LUND UNIVERSITY
School of Economics and Management

Master in Economic Development and Growth

The Impact of Over-education and Under-education on Earnings: Egypt in a Post-Revolutionary Era

Nada Bedir
nada.bedir.390@student.lu.se

Abstract: Previous research has mainly addressed the impact of over/under-education on wages in developed countries. Therefore, this paper attempts to fill the void gap in the literature by empirically examining the impact of over-education and under-education on wages in a developing country, namely Egypt, using the 2012 Egyptian Labor Market Panel Survey (ELMPS). In order to do so, actual years of educations are divided into years of required-education, over-education, and under-education using the realized matches approach. Two modified specifications of the semi-logarithmic Mincer equation are used to estimate the returns: the ORU specification proposed by Duncan and Hoffman (1981) and the dummy variables specification proposed by Verdugo & Verdugo (1989). We contribute to the literature by employing an instrumental variable approach in order to account for the endogeneity of the three components of education. Our results indicate that using the conventional OLS method leads to an under-estimation of the returns to over-education because of ability bias. It is found that returns to over-education are positive and in fact they are higher than returns to adequate education, which contradicts previous literature findings. It is also found that there is a tradeoff between over-education and years of experience.

Key words: education-occupation mismatch, over-education, under-education, instrumental variables, Father's education, Egypt

EKHM52
Master thesis, second year (15 credits ECTS)
June 2014
Supervisor: Anders Nilsson
Examiner: Jonas Helgertz

Acknowledgements

To my Supervisor, Anders Nilsson, thank you for your time, excellent guidance, and insightful advice. It was really a privilege to have you as a supervisor.

To the wonderful MEDEG family, thank you for a once in a life time experience. These two past years would have never been the same without you all. I wish you all the success in the future.

Special thanks to Elisavet Avgeropoulou, Maria Miguel, and Ximena Xativa for their valuable comments and suggestions.

And most importantly to my parents, thank you for all your support, love, and care. Without you I would not have been where I am today.

Any errors or mistakes are my sole responsibility.

Table of Contents

List of Tables and Figures

1. Introduction

Understanding the relationship between education and labor markets is important not only to students but also to parents, educators, economists, and policy makers. There is no doubt that the quality of education and the number of years of schooling has a great impact on the future career of a person. The risk of being unemployed or even employed in a job that does not match ones education, especially after long years of schooling, has both negative economic and social consequences on graduates and their societies.

Matching between a person's job and his/her education and skills is important to realize the human capital stock of the work force in the most efficient way. The existence of education-occupation mismatches has negative effects on the economy (McGuinness, 2006). On the macro level, national welfare is potentially reduced as the skills of workers are not fully utilized. Tax revenues that are allocated to education "are also being wasted on equipping individuals with non-productive education" (Mcguinness, 2006). On the firm level, Tsang (1987) studied the effect of over-education on labor productivity and found that over-education of workers has negative effects on output. On the individual level, overeducated workers are not utilizing their full investment in human capital and studies have found that they tend to receive lower returns in comparison to those who are adequately matched (Hartog, 2000).

The concepts of over/under-education refer to whether the individual possesses a level of education that is above or below the required level of education to perform a certain job (Hartog, 2000). In his book "the Overeducated American", Richard Freeman (1976) was the first to discuss the over-education phenomenon using aggregate level data. He concluded that the decline of returns to schooling among new entrants to the labor market is due to the fact that the supply of college graduates is higher than the demand for them in the labor market. Studies on over/under-education were revitalized when Duncan and Hoffman (1981) published their paper using individual level data from the U.S. to empirically test the incidence and the returns to over-education.

Currently, there exist a substantial number of studies that explore the returns on over- and under-education.[1] Most of these studies, however, focus on developed countries. Few studies explored this area in developing economies, such as the study

[1] For comprehensive review see Hartog (2000), McGinness (2006), and Leuven & Oosterbeek (2011)

done by Quinn & Rubb (2004) on Mexico and Abbas (2008) on Pakistan. The empirical findings of these studies agree that the returns to over-education are positive and significant with the size of the returns bring less than returns to required level of education. They also find that the returns to under-education to be negative with the magnitude being smaller than the returns to required-education in absolute terms (Hartog, 2000).

The aim of this thesis is to examine how the Egyptian education system is matched to the demand for workers in the labor market utilizing the recent Egyptian labor market survey of 2012 in order to formulate policies that suit the needs of the country. It is relevant to start considering the effects of over/under-education in developing countries and particularly in Egypt for several reasons. Firstly, since improving the level of human capital is known to contribute positively to growth, many developing countries have implemented policies to increase enrollment rates and the educational level. In Egypt, the literacy rate among adults (+15) increased from 44.3% in 1986 to 77.9% in 2011 (UNESCO, 2012). Literacy rates are even higher among youth (15-25) as it increased from 63.3% to 89.9% in 2011 (UNESCO, 2012). Secondly, as of 2006, the incidence of over-education in the private sector was estimated to be around 12.1% while the incidence of under-education was estimated to be around 29.7% (El-Hamidi, 2008). This means that around 42% of the workers are performing jobs that are not in line with the years of education they received. Therefore, it is important to measure the effects of such mismatch on the worker's wages in the labor market. Thirdly, in the wake of the January 25th revolution of 2011, it is important to re-assess the labor market and educational conditions in the Egyptian Labor Market in order to formulate suitable policies that can improve educational outcomes and create decent job opportunities.

Using the Egypt Labor Market Panel Survey of the year 2012, the paper shall answer the following question; *what are the returns to over- and under-education?* This question is considered important because the impact of education on earnings has vital implications for educational policies and individual investment decisions. The analysis is conducted on a sub-sample of wage earners. individuals who are out of the labor force, self-employed, agricultural sector workers, part-times workers, and those whose wages are below the minimum wage are excluded.

The paper shall contribute to the literature by estimating the returns when investments in education are endogenous using an instrumental variable approach. Estimating the returns to over/under-education using the ordinary OLS approach may produce biased results.[2] Education can be considered endogenous because there is a simultaneous relationship between investment in education and expected wages. This is because years of education affect wages and also the decision to invest in education is affected by the expected wages in the future (Card, 2001). Another reason why education can be endogenous is due to the existence of a third factor, such as ability, which is correlated both with education and wages. Not accounting for ability in this case can bias the OLS results as well and lead to inconsistent estimates (Card, 2001).

The paper shall also consider different theoretical frameworks in order to better understand the phenomenon of over/under-education. Previous studies have mainly considered the human capital theory (Becker, 1965), job-competition model (Thurow 1975), and assignment models (Sattinger, 1993). Empirical results usually provide support to the assignment model interpretation. In addition to the previous theories, our paper contributes to the literature by explaining the phenomenon of over/under-education using imperfect and segmented labor market theories.

The paper is structured as follows; Section two discusses the definitions of education-occupation mismatches. Section three provides an overview of the Egyptian labor market and educational system. Section four explores the theoretical framework. Section five reviews previous literature from developed and developing countries. Section six introduces the methodology. Section seven presents the data used to conduct the empirical study along with the descriptive statistics. Section eight reports and discusses the results. Section nine concludes.

[2] The type of biases that can result from estimating the returns using OLS will be discussed in details in section

2. Defining Education-Occupation Mismatch

2.1 Definition

There are two types of education-occupation mismatches: horizontal mismatch and vertical mismatch (Piracha & Vadean, 2012). Horizontal mismatch occurs in cases in which the field of education differs from the job one performs (Hartog, 2000). On the other hand, vertical mismatch occurs when the required level (years) of education for a specific job is different from the actual level of education of the worker[3] (Piracha & Vadean, 2012). The analysis in this paper is based on the vertical mismatch definition due to data availability.

Under the vertical mismatch definition, the required level of education refers to the number of years of education that are needed to perform a certain job. The employee is considered overeducated if the level of education required for the occupation is lower than his/her level of education. In contrast, the employee is undereducated if the level of education required to perform the job exceeds his/her level of education.

Based on previous studies, there are three methods that can be used to estimate the required years of education for each occupation, namely, the job analysis method, the self-assessment method, and the realized matches method (Hartog, 2000). The estimated required level of education is then compared to the actual level of education of the worker in order to determine the years of over-education and under-education. The following sub-section explores the three methods in more details.

2.2 Measuring the Mismatch

As mentioned above, there are three different methods that can be used to measure the mismatch. The choice of the method is usually based on the availability of data.

2.2.1 Job Analysis Method (JA)

It is considered an objective method to measure mismatch because it is based on a "systematic evaluation by professional job analysts who specify the required level (and type) of education" (Hartog, 2000). Professional analysts determine the required years of

[3] The terms "required level of education" and years of education are used interchangeably

education based on their job title. The best examples of this method are the dictionary of occupational titles (DOT) of the United States and the Standard Occupational Classification System in the UK. DOT was used by Hartog (1980) and a similar Rumberger (1987) in their mismatch studies.

One of the criticisms of this approach is that having the same job title does not necessarily mean that workers are performing the same tasks and may even require different set of experiences and different educational levels (McGuinness, 2006).

In addition, it is likely that the required level for a certain job changes over time due to changes in the relative supply of educated workers in the economy. For example, Mason (1996) mentions that with increased educational levels, college graduates are being employed in mid-clerical positions where in the past these positions were usually held by workers with secondary level of education. Therefore, the required level of education needs to be updated regularly by analysts. Otherwise it would be considered outdated. Hartog (2000) mentions that different editions of the DOT, that were carried out over long intervals copy previous analyses with only slight modifications which might lead to unreliable results. It is also considered expensive to implement on a large scale (Sala, 2011).

2.2.2 Worker's Self-assessment (WA)

The self-assessment method is based on workers' own valuation of the level of education required to perform their job (Hartog, 2000). It is, therefore, considered very subjective as it depends on the employees' perceptions. This method has been used by Duncan and Hoffman (1981) and Sicherman (1991) among others.

Workers are asked about their perception of the required level to perform the job and then the difference between their actual level of education and the required level will determine whether the worker is over/under-educated. Studies based on this method have used different questions in order to inquire about workers' views. For example, Duncan & Hoffman (1981) ask the following question: how much formal education is required to get a job like yours? On the other hand, Alba-Ramirez (1993) asks the following: what kind of education does a person need in order to perform your job?

By examining these two questions, one can infer that the first question focuses on recruitment standards while the second question focuses on the requirements for job performance. This means that the same person can give different answers to different questions which makes the use of this method questionable at times. The positive aspect of this method is that it provides up to date information about current education requirements (Sala, 2011). On the other hand, there are two sources of potential biases when using this method. Firstly, workers may report higher years of education required to perform their jobs rather than the true required years in order to "inflate the status of their position" (Hartog, 2000). Another bias can be generated if the requirements of the job have changed over time and the employee hired before the change is not affected by such a change (McGuinness, 2006).

2.2.3 Realized matches method (RM)

In this paper, the realized matches technique will be used due to data availability. It is an objective statistical method that can be estimated using two approaches. In the first approach, the mean level of schooling is calculated for workers in each occupation. According to Verdugo and Verdugo (1989), "workers whose educational level is at least one standard deviation above mean schooling is considered over-educated and workers whose education level is at least one standard deviation below are considered under-educated." Those whose educational level is between plus or minus one standard deviation are considered to be adequately educated (Verdugo and Verdugo, 1989)

The second approach employed by Cohn and Khan (1995) and Kiker et al. (1997) differs from the first one in that it uses the mode of level of schooling instead of the mean. It also does not consider standard deviations. Workers who are above the mode value are considered over-educated while those below the mode value are considered under-educated. Kiker et al. (1997) prefers the use of mode as the mean value is sensitive to outliers.

This method reflects the actual educational level of workers in each occupation "as determined by hiring standards and labor market conditions" rather than the job requirements per se (Hartog, 2000).

The main criticism for this method is related to the "arbitrary nature of the cut-off point" which can lead under/over-estimation of the incidence of over/under-education **(**McGuinness, 2006). For example, in the case of an existence of high number of over-educated workers in a certain job, this would result in a higher cut-off point and therefore the number of over-educated workers would be underestimated.

3. The Case of Egypt

Before examining the theories that explain potential reasons behind the existence of education-occupation mismatches, we shall provide an overview of the Egyptian Labor Market and the educational system.

3.1 The Labor Market

Being a developing country, Egypt seeks to develop a well-functioning labor market in order to achieve economic and social progress. In today's globalized world where the required labor market skills are constantly changing one of the main challenges facing the Egyptian labor market is the need to create new jobs. The country's economic performance has been unstable in the last decade. However, even in periods of satisfactory growth rates, the pace of job creation was not enough to meet the increased the number of new entrants mainly due to two problems (Hassan, M. & Sassanpour, C., 2008). The first one is associated with the demographic youth bulge which poses great challenges for the labor market in Egypt as it is unable to provide sufficient labor demand in the formal public and private sector (Assaad & Roudi-Fahimi, 2007).

The second problem is related to the skill mismatch between the supply and the demand in the labor market. There is a strong need to reform education and training to equip job seekers with the skills and knowledge needed to succeed in getting a suitable job in the labor market. It is imperative to develop the educational and training systems in order to meet the new needs of the labor market.

These structural problems facing the Egyptian labor market have accumulated over time due to certain policies that were in place as well as several macro factors.

Starting in 1961, the government in Egypt undertook large public investments as part of the nationalization process and implemented a guaranteed employment policy that

promises employment for recent graduates in the public sector (Hassan, M. & Sassanpour, C., 2008). One of the consequences of this policy is the expansion of the public sector. The policy also caused unemployment to decrease to very low levels reaching around 2% as it absorbed a substantial number of new entrants (Hassan, M. & Sassanpour, C., 2008). Nevertheless, this policy proved to be unsustainable as it was associated with high costs and inefficiencies. Starting 1974, due to the introduction of the open door policy[4] and the rise in oil prices, the Egyptian economy experienced high levels of growth which reached around 10% after the mid-seventies. Nevertheless, unemployment was rising during this period. According to the World Bank Databank (2014), in the year 1980, unemployment reached around 5.2%. During the late 80s and the early 90s, GDP growth slowed down due to several factors including external shocks, decline in investments, and high external debt. During the 1990s, unemployment ranged between 8-10%. Even when the growth rate started picking up again starting in 2003/2004 and up until 2007, unemployment kept rising, and the economy was not able to create enough jobs to meet the increased demands for new entrants.

The increasing unemployment rate and the lack of suitable/decent job opportunities was one of the main drivers of the 2011 revolution. According to Green (2011), "two-thirds of Egyptians are under 30, and each year 700,000 new graduates chase 200,000 new jobs." Nevertheless, after the revolution of 2011, the Egyptian economy is facing economic stagnation with an increasing public deficit, weak economic growth, high unemployment, etc. This also contributed to further slowdown in job creation.

3.1.1 Labor Market Characteristics in the year of 2012

According to the World Bank Databank (2014), the labor force participation[5] compromises around 49% of the population in 2012. The male labor force participation rate is around 79% which is significantly higher than female participation rate which only accounts for 24% in the year 2012. The lower percentage of female labor force

[4] The policy encouraged private investment and FDI inflows.
[5] According to the World Bank Databank (2014), "the Labor force participation rate is defined as the proportion of the population ages 15 and older that is economically active."

participation can be explained by traditions and also by women leaving the workforce earlier in order to start a family, etc. (Assad and Krafft, 2013).

The total unemployment rate was estimated to be around 11.9% in 2012. The issue of unemployment is mainly connected to finding the first job. It is severely acute among the youth in the country and especially those with higher education. Women suffer the most from unemployment. Also those who live in urban areas suffer more compared to those living in rural areas. Underemployment was estimated to be 2.6% in 2006 and has increased considerably to reach 9.3% of the labor force in 2012. This increase is expected, however, due to the decline in employment opportunities in the aftermath of the economic crisis that was a consequence of the 2011 revolution.

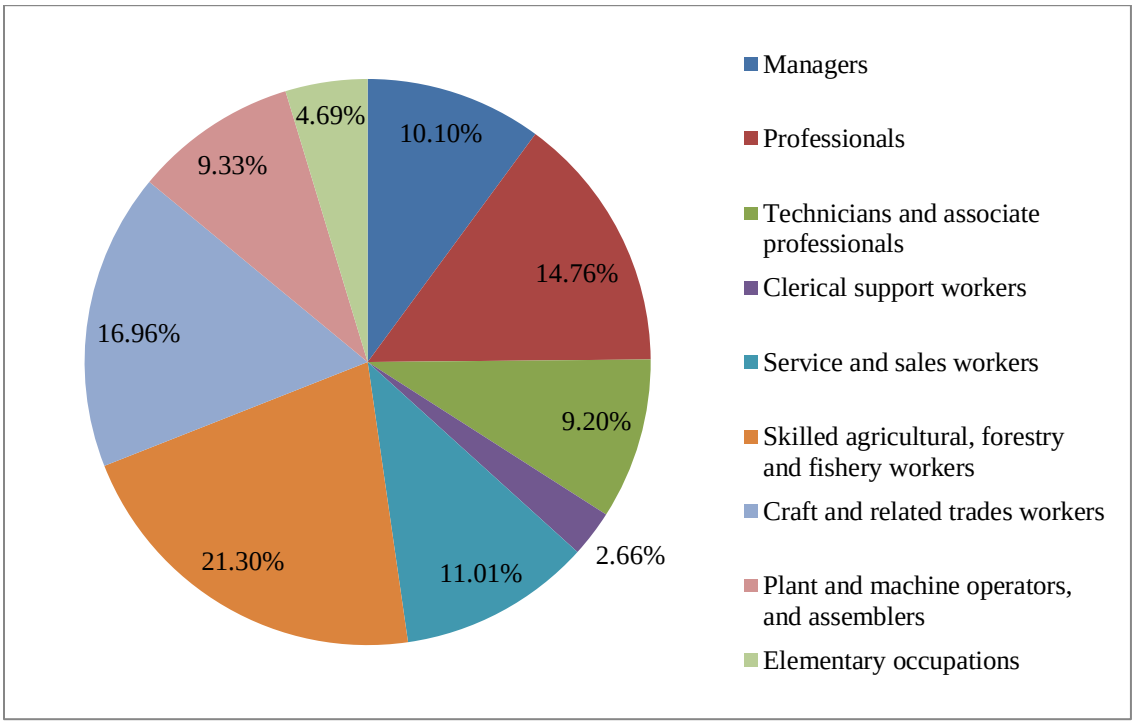
Underemployment[6] is also a major problem in the labor market due to the existence of skill mismatch as well as the overstaffing in the government sector. Estimates from the Egypt Labor Market panel surveys show that visible unemployment rate is on the rise in Egypt (Assaad & Krafft, 2013).

The informal sector in Egypt is very active and acts as an "important shock absorber to the formal sector" (Hassan & Sassanpour, 2008). The informal labor market is estimated to represent around 40 to 60% of total employment and it accounts for 40% of the private sector GDP (Hassan & Sassanpour, 2008). Many of the poor and the uneducated seek jobs in the informal sector. The informal sector provides low-pay jobs and low-security. Unfortunately, very limited data are available on the informal sector in Egypt and therefore it is hard to conduct empirical studies to assess the impact of working in the informal sector on wages, etc.

Figure (1) below presents the composition of Egyptian workers by Occupational Categories as reported in the Egyptian Labor Market Survey (ELMPS) , the agricultural, forestry, and fishery workers constitute the largest category with around 21.30% followed by craft workers with 16.96% and then professionals with 14.76%. The smallest percentage of workers belongs to the clerical support category followed by technicians.

[6] Underemployment "occurs when an individual works less than full time (less than 40 hours per week) because of insufficient employment opportunities"

Figure (1): Composition of Workers by Occupational Categories



Source: ELMPS (Egypt Labor Market Panel Survey) 2012

3.2 Educational System

The educational system in Egypt can be divided into four main stages/levels (Clark, 2013). Firstly, the ***primary schooling level*** which lasts for 6 years. Students usually enter primary schools when they are 6 years old and leave when they are 12 years old. Some, however, start school at a later age. At grades 3 and 6, students have to take a standardized national exam in order to continue to the next level. After passing the qualified exam of grade 6, students move to the ***preparatory schooling level*** which lasts for 3 years. It is worth noting that the compulsory schooling in Egypt is 9 years (6 years of primary education + 3 years of preparatory education). At the end of the preparatory level, students undertake national comprehensive examinations. Upon the successful completion of the exam, students are awarded a Basic Education Certificate after which they are eligible to register for the ***secondary schooling level***. The grades determine whether a student is eligible to apply for general secondary schools, which lasts for 3 years, or technical/vocational secondary schools, which can be either 3 or 5 years for

those with excellent grades. Students with higher grades are the ones that are allowed to get into general secondary schools and they get to choose either the scientific or the literary track (British Council, 2013). Students enrolling in vocational/technical schools have the option to choose from the following three fields: Industrial, commercial, and agricultural. Technical schooling is considered a "less academic track that does not make a student eligible for academic enrollment" (Clark, 2013). The ministry of education is working with the UN and other international organizations to improve the training quality of vocational schooling (British Council, 2013). Upon successful completion of the secondary stage, students receive the General Secondary Education Certificate or Technical/vocational Secondary Education Certificate (Hasan, n.d.). For these three stages, education is administrated by Ministry of Education.

On the other hand, the ***higher educational system*** is administrated by the ministry of higher education. The number of years spent at university level varies by specialization and ranges from 4 to 7 years. As of 2012, there exist around 20 public universities and 23 private universities in Egypt. The public universities and institutions are funded by the state (UNESCO, 2012).

It is worth noting that government expenditure on education as a percentage of GDP is reported to be around 3.8% in the year of 2008 which is much lower when compared to expenditures of similar countries in the region (World Bank, 2014). For example, in Tunisia and Morocco, government expenditure reached 6.2% and 5.5% of the GDP, respectively (World Bank, 2014). This explains the low quality of education that exists in government schools and the deteriorating qualities of government universities that we are witnessing in the recent decades. This also contributed to producing students that are not equipped with the necessary skills and education needed to meet the demands of the labor market, especially in the private sector.

4. Theoretical Framework

Existing labor market theories are used to explain the existence of mismatches in the labor market. Most of the previous studies provide an explanation for the phenomenon of over/under-education by examining three main theories: human capital

theory, job competition, and assignment models. The following section provides an overview of these three labor market theories/models that can be used to explain the phenomenon of over/under-education. In addition, we explore theories of career mobility, labor market segmentation, and labor market imperfections that further deepen our understanding of why such mismatches would exist in the labor market.

4.1 Human Capital Theory

The human capital theory developed by Becker (1964) suggests that workers are paid according to their marginal productivity which is determined by the level of human capital[7]. Firms adapt their technology to respond to changes in the relative supply of educated individuals in the labor market. Firms respond to alternative production techniques that emerge as well as changes in the relative supply of labor in order to minimize costs and maximize profits. If the supply of skilled labors increase, this would cause relative decrease in wages of skilled labor and consequently, firms would adjust their technologies to utilize the abundance of skills labors at lower costs (Duncan & Hoffman, 1981).

Under the classical human capital theory, the existence of education-occupation mismatches can be explained by inefficiencies that occur in the labor market (McGuinness, 2006). Over-education does not necessarily mean that the human capital theory does not hold but rather it might be an indication that firms require time in order to adjust their production technologies to fully utilize the human capital stock (McGuinness, 2006). This suggests that over-education is considered to be a short-run phenomenon under the human capital theory that is supposed to disappear eventually as firms respond to changes in the relative supply. However, if the mismatch lasts for a longer time proves to be a long run phenomenon, this would suggest that the labor market adjusts slowly to changes in relative supply of educated workers which would be considered as a challenge for the neoclassical model.

Based on the classical human capital theory, Mincer (1974) developed an empirical model in which the amount of attained schooling has a direct effect of wages.[8] Nevertheless, as can been in the literature review section, empirical evidence stand

[7] Human capital consists of formal education, on-job trainings, and/or experience (Becker, 1964)
[8] Discussed in details in the Empirical Model Section

against a simple human capital model in which wages are only affected by the actual level of education and experience. Results of the empirical studies are considered consistent with slightly adaptable versions of the Mincer equation (Hartog, 2000).

4.2 Job Competition Model

Another theory that can help explain education-occupation mismatch phenomenon is the "Job Competition" model. In this model developed by Thurow (1975), there is a job queue and a person queue. For each job, there are specific sets of skills and characteristics that are needed to perform the job. The position of each individual in the queue is determined by his/her education and experience which allows the employer to assess the cost of training needed for each individual. The higher the education level and years of experience, the higher the position of the potential employee is in the queue. If individuals are first in the queue and obtain a job then, according to this model, their wages will be determined by the job characteristics. Employers prefer to choose over-educated individuals in order to lower the cost of training even if they will not utilize the additional education these employees poses. In the case in which jobs structure remain the same and the private return to education remain high; individuals may acquire more years of education in order to be standing in a higher position in the queue which improves their chance in getting a job. However, due to the rigid job structure and the fixed job skills requirement, "the underutilization of worker skills will persist" (Tsang and Levin, 1985). According to the job-competition model wages are determined only by the requirements for the job.

4.3 Assignment Model

The Assignment Model developed by Sattinger (1993) can be considered as a combination of the human capital and job competition model. According to Sattinger (1993), "all assignment models specify jobs or sectors available to workers, the relevant differences between workers, the technology relating job and worker characteristics to output, and the mechanisms that assign workers to jobs."

The assignment model suggests that the marginal product of workers, in other words the wage rate, is determined by their human capital as well as the occupation characteristics. Earnings are also linked to "the availability and/or quality of jobs in the

economy" (McGuinness, 2006). Unlike the job competition model, "the allocation process is not merely a lottery" (Sala, 2011), but rather workers are motivated to select certain jobs based on income/utility maximization. Therefore, jobs are not only allocated based on worker's characteristics but also on the job's characteristics.

In other words, the amount of skills supplied by the worker (i.e. the actual years of education), is determined by the costs and expected benefits from investment in education (Groot & Van Den Brink, 1997). On the other hand, the job requirements are the determinants of whether the skills of workers are fully utilized or not (Groot & Van Den Brink, 1997). Therefore, required education, over-education, and under-education are determined by both workers' and the job' characteristics. For example, if an over-educated worker is employed in a job that imposes a ceiling on productivity then he/she may earn less compared to others with the same level of education that were assigned to a job that matches their level of education (Abbas, 2008).

As mentioned before, empirical findings mainly support the assignment model explanation for the existence of mismatch (Hartog, 1988; Hartog & Oosterbeek 1988; Sloane, Battu, & Seaman, 1995).

4.4 Career Mobility Theory

Sicherman and Galor (1990) formulated a theory of career mobility that provides two possible explanations for the phenomenon of over-education. Under the first scenario, workers initially accept an entry level job that requires less years of education in order to obtain experience and hence get promoted later in their career (Büchel & Mertens, 2000). The empirical findings of the paper confirms this hypothesis and finds that over-educated employees are more likely to get promoted within the next year than adequately educated workers (Robst, 2005). The second scenario suggests that over-educated workers may not be considered over-qualified for a particular job if other components of human capital are considered in addition to years of schooling which include experience and on job training (Robst, 2005). Nevertheless, empirical results of Sicherman and Galor (1990) suggest that "overeducated workers have less experience, training, and tenure than adequately educated workers (Robst, 2005). It is important to note that the theory of career mobility does not provide a thorough explanation of the

mismatch phenomenon as it only provides an explanation for the existence of over-education rather than under-education (Büchel & Mertens, 2000). In addition, very few empirical studies have been conducted using this theory in order to explain the mismatch phenomenon in the labor market (Büchel & Mertens, 2000).

4.5 Imperfect and Segmented Labor Markets

As mentioned earlier, Becker's neoclassical theory assumes that labor markets are competitive and that workers are paid according to their marginal product. In reality, however, this is not the case due to several reasons including 1) the existence of information asymmetries between employers and workers 2) transaction costs that arise from matching and job search 3) externalities due to the production and consumption of goods and services that "are not subject to market interactions" (Boeri & Ours, 2008). We shall explore alternative theories that can be used to better understand the functioning of the labor market in presence of asymmetrical information and/or segmentation.

4.5.1 Screening Hypothesis

According to Spence's (1973) screening hypothesis, there exists imperfect information in the labor market and therefore education is used by employers as a signal to identify the most productive workers. Therefore, individuals decide to invest more in education in order to distinguish themselves from others. This would suggest an increase in the level of education of the labor force in relation to the level required by the labor market; according to Tsang and Levin (1985), "if job requirements do not change, the private rate of return from educational investment can stay high and provide continuing incentives for investment in education". If education is used as a signal then over-educated workers will be favored in relation to under-educated workers because they are considered more productive.

4.5.2 Segmented Labor Market Theory

The segmented labor market theory was firstly developed during the early 1970s (Dickens & Lang, 1992). The theory suggests that the labor market consists of different sub-groups -or segments (Dickens & Lang, 1992). According to Fields (2009), a segmented labor market is characterized by two main elements. Firstly, access to jobs is

limited in the economy in the sense that there exists a higher supply compared to the demand of jobs.[9] All qualified workers queue for these highly paid jobs but job opportunities are limited. Therefore, those who are still qualified but are unable to get a matching job would either settle for one with lower pay or become unemployed (Fields, 2009). Secondly, different segments have their own wage and employment policies.

The main criticism from the segmentation theory towards the human capital theory is that market clearing does not occur since "access to some sectors is subject to non-price rationing" (Dickens & Lang, 1992).

The main version of the segmented labor market theory is referred to as Dual labor market theory (Dickens & Lang, 2001). In this version the labor market consists of two segments, the primary and the secondary sectors.[10] In the primary sector, workers are employed in jobs that provide high wages, high returns to education and training, good working conditions, as well as promotion opportunities (Dickens & Lang, 2001). Relationships between employers and employees in this primary sector are usually governed by union contracts or any form of labor law (Dickens & Lang, 2001). On the contrary, the secondary sector is characterized by low pay employment opportunities, bad working conditions, low promotional prospects, and lack of formal contracts –usually short-term employment- (Dickens & Lang, 2001).

Firms that exist in the primary sector are usually referred to as core firms[11] while those that exist in the secondary sector are called periphery firms[12]. Nevertheless, it is unlikely that only high skilled workers exist in core firms while low skilled workers exist in periphery firms. It is more realistic to consider that in the primary sector, core firms consist of a higher percentage of primary workers compared to secondary workers and vice versa (Dickens & Lang, 1992).

The efficiency wage theory can be used to better understand the implications of labor market segmentation (Dickens & Lang, 1992). It can help us answer the question of why employers would not lower the wages when they are faced with excess labor supply.

---

[9] Particularly, there is excess supply for jobs that provide higher wages (Fields, 2009).
[10] The primary and secondary sector can refer to the formal and informal sectors in the economy, respectively
[11] The following industries are usually dominated by core firms: construction, manufacturing, transportation, mining, etc. (Dickens & Lang, 1992)
[12] The following industries are usually dominated by periphery firms: agriculture, textiles, etc. (Dickens & Lang, 1992).

To start with, the efficiency wage theory argues that workers' productivity is not only determined by their level of human capital and ability but also by the wages they receive (Katz, 1986). Therefore employers, at least in the primary sector, prefer to pay some of their workers higher wages compared to the market in order to increase their productivity. In addition, higher wages can be used as a mechanism to increase the cost of job loss and therefore avoid shirking from the employee side (Shapiro & Stiglitz, 1984). In this case imperfect monitoring from the side of the employer make them pay their workers above the market price and therefore wages are used as a discipline device (Shapiro & Stiglitz, 1984).

Under this theory, one can infer that employers pay over-educated workers higher because they consider them a valuable asset. Therefore, higher wages can encourage the over-educated worker to be more productive. Also, in the case of developing countries where the incidence of under-education is expected to be higher compared to over-educated workers, employers may provide over-educated workers with higher wages as they are considered a scare category in the market to reduce the possibilities of them leaving the job.

In conclusion, under the classical human capital theory, labor markets are competitive and employees are paid according to their marginal product. However, the existence of mismatches is a sign of existence of inefficiencies, imperfections and/or segmentations in the labor market.

5. Literature Review

This section examines the previous literature concerned with estimating the impact of over/under-education on wages. We shall focus on the method used and the main findings of each study.

To start with, Duncan and Hoffman (1981) use U.S. data from the Panel Study of Income Dynamics which is a longitudinal survey that provide information on a nationally representative sample of workers. The authors apply self-assessment method where individuals are interviewed and asked about the educational level required to perform their jobs. The reported required years of schooling by each individual is compared with his/her actual education attainment to determine the years of over/under-schooling. This

paper is the first to estimate the returns to over-education using a modified Mincer approach in which actual years of education are decomposed into the required years of education for a specific occupation and years of over-education or under-education. The authors find that returns to over-education are positive and significant but they represent only half of the size of the returns to required schooling.

Rumberger (1987) uses U.S. data from the 1969 Survey of Working Conditions and the 1973 and 1977 Quality of Employment Survey. The author estimates returns to over-schooling using the workers' self-assessment method for the 1969 survey and the 1977 survey. He also estimates returns to schooling using the job analysis method as he has the Dictionary of Occupational titles for the year of 1973. The following are the main findings of his study. Similar to the previous study, it is concluded that returns to over-education are lower than required education. The author related this to the fact that over-educated workers may not fully utilize their education on the job. For the year 1973, similar returns to required and over-schooling were obtained using both the subjective and the objective methods. The returns to over-education vary across different occupational categories and between males and females. For example, in some occupation required schooling is rewarded but surplus schooling is not.

Verdugo and Verdugo (1989) introduce the realized matches technique, explained earlier, to examine over/under-education. Using U.S. data for a sample of white males from the 1980 census, the authors estimate two models. The first model is estimated using the full sample while the second is an occupation-specific one. The authors estimate a different model from the one proposed by Duncan and Hoffman (1981). They use two dummy variables that represent over-education and under-education, respectively. In the first model, eight dummy variables that represent occupations are included. They find that across occupations when controlling for education, overeducated workers earn less than those who are adequately educated. They also find that undereducated workers earn more than overeducated ones. When breaking the sample according to occupation, the results were the same for five occupations out of eight. The authors provide three possible explanations for the results obtained. Firstly, workers who are overeducated are not necessarily more productive. Secondly, undereducated workers may be excellent performers on the job which explains why the

employer chose to hire them. Thirdly, over-education is more likely to exist in jobs where mean education is relatively low. It is important to note that the way the author's interpreted the signs of the coefficients was proven to be wrong by Cohn and Kahn (1995); an issue that shall be discussed in the empirical model section.

Moreover, Groot & Brink (1997) utilize the 1991 British Household Panel Survey to study the degree and the returns to over-education. The main contribution of the paper is the fact that it addresses the endogeneity of schooling when estimating over-education. To measure the returns to education, they estimate three different models. The first is the conventional approach (i.e. OLS) in which education is treated as an exogenous variable. The second model involves the use of instrumental variables to account for the endogeneity of required-over- and under-education. The two instruments used in the analysis are the quarter of birth as well as changes in the compulsory schooling age. Thirdly, the authors estimate an ordered probit model in order to correct for sample selection bias. The authors conclude that previous studies may have overestimated the returns of education due to ability and sample selection biases. The IV model results suggest that returns to over-education are lower than that found using OLS. In addition, sample selection model results show that over-educated workers receive lower returns than adequately educated ones. Another important result is that over-education is a temporary phenomenon and is part of the adjustment period.

Kiker, Santos, and Oliveria (1997) use a unique dataset that consists of over 30,000 workers which was collected by the Portuguese Ministry of Labor in the year of 1991. The authors use three different measures of over/under-education. Firstly, the mean measure in which the over-educated are those who are one SD above the mean and under-educated are those one SD below the mean; secondly, the mode measure in which those who are over-educated are above the mode of the education required for occupation and those below the mode are under-educated; and finally, the job analysis method. They estimate two earnings equations; one based on the Verdugo and Verdugo specification and another based on the Duncan and Hoffman Specification. The results obtained regarding the effect of over/under-education on the earnings of workers in Portugal are similar to previous findings. Workers who are over-educated receive more than their adequately educated counterparts but the size of the return is smaller. In addition, under-

education is rewarded negatively in the labor market. They also find that "there is a trade-off between education and other forms of human capital. More educated workers may require less on-the-job training, while less educated workers may supplement their formal education with job experience, allowing workers with different years of education but similar levels of human capital to perform the same jobs."

Vahey (2000) uses the National Survey of Class Structure and Labor Process in Canada for the year of 1982. Vahey used self-assessment method since workers were asked about their actual education attainment as well as the education requirements for their jobs. The author considers an earnings equation in which required-education, over-education, and under-education are dummy variables. Vahey does not find that over-educated workers receive lower earnings than adequately educated ones. He also finds that returns to over/under-education are sensitive to gender and educational requirements of the employee. For females, there are no evidence for significant returns to either over or under-education. For males, it depends on the educational requirement of the job. He found that there exists positive returns to over-education, only when jobs that require a bachelor degree.

Voon and Miller (2005) use the 1996 Census of Population and Housing Household Sample File in order to study the consequences of over/under-education on earnings The authors use the realized matching method to calculate the jobs' educational requirements. They find that over-education is higher among men compared to women while under-education is higher among women. With regard to the impact of mismatch on earnings, workers who are over-educated receive "a modest payoff" to the years of over-education. On the other hand, those who are under-educated earn less compared to those who are adequately educated.

As can be seen from the presented literature so far, most of the studies on the wage consequences of over/under-education on earnings have been examined using surveys from developed countries. To the best of our knowledge, there are only two journal papers that estimate the returns to over/under-education in developing countries.

Firstly, Quinn and Rubb (2006) explore the effect of education-occupation mismatch on wages and productivity in the Mexican Labor Market. The authors use the data of the Mexican Migration Project from 1987 to 1997. The dataset includes 100 occupation

categories with a sample of 4945 men. For estimating the returns on wages, required education is measured by calculating the mode of the years of schooling for all individuals in a specific job category. The authors find that the results obtained are consistent with those of developed countries. they suggest that policy makers should not only focus on seeking to increase educational attainment level in the society but also increase the occupational level in order to better utilize human capital stock in the society.

Secondly, Abbas (2008) uses the Pakistan Integrated Household Survey and the Pakistan Social and Living Standards Measurement Survey to estimate the incidence of education-occupation mismatch and the returns to over/under-education for the period of 1998 to 2004. The author employs the realized matches technique to estimate required education. Main findings agree with the results from developed countries as over-education is found to have positive returns wherease under-education have negative returns. Also the returns to over-education are lower compared to returns to required-education.

On Egypt, there is only a minor study conducted by El-Hamidi (2008) that looked at the effects of education-occupation mismatches on wages over two points of time 1998 and 2006 for the private sector only. The author used the realized match technique to calculate over-education, under-education, and required education. Using the ORU specification proposed by Duncan and Hoffman (1981), the author estimates the effect of over/under-schooling on wages for three separate occupational categories which are: the professional category, blue-collar, and white-collar categories. The coefficients of the three educational components are found to be insignificant separately. The author's based the results relying on the joint significance of the three components. El-Hamidi (2008) finds that the over-education in Egypt is rewarded higher compared to adequate education, which is not consistent with the findings of the results from the developed world and even developing countries. On another note, El-Hamidi (2008) considers education to be exogenous as she conducts her analysis using the conventional OLS method. This raises specific concerns of potential biases of the results. Therefore, the use of instrumental variables can help verify the validity of these results.

Based on the findings from the previous literature, we can formulate the following hypotheses: (1) the returns to required education and over-education are positive while the returns to under-education are negative. (2) Although the return to over-education is positive, it is smaller in magnitude compared to returns on required-education. (3) Even though the return to under-education is negative, it is smaller in absolute terms compared to the return to required education.

6. Methodology

In order to investigate empirically the effects of over/under-education on wages, we shall employ two different methods which are considered modified specifications of the semi-logarithmic Mincer equation (1974). The first one is the ORU specification[13] introduced by Duncan and Huffman (1981) and the second one is the dummy variable specification introduced by Verdugo and Verdugo (1989).

According to Mincer's framework, the log of wages is composed of years of completed education (E), years of work experience (EXP), and experience squared (EXP^2)[14]:

$$\ln(w_i) = \beta_0 + \beta_1 E_i + \beta_2 EXP_i + \beta_3 EXP_i^2 + u_i \quad (1)$$

Where *u* is an error term that includes all the unobserved characteristics of the individual and is assumed to be uncorrelated with the other explanatory variables.

$$E(U \mid E, EXP) = 0 \quad (2)$$

As mentioned before, this model is derived from the human capital theory which implies that the returns to education are solely determined by the supply side characteristics (Balcar, 2012). Individuals invest in human capital which consists of education and experience to maximize their lifetime earnings.

In order to estimate the effects of over/ under-education on wages, the education variable in equation (1) is decomposed into required education for a certain job (E_r), over-education (E_o), and under-education (E_u) as can be seen in equation (3).

$$E = E_r + E_O - E_u \quad (3)$$

[13] ORU is an abbreviation for Over-education, Required Education, and Under-education
[14] Diminishing returns to experience are accounted for using a quadratic term (EXP2)

Where:

$$E_o = \begin{cases} E - E_r, & E > E_r \\ 0, & otherwise \end{cases}$$

$$E_u = \begin{cases} E_r - E, & E_r > E \\ 0, & otherwise \end{cases}$$

Therefore, we can re-write the equation as:

$$\ln(w_i) = \beta_0 + \beta_1 E_i^r + \beta_2 E_i^o + \beta_3 E_i^u + \beta_4 EXP_i + \beta_5 EXP_i^2 + \beta_6 X_i + u_i \quad (4)$$

Where $\ln(w_i)$ is the log of monthly wages for individual i and x is a vector of control variables that shall be discussed in the next sub-section.

It is important to note that over/under-education variables are measured in years. According to McGuinness (2006), "the overeducated are being compared to people doing the same job who are not overeducated and who have less education." This means that the coefficient $\boldsymbol{\beta_2}$ represents the returns to over-education controlling for under-education and required education along with other controls. It follows that $\boldsymbol{\beta_3}$ represents the returns to under-education controlling for over-education and required education along with other controls.

The Human Capital Theory shall hold if the returns to over-education, under-education, and required-education are not significantly different from zero. However, if there are positive returns to over-education and negative returns to under-education, one would expect the coefficient $\boldsymbol{\beta_2}$ to be positive and $\boldsymbol{\beta_3}$ to be negative. Such a result would support an assignment model. Another plausible scenario is when the coefficient on required education is significant but over-education and under-education coefficients are found to be insignificant (i.e. $\boldsymbol{\beta_2 = \beta_3 = 0}$). This then would support the Job Competition Model in which the requirements of the job are the sole determinants of wages.

The second method used by Verdugo and Verdugo[15] (1989), defines over-education and under-education as dummy variables. Consider the following modification in the Mincerian Equation:

$$\ln(w_i) = \propto_0 + \propto_1 OE_i + \propto_2 UE_i + \propto_3 EXP_i + \propto_4 EXP_i^2 + \propto_5 X_i + u_i \quad (5)$$

Where OE is a dummy variable taking the value of 1 if a worker is over-educated and zero otherwise. Similarly, UE is a dummy variable that takes the value 1 if the worker is under-educated and zero otherwise. The coefficients ($\propto_1$ & $\propto_2$) can be interpreted as the average wage effects of over-education and under-education in comparison to the adequately educated workers, respectively. Verdugo & Verdugo (1989) and Cohn and Kahn (1995) found that the coefficient of $\propto_1$ is negative while $\propto_2$ is positive. However, they differed in the interpretation. Verdugo & Verdugo (1989) argued that the negative sign on the over-education dummy implies that over-educated workers receive negative returns. On the other hand, using the same data, Cohn and Kahn (1995) showed that Verdugo & Verdugo (1989) have misinterpreted the results. Cohn and Kahn (1995) argue that the negative coefficient means that over-educated workers earn less than adequately educated workers and that the return to over-education is not negative.

In these two previous models, wages are determined by demand-side characteristics; the education required to perform the job, and the deviations from the required years of schooling (i.e. years of over-education and under-education).

The conventional way to estimate equations (4) and (5) is to run an OLS regression. In this case, the variables representing over-education, under-education, and required education are considered exogenous. Nevertheless, using OLS raises a couple of concerns when trying to assess the causal effect of education on wages. There are three main biases that could result from using OLS to estimate returns to education (Card, 1999). Firstly, it is very likely that there exists an omitted variable bias where education and earnings are correlated due to some other factor, in the error term that is not accounted for in the regression (i.e. ability bias). If ability is positively correlated with

[15] V & V hereafter

education and wages, then omitting this variable would cause an upward bias. Therefore, it is important to control for ability when measuring the effect of education on wages.

Secondly, the simultaneous causality bias which occurs when there is a two-way relationship between education and wages (Card, 1999). Meaning that, the decisions of investment in education is affected by the wages received and, on the other hand, years of education have an effect on wages received.

Thirdly, there is also a possibility of existence of bias due to a measurement error in years of schooling. Previous evidence suggests that OLS coefficients are usually biased downwards due to measurement errors (Harmon and Walker, 1994 as cited in Groot and Brink, 1997).

Consider the following equation where Ability (ABIL) is added as an explanatory variable:

$$\ln W = \beta_0 + \beta_1 E + \beta_2 ABIL + \beta_3 EXP + \beta_4 EXP^2 + u \quad (7)$$

The problem that usually arises in this case is the fact that ability is unobserved so we are unable to include it in the regression. It includes skills, motivation, energy, and other inherited characteristics. Under omitted variable scenarios, the OLS estimates are considered to be inconsistent. Nevertheless, many empirical studies have explored this issue and tried to assess the size of ability bias and obtain consistent estimates using different econometric techniques.

First of all, the use of proxies such as IQ and/or cognitive skills test scores has been suggested to account for unobservable ability. A good proxy need to be correlated with ability (Wooldridge, 2009). Unfortunately, a good proxy is not always available.

The second solution is to use fixed effects or first differencing methods in which we "estimate the time-varying independent variables in the presence of time-constant omitted variables" (Wooldridge, 2009). For example, Ashenfelter and Kureger (1994) use data of 149 identical twins in order to estimate the effects of education on earnings using the differencing methodology. The intuition behind using identical twins is the fact that twins are assumed to have similar abilities. Using this method is only possible when working with panel data over time. in addition, the independent variable in which we are

interested in studying should vary over time as these methods eliminate time-constant variables.

The third solution, which is the one adopted in this paper, is the instrumental variable approach.[16] An instrument (**z)** is valid when it influences the dependent variable only through the endogenous independent variable. To describe the instrumental variable approach, consider the following multiple regression equation:

$$\boldsymbol{y_1 = \beta_0 + \beta_1 y_2 + \beta_1 x + u} \qquad \textbf{(8)}$$

Where $\boldsymbol{y_1}$ and $\boldsymbol{y_2}$ are endogenous and x is a set of exogenous variables.

Since the Cov (y_2 ,u) ≠ 0, estimating equation (8) using OLS will produce biased and inconsistent results.

In order for an instrument to be valid, it needs to satisfy two requirements (Wooldridge, 2009). Firstly, the IV should not be correlated with the error term (u) which means that our instrument is exogenous.

$$\textbf{cov } (\boldsymbol{y_2}, \textbf{z}) = \textbf{0} \qquad \textbf{(9)}$$

Secondly, the IV needs to be correlated with the endogenous variable which implies that our instrument is relevant.

$$\textbf{Cov } (\boldsymbol{y_2}, \textbf{z}) \neq \textbf{0} \qquad \textbf{(10)}$$

It is only possible to test if the instrument satisfies the second assumption. In order to do so, we run an OLS regression with the endogenous variable being the dependent variable and the IV being the independent variable along with other controls used in the original OLS equation.

$$\boldsymbol{y_2 = \pi_0 + \pi_1 x + \pi_2 z + v} \quad \textbf{(11)}$$

If we are able to reject the null hypothesis ($H_0 : \pi_2 = 0$), we can conclude that our instrument satisfies the second assumption. This regression is usually referred to as the first stage regression.

[16] IV thereafter

We use the two-stage least square (2SLS) method in order to obtain our IV estimates. The first step involves running a regression using equation (11) and predicting the values of that regression $\hat{y}$. The second stage is to run an OLS regression using the predicted value $\hat{y}$ along with the other controls (Wooldridge, 2009).

It is important to make sure that the explanatory variable is endogenous. This is because if the explanatory variable is found to be exogenous, 2SLS estimators would be less efficient than the OLS ones (Wooldridge, 2009). Therefore, it would be better to estimate our equation using OLS instead. Hausman (1978) introduced a test that compares the OLS estimates to the 2SLS estimates. Under the null hypothesis OLS is efficient and consistent. Therefore, if the difference is statistically significant, this indicates that our explanatory variable is endogenous and therefore the use of 2SLS is favorable because it is consistent.

6.1 Control Variables

In addition to experience and experience squared, several control variables shall be included in our two specifications (equations 4 and 5) in order to account for the variations that can affect earnings. Firstly, we include a dummy variable that represents marital status of the worker which takes the value of 1 if the worker is married or zero if the worker is not married or divorced. The marital status is expected to have positive effects on wages. We also include a dummy variable that represents region of residence to account for differences in return to education between rural and urban areas. Furthermore, a dummy variable which indicates whether a person works in the private sector or not is included to account for differences in returns to education between public and private sectors. Three dummy variables are also included to represent the size of the firm of the individual's current occupation. Lastly, three dummy variables are included to represent the occupational categories[17]. The first category labeled "professional workers category" which includes managers and professionals. The second category labeled "white collar workers" includes technicians and associate professionals, and clerical support workers. Lastly, the third category labeled "blue-collar workers" includes craft

[17] Following El-Hamidi (2008)

& related trade workers, elementary occupations, plant and machine operators & assemblers, and service & sales workers.

6.2 Instrumental Variables

There are several variables that have been used in previous literature as IVs for education. Based on data availability, we shall use father's education as an instrument. Father's education can be considered as an instrument since this variable is highly correlated with the educational outcome of the individual and only affects wages through affecting the individual's education (Griliches, 1979 as cited in Card, 1999).

Generally, empirical studies adopting IV approach find that IV estimates are greater than OLS estimates which means that OLS estimates are biased downwards (Trostel, Walker & Woolley 2002). For example, Card (1999, p. 1842), using parental education as an IV results in estimates that are at least 15% higher than OLS estimates. Trostel, Walker, & Woolley (2002) find that returns to education are over 20% higher when using parental education as an IV. One explanation for this can be that parental education may act as a substitute for education. As mentioned by Lindgren (2005), "poorly endowed families would be more dependent on schooling resources for their achievement."

Each endogenous variable requires at least one instrument. Since we divide education into three different components (required education, over- and under-education), we shall also divide our instrument, father's education, into three components father's required-education, over-education, under-education to instrument for each endogenous component. We used the realized matches approach to estimate father's required education, over-education, and under-education.

Of course there exist limitations to the IV method when estimating returns to education. Some studies argue that parental education may not be exogenous as it might be correlated with earnings. In addition, if returns to education are heterogeneous then the coefficients obtained may not produce consistent estimators. Wooldridge (2009) states that that the use of father's education as an instrument is questionable since it might be correlated with the worker's ability.

6.3 Empirical Model Limitations

Sample selection bias is one of the main problems encountered when estimating returns to education. This problem arises because wages are observed only for those individuals who are employed (Card, 1999). This means that returns to education estimated do not reflect the population wage function as the sample is not representative of the entire population. Estimating returns to education using OLS and adding the dummy variable representing labor force participation as a regressor, along with other controls will still produce biased and inconsistent results. This is because Participation in the labor force is considered endogenous as it is a personal choice which depends on many factors (Card, 1999).

Heckman (1979) proposed a two-step sample correction method in which sample selection bias is treated as an omitted variable. The first step is to estimate a probit model with the dependent variable being a dummy representing whether the individual works or not, and calculate the inverse Mills ratio[18]. The second step is to estimate the output equation (i.e the OLS equation)[19] with wages as the dependent variable and include the inverse mills ratio as a regressor.

It is important to note that our analysis only considers full-time workers in the formal sector in the economy and excludes those who are out of the labor force and those who are employed in the agricultural sector, which raises statistical problems related to sample selection bias. Dealing with this issue empirically is out of the scope of this paper due to data constraints and is encouraged for future research.

7. Data

The data used in this analysis is obtained from the 2012 Egypt Labor Market Panel Survey (ELMPS), which is conducted by the Economic Research Four (ERF) in cooperation with the Central Agency for Public Mobilization and Statistics (CAPMAS). The survey offers an opportunity to examine the labor market conditions in the aftermath of the 2011 revolution.

[18] The inverse Mills ratio represents the omitted variable
[19] OLS can be replaced by 2SLS in case the wage equation includes endogenous controls

The 2012 ELMPS is a nationally representative survey which includes a total sample of 12,060 households and 49,186 individuals (Assad and Krafft, 2013).[20] Wages are reported for 10,275 individuals in the sample. The analysis is conducted on a sub sample of wage earners. We Exclude from the analysis individuals who are out of the labor force, above 65 years old (retired) and under 15 years old (children). We also exclude those employed in the agricultural sector, those who are self-employed, part-times workers, and those whose wages are below the minimum wage.[21] The reason behind excluding these categories is that high percentage of them works in the informal economy. Our analysis targets those who are employed in the formal public and private sector. The remaining number of observations is therefore 8442.

Unfortunately the data set does not include proxies for ability and school quality which are theoretically relevant to include in the earnings model. The definitions of the variables are included in table (1A) in the appendix.

As mentioned before, due to data availability, the method we employ to conduct our analysis is the realized match method. The mean for each occupation is calculated followed by the standard deviations for all workers. When educational attainment for the employee is more than one standard deviation above the mean, this means that the employee is over-educated. On the other hand, if the education of the employee is less than one standard deviation below the mean then the individual is considered under-educated. An individual is adequately educated if his/her years of education are within one standard deviation above or below the mean.

7.1 Descriptive Statistics

This section presents a descriptive analysis of the 2012 ELMPS. Firstly, we present information about the composition of the household survey based on gender, area of residence, marriage, employment sector, etc. Secondly, we analyze the incidence of mismatch among the selected sample of wage earners. Lastly, since we are using father's adequate-education, over- and under-education as instruments, it would be interesting to look at the incidence of mismatch among fathers in the sample.

[20] Details on the survey can be found in Assad and Krafft (2013).
[21] The minimum wage in the year 2012 was 700 Egyptian Pounds (Daily News Egypt, 2011).

To start with, our sample is made up of 83.61% males and only 16.39% females. As mentioned before, this reflects the low presence of women in the labor market compared to men. Workers living in urban areas represent around 55.56% of our sample while those residing in rural areas represent around 44.44%. Workers who are married represent 78.81% of the sample while those who are not married make up only 21.19%.

Our sample includes more people who work in the private sector compared to the public sector. Around 65% are employed in the private sector while only 35% are employed in the public sector. More than half of the employees, 58.8%, work in small companies, while only around 15.04% work in medium enterprises and 26.14% work in big enterprises.

The mean monthly wage is 1479 Egyptian Pounds with the minimum wage being 700 Egyptian Pounds and the maximum wage is around 35,000 Egyptian pounds. The mean experience is around 17 years with the minimum being zero and the maximum being 49 years.

Regarding the incidence of mismatch, the table below presents the incidence of mismatch by gender for the year 2012. It indicates that males experience higher incidence of mismatch compared to females. This result is consistent with the results of Voon and Miller (2005) for Australia, Groot (1997) for UK, and McGoldrick and Robst (1996) for the U.S. The share of under-educated individuals seems to be higher compared to over-educated individuals among both females and males. Around 74% of the workers are adequately educated.

Table (1): The incidence of mismatch by gender

	Male	Female	Total
	%	%	%
Over-Educated	**11.40**	**8.00**	**10.80**
Under-Educated	**15.40**	**14.30**	**15.20**
Adequately-Educated	**73.20**	**77.60**	**74.00**

Source: author's calculations using ELMPS 2012

It would be interesting to look at the incidence of mismatch by years of experience in order to see whether the incidence of mismatch is associated with the years of experience. By looking at the estimates in the table below, it appears that as years of experience increase, the incidence of over-education decreases. This finding supports the

theoretical argument that over-education is a temporary situation that is reduced by the number of years of work experience and/or job training. On the other hand, it is clear that the percentage of under-educated workers increases with experience. This indicates that a year of experience is used by under-educated workers to compensate for the lack of education.

Table (2): The incidence of mismatch by years of experience

Years of Experience	Over-Educated	Under-Educated	Adequately Educated
1-5	20.00	7.40	72.60
6-10	15.20	6.10	78.70
11-20	9.90	13.60	76.50
20-30	6.30	18.10	75.60
30+	4.30	31.50	64.20

Source: Author's calculations using ELMPS 2012

It is also important to look at the incidence of mismatch by occupational category as reported in the table below. The professional category includes the lowest percentage of over-educated workers with only 1.7% being overeducated, followed by 13.8% for the Blue Collar Workers, & 21.3% for the White Collars Workers. The low incidence of over-education found among workers in the professional category can be explained by the fact that workers in these occupations are expected to have obtained advanced degrees. Mismatch is expected to be higher among the white collar and the blue collar category where the level of education required is not as high when compared to professional workers.

The blue collars category includes the highest percentage of under-educated with around 19.6%. This can be expected as workers in this category don't require a high level of education compared to other categories but rather higher level of experience.

Table (3): The Incidence of Mismatch by occupational group

	Professionals	White Collars	Blue Collars
	%	%	%
Over-educated	1.7	21.3	13.8
Under-Educated	12.2	7.7	19.6
Adequately-Educated	86.1	71	66.5
Total	100	100	100

Source: Author's calculations using ELMPS 2012

Lastly, the incidence of the father's educational mismatch seems to differ from worker's mismatch in the sample. Looking at table (4), it appears that around 61.93% of the fathers are under-educated while only 15.95% are adequately educated and 22.12% are over-educated.

Table (4): The incidence of mismatch (father's education)

	Percentage
Over-educated	22.12
Under-educated	61.93
Adequately-educated	15.95

Source: Author's Calculations using ELMPS 2012

8. Results

Table (4) reports the results using the ORU and V& V specifications for workers in all occupations. The results below confirm the hypothesis that years of over-education, under-education, and adequate-education are not rewarded equally in the labor market.

When using the ORU specification, we find that holding everything else constant, the rate of return to over-education is positive and constitute around 7.08% while the rate of return to under-education is negative and constitute around 1.69%.

On the other hand, using the V & V definition, we find that holding everything else constant, over-educated workers earn 14.9% more than adequately educated workers. On the other hand, under-educated workers earn 6.12% less compared to adequately educated workers.

Looking at the control variables, experience is found to increase wages by around 0.8% while there seems to be no diminishing returns to experience as the coefficient on the experience squared variable is found to be insignificant. It also appears that, on average, married workers receive 3.7% higher monthly wages in comparison to unmarried workers. As expected, females tend to earn lower wages in comparison to men. A female worker is expected to earn around 15% lower wages compared to a male worker. Those working in the private sector receive 11.5% higher wages compared to those working in the public sector, holding all other controls constant. Working in a large

firm would lead to about 14.2% increase in wages in comparison to working in medium sized firms. [22]

Table (5): OLS regression

	ORU		**V&V**	
VARIABLES	ln_mnthwg	std. error	ln_mnthwg	std. error
Er	0.0147***	(0.00317)		
Eo	0.0708***	(0.00980)		
Eu	-0.0169**	(0.00740)		
OE			0.149***	(0.0207)
UE			-0.0612***	(0.0235)
exp	0.00772***	(0.00210)	0.00802***	(0.00211)
exp2	-2.00e-05	(4.82e-05	-4.09e-05	(4.82e-05)
married	0.0363**	(0.0167)	0.0372**	(0.0168)
urban	-0.0691***	(0.0128)	-0.0726***	(0.0128)
occup_p	0.157***	(0.0205)	0.208***	(0.0186)
occup_b	-0.0757***	(0.0221)	-0.109***	(0.0199)
small	0.0365*	(0.0193)	0.0332*	(0.0193)
big	0.139***	(0.0172)	0.142***	(0.0172)
private	0.118***	(0.0176)	0.114***	(0.0176)
female	-0.149***	(0.0183)	-0.150***	(0.0184)
Constant	6.796***	(0.0543)	6.977***	(0.0355)
Observations	5,891		5,891	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As mentioned before, since education is endogenous, we estimate the previous two models using the (2SLS) IV approach. For the ORU specification, we use father's required-education, over- and under-education to instrument for Er, Eo, and Eu, respectively. On the other hand, for the V & V specification we use father's over-education and under-education to instrument for OE and UE, respectively. It is important to note that, the results from the first stage regressions[23] show that our instruments are jointly significant which means that they are exogenous.

Table (5) reports the IV results, which are obtained using 2SLS. The results show that the OLS estimates are biased downwards. From the ORU specification, it is found

[22] The variable "Medium size firms" is the base group in this case; it is not included in the regression to avoid multicollinearity.

[23] Tables 2A and 3A in the appendix report the results of the first stage regression for the ORU and the V & V specification, respectively.

that over-education contribute to around 61.4% increase in the rate of return, keeping everything else constant. Required education leads only to 2.8% increase in the rate of return and the under-education coefficient is found to be insignificant.

Similarly, using V & V approach, over-education appears to lead to around 66.8% rise in wages compared to adequate education. Under-education on the other hand becomes insignificant. Working in the professional category result in increasing wages by 30% compared to the white collars category while working in the blue collar category results in decrease in wages by 7.14%.

Next, the Hausman Test is applied in order to check for the endogeneity of required-education, over-education and under-education variable. As can be seen in table (5), we reject the null that the OLS coefficient is consistent and efficient. This shows that the IV results are consistent and should be considered rather than the OLS.

Table (6): IV estimates

	ORU		**V & V**	
VARIABLES	ln_mnthwg	std. err.	ln_mnthwg	std. err.
Er	0.028***	(0.00186)		
Eo	0.614***	(0.232)		
Eu	-0.312	(0.338)		
OE			0.668**	(0.294)
UE			-0.210	(0.340)
Exp	0.0208***	(0.00635)	0.0142***	(0.00380)
exp2	-0.000274*	(0.000163)	-9.38e-05	(8.99e-05)
Married	0.00183	(0.0302)	0.0235	(0.0218)
Urban	-0.0689	(0.0517)	-0.0265	(0.0197)
occup_p	0.228	(0.815)	0.300***	(0.0654)
occup_b	0.132	(0.542)	-0.0714***	(0.0235)
Small	0.0915**	(0.0372)	0.0700***	(0.0230)
Big	0.179***	(0.0368)	0.148***	(0.0187)
Private	0.112**	(0.0531)	0.120***	(0.0269)
Female	-0.128***	(0.0222)	-0.130***	(0.0205)
Constant	5.635***	(1.912)	6.725***	(0.108)
Observations	5,636		5,636	
Hausman Test				
Chi2	80.38		27.07	
p-value	0.0000		0.0025	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

8.1 Discussion

The analysis is based on the IV results since its estimates were found to be consistent. Our results contradict the previous empirical studies conducted on developed countries as well as the few studies conducted on developing countries.

The returns to required education and over-education are positive which agrees with the previous findings in the literature. However, the returns to over-education are found to be much larger in magnitude compared to returns on required-education. On the other hand, looking at the OLS estimation, under-education is found to have negative effects on earnings but once IV is applied this effect becomes insignificant.

There are two possible explanations for the positive effects of over-education on wages. The first one is connected to the Spence's (1973) screening hypothesis which suggests that education is used as a signal in the labor market. Due to asymmetric information employers use education as a signal of productivity. Thus, employers view over-educated workers as more productive workers and therefore reward them with higher wages. According to Kampelmann & Rycx (2012), differences in wages mirror differences in productivity. Secondly, as suggested by the efficiency wage theory, employers may decide to pay over-educated workers higher wages in order to reduce the possibility of shirking as over-educated workers are relatively scare compared to the other categories.

It is important to note, however, that the coefficient on over-education is larger than 60% (61.4% using the ORU specification and 66.8% using the V & V) which is considered unreasonable and should not be interpreted at face value since it implies that one extra year of over-education would lead to more than 60% returns. Obtaining such a high coefficient can be due to the fact that we use father's required, over- and under-education as instruments. As mentioned previously, using parental education as an instrument is likely to produce IV estimates that are at least 15% higher (Card, 1999) or even over 20% higher (Trostel, Walker, and Woolley, 2002) compared to the OLS estimates. Also, as mentioned before, the use of father's education as an instrument has been questioned due to the fact that it maybe not be exogenous (Wooldridge, 2009). If father's education has an additional effect on the worker's wage other than through the

worker's education (i.e. through personal networks), the size of the coefficient on the education variables could be really big and bias the IV results. This means that the IV estimates would no longer show the effect of over-education on wages, but rather the combined effect of over-education and father's education. Therefore, for future research, when more data are available, it is encouraged to consider alternative instruments to avoid the problems associated with the use of parental education. Examples of instruments that has been used in previous research include, but are not limited to, geographic proximity to schools, tuition costs, and minimum school-leaving age (Card, 2001).

On another note, finding that coefficients of over-education and required education are significant in the ORU specification and the coefficient of over-education is significant in the V & V specification implies that the returns to required, over- and under-education are not equal. Hence, similar to findings in previous literature, the Egyptian Labor Market follows an assignment model in which wages are allocated based on a combination of workers' characteristics and job characteristics.

9. Conclusions

The paper empirically investigated one of the main consequences of the over/under-education phenomenon in the Egyptian labor market, which is its effect on wages. Using the 2012 Egypt Labor Market Panel survey, we estimated the returns to over-education, under-education, and required education. Due to data availability, the paper considers the vertical mismatch definition which occurs when the required level of education is different from actual level of education. The realized matches method was used to estimate the required years of education.

The paper also contributed to the literature by addressing the issue of the endogeneity of the different components of education, required-education, over-education, and under-education, by estimating the returns using instrumental variable approach. The instruments used were father's required-education, father's over-education, and father's under-education. Two empirical models were utilized to conduct

the empirical estimation: the ORU specification proposed by Duncan and Hoffman (1981) and the dummy variables specification proposed by Verdugo & Verdugo (1989).

The results obtained in this paper should be considered with caution for two main reasons. Firstly, the existence of a sample selection bias as the study focuses only on a sub-sample of wages earners as we exclude those who are out of the labor force, agricultural sector's workers, self-employed workers, part-times workers, and those whose wages are below the minimum wage. It is encouraged for future research to utilize the Heckman two-step method to avoid sample selection bias. Secondly, the very high coefficient obtained on the over-education variable after implementing the IV approach which is most likely the result of using father's required, over-, and under-education as instruments. Therefore, it is also encouraged to utilize alternative instruments, once available, in future research such as proximity to school, tuition costs, etc.

By analyzing the incidence of the mismatch in the Egyptian economy, it is evident that the percentage of mismatched workers among males is higher than females. It is found that there is a trade-off between years of experience and the incidence of over-education. The incidence of under-educated workers seems to increase with experience which is an indication that years of work experience compensate for the lack of education. The lowest incidence of over-education is found among professionals while the highest incidence of under-education is found among blue-collars.

The results of the returns estimation suggest that the situation in the Egyptian economy is unique compared to other developed and developing countries experiences. Though previous literature find that over-education has positive return that is smaller in magnitude compared to the return of required-education; we find that over-education has larger positive effects on wages compared to adequate-educated workers and under-educated workers. OLS regression also shows that under-education has negative effects on wages. After applying the IV approach, it appears that the OLS regression under-estimates the returns to over-education due to the ability bias. In addition, the coefficient on under-education becomes insignificant. There are two possible explanations for the high returns of over-education. Firstly, consistent with the signaling hypothesis, over-education is an indication to the employer that the worker is more productive compared

to the adequately educated worker and the under-educated worker and therefore they receive higher returns. Secondly, following the efficiency wage theory, employers may decide to pay over-educated workers higher wages in order to reduce the possibility of shirking as over-educated workers are relatively scare compared to the other categories.

Since the labor market provides high rewards to extra years of over-education, this may cause an increase in wage inequality in Egypt. The policy implication, therefore, would be to increase average wages in order to avoid wage dispersion. This can be done by implementing policies that lead to an increase in educational attainment and occupational levels at the same time.

It was also found that wages in the Egyptian economy do not depend only on job requirements (i.e. job competition model) or individual's characteristics (i.e. human capital hypothesis) but rather a combination of both. This is due to the fact that returns to required-, over-, and under-education are unequal. This means that an assignment model interpretation is relevant to the Egyptian case which is coherent with findings from previous studies.

References

Abbas, Q (2008). Over-education and under-education and their effects on Earnings: evidence from Pakistan,1998–2004. *SAARC Journal of Human Resource Development* 4, 109–125.

Assaad, R. & Roudi-Fahimi, F. (2007). *Youth in the Middle East and North Africa: Demographic Opportunity or a challenge?* Washington, DC: Population Reference Bureau.

Assaad, R., & Krafft, C. (2013). The Egypt labor market panel survey: introducing the 2012 round. *IZA Journal of Labor & Development, 2*(1), 8.

Assaad, R., & Krafft, C. (2013). The evolution of labor supply and unemployment in The Egyptian economy: 1988-2012. In *Economic Research Forum Working Paper Series* (No. 806).

Balcar, J. (2012). Supply Side Wage Determinants: Overview of Empirical Literature. Review of Economic Perspectives, 12(4), 207-222.

Becker, G. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* New York: Columbia University Press.

Berg, I. (1970). *Education and Jobs: The Great Training Robbery*. Boston, MA: Beacon Press.

Boeri, T. & Ours, J. V. (2008). *The economics of imperfect labor markets*. Princeton University Press. Princeton: New Jersey.

British Council (2013). *Education in Egypt.* Retrieved from http://www.britishcouncil.org/learning-skills-for-employability-egyptian-education.htm#the

Büchel, F. & Mertens, A. (2000). Overeducation, undereducation, and the theory of Career Mobility. *Applied Economics, 36*(8), 803-816.

Clark, N. (2013). *Education in Egypt.* World Education News and Reviews. Retrieved from http://wenr.wes.org/2013/11/education-in-egypt/

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, 3, 1801-1863.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica, 69*(5), 1127-1160.

Cohn, E., & Khan, S. (1995). The wage effects of over schooling revisited. *Labor Economics, 2*, 67-76.

Daily News Egypt. (2011, October 27). Egypt discussing private sector minimum wage. *Daily News Egypt.* Retrieved from http://www.dailynewsegypt.com/2011/10/27/egypt-discussing-private-sector-minimum-wage/

Dickens, W. T., & Lang, K. (1992). Labor market segmentation theory: reconsidering the evidence. *NBER Working Paper.* Retrieved from http://www.nber.org/papers/w4087

Dickens, W. T. & Lang, K. (2001). A test of dual labor market theory. *The American Economic Review,* 75(4).

Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. Economics of Education Review, 19, 179–198.

Duncan, G. J. and Hoffman, S. (1981). The incidence and wage effects of over-education, *Economics of Education Review, 1*, 75–86.

El-Hamidi, F. (2008, November). *Education-Occupation Mismatch and the effect on wages of Egyptian Workers.* Paper presented at 15th Annual Conference of the Economic Research Forum: Equity and Economic Development, Cairo.

Fields, G. S. (2009). Segmented Labor Market Models in Developing Countries. Cornell University, ILR School. Retrieved from http://digitalcommons.ilr.cornell.edu/articles/162/

Freeman, R. B. (1976). The overeducated American. New York: Academic Press.

Green, D. (2011). What caused the revolution in Egypt? *The Guardian*. Retrieved from http://www.theguardian.com/global-development/poverty-matters/2011/feb/17/what-caused-egyptian-revolution

Groot, W. & Brink, M. V. D. (1997). Allocation and the returns to over-education in the UK. *Education Economics, 5*(2). 169-183.

Hasan, A. E. (n.d.) *Education in Egypt.* Retrieved from http://wwwdb.in.tum.de/teaching/ws1112/hsufg/Egypt/Egypt/EducationSystem.html

Hassan, M., & Sassanpour, C. (2008). Labor market pressures in Egypt: Why is the unemployment rate stubbornly high?. *With Mohamed Hassan, Journal of Development and Economic Policies*, 10(2).

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, *46*(6), 1251-1271.

Harmon, C., & Walker, I. (1994). Estimates of the economic return to schooling for the United Kingdom. *The American Economic Review, 85*(5), 1278-1286.

Hartog, J. (2000). Over-education and earnings: where are we, where should we go?. *Economics of education review, 19*(2), 131-147.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153-161.

Katz, L. F. (1986). Efficiency wage theories: A partial evaluation. In *NBER Macroeconomics Annual 1986, Volume 1* (pp. 235-290). MIT Press.

Kiker, B. F., Santos, M. C., & De Oliveira, M. M. (1997). Overeducation and undereducation: evidence for Portugal. Economics of Education Review, 16(2), 111-125.

Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education*, *4*, 283-326.

Lindgren, M. (2005). Does schooling make sense? A household perspective on the returns to schooling for self-employed, farmers and employees in Egypt (Doctoral dissertation, Lund University).

Mason, G. (1996). Graduate utilization in British industry: the initial impact of mass higher education. *National Institute Economic Review, 156*, 93-103.

McGoldrick, K., & Robst, J. (1996). Gender differences in overeducation: a test of the theory of differential overqualification. *The American Economic Review*, 280-284.

McMillen, D. P., Seaman, P. T., & Singell Jr, L. D. (1999). A mismatch made in heaven: a hedonic analysis of overeducation and undereducation. *Southern Economic Journal*, 901-930.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys, 2*(3), 387-418.

Mehta, A., Felipe, J., Quising, P., & Camingue, S. (2011). Overeducation in developing economies: How can we test for it, and what does It mean? *Economics of Education Review* 30, 1334-1347.

Mincer, J. A. (1974). Schooling and earnings. In *Schooling, experience, and earnings* (pp. 41-63). Columbia University Press.

Piracha, M. & Vadean, F. (2012). Migrant educational mismatch and the labour market. *IZA. Discussion Paper No. 6414.*

Quinn, M. and Rubb, S. (2006). Mexico's labor market: The importance of education-occupation matching on wages and productivity in developing countries. *Economics of Education, 25,* 147-156.

Robst, J. (1995). Career mobility, job match, and overeducation. *Eastern Economic Journal, 21*(4), 539-550.

Robst (2007). Education and job match: the relatedness of college major and work. *Economics of Education Review,* 26: 397-407.

Rumberger, R. W. (1987). The Impact of Surplus Schooling on Productivity and Earnings. *Journal of Human resources, 22*(1).

Sala, G. (2011). Approaches to skills mismatch in the labour market: a literature review. *Politica y Sociedad, 48*(2).

Sicherman N., & Galor O., (1990). A Theory of Career Mobility. J*ournal of Political Economy,* 98(1), 169-192.

Sloane, P. J. (2003). Much ado about nothing? What does the overeducation literature really tell us. *Overeducation in Europe. Current issues in theory and policy,* 11-45.

Spence, M. (1973). Job market signaling. *The quarterly journal of Economics, 87*(3), 355-374.

Stiglitz, J. E. (1976). The efficiency wage hypothesis, surplus labor, and the distribution of income in LDCs. *Oxford Economic Papers, 28*(2), 185-207.

Shapiro, C. & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review,* 74(3), 433-444.

Thurow, L. C. (1975). *Generating Inequality*. New York, NY: Basic Books.

Trostel, P., Walker, I., & Woolley, P. (2002). Estimates of the economic return to schooling for 28 countries. *Labour economics*, *9*(1), 1-16.

Tsang, M. (1987). The impact of underutilization of education of productivity: a case study of the U.S. Bell companies. *Economics of Education Review, 11,* 239–234.

Tsang, M. & Levin, H. (1985). *Economics of Education Review, 4*(2), 93-104.

UNESCO (2012). Country Profiles: Egypt. Retrieved from http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=121&IF_Language=en&BR_Country=2200

Vahey (2000). The great Canadian training robbery: evidence on the returns to educational mismatch. *Economics of Education Review, 19*, 219-227.

Verdugo, R. & Verdugo, N. (1989). The impact of surplus schooling on earnings: some additional findings. *Journal of Human Resources*, pp. 629–643.

Wooldridge, J. M. (2009). Introductory econometrics: a modern approach.[Mason, Ohio], South Western cengage learning. *Authors' affiliations Olaf Giebel, Barbara Breitschopf Fraunhofer Institute for Systems and Innovation Research (Fraunhofer ISI) Competence Center Energy Policy and Energy Systems.*

World Bank (2014). *Labor Force Participation Rate.* Retrieved from http://data.worldbank.org/indicator/SL.TLF.CACT.ZS/countries/1W?display=graph

Appendix

Table 1A: Definitions of the Variables

Variable	Description
ln W	Natural log of monthly wages
Eo	Continuous that represents over-education
Eu	Continuous Variable that represents under-education
Er	Continuous variable that represents required education
OE	Dummy Variable takes the value of 1 if worker is over-educated and 0 otherwise
UE	Dummy Variable takes the value of 1 if worker is over-educated and 0 otherwise
Female	Dummy variable that represents gender, takes the value 1 if the person is a female and value 0 if a person is a male
Married	Dummy variable that represents marital status, takes the value 1 if married and 0 if not married (including widowed or divorced)
Exp	Work Experience
Exp2	Work Experience Squared
Private	Dummy variable that represents sector of employment, takes the value 1 if the person works in the private sector and zero if the person works in the public sector
Occup_p	Takes the value 1 if professionals, technicians and associate professionals; 0 otherwise
Occup_w	Takes the value 1 if Managers, technicians & associate professionals, clerical support workers, service and sales workers; 0 otherwise
Occup_b	Takes the value 1 if Craft and related trade workers, plant and machine operators and assemblers, and elementary occupations; 0 otherwise
Region	Dummy variable which takes value 1 if urban and 0 if rural
Small	takes the value 1 if the firm size ranges from 1-24 and zero otherwise
Medium	takes the value 1 if the firm size ranges from 25-99 and zero otherwise
Big	takes the value 1 if the firm size is greater than 100 and zero otherwise
f_req	Father's required education
f_over	Father's over education
f_under	Father's under education

Table 2A: First Stage using ORU specification

Variables	adeq_educ	over_educ	under_educ
exp	-0.0458**	-0.0165***	-0.0132**
	(0.0189)	(0.00262)	(0.00566)
exp2	-0.000913**	0.000200***	0.00109***
	(0.000420)	(5.81e-05)	(0.000125)
married	-0.512***	0.0699***	0.0256
	(0.158)	(0.0219)	(0.0473)
urban	0.0996	-0.114***	0.312***
	(0.119)	(0.0165)	(0.0357)
female	-0.246	-0.0731***	0.483***
	(0.165)	(0.0228)	(0.0492)
occup_p	5.048***	-0.293***	0.294***
	(0.181)	(0.0250)	(0.0540)
occup_b	-1.263***	0.0229	-0.485***
	(0.190)	(0.0263)	(0.0568)
small	-1.286***	-0.0982***	0.562***
	(0.196)	(0.0271)	(0.0586)
big	-0.682***	-0.0315	0.246***
	(0.185)	(0.0257)	(0.0554)
private	-2.083***	0.0241	0.854***
	(0.174)	(0.0240)	(0.0519)
fadeq_educ	0.0564**	-0.00640**	-0.00803
	(0.0231)	(0.00319)	(0.00689)
fover_educ	-0.0528	0.0248***	-0.0180
	(0.0534)	(0.00738)	(0.0159)
funder_educ	-0.363***	-0.0791***	0.286***
	(0.0641)	(0.00887)	(0.0192)
Constant	10.93***	0.729***	-1.273***
	(0.341)	(0.0471)	(0.102)
Observations	7,630	7,621	7,630
R-squared	0.356	0.078	0.201
F test model	18.18	39.46	97.19
P-value of F model	0	0	0

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3A: First stage using V & V specification

Variables	dover_educ	dunder_educ
exp	-0.00932***	0.000971
	(0.00105)	(0.00119)
exp2	0.000120***	0.000158***
	(2.34e-05)	(2.65e-05)
married	0.0265***	0.00546
	(0.00881)	(0.00998)
urban	-0.0494***	0.0525***
	(0.00663)	(0.00752)
occup_p	-0.185***	0.0271**
	(0.0100)	(0.0114)
occup_b	-0.0652***	-0.0367***
	(0.0106)	(0.0120)
small	-0.0385***	0.109***
	(0.0109)	(0.0124)
big	-0.0112	0.0387***
	(0.0103)	(0.0117)
private	0.0106	0.154***
	(0.00966)	(0.0109)
female	-0.0351***	0.0975***
	(0.00916)	(0.0104)
fover_educ	0.0141***	-0.00103
	(0.00288)	(0.00326)
funder_educ	-0.0302***	0.0604***
	(0.00333)	(0.00378)
Constant	0.396***	-0.264***
	(0.0186)	(0.0210)
Observations	7,621	7,630
R-squared	0.108	0.189
F test model	66.39	135.5
P-value of F model	0	0

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1