\lambda_2=....=\lambda_{T-1}= 0$.

The RRSS is obtained from the OLS estimation of the equation (5.3.7) and the URSS is obtained from the within estimation of the same equation with individual and time specific effects. Then, the F statistics needed to test H_0^4 is given by $F_4=\dfrac{(RRSS - URSS)/(N+T-2)}{URSS/((N-1)(T-1)-1)} \sim F_{(N+T-2),\ (N-1)(T-1)-1}$.[47]

The null hypothesis is rejected whenever $F_4 > F_{(N+T-2),\ (N-1)(T-1)-1}$ implying that the within estimation is preferred to the OLS estimation of the equation (5.3.7) since at least one of the individual and time specific effects exist. On the other hand, the non-rejection of the hypothesis H_0^4 reflects the fact that the OLS estimation of (5.3.7) is the better one.

[47] The F statistics is from Baltagi (2001:32).

In the case of the rejection of H_0^4, one could check for significance of individual and time specific effects separately. For instance, after testing H_0^4, the existence of individual specific effects given the presence of time specific effects could be tested. In this regard the hypothesis of H_0^5 will be tested in Chapter 7, where $H_0^5: \mu_1=\mu_2=\ldots=\mu_{N-1}=0$ such that $\lambda_t \neq 0$ for $t=1,2,\ldots,T-1$.

The RRSS could be obtained from the within estimation of the equation (5.3.7) with time specific effects, that is, it could be obtained from the OLS estimation of $(y_{it} - \bar{y}_{t.}) = \beta(x_{it} - \bar{x}_{t.}) + (v_{it} - \bar{v}_{t.})$, and the URSS is the same as the one for the test of the null hypothesis H_0^4 (Baltagi, 2001:33). The rejection of the null hypothesis depends on the condition that $F_5 > F_{(N-1),(N-1)(T-1)-1}$ where $F_5 = \dfrac{(RRSS - URSS)/(N-1)}{URSS/((N-1)(T-1)-1)} \sim F_{(N-1),(N-1)(T-1)-1}$.[48]

The rejection of H_0^5 implies that given the presence of time specific effects, the individual specific effects, captured in μ_i's, do exist; while the nonrejection of the null hypothesis of H_0^5 supports the absence of individual specific effects.

Similarly, one could test for the presence of time specific effects given the existence of individual specific effects. In this regard the hypothesis of H_0^6 will be tested in Chapter 7, where $H_0^6: \lambda_1=\lambda_2=\ldots=\lambda_{T-1}= 0$ such that $\mu_i \neq 0$ for $i=1,2,\ldots,N-1$.

URSS could still be obtained from the within regression of (5.3.7) with individual and time specific effects; however, RRSS is obtained from the within estimation of the equation (5.3.7) with individual specific effects, that is, it is obtained from the OLS estimation of $(y_{it} - \bar{y}_{i.}) = \beta(x_{it} - \bar{x}_{i.}) + (v_{it} - \bar{v}_{i.})$ (Baltagi, 2001:33). Then, the F

[48] The F statistics is from Baltagi (2001:33).

statistics required to test the hypothesis of H_0^6 is given by $F_6 = \frac{(RRSS - URSS)/(T-1)}{URSS/((N-1)(T-1)-1)} \sim F_{(T-1),\,(N-1)(T-1)-1}$.[49]

The rejection of the null hypothesis H_0^6 is again requires $F_6 > F_{(T-1),\,(N-1)(T-1)-1}$; and in the case of such inequality, the time specific effects, captured in λ_t's, do exist given the presence of the individual specific effects. On the other hand, the non-rejection of H_0^6 supports the absence of time specific effects.

The rejection of H_0^4 (the null hypothesis of the absence of individual and time specific effects) implies that there exist time or individual specific effects; so in the case of rejecting H_0^4 one could prefer the within estimation of a given equation rather than OLS estimation of the same equation. In that case, it could be determined to use one-way (including only individual or only time specific effects) or two-way (including both individual and time specific effects) error component models from the results of H_0^5 and H_0^6.

5.3.2.2. Random Effect Estimation

In this study, the random effect method of estimation will be performed as well as the fixed effect estimation. So, consider the equation (5.3.7) with the assumption of individual and time specific effects being random variables, rather than fixed parameters. First of all, it should be emphasized that the main assumption required for the random effect estimation of the equation (5.3.7) is the independency of the explanatory variables with μ_i, λ_t and v_{it} for all i and t. In a two-way error component model, it is assumed that μ_i, λ_t and v_{it} distributed normally with zero means and constant variances σ_μ^2, σ_λ^2 and σ_v^2 respectively.[50]

---

[49] The F statistics is from Baltagi (2001:33).

[50] Because $var(u_{it}) = \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2$ for all i and t, the variances σ_μ^2, σ_λ^2 and σ_v^2 are called variance components and such a model is sometimes referred to as variance-components (error-component) model.

Since u_{it} and u_{js} both contain μ_i and λ_t, the residuals of (5.3.2) are correlated (Hsiao, 1986:34). Therefore, as previously mentioned, the estimation of this equation could be possible with a feasible GLS estimation which eliminates the serial correlation in the error term, u_{it}, of the equation (5.3.2). In order to obtain the GLS estimation of the regression coefficients, one needs the variance-covariance matrix of the residuals, which is given by $V=\sigma_\mu^2(I_N \otimes J_T)+\sigma_\lambda^2(J_N \otimes I_T)+\sigma_v^2(I_N \otimes I_T)$, where I_N and I_T are the identity matrices of dimension N and T; and J_N and J_T are the matrices of ones of dimension N and T respectively.[51] Then, the spectral decomposition of the variance-covariance matrix is $V = \sum_{i=1}^{4} \delta_i Q_i$ where δ_i is the distinct characteristic roots of V and Q_i is the corresponding eigenprojectors for i =1, 2, 3, 4; and $\sigma_v V^{-1/2} = \sum_{i=1}^{4} (\sigma_v/\delta_i^{1/2}) Q_i$ (Baltagi, 2001: 34).[52]

Then, a proper transformation of the equation (5.3.2) in the matrix form could be obtained by multiplying it with the matrix by $\sigma_v V^{-1/2}$, that is, a proper transformation of this equation could be given by

$$y_{it}^* = \alpha^* + \beta x_{it}^* + u_{it}^* \qquad (5.3.9)$$

where, $y_{it}^*=y_{it}-\theta_1 \bar{y}_{i.} - \theta_2 \bar{y}_{.t} + \theta_3 \bar{y}_{..}$, $x_{it}^*= x_{it}-\theta_1 \bar{x}_{i.} - \theta_2 \bar{x}_{.t} + \theta_3 \bar{x}_{..}$, and $u_{it}^*=u_{it}-\theta_1 \bar{u}_{i.} - \theta_2 \bar{u}_{.t} + \theta_3 \bar{u}_{..}$. In that case, $\theta_1 = 1-(\sigma_v/ \delta_2^{1/2})$, $\theta_2=1-(\sigma_v/ \delta_3^{1/2})$, and $\theta_3 = \theta_1+\theta_2+(\sigma_v/ \delta_4^{1/2})-1$. The error terms in the equation of (5.3.9) are serially uncorrelated and OLS estimation of it simply produces the GLS estimator for β (Baltagi, 2001: 34).

---

[51] $\otimes$ denotes the Kronecker product.

[52] In that case, the eigenvalues of the variance-covariance matrix are given by $\delta_1 = \sigma_v^2$, $\delta_2 = T\sigma_\mu^2 + \sigma_v^2$, $\delta_3 = N\sigma_\lambda^2 + \sigma_v^2$ and $\delta_4 = T\sigma_\mu^2 + N\sigma_\lambda^2 + \sigma_v^2$; and the corresponding eigenvectors are given by $Q_1 = E_N \otimes E_T$, $Q_2 = E_N \otimes \bar{J}_T$, $Q_3 = \bar{J}_N \otimes E_T$ and $Q_4 = \bar{J}_N \otimes \bar{J}_T$, where E_N and E_T are $I_N - \bar{J}_N$ and $I_T - \bar{J}_T$, and $\bar{J}_N$ and $\bar{J}_T$ are J_N/N and J_T/T respectively (Baltagi, 2001: 12-34).

Since the variance components are unknowns, the estimation of δ_i's are required for the GLS estimation of the equation (5.3.2); and after estimating δ_i's one could obtain the estimates for θ_i's. The best quadratic unbiased estimators of δ_i is given by $\hat{\delta}_i = u'Q_iu/tr(Q_i)$, for i = 2, 3, 4, and this can be obtained by replacing the true disturbances with OLS or within residuals (Baltagi, 2001: 34).[53]

As in the case of fixed effect estimation, one could test whether the random effect estimation is better than OLS estimation for the equation (5.3.7); and in this study the joint significance of the variances of individual and time specific effects, σ_μ^2 and σ_λ^2, will be tested for our models. The joint significance of the variances implies that at least one of the individual or time specific effects exist, and hence it indicates that the usage of random effect estimation method is more preferable than the OLS estimation. The joint significance of the variance components will be tested with a Lagrange Multiplier (LM) test which is introduced by Breusch and Pagan (1980).[54]

After the Breusch-Pagan LM test supports the estimation of the equation (5.3.7) with method of random effects, the individual significance of the variance components could be tested. The significance of σ_μ^2 and σ_λ^2 will be tested for two models being estimated in this thesis with appropriate LM statistics in Chapter 7. In this regard, the significance of σ_μ^2 could be tested with the hypothesis H_0^7:σ_μ^2=0 such that σ_λ^2>0.

In that case, because the variance of individual specific effect is nonnegative, the alternative hypothesis is one-sided. In addition, the assumption of σ_λ^2 being positive is important in the sense that the exclusion of such condition in the null hypothesis implicitly assumes that time specific effect does not exist. Then, the hypothesis H_0^7 could be tested with the LM statistics which is defined by

---

[53] In addition to the way that has been just defined, many other feasible GLS estimators of β are available in literature.

[54] The derivation of the test statistics will be stated in the next section.

$$\mathrm{LM}_\mu = \frac{\sqrt{2}\tilde{\sigma}_2^2\tilde{\sigma}_V^2}{\sqrt{T(T-1)[\tilde{\sigma}_V^4 + (N-1)\tilde{\sigma}_2^4]}}\,\tilde{D}_\mu$$

where, $\tilde{D}_\mu = T/2\left\{\frac{1}{\tilde{\sigma}_2^2}\left[\frac{\tilde{u}'(\bar{J}_N \otimes \bar{J}_T)\tilde{u}}{\tilde{\sigma}_2^2} - 1\right] + \frac{N-1}{\tilde{\sigma}_V^2}\left[\frac{\tilde{u}'(E_N \otimes \bar{J}_T)\tilde{u}}{(N-1)\tilde{\sigma}_V^2} - 1\right]\right\}$

with $\tilde{\sigma}_2^2 = \tilde{u}'(\bar{J}_N \otimes I_T)\tilde{u}/T$, $\tilde{\sigma}_v^2 = \tilde{u}'(E_N \otimes I_T)\tilde{u}/T(N-1)$, and the estimated disturbances $\tilde{u}$ represents the one-way GLS residuals. LM_μ is asymptotically distributed as N(0,1) under the null hypothesis (Baltagi, 2001: 62). If the test statistics LM_μ is greater than the critical value, then the null hypothesis will be rejected. Moreover, the rejection of hypothesis H_0^7 implies that there exists an individual specific effect given the presence of the time specific effect.

Similarly, the existence of time specific effect could be tested given the presence of the individual specific effect. In that case, the corresponding hypothesis is H_0^8: $\sigma_\lambda^2 = 0$ such that $\sigma_\mu^2 > 0$. As in the previous case, the alternative hypothesis is one-sided, since the variance of time specific effect must be positive provided that such effect exists. One could test the hypothesis H_0^8 with the LM statistics which is defined by

$$\mathrm{LM}_\lambda = \frac{\sqrt{2}\tilde{\sigma}_1^2\tilde{\sigma}_V^2}{\sqrt{N(N-1)[\tilde{\sigma}_V^4 + (T-1)\tilde{\sigma}_1^4]}}\,\tilde{D}_\lambda$$

where, $\tilde{D}_\lambda = N/2\left\{\frac{1}{\tilde{\sigma}_1^2}\left[\frac{\tilde{u}'(\bar{J}_N \otimes \bar{J}_T)\tilde{u}}{\tilde{\sigma}_1^2} - 1\right] + \frac{T-1}{\tilde{\sigma}_V^2}\left[\frac{\tilde{u}'(\bar{J}_N \otimes E_T)\tilde{u}}{(T-1)\tilde{\sigma}_V^2} - 1\right]\right\}$

with $\tilde{\sigma}_1^2 = \tilde{u}'(I_N \otimes \bar{J}_T)\tilde{u}/N$, and $\tilde{\sigma}_v^2 = \tilde{u}'(I_N \otimes E_T)\tilde{u}/N(T-1)$. LM_λ is asymptotically distributed as N(0,1) under the null hypothesis (Baltagi, 2001: 62). If the test statistics LM_λ is greater the critical value, then the null hypothesis of H_0^8 will be rejected. Furthermore, the rejection of the hypothesis H_0^8 implies that there exists time specific effect given the presence of the individual specific effect.

In order to determine the usage of one-way or two-way error component models within the random effect estimation method, it could be useful to test both hypotheses of H_0^7 and H_0^8 in addition to the joint significance test. However, the test for joint significance of the variance components is the first step, and the LM test developed by Breusch and Pagan (1980) is an appropriate test for a hypothesis that assumes the variance components of the individual and time specific effects are zero.

5.3.2.3. Breusch Pagan Test

Breusch and Pagan (1980) derive an LM test for a two-way random error component model in order to test the hypothesis of H_0^9: $\sigma_\mu^2 = \sigma_\lambda^2 = 0$.

In an error-component model, the zero variance components of individual and time specific effects imply that those effects do not exist; since the zero mean of those effects, which is assumed under the model, together with zero variances means that all such variables are not different from zero. Thus, the rejection of the null hypothesis of H_0^9 implies the existence of unobservable effects which are assumed to be random.

The LM statistics, which will be calculated for our model in Chapter 7, is needed to test H_0^9 and known as the Breusch-Pagan LM statistics. It is given by the sum of following two LM statistics: LM_1 and LM_2, where

$$LM_1 = \frac{NT}{2(T-1)}\left[1-\frac{\tilde{u}'(I_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}}\right]^2 \text{ and } LM_2 = \frac{NT}{2(N-1)}\left[1-\frac{\tilde{u}'(J_N \otimes I_T)\tilde{u}}{\tilde{u}'\tilde{u}}\right]^2$$

with $\tilde{u}$ being the OLS residuals. That is, $LM_{BP}=LM_1+LM_2$. Under the null hypothesis of H_0^9, the LM_{BP} statistics is asymptotically distributed as χ^2_2. The computation of this test statistics is easy since it includes only the OLS residuals, and many software packages are able to implement the Breusch and Pagan LM test for random effects (Baltagi, 2001: 58-60).

It should be noted that the components of LM_{BP} (LM_1 and LM_2) correspond to the LM statistics obtained when one desires to test the hypotheses of H_0^{10}:σ_μ^2=0 and H_0^{11}:σ_λ^2=0 respectively. Those statistics are each asymptotically distributed as χ_1^2 under H_0^{10} and H_0^{11} respectively (Baltagi, 2001: 59). The statistics LM_1 and LM_2 could be obtained in the case of testing H_0^{10} and H_0^{11}; however, those are not equal to LM_μ and LM_λ because H_0^{10} and H_0^{11} ignore the variance components of σ_λ^2 and σ_μ^2 are positive, that is, contrary to the null hypotheses of H_0^7 and H_0^8, the conditions σ_λ^2>0 and σ_μ^2>0 are not included in null hypotheses of H_0^{10} and H_0^{11} respectively.

If $LM_{BP} > \chi_2^2$ at a proper significance level, then one rejects the null hypothesis of H_0^9 implying that the random effect estimation for a given equation is preferred to the OLS estimation of the same equation due to the existence of individual or time specific effects which are assumed as random.

5.3.2.4. Hausman Test

After testing fixed and random effect estimations versus OLS estimation, one could test which of the two methods is more appropriate for the estimation of a given model provided that both fixed and random effects estimation are preferred to OLS estimation. Deciding between the usage of fixed and random effects depends on whether the unobservable effects are best viewed as parameters or random variables. In the case where the unobservable heterogeneity could not be seen as random variables from a large population, fixed effect estimation method could be more appropriate. In addition, even if the unobservable variables are assumed to be random, this does not mean that they are in fact random. The explanatory variables and the unobservables should be uncorrelated in order to make random effect estimation; that is, the correlation of those variables, which hampers the usage of random effect estimation method, makes the fixed effect estimation method more appropriate. In fact, if such a correlation exists, random effect estimation will produce inconsistent estimators (Wooldridge, 2002:452-453).

In that regard, the test of whether fixed or random effect estimation is more appropriate could be a test of whether the explanatory variables and the unobservables are correlated (Wooldridge, 2002: 453). The null hypothesis of the absence of correlation between the explanatory variables and the unobservable variables is given by H_0^{12}: $E(u_{it}/x_{it}) = 0$ versus H_a^{12}: $E(u_{it}/x_{it}) \neq 0$, where the u_{it} is the composite error of the equation (5.3.2) (Baltagi, 2001: 65). Under the assumption that the x_{it} and v_{it} are uncorrelated, $E(u_{it}/x_{it}) = 0$ implies that the x_{it} is uncorrelated with the unobservable variables.

Hausman (1978) suggests a test statistics for the hypothesis of H_0^{12} which is based on the difference between the fixed effect and GLS estimator for β, and it is given by $m = \hat{q}' \left[\mathrm{var}(\hat{q}')\right]^{-1} \hat{q}$, where, $\hat{q}$ is the difference between the within estimator and GLS estimator of β; that is, $\hat{q} = \hat{\beta}_{GLS} - \widetilde{\beta}_{within}$.[55] The test statistics m is asymptotically distributed as χ_1^2 since the regression equation being interest has only one explanatory variable. On the other hand, when there exist more than one explanatory variable, the test statistics introduced by Hausman (1978) is asymptotically distributed as χ_k^2, where k is the number of explanatory variables (Baltagi, 2001: 65-66).

In that case, if m is greater than the critical value, then one is able to reject the null hypothesis of no correlation between the explanatory variable and the unobservables. That is, one may conclude that the random effect estimation method is less

[55] The F statistics is from Baltagi (2001:66).

appropriate and it is better to use the fixed effect estimation method to make more appropriate implications about the coefficient estimates for the models of interest.

CHAPTER 6

DATA SOURCE AND VARIABLES

In this chapter, firstly the data source used in the empirical part of this thesis will be mentioned. Namely, the variables employed in the model of Knowles (1997) and the augmented Solow model with different levels of education, which are estimated in this study, will be presented.

6.1. Data Source

The main data source employed in the present study is the population censuses demonstrating the social, demographic and economic structure of Turkey. They are conducted by TÜİK. The first population census was carried out in 1927 and second was in 1935. After 1935, the censuses were carried out in every five-year periods up to 1990 and decennially after that time (Devlet İstatistik Enstitüsü, 2003). In our work, the census data covering the years 1975, 1980, 1985, 1990 and 2000 for education variables are used.[56] Those variables are available at the province level and the province level data will be used to construct a panel in this study. There were 67 provinces in 1975, 1980 and 1985; 71 provinces in 1990 and 81 provinces in 2000. Hence, to avoid the use of an unbalanced panel, the data are redesigned to include only 67 provinces in this study.

Moreover, the data for nominal GDP series in the year 1975 are obtained from Özötün (1980) and the data for nominal GDP series in the years 1980 and 1985 are obtained from Özötün (1988). In addition, TÜİK provides the nominal GDP series for the years 1990 and 2000. However, in this study, the real GDP (1987 based) for

[56] Since the data is not available for the variables used in this work in the years before 1975 and after 2000, this study is restricted to the years 1975, 1980, 1985, 1990 and 2000.

the years 1975, 1980, 1985, 1990 and 2000 are used and so those series are constructed from the nominal GDP series.

The other sources of data are from the publications of Turkish Electricity Authority (TEK) and from the Turkish Electricity Distribution Company (TEDAŞ) which provide the physical capital variable employed in the empirical work. The physical capital variable is also available at the province level for the years 1975, 1980, 1985, 1990 and 2000.

6.2. Variables

In order to determine the determinants of output level of Turkey in this study the model introduced by Knowles (1997) and the augmented Solow model with different levels of education are estimated; a human capital variable, in the form of education and a physical capital variable are required. Hence, in this section, those variables used in the estimation of our models will be presented.

In this study, it is planned to explain the determinants of real GDP (at constant 1987 prices) and real GDP per workforce (at constant 1987 prices) in Turkey with the estimates depending on the model of Knowles (1997) and the augmented Solow model respectively. Since the province level data used in the present study, the real GDP (and real GDP per workforce) could be called as real Gross Provincial Product (GPP) (and real GPP per workforce). Thereafter, real GDP (and real GDP per workforce), which are available at the province level, are called as real GPP (and real GPP per workforce).

In the empirical models estimated in this study, the human capital effect will be measured by using the *educational attainments* at the province level from the censuses of population in 1975, 1980, 1985, 1990 and 2000. The educational attainment of the labor force aged 12 and over will be differentiated with respect to the highest level of schooling completed. In the original data obtained from TÜİK, the labor force is divided into six groups with respect to their educational attainment: illiterate, literate without diploma, primary, middle and high school graduates and higher education graduates. The educational attainment of the labor force in this study is simply differentiated into four main groups: nongraduates, basic, secondary and higher level graduates. The "nongraduates" in the labor force are illiterates and the literates without a diploma (L_1), and the graduates from "basic education" are the ones in the workforce completing primary or middle school (L_2) which is a total of 8 years of schooling. The labor force that completes high school and higher education are named as "secondary level" (L_3) and "higher level" (L_4) respectively.

In this study, the industrial electricity consumption, which excludes commercial, residential and service sector electricity consumption, is used to proxy for the physical capital variable. The main reason for using this as a physical capital proxy is that the panel data is desired to be constructed in the present study and the industrial electricity consumption, which is available at the province level in the years 1975, 1980, 1985, 1990 and 2000, is the best choice among the limited choices. The industrial electricity consumption for each province is available in each year. Nevertheless, the industrial electricity consumption for the provinces Bingöl and Hakkari are approximately zero in 1975. This could be due to the fact that main production activity in these provinces based mainly on agriculture which uses little electricity. In addition, when the models in this study are estimated, the natural logarithm of each variable is used. Hence, most of the estimations in this study are based on 65 provinces rather than 67. Indeed, the natural logarithm of zero is undefined for the provinces Bingöl and Hakkari and therefore those observations are excluded from the regression estimates where logarithms of the variables are taken.

CHAPTER 7

ESTIMATION RESULTS

In this chapter, the model of Knowles (1997) and the augmented Solow model with different levels of education introduced in the section (2.3) will be estimated. Firstly, the results of the hypothesis tests for poolability for both models will be put forward. Then, the models will be estimated with the methods of single OLS, pooled OLS, and fixed and random effects estimations. Simultaneously, the hypothesis tests previously introduced in the section (5.3.2) will be performed in order to determine a proper way of estimation for each model.

7.1. Results of Hypothesis Tests for Poolability

First of all, consider the Cobb-Douglas production function that Knowles (1997) uses, $Y = AK^{\alpha}L_1^{\beta 1}\ L_2^{\beta 2}\ L_3^{\beta 3}\ L_4^{\beta 4}$. In our case, Y represents real GPP at constant 1987 prices, K is the physical capital proxied by the industrial electricity consumption, A represents the level of technology, L_1 is the number of persons in the labor force, aged 12 years and older, who have no formal schooling and L_2, L_3 and L_4 are those in the labor force who complete the basic, secondary and higher level of education respectively.[57] Taking the natural logarithm of both sides of this equation one gets, $\text{In}Y=\text{In}A+\alpha\text{In}K+\beta_1\text{In}L_1+\beta_2\text{In}L_2+\beta_3\text{In}L_3+\beta_4\text{In}L_4$. Then, within the panel data framework this equation could be written as

$$\text{In}Y_{it} = \text{In}A_{it}+\alpha\text{In}K_{it}+\beta_1\text{In}L_{1it}+ \beta_2\text{In}L_{2it}+\beta_3\text{In}L_{3it}+\beta_4\text{In}L_{4it}+ v_{it} \qquad (7.1.1)$$

[57] For the variables proxying different levels of education in the Knowles's study, it is enough for a worker to enter a given level of schooling, not necessarily complete, to characterize that level; however, in our study, in order to characterize a level, one should complete that level. Hence, the education variables used in this thesis represent a more qualitative dimension of schooling relative to those employed by Knowles.

where i is the provinces of Turkey and t is the time period, for i =1, 2,….,65 and t =1975, 1980, 1985, 1990, 2000. Assume that the level of technology grows at a constant rate of g for each province, that is, $A_{it} = A_i(0)e^{gt}$ for i = 1,2,…,65, where $A_i(0)$ is the initial level of technology for the province i. Then, $\ln A_{it} = \ln A_i(0) + gt$ implies that[58]

$$\ln Y_{it} = \ln A_i(0) + gt + \alpha \ln K_{it} + \beta_1 \ln L_{1it} + \beta_2 \ln L_{2it} + \beta_3 \ln L_{3it} + \beta_4 \ln L_{4it} + v_{it} \qquad (7.1.2)$$

The equation (7.1.2) could be viewed as a two-way error component regression model in the form $y_{it} = \alpha + \beta x_{it} + \mu_i + \lambda_t + v_{it}$, where y_{it} is the natural logarithm of the real GPP; x_{it} is a 5×1 vector of explanatory variables, natural logarithm of the amounts of labor force with no formal schooling and with schooling at basic, secondary and higher levels, and the natural logarithm of industrial electricity consumption and β is a 1×5 vector of coefficients. The individual and time specific effects, μ_i and λ_t, could be captured in the term $\ln A_i(0)$, the natural logarithm of initial technology specific to each province, and the term gt respectively. Then, the equation (7.1.2) could be estimated with two-way fixed or random effect estimation methods. However, those estimations are based on the assumption of poolability.

The homogeneity of slope coefficients and intercept term of equation (7.1.2) could be tested for poolability with the F tests described in section (5.2). Then, for simplicity, assume that the parameters in this equation are constant over time and variable across individual units, that is, assume the equation (7.1.2) is in the form $y_{it} = \alpha_i + \beta_i x_{it} + u_{it}$, where u_{it} is the composite error and equals to the sum $\mu_i + \lambda_t + v_{it}$.

Now, consider the hypothesis of $H_0^1: \beta_1 = \beta_2 = ... = \beta_{65}$. In that case, the RRSS, as described in the section (5.2), could be obtained from the sum of RSS_i of the OLS estimation of $y_{it} - \bar{y}_{i.} = \beta_i(x_{it} - \bar{x}_{i.}) + (u_{it} - \bar{u}_{i.})$ for each province i = 1,2,…,65. However,

[58] In the section (7.2.1), the equation (7.1.2) will be estimated with a constant term, that is, when the equation (7.1.2) estimated one should remember that the constant term is included.

for each equation the degrees of freedom is not enough in time dimension for the estimation; since there are 5 observations and 6 parameters to be estimated for each regression, and this contradicts the rule N>k+1 for a regression estimate of N observation with k explanatory variable (because 5<6).

In addition, there is a similar problem of computing the F statistics, F_3, for the test of the null hypothesis H_0^3: $\alpha_1 = \alpha_2 = \ldots = \alpha_{65}$ and $\beta_1 = \beta_2 = \ldots = \beta_{65}$. Since the unrestricted model for this hypothesis is the same as in the previous case, it is also not possible to calculate F_3 due to the reason that the degrees of freedom is not enough in time dimension for the individual equations $y_{it} - \overline{y}_{i.} = \beta_i(x_{it} - \overline{x}_{i.}) + (u_{it} - \overline{u}_{i.})$.

However, it is possible to test the null hypothesis H_0^2: $\alpha_1 = \alpha_2 = \ldots = \alpha_{65}$ such that $\beta_1 = \beta_2 = \ldots = \beta_{65}$. In that case, the unrestricted model is $y_{it} = \alpha_i + \beta x_{it} + u_{it}$, for i = 1,…,65 and t = 1975, 1980, 1985, 1990, 2000. So, URSS for the hypothesis H_0^2 is obtained from the one-way within estimation of the equation (7.1.2) with individual specific effects. Then, URSS=9.5828. Moreover, RRSS is obtained from the pooled OLS estimation of the same equation, and RRSS= 30.6606. Then, the test statistics, defined in the section (5.2), is given by

$$F_2 = \frac{(30.6606 - 9.5828)/(65 - 1)}{9.5828/(65 * (5 - 1) - 5)} = \frac{21.0777/64}{9.5828/255} = 8.76 \sim F_{64,\,255}.$$

Since F_2 is greater than $F_{64,\,255}$, the null hypothesis of H_0^2 is rejected at 5% significance level implying that given the equality of slope coefficients, the intercept terms are not same across individual units. On the other hand, although the equality of the intercept term across provinces is rejected, this conclusion supports the usage of one-way within estimation of the equation (7.1.2) with individual specific effect rather than the pooled OLS estimation of the same equation.

Now, consider the equation (7.1.2) with the assumption that the parameters in this equation are constant through individual units but vary across time. In that case, it is possible to test the hypothesis H_0^{13}: $\alpha_{1975} = \alpha_{1980} = ... = \alpha_{2000}$ such that $\beta_{1975} = \beta_{1980} = ... = \beta_{2000}$. In that case, RRSS is obtained from the pooled OLS estimation of the equation (7.1.2) and RRSS= 30.6606. In addition, URSS is obtained from the one-way within estimation of the same equation with time specific effects and URSS=21.9844 (Hsiao, 1986:18). Then, the F statistics, defined in the section (5.2), is given by

$$F_2' = \frac{(30.6606 - 21.9844)/(5 - 1)}{21.9844/(5 * 64 - 5)} = \frac{8.6762/4}{21.9844/315} = 31.08 \sim F_{4,\,315}.$$ [59]

Since F_2' is greater than $F_{4,\,315}$, the null hypothesis of H_0^{13} is rejected at 5% level of significance implying that given the equality of slope coefficients, the intercept terms are not same across time. This conclusion supports the usage of one-way within estimation of (7.1.2) with time specific effects rather than the pooled OLS estimation of this equation. Furthermore, the rejection of the hypotheses H_0^{2} and H_0^{13} supports the presence of both individual and time specific effects in the equation (7.1.2) given the equality of the slope terms. On the other hand, in the following sections, additional tests will be performed in order to test for the existence of those effects within this model.

Secondly, the above hypotheses will also be tested for the augmented Solow model with different education levels, which is constructed in the section (2.3). So, consider the equation (2.3.3) in the panel data framework, which is given by[60]

---

[59] The F statistics is from Hsiao (1986:18).

[60] In the section (7.2.2), the equation (7.1.3) will be estimated with a constant term, that is, when the equation (7.1.3) estimated one should remember that the constant term is included.

$$\text{Iny}_{i,t}=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}[\text{In}(s_k)_{i,t}-\text{In}(n+g+\delta)_{i,t}]+e^{-\lambda\tau}\text{Iny}_{i,t-1}+\beta_1\text{In}(L_1/L)_{i,t+}\beta_2(L_2/L)_{i,t}+\beta_3$$

$$\text{In}(L_3/L)_{i,t}+\beta_4\text{In}(L_4/L)_{i,t}+(1-e^{-\lambda\tau})\text{InA}_i(0)+g(t-e^{-\lambda\tau}(t-1))+v_{it} \quad (7.1.3)$$

where i=1,2,…,65 and t=1980, 1985, 1990 and 2000. In this model, y represents the real GPP per workforce; s_k refers to the physical capital variable, being proxied by the industrial electricity consumption; n is the population growth rate between two subsequent time periods; g is the growth rate of technology; δ is the depreciation rate of physical capital; λ equals (1-α)(n+g+δ); τ is the difference between the two subsequent time periods, that is τ=5 for t=1980,1985, 1990, and τ=10 for t=2000; α is the share of physical capital in total output; L is the total labor force and the others are the same as defined above. In this study, as Mankiw et al. (1992) and Islam (1995) assume g+δ is taken as 0.05. In addition, under the assumption of poolability, the equation (7.1.3) could be written in the form $y_{it} = \alpha+\gamma y_{it-1}+\beta x_{it}+\mu_i+\lambda_t+v_{it}$, where x_{it} is a 5×1 vector of explanatory variables, and μ_i and λ_t correspond to the terms $(1-e^{-\lambda\tau})\text{InA}_i(0)$ and $g(t-e^{-\lambda\tau}(t-1))$ respectively.

Before starting to give the results of the hypothesis tests, it should be emphasized that the equation (7.1.3) is estimated, by each method, under the restriction of the coefficients of Ins_k and In(n+g+δ) being the same in magnitude but opposite in sign, since each estimation method of the equation (7.1.3) accepts this restriction.

Then, assuming that the parameters in the equation (7.1.3) are constant over time but variable across individual units, consider the hypothesis of H_0^2: $\alpha_1 =\alpha_2 =\ldots= \alpha_{65}$ such that $\beta_1 =\beta_2 =\ldots= \beta_{65}$. The F statistics, which is defined in the section (5.2), is given by

$$F_2=\frac{(9.9446-4.8969)/(65-1)}{4.8969/(65*(4-1)-6)}=\frac{5.0476/64}{4.8969/189}=3.04 \sim F_{64,189}$$

where RRSS=9.9446 is obtained from the pooled OLS estimation of the equation (7.1.3) and URSS=4.8969 is obtained from the one-way within estimation of the same equation with individual specific effects. Since F_2 is greater than the critical value, the null hypothesis of H_0^2 is rejected at 5% significance level for the augmented Solow model implying that given the equality of slope coefficients the intercept terms are not the same across individual units. On the other hand, as in the previous model, this conclusion supports the usage of one-way within estimation of (7.1.3) with individual specific effects rather than the pooled OLS estimation of the same equation.

Now, consider the equation (7.1.3) with the assumption that the parameters in this equation are constant through individual units but vary across time. In that case, it is possible to test the following hypothesis $H_0^{13}:\alpha_{1975}=\alpha_{1980}=...=\alpha_{2000}$ such that $\beta_{1975}=\beta_{1980}=....=\beta_{2000}$. RRSS is obtained from the pooled OLS estimation of the equation (7.1.3) and URSS is obtained from the one-way within estimation of the same equation with time specific effects (Hsiao, 1986:18). Then, the F statistics, defined in the section (5.2), is given by

$$F_2' = \frac{(9.9446-8.0319)/(4-1)}{8.0319/((65-1)*4-6)} = \frac{1.9127/3}{8.0319/250} = 19.84 \sim F_{3,\,250}.$$

Since F_2' is greater than $F_{3,\,250}$, the null hypothesis of H_0^{13} is rejected at 5% level of significance implying that given the equality of slope coefficients, the intercept terms are not same across time. This conclusion supports the usage of one-way within estimation of (7.1.3) with time specific effects rather than the OLS estimation of the same equation. Furthermore, the rejection of the hypotheses H_0^2 and H_0^{13} supports the presence of both individual and time specific effects in the equation (7.1.3) given the equality of slope terms. On the other hand, in the following sections, additional tests will be performed in order to test for the existence of those effects within this model.

In addition, due to the same reasoning in the previous model, the null hypotheses of H_0^1 and H_0^3 could not be tested for the augmented Solow model. In that case, there are 4 observations and 6 parameters to be estimated; and this contradicts the rule of $N>k+1$ (since $4<6$). On the other hand, in the following section the model of Knowles (1997) and the augmented Solow model will be estimated under the assumption of the equality of slope coefficients across time and individual units for those models even though this could not be tested.

7.2. Model Estimation

In this section, firstly the estimation results for the model of Knowles (1997) and then the estimation results for the augmented Solow model with different levels of education will be discussed. Simultaneously, in order to find plausible means of estimation the results of the hypothesis tests for those models, which are introduced in the section (5.3.2), will be presented.

7.2.1. The Estimation Results Based on the Model of Knowles

The results of the estimations of the equation (7.1.2) based on single OLS, pooled OLS, and fixed and random effects methods will be presented in this section. At the same time, the hypothesis tests about the selection of a proper estimation method for this equation will be performed.

7.2.1.1. Single OLS Estimation Results

Consider the Knowles's (1997) model of

$$\mathrm{In}Y = \mathrm{In}A+\alpha \mathrm{In}K+\beta_1 \mathrm{In}L_1+ \beta_2 \mathrm{In}L_2+\beta_3 \mathrm{In}L_3+\beta_4 \mathrm{In}L_4+\varepsilon \qquad (7.2.1)$$

Knowles estimates the equation (7.2.1) by using OLS for a single time period of 1985. Since it is not possible to distinguish the impacts caused by individual and time specific variables with a single time period estimation, in his estimation those effects arising from the level of technology are captured in error term. In our work, this handicap is tried to be eliminated by using panel data estimation techniques. On the other hand, in order to be able to compare the results of each estimation method that will be obtained in this section (7.2.1), the single OLS estimation covering the period of 1975-2000 and the OLS estimations for each year, 1975, 1980, 1985, 1990 and 2000, will also be performed. In the OLS estimations for single years, the dependent variable is the natural logarithm of real GPP (InGPP); the explanatory variables are natural logarithms of the amounts of labor force without formal schooling (Innonrgd) and with schooling at basic (Inbasic), secondary (Insecd) and higher (Inhigher) levels, and the natural logarithm of industrial electricity consumption (InIND). Furthermore, in the OLS estimation of equation (7.2.1) covering the period of 1975-2000, the dependent and the independent variables are taken as averages of those variables over 1975-2000. Then, the OLS estimation results of the equation (7.2.1) are given in Table 8.

As can be seen from Table 8, the physical capital variable has a positive effect on real GPP and its effect is highly significant for each equation. Indeed, the elasticity of physical capital with respect to real GPP is 0.10, 0.14, 0.23, 0.21 and 0.16 in the years 1975, 1980, 1985, 1990 and 2000 respectively. Moreover, in the equation covering the period of 1975-2000, the estimated coefficient of physical capital variable is 0.20, which is highly significant, and the standard deviation of this variable is 1.69 implying that a one-standard deviation increase in industrial electricity consumption raises real GPP by 34.5 percent.

Table 8. The Regression for lnGPP by Single OLS Estimations

	1975[1]	1980	1985[1]	1990	2000	1975-2000
lnnongrd	0.077 (0.084)	-0.060 (0.072)	-0.192* (0.068)	-0.035 (0.080)	-0.142** (0.071)	-0.094 (0.060)
lnbasic	0.074 (0.152)	0.168 (0.156)	0.379*** (0.193)	0.264 (0.199)	0.524* (0.168)	0.283** (0.137)
lnsecd	0.604* (0.194)	0.268*** (0.160)	0.091 (0.261)	0.044 (0.230)	-0.024 (0.207)	0.040 (0.206)
lnhigh	0.131 (0.128)	0.409** (0.155)	0.432*** (0.226)	0.561** (0.239)	0.564* (0.207)	0.546* (0.204)
lnIND	0.102* (0.029)	0.145* (0.026)	0.229* (0.044)	0.209* (0.034)	0.163* (0.032)	0.204* (0.026)
Constant	3.469* (1.031)	3.811* (1.037)	3.207* (1.087)	2.354** (0.984)	1.474** (0.731)	3.062* (0.839)
Observations	65	67	67	67	67	67
R-squared	0.946	0.944	0.942	0.947	0.960	0.969

[1] The heteroscedasticity-robust standard errors in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

One of the interesting results obtained from the OLS estimations of the equation (7.2.1) is that the effect of the labor force without formal schooling enters each regression, except in 1975, negatively. On the other hand, its effect is mostly found to be insignificant. Moreover, the basic, secondary and higher education variables enter positively in each equation, except that the secondary education variable in 2000, is insignificantly negative. The amount of the higher education graduates in labor force has the greatest impact on real GPP in 1980, 1985, 1990 and 2000; and its effect is also significant in those years. Indeed, in 1980, 1985, 1990 and 2000, the estimated coefficients for the higher education variable are 0.41, 0.43, 0.56 and 0.56; and the standard deviations of this variable are 0.97, 0.92, 0.95 and 0.97 respectively. This implies that a one-standard deviation increase in the number of college graduates in labor force leads to 39.6 percent increment in real GPP in 1980, 39.7

percent increase in 1985, 53.2 percent increment in 1990 and 54.8 percent increase in 2000.

In the equation of (7.2.1) covering the period of 1975-2000, each coefficient for the education of the labor force, except for the workforce without formal schooling, is found to be positive. Moreover, the basic and higher education variables are also found to be significant. Indeed, the estimated coefficients for the labor force having basic and higher levels of schooling are 0.28 (significant at 5%) and 0.55 (significant at 1%) respectively. As in the single OLS estimations for the years 1980, 1985, 1990 and 2000, the higher level of schooling, among all levels, has the largest impact on real GPP. The estimated coefficient of the higher education variable is 0.55 implying that a one percent increase in the number of labor force who has college degree raises the real GPP by 0.55 percent.

Knowles (1997) calculates the marginal productivity of each level of schooling when interpreting the marginal contributions to real GDP of each level. The estimated coefficients obtained from OLS estimation do not produce those marginal productivities directly; however, the marginal productivities could be obtained from the estimated coefficients, β_i. Since $\beta_i = d\ln Y/d\ln L_i$ and $d\ln Y/d\ln L_i = \frac{dY/Y}{dL_i/L_i}$; $MP_{Li} = dY/dL_i = Y/L_i * \beta_i$ where MP_{Li} represents the marginal productivity of L_i for i=1, 2, 3, 4. The marginal productivities of the labor force with different levels of schooling for the model of Knowles are calculated in this study and the results are presented in Table 9.[61] In addition, in order to make a comparison between the results obtained by Knowles, for each level of education the marginal productivities, which are calculated for the whole sample of the model in his article, are given in Table 9. Table 9 presents similar results in this study with the Knowles's original model. In fact, the labor force having higher level of education has the greatest contribution to real GDP at the margin in both models.

[61] The marginal productivities obtained from the model of Knowles (1997) are based on the single OLS estimation of our model covering the period of 1975-2000.

Tablo 9: The Marginal Productivities Obtained from Single OLS Estimation for the Model of Knowles and from the OLS Estimation for the Knowles's Original Model (1)

	MP_{L1}	MP_{L2}	MP_{L3}	MP_{L4}
Results obtained by Knowles's original model	354	4,308	6,309	8,197
Results obtained in this study from the model of Knowles	0	1	0	31

(1) L_1, L_2, L_3 and L_4 represent Innongrd, Inbasic, Insecd and Inhigh respectively for our model; and represent labor force with no formal schooling, with schooling at primary, secondary and tertiary level of schooling respectively for the Knowles's original model.

7.2.1.2. Pooled OLS Estimation Results

In the previous sub section, the OLS estimations of the Knowles's (1997) model for the years 1975, 1980, 1985, 1990 and 2000, and for the period 1975-2000 are given. Additionally, in this subsection, the results of the pooled OLS estimation of the equation (7.1.2), which are available in Table 10, will be presented. It could be seen from Table 10 that the number of labor force with no formal schooling and with schooling at the college level have negative impact on real GPP whereas the labor force with schooling at the basic and secondary levels have positive one. In addition, the only education variable whose effect is found to be insignificant is the no schooling variable; that is, each education variable, except for the Innongrd, has a significant effect on real GPP. The most interesting result from this estimation method is that having higher educational attainment for the labor force reduces the real GPP. Nevertheless, the significance of the variable Inhigh is at very low level. Furthermore, when the dependent variable is taken as the natural logarithm of real GPP per workforce, InGPPperworkforce, the effect of the higher education becomes insignificant. The results of the single OLS, pooled OLS, and random and fixed effect estimations for the regression of the natural logarithm of the real GPP per worker on the variables Innonrgd, Inbasic and Inhigher, and InIND are given in Appendix A. Moreover, when the model is extended to include regional dummies for

Turkey, the impact of the higher education variable also becomes insignificant (see Table 15)

Table 10. The Regression for lnGPP by Pooled OLS Estimation

lnnongrd	-0.019 (0.036)
lnbasic	0.675* (0.067)
lnsecd	0.400* (0.081)
lnhigh	-0.115*** (0.061)
lnIND	0.169* (0.017)
Constant	0.448 (0.406)
Observations	325
R-squared	0.914

Standard errors are in parentheses
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Contrary to the results obtained from the single OLS estimations of the equation (7.2.1), the basic educational attainment of the labor force, among all, has the greatest impact on real GPP. The estimated coefficient of the basic education variable, which is highly significant, is 0.68 and the standard deviation of this variable is 0.75 implying that a one-standard deviation increase in the number of the labor force with basic level of schooling increases the real GPP by 50.5 percent.

The Breusch-Pagan/Cook-Weisberg test for heteroscedasticity is performed for the pooled OLS estimation of the equation (7.1.2). The resulting test statistics, which is distributed as χ_1^2, is 0.40. Since it is smaller than the critical value, the disturbances are found to be homoskedastic at 5% significance level. On the other hand, the RESET test (using powers of the fitted values of the dependent variable, InGPP) for the model misspecification is performed and the null hypothesis of having no omitted variables is rejected. Nonetheless, this conclusion supports our preference of the panel data estimation against OLS estimation as it is possible to overcome the difficulties arising from the possible omitted variables with the usage of the panel data estimation techniques.

7.2.1.3. Fixed Effect Estimation Results

In this sub section, the results of the fixed effect estimations of the equation (7.1.2) will be presented. The results of the hypothesis tests which are performed to choose an appropriate estimation technique for this equation will be given.

First of all, consider the equation (7.1.2) with the assumption that the individual and time specific effects in this equation are fixed parameters to be estimated. In order to determine whether those effects exist, the null hypothesis of H_0^4: $\mu_1 = \mu_2 = \ldots = \mu_{64} = 0$ and $\lambda_{1975} = \lambda_{1980} = \ldots = \lambda_{1990} = 0$ could be tested. In that case, the RRSS is obtained from the OLS estimation of the equation (7.1.2) and RRSS=30.6606. In addition, the URSS is obtained from the two-way within estimation of the same equation and URSS=6.8839. Then, the F statistics, defined in the section (5.3), is given by

$$F_4 = \frac{(30.6606 - 6.8839)/(65 + 5 - 2)}{6.8839/((65 - 1)(5 - 1) - 5)} = \frac{23.7767/68}{6.8839/251} = 12.75 \sim F_{68,\,251}$$

Since F_4 is greater than $F_{68,\,251}$, the null hypothesis H_0^4 is rejected at 5% level of significance. This implies that at least one of the individual or time specific effects exist, that is, the fixed effect estimation of equation (7.1.2) is preferred to the OLS estimation of the same equation. Then, one could test the existence of individual and time specific effects separately. Firstly, consider the null hypothesis of H_0^5: $\mu_1 = \mu_2$

$=\ldots= \mu_{64} = 0$ such that $\lambda_t \neq 0$ for t =1975, 1980, 1985, 1990. In that case, the RRSS is obtained from the within estimation of the equation (7.1.2) with time specific effects and it is equal to 21.9844. Since the unrestricted model is the same as in the previous case, URSS is 6.8839. Then, the F statistics, defined in the section (5.3), is obtained by

$$F_5 = \frac{(21.9844 - 6.8839)/(65-1)}{6.8839/((65-1)(5-1)-5)} = \frac{15.1005/64}{6.8839/251} = 8.60 \sim F_{64,\,251}$$

F_5 is larger than $F_{64,\,251}$ implying that the null hypothesis of H_0^5 is rejected at 5% significance level and hence that given the presence of time specific effects the individual specific effects exist for this model. Secondly, consider the null hypothesis of H_0^6: $\lambda_{1975} =\ldots=\lambda_{1990}= 0$ such that $\mu_i \neq 0$ for i =1,2,…,64. The unrestricted model is the same as in the case of H_0^4 and H_0^5, so URSS equals to 6.8839. The RRSS is obtained from the within estimation of the equation (7.1.2) with individual specific effects and it is equal to 9.5828. Then, the F statistics, defined in the section (5.3), is given by

$$F_6 = \frac{(9.5828 - 6.8839)/(5-1)}{6.8839/((65-1)(5-1)-5)} = \frac{2.6989/4}{6.8839/251} = 24.60 \sim F_{64,\,251}$$

The null hypothesis of H_0^6 is rejected at 5% level of significance because the F_6 is greater than $F_{64,\,251}$. This means that individual specific effects exist given the presence of time specific effects. This, together with the previous conclusion, implies that both individual and time specific effects exist; that is, the two-way fixed effect estimation of the equation (7.1.2) could be preferred to the one-way within estimations of the same equation with only time or only individual specific effects.

In the following, the results of the one-way within estimations of the equation (7.1.2) together with the results of the two-way fixed effect estimation of the same equation will be discussed in order to see the overall picture arising from the fixed effect estimation for the model of Knowles (1997). Firstly, consider the one-way fixed effect estimation of this equation with individual specific effects. Then, the results of within estimation of the equation (7.1.2) with individual specific effects are given in the first column of Table 11. As can be seen from Table 11, the one-way within estimation of this equation with individual specific effects, contrary to the previous methods, leads an insignificant coefficient for the physical capital variable. Yet, each education variable for the workforce has a significant impact. Moreover, the schooling of the labor force at the basic and secondary levels positively affect real GPP. On the other hand, the numbers of the labor force without formal schooling and with schooling at higher levels have, negative impact on the income level. However, as in case of pooled OLS estimation of the equation (7.1.2), the significance of the explanatory variable lnhigh is at very low level; and when the dependent variable is taken as the natural logarithm of real GPP per workforce in the same model the effect of higher education variable becomes insignificant (see Appendix A).

In the case of one-way within estimation with individual specific effects, the level of schooling that affects the real GPP most is the secondary level. The estimated coefficient of the variable lnsecd is 0.45 implying that a one percentage rise in the number of the labor force having secondary level education increases the real GPP by 0.45 percent.

Secondly, consider the one-way fixed effect estimation of the equation (7.1.2) with time specific effects. Then, the results of the one-way within estimation of this equation with time specific effects, which are given in the second column of Table 11, differs from the results obtained from one-way within estimation of the same equation with individual specific effects in the sense that the physical capital variable is now significant and the educational attainment of the workforce at higher level affects real GPP positively. In that case, in fact, a one percentage increase in the number of the labor force with college degree raises real GPP by 0.14 percent.

Table 11. The Regression for lnGPP by Fixed Effect Estimations

	(1)	(2)	(3)
lnnongrd	-0.239* (0.081)	-0.096* (0.034)	-0.194 (0.132)
lnbasic	0.385* (0.095)	0.397* (0.073)	0.318*** (0.174)
lnsecd	0.453* (0.075)	0.442* (0.093)	0.501* (0.115)
lnhigh	-0.073*** (0.043)	0.145*** (0.081)	-0.052 (0.070)
lnIND	0.004 (0.021)	0.159* (0.014)	0.014 (0.018)
Constant	7.349* (1.006)	2.005* (0.427)	8.254* (1.132)
Observations	325	325	325
Number of id/year	65	5	-
R-squared	0.823	0.929	0.981

Model (1) refers to the one-way within estimation with individual specific effects; model (2) refers to the one-way estimation with time specific effects and model (3) refers to the two-way within estimation.
Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Moreover, the schooling of the workforce at basic and secondary levels have positive impact with high significance. The estimated coefficients of the variables for the basic and secondary schooling are 0.40 and 0.44, and the standard deviations of those variables are 0.75 and 1.00 respectively. This implies that a one-standard deviation raise in the number of the labor force with basic and secondary schooling increases the real GPP by 29.7 and 44.3 percent respectively. In addition, it is clear from Table 11 that the number of the workforce with secondary schooling has the greatest impact on real GPP.

The one-way within estimations of the equation (7.1.2) produces different results for the impacts of higher education of the workforce and of the physical capital. So, in order to get rid of this ambiguity regarding those effects and to see a more convincing picture, the two-way within estimation of the same equation could be useful since; as mentioned in the beginning of this subsection (7.2.1.3), the two-way within estimation of the equation (7.1.2) is preferred to the one-way estimations.

The two-way fixed effect estimation results of the equation (7.1.2) are given in the last column of Table 11. It can be seen from Table 11 that the estimated coefficient of the physical capital variable is insignificant like in the case of one-way within estimation of the same equation with time specific effects. The estimated coefficient of the variable for the higher level of schooling of labor force is also found to be insignificant. Moreover, although the impact of the labor force without formal schooling is insignificant, its effect is found to be negative.

As in the one-way within estimations, the schooling of the labor force at the basic and secondary levels significantly affect the real GPP. Indeed, the estimated coefficients of the variables Inbasic and Insecd are 0.32 and 0.50 respectively. This means that a one percentage increase in the number of labor force with educational attainment at the basic and secondary levels raises real GPP by 0.32 and 0.50 percent respectively. In addition, it can be seen from Table 11 that the secondary schooling attainment of the labor force has the greatest impact on real GPP.

7.2.1.4. Random Effect Estimation Results

In the previous section, the results obtained from the fixed effect estimations of the equation (7.1.2) are given. In this section, firstly the results of the hypothesis tests for the existence of random effects will be given and then, the each equation estimated in the previous section will be reestimated with the random effect estimation methods. Finally, in order to determine which of the fixed or random effect estimations is more appropriate the Hausman specification test will be performed.

Then, consider the equation (7.1.2) with individual and time specific effects which are assumed as random. As mentioned in the section (5.3), the existence of individual and time specific effects could be tested by testing the variance components being equal to zero. To do so, in this study, the Breusch Pagan LM test, which is described in the section (5.3), will be performed.

Now, consider the null hypothesis of H_0^9: $\sigma_\mu^2=\sigma_\lambda^2=0$. The LM statistics required to test this hypothesis has two components, LM_1 and LM_2; and those statistics could be obtained from the LM statistics when one tests the hypotheses of H_0^{10}: $\sigma_\mu^2 = 0$ and H_0^{11}: $\sigma_\lambda^2=0$ respectively. So, one-way random effect estimations with individual and time specific effects are performed for the equation (7.1.2) and then the significance of the variance components is tested separately to obtain LM_1 and LM_2. The calculated LM statistics are 96.03 and 242.57 respectively; and the Breusch Pagan LM test statistics, the sum of those two, is obtained by $LM_{BP}=LM_1+LM_2=96.03+242.57=338.60 \sim \chi_2^2$

Since LM_{BP} is greater than the critical value, the null hypothesis of H_0^9 is rejected at 5% significance level implying that at least one of the variance components are significantly different from zero. That is, the random effect estimation of the equation (7.1.2) is preferred to the OLS estimation of the same equation.

Since the random effect estimation is more plausible for this model than the OLS estimation, the next step is the determination of whether the one-way or two-way random effect estimation method is better. To do so, the hypothesis tests for the significance of individual and time specific effects will be performed separately.

Firstly, consider the null hypothesis of H_0^7: $\sigma_\mu^2=0$ such that $\sigma_\lambda^2>0$. Then, required LM statistics is given by $LM_\mu=\dfrac{\sqrt{2}\tilde{\sigma}_2^2\tilde{\sigma}_v^2}{\sqrt{5(5-1)[\tilde{\sigma}_v^4+(65-1)\tilde{\sigma}_2^4]}}\tilde{D}_\mu$, where $\tilde{D}_\mu$, $\tilde{\sigma}_2^2$ and $\tilde{\sigma}_v^2$ are as defined in the section (5.3.2). The values for those statistics are calculated from the one-way random effect estimation of the equation (7.1.2) with time specific effects, and $\widetilde{D}_\mu = 4084.30$, $\widetilde{\sigma}_2^2 = 1.02$ and $\widetilde{\sigma}_v^2 = 0.08$. Then,

$$LM_\mu = \frac{\sqrt{2}(1.02)^2(0.08)^2}{\sqrt{5(5-1)[(0.08)^2 + (65-1)(1.02)^2]}} 4084.30 = \frac{469.3}{36.4} = 12.90$$

Since LM_μ is greater than the critical value, the null hypothesis H_0^7 is rejected at 5% level of significance implying that given the presence of time specific effects the individual specific effects also exist.

Secondly, one can test for the presence of time specific effects in a similar way. Consider the null hypothesis of H_0^8: $\sigma_\lambda^2 = 0$ such that $\sigma_\mu^2 > 0$. In that case, the hypothesis H_0^8 could be tested with the LM statistics given by $LM_\lambda = \frac{\sqrt{2}\widetilde{\sigma}_1^2\widetilde{\sigma}_v^2}{\sqrt{65(65-1)[\widetilde{\sigma}_v^4 + (5-1)\widetilde{\sigma}_1^4]}} \widetilde{D}_\lambda$, where $\widetilde{D}_\lambda$, $\widetilde{\sigma}_1^2$ and $\widetilde{\sigma}_v^2$ are as defined in the section (5.3.2). The values for those statistics are calculated from the one-way random effect estimation of the equation (7.1.2) with individual specific effects, and $\widetilde{D}_\lambda = 49105.92$, $\widetilde{\sigma}_1^2 = 0.35$ and $\widetilde{\sigma}_v^2 = 0.05$. Then,

$$LM_\lambda = \frac{\sqrt{2}(0.35)^2(0.05)^2}{\sqrt{65(65-1)[(0.05)^2 + (5-1)(0.35)^2]}} 49105.92 = \frac{1140.70}{45.50} = 25.04$$

The null hypothesis of H_0^8 is rejected at 5% significance level because LM_λ is greater than the critical value. This means that given the existence of individual specific effects time specific effects also present. The conclusions of the last two hypothesis tests imply that the two-way random effect estimation is preferable to the one-way random effect estimations for the model of Knowles (1997). Even though the two-way random effect estimation is more appropriate for this model with the existing data, both the one-way and two-way random effect estimations results of the equation (7.1.2) will be discussed so as to see the complete picture for the random effect estimations.[62] The results of the one-way and two-way random effect estimations of the equation (7.1.2) are presented in Table 12.

Firstly, consider the results of the random effect estimation of the equation (7.1.2) with individual specific effects which are available in the first column of Table 12. As can be seen from Table 12, the physical capital variable has positive effect on real GPP which is highly significant. In this case, the only problem regarding the effect of educational attainment of the workforce is that schooling of the labor force at higher levels has a negative and significant impact. However, the impact of the higher level education of the workforce becomes insignificant when the dependent variable in the same equation is taken as the natural logarithm of the real GPP per workforce (see Appendix A). Moreover, contrary to the one-way within estimations of the same equation, the effect of the labor force without formal schooling has a positive but insignificant impact on real GPP.

It could be seen from Table 12 that the one-way random effect estimation of the equation (7.1.2) with individual specific effects results in positive estimated coefficients for the variables of basic and secondary education, and those coefficients are also found to be highly significant. Indeed, the estimated coefficients of the variables lnbasic and lnsecd are 0.62 and 0.44, and the standard deviations for those variables are 0.75 and 1.00 respectively. This implies that a one-standard deviation increase in the number of the labor force with basic and secondary levels of schooling leads to 46.64 and 44.53 percent raise in real GPP respectively. Yet, contrary to the one-way fixed effect estimation of the same equation, the schooling of the workforce at the basic level has the greatest impact on the real GPP.

Then, the estimation results from the one-way random effect estimation of the equation (7.1.2) with time specific effects are given in the second column of Table

[62] When performing the random effect estimations, the Wallace and Hussian method estimator for variance components is used.

12. The estimation based on this method produces very similar results that are obtained from the one-way random effect estimation of the same equation with individual specific effects. For instance, as can be seen from Table 12, the industrial electricity consumption significantly affects the real GPP and this effect is positive. Moreover, as in the previous case, the effect of higher education variable on real GPP is significantly negative. Nevertheless, the significance of the estimated coefficient for this variable is at a low level. On the other hand, as in the random effect estimation of the equation (7.1.2) with individual specific effects, the impact of the higher level of schooling becomes insignificant when the dependent variable in the same equation is taken as the natural logarithm of the real GPP per workforce (see Appendix A). Furthermore, the impact of the higher educational level of the workforce will also become insignificant when the regional dummies included in this model (see Table 16).

The impacts of the basic and secondary education of the labor force are also very similar to those obtained in the previous estimation in the sense that the estimated coefficients of the variables lnbasic and lnsecd are very close to those obtained in the one-way random effect estimation of the equation (7.1.2) with individual specific effects. In fact, in this case, the estimated coefficients for those variables are 0.68 and 0.40 respectively (the coefficients obtained from the one-way random effect estimation with individual specific effects are equal to 0.62 and 0.44 respectively). In addition, the schooling level of the workforce that affects the real GPP most is the basic level like in the one-way random effect estimation with individual specific effects.

Table 12. The Regression for lnGPP by Random Effect Estimations

	(1)	(2)	(3)
lnnongrd	0.012 (0.050)	-0.019 (0.036)	-0.154** (0.054)
lnbasic	0.624* (0.081)	0.675* (0.067)	0.498* (0.098)
lnsecd	0.445* (0.073)	0.400* (0.081)	0.589* (0.092)
lnhigh	-0.124** (0.049)	-0.115*** (0.061)	0.008 (0.069)
lnIND	0.105* (0.020)	0.169* (0.017)	0.092* (0.017)
Constant	1.066*** (0.545)	0.448 (0.406)	2.000* (0.529)
Observations	325	325	325
Number of id/year	65	5	-

Model (1) refers to the one-way random effect estimation with individual specific effects; model (2) refers to the one-way estimation with time specific effects and model (3) refers to the two-way estimation.
Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

After getting an idea about the results of the one-way random effect estimation of the equation (7.1.2), one could discuss the results of two-way random effect estimation of this equation. Then, the results obtained from the two-way random effect estimation of the equation (7.1.2) are given in the last column of Table 12. The two-way random effect estimation results of this equation are, in general, different from the results obtained with the one-way random effect estimations. For example, the impact of the nongraduates in the labor force on real GPP is significantly negative in the two-way random effect estimation. In addition, contrary to the one-way random effect estimations, the educational attainment of the labor force at the higher level has positive but insignificant impact on the real GPP.

Moreover, although the effects of the labor force with basic and secondary level of schooling are both significantly positive; the schooling of the workforce at the secondary level has the largest impact on real GPP. To be specific, the estimated coefficients of the variables for the basic and secondary level of schooling are 0.50 and 0.59 respectively meaning that a one percentage increase in the number of the labor force having basic and secondary levels of schooling increases the real GPP by 0.50 and 0.59 percent respectively.

When making inferences from the impacts of different levels of education on real GPP it is important to keep in mind that the two-way random effect estimation of the equation of (7.1.2) is more appropriate than the one-way random effect estimations of this equation. On the other hand, which of the methods of fixed and random effect estimation is more plausible for the model of Knowles (1997) is still an issue to be determined. So, in order to determine the better method the Hausman test, which is described in the section (5.3), will be performed.

Consider the null hypothesis of H_0^{12}: $E(u_{it}/x_{it}) = 0$ versus H_a^{12} : $E(u_{it}/x_{it}) \neq 0$. In order to test this hypothesis, the two-way random and within estimation coefficients of the equation (7.1.2) are required. Then, $\hat{q}$, defined in the section (5.3.2), is the difference between GLS and within estimation coefficients of the explanatory variables in this model. So, $\hat{q} = \hat{\beta}_{GLS} - \tilde{\beta}_{within} = \begin{bmatrix} 0.040 \\ 0.180 \\ 0.088 \\ 0.061 \\ 0.078 \end{bmatrix}$. The calculation of the test statistic, m, introduced by Hausman (1978) also requires the variance-covariance matrix of $\hat{q}$, and this is equal to $\begin{bmatrix} 0.01459 & -0.01508 & 0.00221 & 0.00053 & 0.00002 \\ -0.01508 & 0.02053 & -0.00589 & -0.00139 & 0.00009 \\ 0.00221 & -0.00589 & 0.00474 & 0.00205 & 0.00010 \\ 0.00053 & -0.00139 & 0.00205 & 0.00011 & 0.00001 \\ 0.00002 & 0.00009 & 0.00010 & 0.00001 & 0.00006 \end{bmatrix}$. Then the test statistics of m is obtained as follows $m = \hat{q}' \left[\mathrm{var}(\hat{q}')\right]^{-1} \hat{q} = 98.7 \sim \chi_5^2$.

Since m is greater than the critical value, the null hypothesis of the absence of correlation between the explanatory and unobservable variables is rejected at 5% significance level meaning that the two-way fixed effect estimation of the equation (7.1.2) is preferred to the two-way random effect estimation of the same equation. Furthermore, the Hausman specification test, together with the conclusions from the previous hypothesis tests performed, implies that the two-way fixed effect estimation is the most appropriate way of estimation, among all others mentioned above, for the model of Knowles (1997). Hence, it is meaningful to focus on the results of the two-way within estimation of the equation (7.1.2) for this model and to discuss those results in a more detailed way.

7.2.1.5. Two-way Fixed Effect Estimation

As previously mentioned the results of the two-way fixed effect estimation for the model of Knowles (1997) could be discussed with more attention since it is the most appropriate way of estimation for this model. The estimated coefficients resulted from the two-way fixed effect estimation of the equation (7.1.2) are given in the last column of Table 11. Firstly, consider the impact of the labor force without formal schooling. The two-way within estimation of this model leads to an insignificant effect of those in the labor force having no formal schooling. This consequence may be due to the reason that the workforce having no formal schooling are working in the jobs or in the sectors that have negligible value added to the real GPP.

Secondly, similar to the variable Innongrd, the higher education variable is found to be insignificant. Nonetheless, contrary to the workforce without formal schooling, the labor force with higher education level is not expected to work in the jobs having little value added to real GPP. On the other hand, this result may be due to the low proportion of the ones with college degree in the total workforce. In fact, the percentages of the number of people in labor force with higher education degree are 0.97, 2.91, 2.78, 3.97 and 6.93, which are significantly low, in the years 1975, 1980, 1985, 1990 and 2000 respectively. Another explanation for this result may be due to the fact that the higher education graduates in the workforce could also affect the real GPP through technology. It may be the indirect effect that makes the greater contribution to the output level. The positive correlation, which is highly significant, between the natural logarithm of the initial level of technology, $\text{In}A_i(0)$, and the higher education variable could support our reasoning (see Appendix B). Moreover, the differentiation of human capital in the form of education could be viewed as separating the direct and indirect effects of education capital on real GPP. The knowledge capital used in the production process is mainly obtained through research and development, and so this knowledge may be related more to the workforce having higher level of education rather than the ones having lower levels of schooling. This could suggest that the higher level of education impacts real GPP more indirectly through technological improvement. On the other hand, the labor force with lower levels of education could participate in the production more directly. Thus, the impact of the educational attainment of the workforce at lower levels may be viewed as a direct effect while the impact of the schooling of the labor force at higher levels could be viewed as an indirect effect. This assertion could be supported with the significance of the basic and secondary schooling variables, which shows the direct impacts of those levels of education, and with the insignificance of higher education variable, which indicates the indirect impact of this level of education providing technological improvement.

Thirdly, the impact of the physical capital variable is also found to be insignificant. This consequence may be due to the insufficient variable used for proxying the physical capital, the industrial electricity consumption for the provinces in Turkey. Since our physical capital variable is limited to industrial sectors, the effect of this variable could be underestimated. However, our data is available at the province level; so it is hard to obtain a better proxy for the physical capital and to check whether the physical capital's effect on real GPP is actually insignificant.

Finally, consider the impacts of the educational attainment of the labor force at the basic and secondary levels. As can be seen in Table 11, the secondary school attainment of the workforce has the greatest impact on real GPP. However, the share of the labor force with secondary schooling is much lower relative to the share of the workforce with basic level of schooling. Indeed, while the percentages of the labor force having basic level degree change in the range of 46.4-61.5 percent, the percentages of those having secondary school attainment change in the range of 3.5-15.1. This result may be due to the fact that even though the labor force having high school education has a smaller proportion, its contribution to the real GPP is more than that of the labor force having basic level of schooling.

Moreover, for the two-way within estimation of the equation (7.1.2), the equality of the estimated coefficients of different education variables is tested. To do so, the null hypothesis of H_0[14]: $\beta_1=\beta_2=\beta_3=\beta_4$ is tested for this equation. The resulting test statistics is 7.22 which distributes as $F_{3,\ 251}$ implying that the null hypothesis of the equality of the coefficients for the variables lnnongrd, lnbasic, lnsecd and lnhigh is rejected at 5% significance level. This means that the impacts of different levels of education of the labor force on real GPP are significantly different for the model of Knowles (1997).

Table 13. The Regression for lnGPP by Two-way Fixed Effect Estimation with Different Education Levels

	(1)	(2)	(3)	(4)
lnnongrd	0.276* (0.074)	-0.294** (0.133)	-0.182 (0.131)	-0.194 (0.132)
lnbasic		0.704* (0.139)	0.282*** (0.167)	0.318*** (0.174)
lnsecd			0.481* (0.112)	0.501* (0.115)
lnhigh				-0.052 (0.070)
lnIND	0.024 (0.020)	0.012 (0.019)	0.012 (0.018)	0.014 (0.018)
Constant	10.044* (0.910)	8.509* (0.920)	7.718* (0.909)	8.254* (1.132)
Observations	325	325	325	325
R-squared	0.977	0.979	0.981	0.981

Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%

Furthermore, the same model is reestimated by adding up different education levels step by step, and the resulting estimates are given in Table 13. Firstly, the model including only variable of lnnongrd together with the physical capital variable, model (1), is estimated. Then, the model (2), being obtained when the model (1) is extended to include lnbasic variable, is estimated. Thirdly, in order to get the model (3) variable lnsecd is added to the model (2). Finally, the initial model estimated previously, model (4), is obtained by extending the model (3) to include variable lnhigh. As can be seen from Table 13, in each model except for the model (4), when a new model is obtained by adding a higher education variable to the previous model, the impact of this higher educational variable is found to be greater than the effects of the previous education variable(s). Hence, it could be concluded that instead of using some of those education variables to proxy for education capital, the usage of all education levels simultaneously in a single equation accounts for real GPP in a better way.

7.2.1.5.1. Additional Tests on Two-way Fixed Effect Estimation

In this sub section, heteroscedasticity and serial correlation for the two-way within estimation of the equation (7.1.2) are tested. First of all, if one assumes homoskedastic disturbances or serial correlation when it is not the case, then the resulting estimates will become inefficient though they will still be consistent (Baltagi, 2001:77, 81). Moreover, the standard errors of those estimates will also become biased. So, the Breusch-Pagan/Cook-Weisberg test for heteroscedasticity is performed for the model of Knowles (1997). The resulting test statistics, which is distributed as χ_1^2, is 1.71. Since the test statistics is smaller than the critical value of 3.84, the disturbances in the equation (7.1.2) are found to be homoskedastic at 5% level of significance.

Secondly, the serial correlation in the idiosyncratic error, v_{it}, for the same equation is tested via the prediction of this term for this estimation.[63] For the serial correlation testing, firstly the idiosyncratic error for the two-way within estimation of the equation (7.1.2) is predicted, and the lag values for this predicted term are generated. Then, the equation (7.1.2) is reestimated by inserting the lagged values of the idiosyncratic error into the same equation; and the F test for the significance of the coefficient for the lagged values is performed. The resulting test statistics is 1.77, which is distributed as $F_{1,\ 186}$, implying that the null hypothesis of the estimated coefficient for the lagged values of idiosyncratic error being zero is not rejected. This means that the serial correlation in the idiosyncratic error does not exist.

Furthermore, the RESET test (using powers of the fitted values of the dependent variable, InGPP) for the model misspecification is performed and the null hypothesis of having no omitted variables is rejected. This conclusion implies that the omitted variables for this model still exist. So, more explanatory variables such as additional

[63]In the two-way within estimation of the equation (7.1.2), the idiosyncratic error could be distinguished from the individual and time specific effects, μ_i and λ_t, since this equation is estimated by using dummy variables for individual units and time periods.

dummies may be required to be included in the model. On the other hand, in the following section the same model is estimated with including regional dummies for Turkey.

7.2.1.6. Estimation Results of the Model of Knowles with Regional Dummies

In this section, the model (7.1.2) will be extended to cover regional dummies for Turkey and the new model will be estimated with single and pooled OLS, and with panel data techniques. Then, consider the following model[64]

$$\text{In}Y_{it} = \text{In}A_i(0)+gt+\alpha \text{In}K_{it}+\beta_1\text{In}L_{1it}+ \beta_2\text{In}L_{2it}+\beta_3\text{In}L_{3it}+\beta_4\text{In}L_{4it}+\beta_5\text{marmara}+\beta_6\text{aegean}+\beta_7\text{medit}+\beta_8\text{blacksea}+\beta_9\text{central}+ \beta_{10}\text{southeast}+ v_{it} \quad (7.2.2)$$

where the variables marmara, aegean, medit, blacksea, central and southeast are the dummies equal to 1 when a province belongs to the region Marmara, Aegean, Mediterranean, Black Sea, Central Anatolia and Southeast Anatolia respectively. Since seven regions exist in Turkey, six dummy variables are used in the equation (7.2.2); and the base group is the region East, that is, the estimated coefficients of regional dummies will be interpreted relative to the region East.

The single OLS estimation results of the equation (7.2.2) are given in Table 14. The estimation results are found to be similar to those obtained from the single OLS estimations of the equation (7.2.1). For instance, the impact of the industrial electricity consumption on real GPP is still significantly positive in each year and in the period of 1975-2000. Moreover, the higher educational attainment of the workforce affects the real GPP positively and this impact is still significant in each year except in 1975. Yet, the labor force without formal schooling and the labor

---

[64] This equation is obtained from the equation (7.1.2) by adding the regional dummy variables to the equation (7.1.2).

force with basic level schooling affect the real GPP insignificantly in each year. Similar to the results of single OLS estimations of the equation (7.2.1), the schooling level that affects the real GPP most is the secondary level in 1975 and the higher level in the years 1980, 1985, 1990 and 2000. Nonetheless, the impact of the secondary schooling in 1975 and that of higher schooling in 1980, 1985, 1990 and 2000 are found to be greater than the impacts of those variables obtained from the OLS estimations of the same equation without regional dummies.

Moreover, in each year, except in 1990, the region that impacts the real GPP, relative to the region East, most is the region Marmara. In fact, holding the education and physical capital variables being the same across provinces, a province being in the region Marmara affects the real GPP 30.00, 38.80, 16.80 and 50.00 percent more relative to the region East in 1975, 1980, 1985 and 2000 respectively. This seems logical in the sense that the region Marmara is the most industrialized and developed region of Turkey.

The equation (7.2.2) is then estimated with pooled OLS, and the pooled OLS results of this equation are given in Table 15. However, in that case it is not possible to test the poolability of the data. In order to calculate the test statistics being necessary to test the equality of the regression coefficients across time and provinces, the within estimation of the equation (7.2.2) is required. Nevertheless, since the regional dummy variables are time invariant, the first difference of the equation (7.2.2) eliminates the dummy variables; and so the within regression of this equation do not produces different results than those obtained from the within estimation of the same equation without regional dummies. Hence, the usage of the test statistics obtained from the within estimation of the equation (7.2.2) are not plausible. On the other hand, the poolability is assumed before estimating this equation with pooled OLS.

Table 14. The Regression for lnGPP by Single OLS Estimations with Regional Dummies

	1975	1980	1985	1990	2000	1975-2000
lnnongrd	0.088 (0.098)	0.052 (0.107)	-0.168 (0.107)	-0.065 (0.114)	-0.013 (0.108)	-0.063 (0.090)
lnbasic	0.079 (0.161)	-0.017 (0.216)	0.332 (0.219)	0.382 (0.235)	0.303 (0.221)	0.263 (0.179)
lnsecd	0.653* (0.187)	0.357*** (0.198)	0.060 (0.324)	-0.088 (0.237)	-0.086 (0.223)	-0.076 (0.228)
lnhigh	0.098 (0.124)	0.427** (0.163)	0.501*** (0.281)	0.664* (0.230)	0.738* (0.207)	0.681* (0.212)
lnIND	0.083* (0.021)	0.117* (0.028)	0.211* (0.041)	0.134* (0.039)	0.117* (0.035)	0.164* (0.030)
marmara	0.300** (0.132)	0.388** (0.163)	0.168 (0.172)	0.380** (0.162)	0.500* (0.166)	0.272** (0.129)
aegean	0.181 (0.135)	0.366** (0.169)	0.224 (0.167)	0.154 (0.162)	0.272 (0.166)	0.142 (0.130)
medit	0.117 (0.124)	0.226 (0.156)	-0.155 (0.163)	0.072 (0.153)	0.218 (0.152)	0.015 (0.122)
blacksea	0.082 (0.103)	0.105 (0.129)	0.028 (0.129)	0.038 (0.117)	0.237** (0.117)	0.063 (0.095)
central	0.070 (0.108)	0.172 (0.145)	-0.001 (0.145)	0.066 (0.140)	0.184 (0.137)	0.036 (0.109)
southeast	0.294** (0.115)	0.202 (0.131)	0.099 (0.156)	0.381** (0.147)	0.226*** (0.133)	0.191*** (0.111)
Constant	3.149* (0.813)	3.813* (1.019)	3.366* (1.180)	2.431** (0.917)	2.000* (0.719)	3.238* (0.818)
Observations	65	67	67	67	67	67
R-squared	0.958	0.954	0.952	0.960	0.968	0.975

Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

In that case, the pooled OLS estimation results of the equation (7.2.2) are similar to those obtained from the OLS estimation of the equation (7.1.2) in the sense that schooling of the workforce at the basic and secondary levels positively affects the real GPP with high significance. Moreover, as in the case of pooled OLS estimation of the equation (7.1.2), the OLS estimation of the equation (7.2.2) implies that the schooling of labor force at the basic level impacts the real GPP most. Nonetheless, in that case, contrary to the results obtained from the OLS estimation of the equation

(7.1.2) and from the single OLS estimations of the same equation; the high school graduates in the workforce insignificantly affect the real GPP.

In addition, the estimated coefficient of each dummy variable is significantly positive. The positive coefficients of those dummies imply that the impact on real GPP of a province being in the regions Marmara, Aegean, Mediterranean, Black Sea, Central Anatolia and Southeast Anatolia is greater than the effect of a province being in the region East. Furthermore, the provinces in the region Marmara have the greatest impact on real GPP relative to the region East. Indeed, being in the region Marmara relative to East increases the impact of a province on real GPP by 48 percent.

As previously mentioned, the within estimation of the equation (7.2.2) produces the same results with those obtained from the within estimation of the same equation without dummies; because the regional dummy variables are time invariant and the within estimation of the equation (7.2.2) eliminates the effects of those variables. Hence, in that case, it is meaningful to estimate the equation (7.2.2) merely with random effect estimation methods. The hypothesis tests only for the existence of random effects for the equation (7.2.2) will also be performed. In the rest of this section, the conclusions from those tests together with the random effect estimation results will be discussed.

Firstly, in order to test the existence of the individual and time specific effects which are assumed to be random, the Breusch Pagan LM test, which is described in the section (5.3), will be performed. For the null hypothesis of H_0^9, the test statistics LM_{BP} is given by $LM_{BP} = LM_1 + LM_2 = 90.13 + 227.47 = 317.60 \sim \chi_2^2$.

Since LM_{BP} is larger than the critical value, the null hypothesis for the absence of variance components is rejected at 5% significance level implying that the random effect estimation of the equation (7.2.2) is preferred to the OLS estimation.

Table 15. The Regression for lnGPP by Pooled OLS Estimations with Regional Dummies

lnnongrd	0.063 (0.048)
lnbasic	0.526* (0.090)
lnsecd	0.484* (0.084)
lnhigh	-0.092 (0.058)
lnIND	0.122* (0.017)
Marmara	0.480* (0.082)
Aegean	0.346* (0.083)
Medit	0.251* (0.079)
Blacksea	0.138** (0.063)
Central	0.232* (0.071)
Southeast	0.303* (0.073)
Constant	0.538 (0.382)
Observations	325
R-squared	0.926

Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Secondly, the existence of individual and time specific effects will be tested separately. So, in order to test the null hypotheses of H_0^7 and H_0^8, the test statistics of LM_μ and LM_λ, which are asymptotically distributed as N(0,1), are calculated. The calculated values of LM_μ and LM_λ are 12.43 and 23.91 respectively. Since both LM_μ and LM_λ are greater than the critical value, the null hypotheses of H_0^7 and H_0^8 are both rejected at 5% level of significance. This means that the individual and time specific effects exist in the model of Knowles (1997) with regional dummies. Hence, the two-way random effect estimation of the equation (7.2.2) could be favored to the one-way random effect estimations of the same equation.

Finally, the results of the random effect estimations of the equation (7.2.2), which are available in Table 16, will be given. Then, consider the one-way random effect estimation of the equation (7.2.2) with individual specific effects whose results are available in the first column of Table 16. Contrary to the estimation results of the Knowles's (1997) model without regional dummies, the impact of the nongraduates in the labor force is significant. Moreover, the effects of the education at the basic and secondary levels on real GPP are significantly positive. In that case, the labor force having secondary level of education affects the real GPP most. Indeed, a one-percentage increase in the number of the labor force with secondary level of schooling leads to a 0.55 percent raise in real GPP.

Then, consider the one-way random effect estimation of the equation (7.2.2) with time specific effects whose results are available in the second column of Table 16. Contrary to the estimation of the same model without dummies, the estimated coefficient of the higher education variable becomes insignificant. However, the impacts of the basic and secondary levels of education are still significantly positive. Furthermore, the schooling level of the workforce that affects the real GPP most is the basic level of schooling. In fact, a one-percentage rise in the number of the labor force having basic level of schooling increases the real GPP by 0.53 percent.

It is clear from Table 16 that the one-way random effect estimations of the equation (7.2.2) produce different results for different levels of schooling affecting the real GPP. Since the two-way random effect estimation of the equation (7.2.2) is preferred to the one-way random effect estimations, it is better to discuss the two-way estimation results of this equation in order to clarify the differences arising from the results of the one-way random effect estimations of the same equation. The results of the two-way random effect estimation of the equation (7.2.2) are given in the last column of Table 16.

According to the results of the two-way random effect estimation of the Knowles's (1997) model with regional dummies, the impact of the labor force without formal schooling is found to be insignificant. Furthermore, the college graduates in the labor force impacts the real GPP positively but its effect is also insignificant. Similar to the results obtained from the two-way random effect estimation of the equation (7.1.2), the number of the secondary school graduates in the labor force has the highest impact on real GPP. In addition, the regional dummies are all found to be significant and, as usual, the Marmara region relative to the region East has the greatest impact on real GPP. Indeed, being in the region Marmara relative to East raises the effect of a province on real GPP by 53 percent.

It could be emphasized that for the model of Knowles (1997) without regional dummies, the best way of estimation is the two-way within estimation; and for the same model with dummies, the most appropriate way is the two-way random effect estimation. The two-way random effect estimation results of the equation (7.2.2) are similar to those obtained from the two-way within estimation of the same equation without regional dummies. The education variables being significant are at the basic and secondary levels; and the labor force with secondary level of schooling has the greatest impact on real GPP for both models. Indeed, the estimated coefficients for the variables of basic and secondary levels of education are 0.33 and 0.65 respectively for the two-way random effect estimation of the equation (7.2.2); and the estimated coefficients of those variables are 0.32 and 0.50 respectively for the two-way within estimation of the equation (7.1.2). This similarity could imply that the inclusion of the regional dummies into the model of Knowles does not change the main results of this model.

Table 16. The Regression for lnGPP by Random Effect Estimations with Regional Dummies

	(1)	(2)	(3)
lnnongrd	0.103*** (0.055)	0.063 (0.048)	-0.066 (0.076)
lnbasic	0.439* (0.093)	0.526* (0.090)	0.332** (0.128)
lnsecd	0.550* (0.075)	0.484* (0.084)	0.651* (0.095)
lnhigh	-0.113** (0.047)	-0.092 (0.058)	0.026 (0.067)
lnIND	0.075* (0.019)	0.122* (0.017)	0.065* (0.017)
marmara	0.678* (0.111)	0.480* (0.082)	0.529* (0.125)
aegean	0.530* (0.115)	0.346* (0.083)	0.466* (0.124)
medit	0.428* (0.115)	0.251* (0.079)	0.330* (0.120)
blacksea	0.240** (0.094)	0.138** (0.063)	0.229** (0.096)
central	0.363* (0.101)	0.232* (0.071)	0.278** (0.108)
southeast	0.370* (0.114)	0.303* (0.073)	0.438* (0.113)
Constant	1.085** (0.503)	0.538 (0.382)	2.214* (0.516)
Observations	325	325	325
Number of id/year	65	5	5

Model (1) refers to one-way random effect estimation with individual specific effects; model (2) refers to the one-way estimation with time specific effects and model (3) refers to the two-way estimation.
Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

7.2.2. The Estimation Results Based on the Augmented Solow Model with Different Education Levels

In this section, the estimation results of the equation (7.1.3) depending on single and pooled OLS, and fixed and random effects estimations methods will be discussed. At the same time, the hypothesis tests about the selection of proper estimation methods for this equation will be performed.

7.2.2.1. Single OLS Estimation Results

Consider the equation (7.1.3) in the form

$$\text{Iny}_i=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}[\text{In}(s_k)_i-\text{In}(n+g+\delta)_i]+e^{-\lambda\tau}\text{Iny}_i(0)+\beta_1\text{In}(L_1/L)_i+\beta_2(L_2/L)_i+\beta_3\text{In}(L_3/L)_i+\beta_4\text{In}(L_4/L)_i+\varepsilon_i \quad (7.2.3)$$

where $\text{Iny}_i(0)$ corresponds to the initial real GPP per workforce. This equation (7.2.3) could be estimated by OLS in the years 1980, 1985, 1990 and 2000 separately. As it is not possible for OLS estimations to distinguish the impacts of individual and time specific variables in the equation (7.1.3), the single OLS estimations of this equation implies that those impacts arising from the individual and time specific terms are captured in the error term. However, as for the model of Knowles (1997), this obstacle is tried to be eliminated by using panel data estimation techniques. On the other hand, in order to be able to compare the results of each estimation method for the augmented Solow model that will be mentioned in this section (7.2.2), the single OLS estimations of the regression of InGPPperworkforce on Innongrd (the amount of the labor force without formal schooling divided by total labor force), Inbasic (the amount of the labor force with basic level of schooling divided by total labor force), Insecd (the amount of the labor force with secondary level of schooling divided by total labor force), Inhigh (the amount of the labor force with higher level of schooling divided by total labor force), InIND-In(n+g+δ) (the difference between the natural logarithms of the physical capital variable and of the sum n+g+δ) and $\text{InGPPperworkforce}_{t-1}$ (the lagged dependent variable) in 1980, 1985, 1990 and 2000, and in the period covering 1980-2000 will also be performed.

The single OLS estimation results for the augmented Solow model are given in Table 17. Contrary to the results obtained from the OLS estimations of the model of Knowles (1997), the educational attainment of the labor force at each level, except for the basic level of schooling in year 1980, is found to be insignificant. On the other hand, in 1985, the workforce without formal schooling and the workforce having higher level of schooling have significant, which are negative, impact on real GPP per workforce. However, the effects of those variables are significant at very low levels. In addition, as can be seen from Table 17, the schooling of the labor force at the basic level is significantly positive in the period of 1980-2000. Indeed, a one percentage rise in the share of the labor force having basic level of education increases real GPP per workforce by 0.58 percent.

Table 17. The Regression for lnGPPperworkforce by Single OLS Estimations in Augmented Solow Model

	1980	**1985**	**1990**	**2000**	**1980-2000**
lnnongrd	0.037 (0.205)	-0.358*** (0.195)	-0.095 (0.219)	-0.103 (0.127)	-0.074 (0.238)
lnbasic	0.372** (0.182)	-0.252 (0.246)	0.168 (0.343)	0.191 (0.254)	0.583*** (0.317)
lnsecd	0.012 (0.121)	0.144 (0.159)	-0.081 (0.199)	-0.036 (0.152)	0.048 (0.188)
lnhigh	-0.010 (0.098)	-0.261*** (0.154)	0.225 (0.208)	0.103 (0.141)	0.174 (0.194)
lnIND-ln(n+g+δ)	0.052* (0.018)	-0.000 (0.025)	0.095* (0.029)	0.053** (0.023)	0.132* (0.027)
$\text{lnGPPperworkforce}_{t-1}$	0.815* (0.089)	0.990* (0.092)	0.480* (0.101)	0.655* (0.071)	0.195* (0.045)
Constant	-0.249 (0.712)	-1.107 (0.896)	-0.325 (1.096)	-0.242 (0.812)	-0.624 (1.087)
Observations	65	65	65	65	65
R-squared	0.910	0.924	0.839	0.920	0.875

Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

As expected the effect of the physical capital variable is significantly positive in years 1980, 1990 and 2000; and in the period 1980-2000. Since the single OLS estimations are obtained from the restricted equation of (7.2.3), the positive coefficient of the variable lnIND-ln(n+g+δ) implies that the effect of the population growth on real GPP per workforce is negative.

7.2.2.2. Pooled OLS Estimation Results

In this sub section, the equation (7.1.3) is estimated by pooling the existing data in the years 1980, 1985, 1990 and 2000. The results obtained from the pooled OLS estimation of this equation are given in Table 18. Contrary to the single OLS estimation results, the pooled OLS estimation of the equation (7.1.3) produces significant coefficients for each schooling variable, except for the variable lnnongrd. Moreover, the pooled OLS estimation of the augmented Solow model displays different results from those obtained by the pooled OLS estimation of the model of Knowles (1997) using lnGPPperworkforce as the dependent variable. In fact, the educational attainment of the labor force at secondary level has a negative impact and those at higher level has a positive and significant impact on real GPP per workforce within the augmented Solow model (see Appendix A).

In that case, it is interesting that the secondary schooling of the labor force affects real GPP per workforce negatively. Nonetheless, this effect is significant at very low level and it becomes significantly positive when the schooling variable for secondary level is taken as the only education variable in the equation (7.1.3) (see Appendix A). On the other hand, the augmented Solow model with different levels of education produces significantly positive coefficients for the variables lnbasic and lnhigh. Indeed, the estimated coefficients for those variables are 0.28 and 0.19 respectively. The standard deviations of the variables for basic and higher level of schooling are 0.24 and 0.84 respectively implying that a one-standard deviation increase in the share of the labor force having basic and higher levels of education raises the real

GPP per workforce by 6.59 and 15.80 percent respectively. In addition, for the pooled OLS estimation of the augmented Solow model, the level of education of the labor force that affects the real GPP per workforce most is the basic level of education.

Table 18. The Regression for lnGPPperworkforce by Pooled OLS Estimation in Augmented Solow Model

lnnongrd	-0.087 (0.074)
lnbasic	0.277* (0.101)
lnsecd	-0.109*** (0.063)
lnhigh	0.187* (0.063)
lnIND-ln(n+g+δ)	0.055* (0.013)
$\ln GPPperworkforce_{t-1}$	0.640* (0.045)
Constant	-0.065 (0.358)
Observations	260
R-squared	0.880

Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

The pooled OLS estimation of the equation (7.1.3) implies that the impact of the physical capital on real GPP per workforce is significantly positive and its effect is less than those of the education variables for the augmented Solow model.

7.2.2.3. Fixed Effect Estimation Results

In this sub section, the equation of (7.1.3) is estimated with fixed effect estimation methods; and before declaring the results based on those estimations, the conclusions obtained from the hypothesis tests, which are performed to determine whether the within estimation is appropriate for the augmented Solow model, are discussed.

Then, consider the null hypothesis H_0^4: $\mu_1=\mu_2=\ldots=\mu_{64}=0$ and $\lambda_{1980}=\lambda_{1985}=\lambda_{1990}=0$. In that case, the RRSS is obtained from the OLS estimation of the equation (7.1.3) and RRSS=9.9446. In addition, the URSS is obtained from two-way within estimation of the same equation and URSS=4.2061. Then, the F statistics, defined in the section (5.3), is given by

$$F_4 = \frac{(9.9446-4.2061)/(65+4-2)}{4.2061/((65-1)(4-1)-6)} = \frac{5.7385/67}{4.2061/186} = 3.79 \sim F_{67,\,186}$$

Since F_4 is greater than the critical value, the null hypothesis of the absence of parameters μ_i and λ_t for i =1,...,65 and t=1980, 1985, 1990, 2000 is rejected at 5% level of significance. This implies that the individual or time specific effects, which are assumed as fixed parameters, exist for the augmented Solow model. That is, the fixed effect estimation of equation (7.1.3) is preferred to the OLS estimation of the same equation. Then, one could also test the existence of individual and time specific effects for the augmented Solow model separately.

Now, consider the hypothesis of H_0^5: $\mu_1=\mu_2=\ldots=\mu_{64}=0$ such that $\lambda_t \neq 0$ for t=1980, 1985, 1990 and 2000. In that case, the URSS equals 4.2061; and the RRSS which is obtained from the one-way within estimation of the equation (7.1.3) with time specific effects equals 8.0319. Then, the F statistics, defined in the section (5.3), is given by

$$F_5 = \frac{(8.0319-4.2061)/(65-1)}{4.2061/((65-1)(4-1)-6)} = \frac{3.8258/64}{4.2061/186} = 2.64 \sim F_{64,\,186}$$

Since F_5 exceeds the critical value, the null hypothesis of H_0^5 is rejected at 5% significance level implying that given the presence of time specific effects, the individual specific effects also exist for the augmented Solow model. Then, consider the null hypothesis of H_0^6: $\lambda_{1980} = \lambda_{1985} = \lambda_{1990} = 0$ such that $\mu_i \neq 0$ for i=1,2,…,65. URSS is again equals to 4.2061, and the RRSS, which is obtained from the one-way fixed effect estimation of the equation (7.1.3) with individual specific effects, is 4.8969. Then, the F statistics, defined in the section (5.3), is given by

$$F_6 = \frac{(4.8969-4.2061)/(4-1)}{4.2061/((65-1)(4-1)-6)} = \frac{0.6909/3}{4.2061/186} = 10.18 \sim F_{3,\,186}$$

As F_6 is greater than the critical value, the null hypothesis of H_0^6 is rejected at 5% level of significance. This means that given the presence of individual specific effects time specific effects also exist in the augmented Solow model. In addition, the results of the aforementioned hypothesis tests imply that the two-way within estimation of the equation (7.1.3) is the most appropriate way of estimation when the individual and time specific effects are assumed to be fixed parameters. On the other hand, the results obtained from the one-way and two-way within estimations of this equation will be given in order to see the complete picture regarding the fixed effect estimation within the augmented Solow model framework.

The one-way fixed effect estimations of the equation (7.1.3) with individual and time specific effects are given in the first and second columns of Table 19 respectively. As can be seen from Table 19, the one-way fixed effect estimations of this equation imply that the estimated coefficients of each education variable, except for the secondary level, has the same sign in the one-way within estimations with individual and time specific effects. Nonetheless, the significance of those variables changes with respect to the unobservable effects being individual or time specific. In fact, education variables are all significant when the fixed effect estimation includes only individual specific effects, while only the basic level of education is significant when the fixed effect estimation includes only time specific effects.

The one-way within estimation of the equation (7.1.3) with individual specific effect implies that the labor force having no formal schooling and those with educational attainment at the secondary level affect real GPP per workforce negatively. Yet, the impacts of those variables are insignificant for the one-way within estimation of the same equation with time specific effects. Furthermore, the variable lnsecd becomes significantly positive when the schooling variable for secondary level is taken as the only education variable in the equation (7.1.3) (see Appendix A). In addition, the higher level graduates in the labor force has the greatest contribution to real GPP per workforce for the one-way within estimation of this equation with individual specific effects.

In the augmented Solow model, the schooling of the labor force at the basic level positively affect the real GPP per workforce for the one-way within estimations of the equation (7.1.3).

The ambiguity arising from the results of one-way within estimations of the equation (7.1.3) could be eliminated by considering the two-way within estimation of the same equation because two-way within estimation of the equation (7.1.3) is preferred to one-way within estimations of the same equation for the augmented Solow model. The resulting estimates of the two-way fixed effect estimation of this equation are available in the last column of Table 19. In that case, as in the one-way within estimation of the equation (7.1.3) with time specific effects, the two-way fixed effect estimation of the same equation produces insignificant results for the variables lnnongrd and lnsecd. However, the impact of the workforce having higher level of schooling on the real GPP per workforce is significantly positive as in the one-way within estimation with individual specific effects. Indeed, a one-percentage rise in the share of the labor force having higher level of schooling increases the real GPP per workforce by 0.18 percent. On the other hand, contrary to the one-way within estimations, the effect of basic level of education is found to be insignificant.

The impacts of the physical capital and population growth variables are again have the expected signs which are significant for the one-way and two-way within estimation of the equation (7.1.3).

Table 19. The Regression for InGPPperworkforce by Fixed Effect Estimations in Augmented Solow Model

	(1)	(2)	(3)
Innongrd	-0.422* (0.093)	-0.067 (0.069)	-0.127 (0.185)
Inbasic	0.237*** (0.134)	0.234** (0.093)	-0.133 (0.241)
Insecd	-0.275* (0.102)	0.050 (0.073)	0.161 (0.159)
Inhigh	0.461* (0.075)	0.042 (0.069)	0.176*** (0.099)
InIND-In(n+g+δ)	0.055** (0.024)	0.056* (0.012)	0.059** (0.023)
$\text{InGPPperworkforce}_{t-1}$	-0.052 (0.072)	0.691* (0.044)	0.032 (0.070)
Constant	0.668 (0.555)	-0.162 (0.360)	0.854 (0.653)
Observations	260	260	260
R-squared	0.941	0.903	0.949

Model (1) refers to one-way within estimation with individual specific effects; model (2) refers to the one-way estimation with time specific effects and model (3) refers to the two-way estimation.
Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

7.2.2.4. Random Effect Estimation Results

In this sub section, the equation (7.1.3) is estimated with random effect estimations methods, and before stating the results based on those estimations, the conclusions obtained from the hypothesis tests, which are performed to determine whether the random effect estimation is appropriate for the augmented Solow model, will be mentioned.

Firstly, the Breusch Pagan LM test is performed to test the existence of individual and time specific effects being assumed to be random in the augmented Solow model. Since the Breusch Pagan LM statistics equals to the sum of LM_1 and LM_2, which are defined in the section (5.3), this statistics, LM_{BP}, is 196.51=0.57+195.94 where LM_1=0.57 and LM_2=195.94. This implies that at least one of the individual or time specific effects exist as LM_{BP} exceeds the critical value at 5% significance level. That is, the random effect estimation of the equation (7.1.3) is more appropriate than the OLS estimation of the same equation. In addition as stated in the sub section (5.3.2), LM_1 and LM_2, which are distributed as χ_1^2, are obtained in the case of testing H_0^{10}: $\sigma_\mu^2=0$ and H_0^{11}: $\sigma_\lambda^2=0$ respectively. The calculated values for those statistics imply that the null hypothesis H_0^{10} is accepted whereas H_0^{11} is rejected at 5% level of significance because $LM_1=0.57<3.84=\chi_1^2$ and $LM_2=195.94>3.84=\chi_1^2$. Hence, in the augmented Solow model, the time specific effects exist, while the individual specific effects do not present when those effects are assumed as random. That is, the one-way random effect estimation of the equation (7.1.3) with time specific effects is better than the two-way random effect estimation of the same equation.

In that case, since the variance components of individual specific effects are insignificant, the unobservable effects being individual specific do not exist when they are assumed as random. The absence of individual specific effects causes that the one-way random effect estimation of the augmented Solow model with individual specific effects produces exactly the same results with those obtained from the pooled OLS estimation. In addition, the one-way random effect estimation of the equation (7.1.3) with time specific effects also produces the same results with the two-way random effect estimation of the same equation due to the absence of the individual specific effects. Hence, in this section (7.2.2), only the results from one-way random effect estimation of the equation (7.1.3) with time specific effects will be discussed within the augmented Solow model.

Then, the estimation results obtained from the one-way random effect estimation of the equation (7.1.3) with time specific effects are presented in Table 20. In that case, the results are similar with those obtained from the one-way within estimation of the same equation with time specific effects. In fact, in both cases the effects on real GPP per workforce of the labor force having basic level of schooling are significantly positive. Moreover, the physical capital and population growth variables have positive and negative impacts on real GPP per workforce respectively.

It should be noted that the two-way within estimation of the equation (7.1.3) is the most appropriate way of estimation when the unobservable effects are assumed as fixed parameters, and that the one-way random effect estimation of the same equation with time specific effects is the most appropriate method of estimation when the unobservable effects are assumed as random. However, since those two methods produce different results, it is required to determine which of the fixed or random effect estimation is more appropriate for the augmented Solow model. To do so the Hausman test statistics, which is described in the section (5.3), will be calculated.

Table 20. The Regression for InGPPperworkforce by One-way Random Effect Estimation with only Time Specific Effects in Augmented Solow Model

Innongrd	-0.068 (0.067)
Inbasic	0.235* (0.090)
Insecd	0.044 (0.070)
Inhigh	0.047 (0.066)
InIND-In(n+g+δ)	0.056* (0.011)
$\text{InGPPperworkforce}_{t-1}$	0.690* (0.042)
Constant	-0.189 (0.353)
Observations	260
R-squared	0.885

Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Consider the null hypothesis of H_0^{12}: $E(u_{it}/x_{it}) = 0$ versus H_a^{12}:$E(u_{it}/x_{it}) \neq 0$. In order to test this hypothesis, the difference between the estimated coefficients of the explanatory variables for the one-way GLS and the two-way within estimations, $\hat{q}$, is required. So, $\hat{q} = \hat{\beta}_{GLS} - \tilde{\beta}_{within} = \begin{bmatrix} -0.068 \\ 0.235 \\ 0.044 \\ 0.047 \\ 0.056 \\ 0.690 \end{bmatrix} - \begin{bmatrix} -0.127 \\ -0.133 \\ 0.161 \\ 0.176 \\ 0.059 \\ 0.032 \end{bmatrix} = \begin{bmatrix} 0.059 \\ 0.368 \\ -0.117 \\ -0.130 \\ -0.003 \\ 0.658 \end{bmatrix}$. Then, the Hausman test statistics is obtained by m = $\hat{q}' \left[\text{var}(\hat{q}')\right]^{-1} \hat{q} = 158.71 \sim \chi_6^2$. In that case, the null hypothesis of H_0^{12} is rejected at 5% level of significance as this test statistics is greater than the critical value. This means that the two-way within estimation of the equation (7.1.3) is preferred to the one-way random effect estimation of the same equation with time specific effects.

7.2.2.5. Two-way Fixed Effect Estimation

In the previous sub section, it is concluded that the two-way within estimation of the augmented Solow model with different levels of education is the most appropriate way of estimation for this model. So, the results obtained from this method could be discussed further. In that regard, consider the estimation results of the two-way within estimation of the equation (7.1.3).

Then, the schooling of the labor force at each level is found to have insignificant impact on real GPP per workforce except for the labor force having higher level of education. In that case, the insignificance of the variables lnnongrd and lnbasic could be attributed to the reasoning which is concluded from the two-way within estimation of the model of Knowles (1997). In fact, the workforce without formal schooling and the labor force with schooling at the basic level are working in the jobs or in the sectors with limited value added. Moreover, the labor force without formal schooling and those having basic level of schooling affect the real GPP per workforce negatively. On the other hand, the effects of the labor force with secondary and higher levels of schooling are found to be positive. Those conclusions, however, may be misleading and may be resulted from the model specification; because when the dependent variable is taken as the real GPP, not real GPP per workforce, the two-way within estimation of the model of Knowles implies that the effects of the labor force with basic and secondary levels of schooling are significantly positive.

In this model, only the impact on real GPP per workforce of the education variable for the higher level is significantly positively. Moreover, the higher level of schooling has the greatest contribution to real GPP per workforce. In that case, since the estimated coefficient of the higher education variable is found to be significantly positive, it could be concluded that the direct effects on the real GDP per workforce of the labor force having higher level of schooling become apparent in augmented

Solow model. On the other hand, the labor force with higher educational attainment could improve the level of technology used in the production process through research and development. Hence, the relationship between the more educated workforce and the amount of technology may give some idea about the indirect impact of the higher level of education on real GPP per workforce. The correlation between the natural logarithm of the initial level of technology, $InA_i(0)$, and the higher education variable is calculated in Appendix B. This correlation is found to be positive, which is highly significant, for this model implying an indirect effect of the higher education through technological improvement.

Moreover, for the augmented Solow model, the equality of the estimated coefficients for different education variables is tested. To do so, the null hypothesis of H_0[14]: $\beta_1=\beta_2=\beta_3=\beta_4$ is tested for the equation of (7.1.3). The resulting test statistics is 1.98 which distributes as $F_{3,\ 186}$ implying that the null hypothesis of the equality of the coefficients for the variables Innongrd, Inbasic, Insecd and Inhigh is rejected at 12% level of significance in this model. This means that the impacts of different education levels of the labor force on real GPP per workforce are significantly different, and it could be appropriate to differentiate the education capital into different levels in the augmented Solow model.

7.2.2.5.1. Additional Tests on Two-way Fixed Effect Estimation

In this sub section, heteroscedasticity and serial correlation for the two-way fixed effect estimation of the equation (7.1.3) are tested within the augmented Solow model framework. First of all, the Breusch-Pagan/Cook-Weisberg test for heteroscedasticity is performed for this model. The resulting test statistics, which is distributed as χ_1^2, is 0.05; and since this statistics is smaller than the critical value of 3.84, the disturbances in the equation (7.1.3) are homoskedastic at 5% significance level.

Then, the serial correlation in the idiosyncratic error, v_{it}, for the same equation is tested with the same method used in the section (7.2.1). For the serial correlation testing, firstly the idiosyncratic error for the two-way within estimation of the equation (7.1.3) is predicted, and then the lag values for this predicted term are generated. Then, the equation (7.1.3) is reestimated by inserting the lagged values of the idiosyncratic error into the same equation and the F test for the significance of coefficient for the lagged values is performed. The resulting test statistics is approximately equal to zero, which is distributed as $F_{1,121}$, implying that the null hypothesis of the estimated coefficient for the lagged values of idiosyncratic error being zero is not rejected. This means that the serial correlation in the idiosyncratic error does not exist for the augmented Solow model.

Additionally, the RESET test (using powers of the fitted values of the dependent variable, lnGPPperworkforce) for the model misspecification is performed for the augmented Solow model. The resulting test statistics is 2.06, which is distributed as $F_{3,183}$, and since it is smaller than the critical value the null hypothesis of having no omitted variables is not rejected. So, contrary to the model of Knowles (1997), the regional dummies will not be included in the augmented Solow model with different levels of education.

To sum up, in this chapter, firstly the model of Knowles (1997) and the augmented Solow model with different levels of education are tested for the poolability within the panel data framework. Then, those models are estimated separately with the single and pooled OLS, and fixed and random effect estimation methods. At the same time, the hypothesis tests are performed in order to determine appropriate ways of estimation for each model. It is found that the two-way fixed effect estimation is the most appropriate way for both the Knowles's model and the augmented Solow model. In addition, it can be concluded from the results of the two-way within estimation of the model of Knowles and the two-way within estimation of the augmented Solow model that the educational attainment of the workforce above the basic level affects the real GPP (and the real GPP per workforce) in Turkey more than the ones with schooling below the basic levels.

CHAPTER 8

CONCLUSION

One of the significant issues arising from the growth theory is the determination of the determinants of output level and sources of growth for economies. The neoclassical and the endogenous growth theories do support different views regarding this subject. In fact, the supporters of neoclassical theory assert that the economies could grow only at the rate of technological progress in the long run; whereas it is possible for economies to growth perpetually without technological improvement in endogenous growth models. The continuous economic growth that the endogenous theory supports is due to the absence of diminishing returns to factors of production which is assumed by the neoclassical economists. It is the inclusion of the human capital into the production process that eliminates the diminishing returns in the endogenous growth theory. According to the endogenous growth theory, the accumulation of human capital through research and development, learning by doing or knowledge spillovers eliminates the tendency for diminishing returns to the factors of production and makes increasing returns possible.

The human capital which could be defined as the skills, knowledge, attributes and competencies of labor force that generates economic value could be improved by investment in education, training and health (OECD, 2001: 18). Indeed, although the "human capital" is a broad concept, in the literature it is mainly available in the two forms: education and health. As a matter of fact, the more educated and healthier the workforce are; the more productive they become and hence the more they produce. However, this study focuses on the education side of the human capital.

This study contributes the literature in two respects. Firstly, the human capital, in the form of education, is disaggregated into four parts in order to see the effects of different levels of education on real GDP and real GDP per workforce in Turkey. This thesis also puts forward the importance of educational attainment of the labor force at different levels by providing evidence for the case of Turkey. Secondly, the panel data, which is not commonly employed in growth models, is constructed for Turkey and used in the empirical analysis of this thesis. The usage of panel data methods provides an advantage over the simple OLS estimations of the empirical models. In fact, the most important advantage of the panel data is that it allows the specification of the individual and time specific effects.

In this thesis, the model introduced by Knowles (1997) and the augmented Solow model of Islam (1995) which is extended to include different levels of education are estimated in order to examine the relationship between the four levels of education constructed (no schooling, and schooling at basic, secondary and higher levels) and the real GDP (real GDP per workforce for the augmented model) within the panel data framework. Those models are used since they are more applicable to the panel data used in this study.[65] The model of Knowles is constructed directly from the Cobb-Douglas production function. On the other hand, the augmented Solow model, which is constructed from the out-of-steady state behavior of an economy, exhibits the dynamics of the economy.

In the empirical analysis, firstly, the models of Knowles (1997) and the augmented Solow model constructed in the study are tested for the poolability of the data available in both cross section (provinces) and time (year) dimensions respectively. Even though the time dimension (5 periods in the model of Knowles and 4 periods in the augmented Solow model) is not enough to test some of the hypothesis for

---

[65] In most of the studies examining the impacts of different levels of education such as Liu and Armer (1993), time series data is used; and the most appropriate model available for panel data is the model introduced by Knowles (1997). Hence, the Knowles's model is employed in the present study. Moreover, since the augmented model constructed by Islam (1995) is appropriate for panel data, the extension of his model (including different levels of education) is also used in this study.

poolability, the results of the remaining hypothesis tests for poolability support our preference of the usage of panel data methods in the empirical study.

Secondly, in order to examine the impact of different educational levels on real GPP and real GPP per workforce in Turkey and to determine which level of education affects the real GPP most; the model introduced by Knowles and the augmented Solow model with different levels of education are estimated with the single OLS, pooled OLS and fixed and random effect estimation methods.[66] At the same time, in the empirical analysis, some tests are performed so as to determine the most appropriate method of estimation for those models. Finally, the estimation results obtained from the best methods of estimation for the models are examined with additional tests such as tests for autocorrelation and heteroscedasticity in a more detailed way.

According to the results of hypothesis tests performed to determine the most appropriate way of estimation for the model of Knowles (1997), the most appropriate way is found to be the two-way fixed effect estimation. In this model, the two-way fixed effect estimation implies that the labor force without formal schooling and those with schooling at the higher level are found to have no significant impact on real GPP in Turkey. This result could be due to the fact that the workforce having no formal schooling are working on the jobs or sectors with limited value added. Moreover, as in the endogenous growth models, some evidence regarding the indirect effect of the higher educational attainment of the workforce on real GDP through technology is also found in Appendix B. This is indicated by the

[66] It could be mentioned that advanced economies are more likely to allocate more resources in education and hence education capital in those countries would be higher than in developing countries. That is, not only education generates output but also output growth causes higher educational attainment. This could lead a possible problem of endogeneity. However, due to the lack of proper instrumental variables, the method of instrumental variables for our models is not employed in this thesis.

significantly positive correlation between the level of technology and the labor force having higher level of schooling. On the other hand, the labor force having basic and secondary levels of education are found to affect the real GPP positively; and the secondary level is found to have the greatest impact on real GPP in Turkey. Indeed, a one percent increase in the amount of basic and secondary school graduates in labor force leads to 0.32 and 0.50 percent increases in real GPP of Turkey respectively. Furthermore, the absence of the omitted variables is rejected for the model of Knowles and hence this model is extended to include the regional dummy variables for Turkey. Nevertheless, regarding the impact of different levels of education, the estimation of the same model with regional dummies does not produce different conclusions from the estimates obtained with the model without dummies.

Hypothesis tests are performed to determine the most appropriate way of estimation of the augmented Solow model. The test results imply that the best way for this model is the two-way fixed effect estimation. The two-way fixed effect estimation results of the augmented Solow model are different from those obtained from the two-way within estimation of the model of Knowles (1997). For instance, in the augmented Solow model, the impact of the labor force with schooling at the basic and secondary levels on the real GPP per workforce in Turkey are insignificant. Moreover, whereas the workforce with basic level of schooling negatively affects the real GPP per workforce, the labor force with secondary level of schooling has a positive impact. In the augmented Solow model, the higher level of education is significant and positive and it has the greatest impact among all other education levels on real GPP per workforce. In fact, one percent rise in the share of the labor force with higher level of schooling increases the real GPP per workforce by 0.18 percent in Turkey.

In conclusion, even though the models estimated in this study within the panel data framework produce different results regarding the effects of the different levels of education, the empirical study for those models puts forward the importance of the differentiated education capital in determining the determinants of output level in Turkey.[67] As in the case of the developed countries, the educational attainment above the basic level has the greatest contribution to development of Turkey. This chapter concludes with the following suggestions about the educational system in Turkey.

- The number of students enrolled in secondary education after basic education has increased after the introduction of 8-year compulsory education (DPT, 2006:40).[68] The statistically significant and positive impacts of secondary and higher levels of education on the development in Turkey found in this study imply that the compulsory education could be increased from 8 to 12 years in order to increase the enrollment rates in the secondary and higher education. However, increasing only the years of compulsory education is not adequate because to do so the state should provide equal opportunities, regarding the access to education. Hence, the resources allocated to secondary education should be increased in order to provide equal opportunities for those enrolling in secondary level of education. Nevertheless, allocating more resources to secondary education does not mean allocating fewer resources to primary and higher levels of education. On the contrary, more resources could be allocated to each level of education while giving more importance to secondary and higher levels.

- Regarding the importance of secondary and higher levels of education in development of Turkey, the resources allocated to secondary and higher levels of education could be raised by encouraging the private sector to invest in those levels of education. The encouragement of the private sector is important; since the youth

[67] It should be noted that one of the models estimated in this study investigates the impact of education on real GPP (model of Knowles (1997)) and the other model examines the impact of education on real GPP per workforce (augmented Solow model); hence the results of two models are in fact not comparable. On the other hand, when the estimation results of those models are considered as the education capital affecting development level of Turkey, it could be concluded that those models produce different results regarding the impact of different levels of education.

[68] The compulsory education is increased to 8 years with the law No. 4306 dated 18.08.1997 since 1997/98 school year.

population in Turkey is rather large and public resources devoted to education are not large enough to cover the nation's needs.

- Nowadays, the necessity of raising the quality of education, as well as the quantity, is an important issue for the education system in Turkey (DPT, 2006:40). Increasing only the number of students may create the problem of low quality of education. The measures such as development of teacher qualifications, updating of curricula and elimination of the deficiencies of the physical infrastructure could increase the quality of education at each level (DPT, 2007:203). Since the quality of the basic education facilitates the transition of students from primary level to secondary and higher levels of education, the quality of the basic education should be given more importance.

- Since the contribution of higher education to the development of Turkey is the largest among other levels of education, in order to increase the enrollment ratio in higher education scholarships for them could be provided. Moreover, regarding the development of the quality of higher education, as well as the quantity, the guidance and consultancy about selection of proper programs in universities for the students in secondary education could be improved. Those students already enrolled in higher education, who make incorrect decisions about their programs, take the university entrance examination again. This causes inefficiency for both the student and the economy.

REFERENCES

Baltagi, B. H. (2001) *Econometric Analysis of Panel Data*, Second Edition, John Wiley & Sons, Ltd.

Barro, R. J. (1991) "Economic Growth in a Cross Section of Countries", The Quarterly Journal of Economics, Vol. 106 (2): 407-443.

Barro, R. J. (2001) "Human Capital and Growth", The American Economic Review, Vol. 91 (2): 12-17.

Barro, R. J. and Lee, J. W. (1993) "International Comparisons of Educational Attainment", National Bureau of Economic Research Working Paper 4349, Cambridge.

Barro, R. J. and Lee, J. W. (1994) "Sources of Economic Growth", Carnegie-Rochester Conference Series on Public Policy, North-Holland, Vol. 40: 1-46.

Barro, R. J. and Lee, J. W. (1996) "International Measures of Schooling Years and Schooling Quality", The American Economic Review, Vol. 86 (2): 218-223.

Barro, R. J. and Sala-I Martin, X. (2004) *Economic Growth*, Second Edition, MIT, Cambridge.

Bassanini, A. and Scarpetta, S. (2002) "Does Human Capital Matter for Growth in OECD Countries? A Pooled Mean-Group Approach", Economics Letters, 74: 399-405.

Benhabib, J. and Spiegel, M. M. (1994) "The Role of Human Capital in Economic Development Evidence from Aggregate Cross-country Data", Journal of Monetary Economics, pp. 143-173.

Bils, M. and Klenow, P. J. (2000) "Does Schooling Cause Growth?", The American Economic Review, Vol. 90 (5): 1160-1183.

Breusch, T. S. and Pagan, A. R. (1980) "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics" Review of Economic Studies, 47: 239-253.

Cabelle, J. and Santos, S. S. (1993) "On Endogenous Growth with Psychical and Human Capital", The Journal of Political Economy, Vol. 101 (6): 1042-1067.

Cass, D. (1965) "Optimum Growth in an Aggregative Model of Capital Accumulation", The Review of Economic Studies, Vol. 32 (3): 233-240.

Cheng, B. S. and Hsu, R. C. (1997) “Human Capital and Economic Growth in Japan: An Application of Time Series Analysis”, Applied Economics Letters, 4: 393-395.

Dessus, S. (2001) “Human Capital and Growth: The Recovered Role of Education Systems”, World Bank Policy Research Working Paper 2632, Washington.

Deniz, Z. and Doğruel, A. S. (2008) “Disaggregated Education Data and Growth: Some Facts from Turkey and MENA Countries”, Middle East Economic Association Annual Meeting, New Orleans.

Devlet İstatistik Enstitüsü (2003) *2000 Genel Nüfus Sayımı Nüfusun Sosyal ve Ekonomik Nitelikleri*, Ankara.

Devlet Planlama Teşkilatı (2006) *Dokuzuncu Kalkınma Planı (2007-2013)*, Ankara.

Devlet Planlama Teşkilatı (2007) *Ekonomik ve Sosyal Göstergeler (1950-2006)*, Ankara.

Devlet Planlama Teşkilatı (2007) *2008 Yılı Programı*, Ankara.

Doğan, S. and Bozkurt H. (2003) “Eğitim İktisadi Büyüme İlişkisi: Türkiye için Kointegrasyon Analizi”, II. Bilgi, Ekonomi ve Yönetim Kongresi, İzmit.

Güngör, N. D. (1997) “Education and Economic Growth in Turkey 1980-1990: A Panel Study”, METU Studies in Development, Vol. 24(2): 185-214.

Grammy, A. P. and Assane, D. (1996) “New Evidence on the Effect of Human Capital on Economic Growth”, Applied Economics Letters, Vol. 4: 121-124.

Gujarati, D. N. (2003) *Basic Econometrics*, Forth Edition, McGraw-Hill, New York.

Gyimah-Brempong, K., Paddison, O. and Mitiku, W. (2006) “Higher Education and Economic Growth in Africa”, Journal of Development Studies, Vol. 42 (3): 509-529.

Hamilton, J. D. and Monteagudo, J. (1998) “The Augmented Solow Model and the Productivity Slowdown”, Journal of Monetary Economics, pp. 495-509.

Hanushek, E. A. and Kimko, D. D. (2000) “Schooling, Labor-force Quality, and the Growth of Nations” The American Economic Review, Vol. 90 (5): 1184-1208.

Hausman, J. A. “Specification Tests in Econometrics” Econometrica, Vol. 46 (6): 1251-1271.

Hsiao, C. (1986) *Analysis of Panel Data*, Cambridge University Press.

Islam, N. (1995) “Growth Empirics: A Panel Data Approach”, The Quarterly Journal of Economics, Vol. 110 (4): 1127-1170.

Kar, M. and Ağır, H. (2006) “Türkiye'de Beşeri Sermaye ve Ekonomik Büyüme İlişkisi: Eşbütünleşme Yaklaşımı ile Nedensellik Testi, 1926-1994”, Selçuk Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 11: 51 – 68.

Knowles, S. (1997) “Which Level of Schooling has the Greatest Economic Impact on Output?”, Applied Econometrics Letter, pp. 177-180.

Koopmans, T. C. (1965) “On the Concept of Optimal Economic Growth”, Cowles Fouindation Pa, Vol. 4: 225-300.

Lau, J. L., Jamison, D. T. and Louat, F. F. (1991) “Education and Productivity in Developing Countries: An Aggregate Production Function Approach”, World Bank Working Paper No. 612.

Lin, T. (2006) “Alternative Measure for Education Variable in an Empirical Economic Growth Model: Is Primary Education Less Important?”, Economics Bulletin, Vol. 15 (15): 1-6.

Liu, C. and Armer, J. M. (1993) “Education’s Effect on Economic Growth in Taiwan”, Comparative Education Review, Vol. 37 (3): 304-321.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992) “A Contribution to the Empirics of Economic Growth”, The Quarterly Journal of Economics, Vol. 107 (2): 407-437.

McDonald, S. and Roberts, J. (1999) “On Testing Cross-Section Growth Models”, South African Journal of Economics, Vol. 67 (4): 240-254.

McDonald, S. and Roberts, J. (2002) “Growth and Multiple Forms of Human Capital in an Augmented Solow Model: A Panel Data Investigation”, Economics Letters, 74: 271-276.

McMahon, W. W. (1998) “Education and Growth in East Asia”, Economics of Education Review, Vol. 17 (2): 159-172.

Middendorf, T. (2005) “Human Capital and Economic Growth in OECD Countries”, Rheinisch- Westfälisches Institut Discussion Papers 30, Rheinisch-Westfälisches Institut für Wirtschaftsforschung.

Milli Eğitim Bakanlığı (2001-2008) *Milli Eğitim İstatistikleri Örgün Eğitim 2000-2008*, Ankara.

OECD (2001) *The Well-being of Nations: The Role of Human and Social Capital*, Paris.

Oketch M. O. (2006) "Determinants of Human Capital Formation and Economic Growth of African Countries", Economics of Education Review, pp. 554-564.

Öğrenci Seçme ve Yerleştirme Merkezi (1998-2008) *1997-2007 Öğretim Yılı Yükseköğretim İstatistikleri*, Ankara.

Papageorgiou, C. (2003) "Distinguishing Between the Effects of Primary and Post-primary Education on Economic Growth", Review of Development Economics, pp. 622-635.

Petrakis, P. E. and Stamatakis, D. (2002) "Growth and Educational Levels: A Comparative Analysis", Economics of Education Review, Vol. 20: 513-521.

Psacharopoulos G. and Arriagada A. M. (1986) "The Educational Composition of the Labor Force: An International Comparison", International Labor Review, Vol. 125 (5): 561-574.

Ramanathan, R. (2002) *Introductory Econometrics with Applications*, Fifth Edition, South Western.

Ramsey, F. P. (1928) "A Mathematical Theory of Saving", The Economic Journal, Vol. 38 (152): 543-559.

Romer, P. M. (1989) "Human Capital and Growth: Theory and Evidence", National Bureau of Economic Research Working Paper 3173, Cambridge.

Sarı, R. and Soytaş, U. (2006) "Income and Education in Turkey: A Multivariate Analysis", Education Economics, Vol. 14 (2): 181-196.

Saygılı, Ş., Cihan, C. ve Yavan, Z. A. (2006*) Eğitim ve Sürdürülebilir Büyüme: Türkiye Deneyimi Riskler ve Fırsatlar*, Türkiye Sanayicileri ve İşadamları Derneği Büyüme Stratejileri Dizisi No. 7, İstanbul.

Schultz, T. W. (1960) "Capital Formation by Education", The Journal of Political Economy, Vol. 68 (6): 571-583.

Schultz, T. W. (1961) "Investment in Human Capital", The American Economic Review, Vol. 51 (1): 1-17.

Self, S. and Grabowski, R. (2004) "Does education at all levels cause growth? India, a Case Study", Economics of Education Review, No. 23, pp. 47-55.

Solow, R. M. (1956) “A Contribution to the Theory of Economic Growth”, The Quarterly Journal of Economics, Vol. 70 (1): 65-94.

Summers, R. and Heston, A. (1991) “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988” The Quarterly Journal of Economics, Vol. 106 (2): 327-368.

Temple, J. (1999) “A Positive Effect of Human Capital on Growth”, Economics Letters, No. 65, pp. 131-134.

Temple, J. (2001) “Generalization that Aren’t? Evidence on Education and Growth”, European Economic Review, pp. 905-918.

Türkiye Elektrik Dağıtım Anonim Şirketi (Turkish Electricity Distribution Company) (TEDAŞ) (2008) *Turkish Electricity Consumption*.
http://www.tedas.gov.tr

Türkiye Elektrik Kurumu (Turkish Electricity Authority) (TEK) (1985) *Analysis of Turkish Electricity Consumption 1980-1982*, PKD-197 No. 10/1-140, Ankara: Turkish Electricity Authority General Management.

Türkiye Elektrik Kurumu (Turkish Electricity Authority) (TEK) (1991) *Analysis of Turkish Electricity Consumption 1983-1989*, APK-350 No. 10/1-193, Ankara: Turkish Electricity Authority General Management.

Türkiye Elektrik Kurumu (Turkish Electricity Authority) (TEK) (1994) *Analysis of Turkish Electricity Consumption 1990-1992*, APK Ankara: Turkish Electricity Authority General Management.

Türkiye İstatistik Kurumu (Turkish Statistical Agency) (TÜİK) (census years 1975, 1980, 1985, 1990 and 2000) *Generalized Population Censuses*, Ankara: Turkish Statistical Agency.

Türkiye İstatistik Kurumu (Turkish Statistical Agency) (TÜİK) (2008) *Gross Domestic Product of Turkey 1987-2000*, Ankara: Turkish Statistical Agency.
http://www.tuik.gov.tr

Özötün, E. (1980) *İller İtibariyle Türkiye Gayri Safi Yurtiçi Hasılası-Kaynak ve Yöntemler 1975-1978 (Provincial Distribution of the Gross Domestic Product of Turkey-Sources and Methods 1975-1978)*, Publication No. 907, Ankara: State Institute of Statistics.

Özötün, E. (1988) *Türkiye Gayri Safi Yurtiçi Hasılasının İller İtibariyle Dağılımı 1979-1986 (Provincial Distribution of the Gross Domestic Product of Turkey 1979-*

1986), Publication No. 1988/8, İstanbul: İstanbul Chamber of Commerce Research Department.

Wooldridge, J. M. (2002) *Introductory Econometrics: A Modern Approach*, Second Edition, South-Western.

Yaşar, P. (2008) "Alternatif Hasıla Açığı Yöntemleri Tahmin Yöntemleri ve Phillips Eğrisi: Türkiye Üzerine Bir Çalışma", (Devlet Planlama Teşkilatı Uzmanlık Tezi), Ankara.

http://www.tuik.gov.tr/isgucu/Isgucu.do

http://www.yok.gov.tr/istatistikler/istatistikler.htm

APPENDICES

A. TABLES

Table A.1. The Regression for lnGPPperworkforce by Single OLS Estimations in the Model of Knowles

	1975(1)	**1980**	**1985**(1)	**1990**	**2000**	**1975-2000**
lnnongrd	-0.440* (0.079)	-0.499* (0.069)	-0.544* (0.067)	-0.332* (0.077)	-0.365* (0.069)	-0.445* (0.061)
lnbasic	-0.266*** (0.147)	-0.180 (0.148)	-0.040 (0.193)	-0.222 (0.194)	0.020 (0.164)	-0.133 (0.163)
lnsecd	0.494* (0.166)	0.119 (0.152)	0.024 (0.260)	-0.102 (0.224)	-0.222 (0.202)	-0.096 (0.267)
lnhigh	0.078 (0.120)	0.329** (0.147)	0.270 (0.227)	0.463*** (0.232)	0.463** (0.202)	0.418** (0.250)
lnIND	0.109* (0.027)	0.155* (0.025)	0.234* (0.044)	0.219* (0.033)	0.172* (0.031)	0.211* (0.027)
Constant	2.351** (0.971)	2.515** (0.987)	1.705 (1.072)	1.232 (0.957)	0.373 (0.715)	1.900** (0.954)
Observations	65	67	67	67	67	67
R-squared	0.808	0.821	0.828	0.811	0.816	0.888

(1) The heteroscedasticity-robust standard errors in parentheses.
Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Table A.2. The Regression for lnGPPperworkforce by Pooled OLS Estimation in the Model of Knowles

lnnongrd	-0.381* (0.032)
lnbasic	0.191* (0.060)
lnsecd	0.146** (0.072)
lnhigh	-0.057 (0.055)
lnIND	0.176* (0.015)
Constant	-0.337 (0.363)
Observations	325
R-squared	0.761

Standard errors are in parentheses
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Table A.3. The Regression for lnGPPperworkforce by Within Estimations in the Model of Knowles

	(1)	**(2)**	**(3)**
lnnongrd	-0.502* (0.073)	-0.446* (0.031)	-0.461* (0.121)
lnbasic	0.008 (0.086)	-0.032 (0.068)	-0.052 (0.159)
lnsecd	0.185* (0.068)	0.201** (0.085)	0.216** (0.105)
lnhigh	-0.008 (0.039)	0.136*** (0.074)	0.022 (0.064)
lnIND	0.029 (0.019)	0.169* (0.013)	0.038** (0.017)
Constant	4.076* (0.910)	0.878** (0.393)	4.507* (1.036)
Observations	325	325	325
Number of id/year	65	5	-
R-squared	0.706	0.776	0.944

Model (1) refers to the one-way within estimation with individual specific effects; model (2) refers to the one-way estimation with time specific effects and model (3) refers to the two-way within estimation.
Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Table A.4. The Regression for InGPPperworkforce by Random Effect Estimations in the Model of Knowles

	(1)	(2)	(3)
Innongrd	-0.350* (0.045)	-0.381* (0.032)	-0.482* (0.050)
Inbasic	0.171** (0.071)	0.191* (0.060)	0.093 (0.089)
Insecd	0.169* (0.064)	0.146** (0.072)	0.277* (0.083)
Inhigh	-0.049 (0.043)	-0.057 (0.055)	0.047 (0.061)
InIND	0.111* (0.017)	0.176* (0.015)	0.100* (0.015)
Constant	0.004 (0.490)	-0.337 (0.363)	0.628 (0.480)
Observations	325	325	325
Number of id	65	5	-

Model (1) refers to the one-way random effect estimation with individual specific effects; model (2) refers to the one-way estimation with time specific effects and model (3) refers to the two-way estimation.
Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Table A.5. The Regression for InGPPperworkforce by Pooled OLS Estimations in Augmented Solow Model with Different Education levels

Innongrd	-0.237* (0.041)			
Inbasic		0.381* (0.074)		
Insecd			0.110** (0.036)	
Inhigh				0.156* (0.035)
InIND-In(n+g+δ)	0.069* (0.012)	0.064* (0.013)	0.073* (0.013)	0.065* (0.013)
$\text{InGPPperworkforce}_{t-1}$	0.632* (0.044)	0.738* (0.038)	0.729* (0.042)	0.693* (0.042)
Constant	-0.968* (0.167)	-0.416** (0.186)	-0.489** (0.205)	-0.116 (0.234)
Observations	260	260	260	260

Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

Table A.6. The Regression for InGPPperworkforce by One-way Within Estimation with Individual Specific Effects in Augmented Solow Model with Different Education Levels

Innongrd	-0.560* (0.074)			
Inbasic		0.313** (0.146)		
Insecd			0.394* (0.060)	
Inhigh				0.408* (0.047)
InIND-In(n+g+δ)	0.087** (0.025)	0.170* (0.026)	0.095* (0.026)	0.101* (0.022)
$\text{InGPPperworkforce}_{t-1}$	0.022 (0.071)	0.358* (0.066)	0.022 (0.077)	-0.006 (0.069)
Constant	-1.062** (0.361)	-1.883* (0447)	0.558 (0.561)	0.836*** (0.487)
Observations	260	260	260	260

Standard errors are in parentheses.
*** Significant at 10%; ** significant at 5%; * significant at 1%.

B. CORRELATION BETWEEN THE INITIAL LEVEL OF TECHNOLOGY AND THE HIGHER EDUCATION VARIABLE

It should be remembered that the fixed effect estimation methods assume individual and time specific effects are fixed parameters to be estimated. It is possible to estimate the individual and time specific effects separately by using least squares dummy variable estimation method. So, the derivation of the individual specific term for the model of Knowles (1997) and the augmented Solow model with different levels of education will be given within the fixed effect estimation framework.

First of all consider the model of Knowles which is estimated in this study. As concluded in the section (7.2.1), the best mean of estimation for this model is the two-way within estimation. Then, since the initial technology term $InA_i(0)$ is time invariant, it is included in the individual specific effects. One could ignore the other variables which are individual specific and affect the natural logarithm of the real GPP for this model. That is, one could assume that the estimated individual specific effects comprise only the impact of initial technology specific to each province. Then, under this assumption, by estimating the dummy variables for each province with the two-way fixed effect estimation in the model of Knowles, the estimation of the $InA_i(0)$ is obtained from the summation of the product of individual dummies and the estimated coefficients of those dummies.

Then, the correlation coefficient between the estimated values of the initial technology level, which is obtained from the two-way within estimation of the equation (7.1.2), and the natural logarithm of the number of labor force with higher education degree is 0.606, which is significant at 1% level (see Table B.1)

Table B.1. The Correlation Matrix Between $\ln A_i(0)$ and lnhigh Resulting from the Two-way Within Estimation of the Equation (7.1.2)

		lnhigh	lnAi(0)
lnhigh	Pearson Correlation Sig. (2-tailed)	1.000	0.606* 0.000
$\ln A_i(0)$	Pearson Correlation Sig. (2-tailed)	0.606* 0.000	1.000

*Significant at 1%.

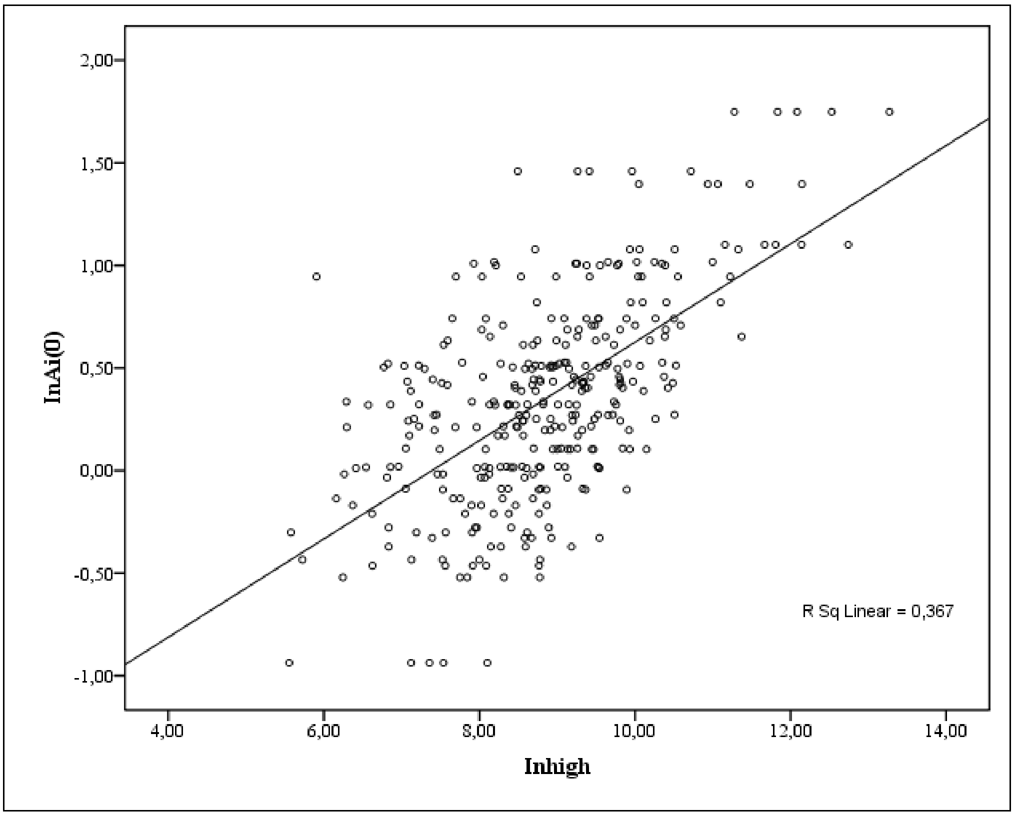

Figure **B.1.** The $\ln A_i(0)$ versus lnhigh with Fitted Line Resulting from the Two-way Within Estimation of the Equation (7.1.2)

Table B.2. The Correlation Matrix between $\ln A_i(0)$ and lnhigh Resulting from the Two-way Within Estimation of the Equation (7.1.3)

		Inhigh	InAi(0)
Inhigh	Pearson Correlation Sig. (2-tailed)	1.000	0.306* 0.000
$\ln A_i(0)$	Pearson Correlation Sig. (2-tailed)	0.306* 0.000	1.000

*Significant at 1%

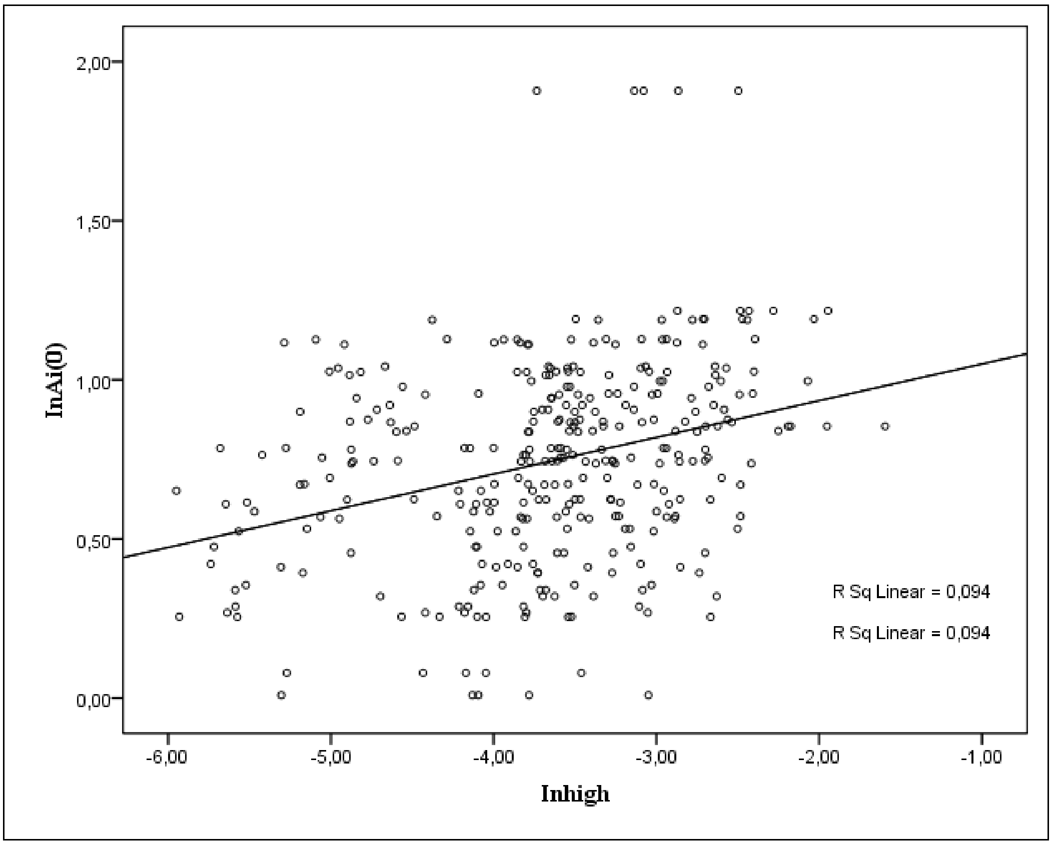

Figure **B.2.** The $\ln A_i(0)$ versus lnhigh with Fitted Line Resulting from the Two-way Within Estimation of the Equation (7.1.3)

Secondly, consider the augmented Solow model with different educational levels. As concluded in the section (7.2.2), the best mean of estimation for this model is the two-way fixed effect estimation. In that case, the estimated values for the individual specific term, $InA_i(0)$, is obtained from the two-way within estimation of the equation (7.1.3) with the same aforementioned method. Then, the correlation coefficient between the estimated values of the initial technology level, which is obtained from the two-way fixed effect estimation of the equation (7.1.3), and the natural logarithm of the number of labor force with higher education degree is 0.306, which is significant at 1% level (see Table B.2)

C. CONVERGENCE

One of the main questions arising from the empirical work on growth is the existence of convergence. Starting from an initial output per capita, a country could reach a point after some time in which the economy has no tendency to diverge. That is, an economy may eventually find itself at its steady state where various quantities grow at a constant rate (Barro and Sala-i Martin, 2004). Such a tendency for an economy is called as convergence. The "convergence" is an important issue in the sense that it gives the chance of studying the economies' long run behavior and examining cross country relationships.

It is generally examined whether countries with initially lower output per capita are able to catch up the ones with higher output per capita. In empirical studies using cross country data, different results with regard to convergence has been obtained, that is, some findings support convergence and some findings do not. Moreover, the existence of convergence has being generally regarded as a support of the textbook Solow model, whereas the absence of it has been considered as evidence of endogenous growth models.

Structurally similar countries with different initial levels of output per capita may end up different growth rates. Countries with lower starting values of real GDP per capita may tend to grow faster than those with having higher initial output per capita that is poor countries are likely to catch up the rich ones. The hypothesis that poor economies growing faster than the rich ones, without conditioning any determinants of the steady state, is called as absolute convergence. In empirical studies some findings have been supported absolute convergence, but some have not. For instance, whereas Barro and Sala-i Martin (2004) find that for 18 countries being members of OECD there is a negative relationship between the 1960 level real per capita GDP and the average annual growth rate of real GDP per capita from 1960 to 2000; the

Solow model estimated by Mankiw et al. (1992) does not produce a significant negative relationship between the growth of output during the period 1960-85 and level of output in 1960. The absence of such convergence leads many researchers to look for more comprehensive tools. In fact, they find that the inverse relationship between the initial GDP per capita and the rate of convergence still hold when some determinants of the steady states such as saving rate, rate of population growth and human capital, are kept constant. Such convergence in empirics refers to conditional convergence, and many findings support it (Barro and Sala-i Martin, 2004). For example, Barro(1991) finds that there is a strong negative relation between the growth of per capita income from 1960 to 1985 and the level of per capita GDP in 1960 when holding various variables including human capital are constant.