Education - Occupation mismatching and labor market segmentation in developing countries: The Colombian case
(Preliminary version)

Paula Herrera-Idárraga[a,b,*] - paulaherrera@ub.edu
Enrique López-Bazo[a,c] - elopez@ub.edu
Elisabet Motellón[a,d] - emotellon@ub.edu

[a] AQR – IREA, University of Barcelona, Avda Diagonal, 690, 08034 Barcelona, Spain, tel.: +34 93 4021824, fax +34 93 4021821
[b] Department of Economics, Pontificia Universidad Javeriana, Carrera 7 No. 40 – 62, Bogotá, Colombia.
[c] European Commission, Joint Research Center (*JRC*), Institute for Prospective Technological Studies (*IPTS*). Calle Inca Garcilaso 3, 41092 Sevilla, Spain.
[d] Universitat Oberta de Catalunya, Avda Tibidabo 39-43, 08035 Barcelona, Spain.

Abstract - In this paper, we explore the connection between labor market segmentation in two urban sectors, a modern protected formal sector and a traditional-unprotected-informal sector, and over-education in a developing country. Informality is thought to have negative implications, primarily through poorer working conditions, lack of social security, as well as low levels of productivity throughout the economy. However, in this paper we consider an aspect which has not been addressed previously, that informality could also affect the way workers match their acquired education with the one required to perform their job. We first adapt a model, which illustrates that a segmented labor market can lead to over-education. Then we estimate two types of empirical models in order to test the relationship between over-education and informality: a simple univariate probit model and a probit model with selection. Empirical results from Colombia suggest that labor market segmentation, rather than having a direct effect, has a subtler impact on over-education. Some unobservable factors that affect the probability of working in the formal sector affect also the probability of being over-educated, whereas this mechanism does not exist in the case of the informal sector.

[*] Corresponding author: E-mail address: paulaherrera@ub.edu

1. Introduction: A brief literature review and motivation of the study

In this paper, we explore the connection between labor market segmentation in two urban sectors, a modern protected formal sector and a traditional-unprotected-informal sector, and over-education. Informality is thought to have negative implications, primarily through poorer working conditions, lack of social security coverage in health and/or old age, as well as low levels of productivity throughout the economy. However, this paper considers an aspect which has not been addressed previously, which is that informality could also affect the way workers match their acquired education with the one required to perform their job. An individual worker is said to be overeducated if she has acquired more education than what it is required to perform her job. Over-education is often taken to imply that resources are wasted, because over-educated workers earn lower returns on their investment relative to similarly educated individuals whose jobs match their education. Our assumption is that over-education is not independent of market segmentation into a formal and an informal sector.

There is now a substantial body of literature that addresses the phenomenon of over-education for developed countries[1]. An increase amount of this literature is devoted to provide an explanation for over-education that is consistent with one of the theoretical labor market frameworks: Human Capital Theory (Becker, 1964), the Job Competition Model (Thurow, 1975) or the Assignment Models (Tinbergen, 1956). The majority of studies tend to support the Assignment Interpretation; within this framework the earnings will depend to some extent on both the individual and the job characteristics. These models also imply that there is no reason to expect that wage rates will be only correlated to acquired schooling or other individual attributes (Human Capital Theory), neither should be expected that individual's productivity and hence their earnings will be determined only by the job characteristics (Job Competition Model). In addition, a number of studies have also estimated the effects of over-education on earnings. These studies show that overeducated workers tend to earn higher returns to their years of schooling than co-workers who are not over-educated, but lower returns than workers with similar education who work in jobs that require the level of education that they possess. It is worth noting that the choice of the method to measure over-education has a large effect on the incidence found but not on the pay penalty associated with over-education (Groot and Maasen van den Brink 2000).

To the best of our knowledge, few studies had investigated the over-education phenomenon for developing countries. Quinn and Rubb (2006) study the phenomenon for Mexico, Abbas (2008) for Pakistan and Mehta et al. (2011) for India, Mexico, the Philippines and Thailand. One reason for the lack of studies in developing countries could be due to data limitation in order to identify the education levels required for specific jobs. On the other hand, despite the increase in the past decades in the average schooling attainment in developing countries, the average presented in these economies is less than the one presented in high-income countries. In Latin American and the Caribbean Countries the average educational attainment for those females and males between 21 and 24 years old were 9.6 years and 9.3 respectively (Duryea et al. 2007). Meanwhile the average for the OECD countries is 12.5 for males and 12.8 for females between 25 and 34 years (OECD Education at a Glance, 2010). The fact that the educational attainment remains low in developing countries makes the over-education phenomenon somewhat contradictory for

---

[1] Duncan and Hoffman (1981), Verdugo and Verdugo (1989), Sicherman (1991), Tsang et al. (1991), McGoldrick and Robst (1996) studied the phenomenon for the United States; Alpin (1998), Green et al. (1999), Dalton et al. (2000) and Chevalier (2003) for UK; Hartog and Oosterbeek (1998) and Groot and Massen van den Brink (2000) for Holland; Bauer (2000) and Buchel et al. (2000) for Germany; Kiker et al. (1997) and Mendes de Oliveira et al. (2000) for Portugal; Alba-Ramirez (1992) for Spain. For an extensive review of over-education in developed countries see McGuinness (2006).

these economies. Nevertheless Quinn and Rubb (2006) and Mehta et al (2011) find evidence that over-education exists in developing countries and that the incidence of over education is similar to that presented in developed ones.

Given the differences between labor markets in developed and developing economies, it is plausible that the explanatory factors of being over-educated may differ. For instance, labors markets of developing economies are characterized by the existence of highly informality. In the particular case of Latin America, the informal sector employs between 30% and 70% of the urban work force (Maloney, 2004), embracing a variety of heterogeneous activities, such as self-employment entrepreneurs, salaried workers of large and small firms, and unpaid domestic workers. Beside the well-known negative implication of informality, primarily through poorer working conditions, a segmented labor market (in a formal and an informal sector) could also affect the way workers match their acquired education with the one required to perform their job. As Berry and Sabot (1978) affirmed, *one of the inefficiencies associated with segmentation, more difficult to document but possibly imposing greater resource costs on the economies of developing countries, involve the failure of the market to move the "right" resources into high wage sectors, a failure commonly described by the term "mismatch"*. Building on this statement, our assumption is that the study of over-education in a developing economy with a large informal sector must include as a key element the role played by this type of segmentation.

One developing country, which is characterized for the presence of high informality in its labor market, is Colombia. Colombian urban informal sector is an interesting case to study for several reasons. First, informality has been at the center of the economic debate due to the high levels prevalent in the country. According to Firm Size and Occupation criterion, informality was constant around 52% from 1984 to 1996, grew steadily between 1996 and 2001 to 56%, and remained at this level until 2006 (Mondragón-Vélez et al., 2010). Second, previous studies have found that in Colombia over-education exist (Mora, 2005 and Castillo, 2007). Finally the studies about informality in Colombia and other developing countries focus primarily on the size of the informal sector and on the effects of labor market rigidities on employment, wages and its distribution, and on the probability to become informal[2]. However little attention has been paid to the effects of a large informal sector in the way workers match their education with the one required to perform their job. If labor market segmentation may lead to education-occupation mismatches, then it can also affect the allocation of resources in the educational system, one manifestation of which is the bias toward academic training (Berry and Sabot, 1978).

The aim of this paper is, therefore, to extend the over-education literature for developing countries by exploring the connection between labor market segmentation and over-education. We first adapt a model, which illustrates that a segmented labor market in two urban sectors, a modern protected formal sector and a traditional-unprotected-informal sector, can lead to over-education. Then, we estimate two types of empirical models for testing this relationship: a simple univariate probit model and model with selection. We search if there is any difference in the incidence of over-education between the formal and informal sector, and then if this difference exist, we implement a decomposition of the gap in the incidence of over-education between the formal and the informal sector. The decomposition allows us assessing the contribution of differences in observable characteristics and in its effects to the over-education gap. We hypotheses that in the

---

[2] Magnac (1991), Nuñez (2002), Maloney and Nuñez (2004), Floréz (2002), Kugler and Kugler (2009) and Mondragón-Vélez et al. (2010) for Colombia; Pradhan and van Soest for Bolivia (1995); Pratap and Quintin for Argentina (2006); Tansel (1999) for Turkey and Gong and Van Soest for Mexico.

presence of a large informal sector in developing countries, it is possible that workers at the formal sector end up being more overeducated than informal workers in order to seek for better work conditions, after controlling for other characteristics. An alternative hypothesis is that a high skill worker who is unable to get a high skill job in the formal sector may accept a low skill wage in a low skill job in the informal sector for which he or she is over-educated. Nevertheless these considerations can only constitute one of the plausible explanations of why over-education may exist in developing countries. It is known that in developing countries information about jobs and access to employers depends on the personal contacts of the individual (Tenjo, 1990). Given the importance of these informal channels, through which job search takes place in developing countries, it is probable that education mismatching occurs for those individual who don't have access to these networks.

The rest of the paper is organized as follows. The next section sketches the theoretical framework for our empirical analysis. Details on the data and some selected descriptive are given in section 3, while the empirical approach is presented in section 4. Section 5 summarizes the results regarding the estimates of the empirical models and the decomposition exercise and, finally, section 6 concludes.

2. Theoretical framework: search frictions in the labor market, labor market segmentation and over-education

It is possible to adapt Charlot and Decreuse (2005) model to illustrate that a segmented labor market in two urban sectors, a modern protected formal sector and a traditional-unprotected-informal sector, can lead to over-education. The model is based on three main figures. First, workers can direct their search into the formal or informal sector, second the formal sector is intensive in high skilled workers and the informal sector is intensive in low skilled workers. Third, workers are heterogeneous; there are two schooling levels, high and low. Variables are indexed by k and j, where $k = h$ stands for high and $k = l$ for low and $j = f$ stands for formal sector and $j = i$ for informal sector. The output y_j of a match depends on the worker's characteristic a (ability) and on the sector-specific productivity parameter A_j (technology). The technology currently in use in the formal sector is more efficient than that used in informal sector $(A_f > A_i)$.

Vacant jobs and unemployed workers are brought together in pairs through an imperfect matching process. The total number of employer-worker contacts M_j on sector j is given by the following matching technology:

$$M_j \equiv M_j(u_j, v_j) \tag{1}$$

Where u_j is the number of unemployed and v_j denote the numbers of vacancy on sector j. The function $M_j(u_j, vj)$ is assumed to be twice differentiable, increasing in its arguments and to exhibit constant returns to scale. The labor sector tightness is determined by the ratio between the vacancy rate and the number of unemployed ($\theta_j = v_j/u_j$). Because of constant returns to scale, it is possible to write the rate at which vacancies are filled on sector j as $M_j/v_i \equiv m_j(\theta_j)/\theta_j$, a decreasing function of θ_j. Similarly, the unemployed meet jobs in a sector j at rate $M_j/u \equiv m_j(\theta_j)$, an increasing function of θ_j. Firms face two possible situations to fill or to hold a vacancy. It is assumed that holding a vacancy induces a flow cost $\gamma_j > 0$. Vacancies may become filled with probability $m_j(\theta_j)/\theta_j$ and bring a revenue to the firm equal to $y_i(a) - w_i(a)$, where $w_i(a)$ denotes the wage of a worker endowed with ability a.

If it is assumed, for simplicity, no job quits and no unemployment benefits, then job seekers move out of unemployment with probability $m(\theta_i)$ and enjoy a gain from of state change equal to the difference between the asset value of being employed, $W_i(a)$, and of being unemployed, $U_i(a)$. On the other hand, firms gains are a function of the asset values of holding a vacancy, V_i, and a filled job, $J_i(a)$, in sector i. It is assumed that holding a vacancy induces a flow cost $\gamma > 0$. Vacancies may become filled with probability $m(\theta_i)/\theta_i$ and bring a revenue to the firm equal to $y_i(a) - w_i(a)$, where $w_i(a)$ denotes the wage of a worker endowed with ability a.

On the workers side, it is assumed for simplicity no unemployment benefits. Job seekers move out of unemployment with probability $m_j(\theta_j)$ and enjoy a gain from of state change equal to the difference between the asset value of being employed, $W_i(a)$, and of being unemployed $U_j(a)$.

There is a rent sharing process between workers and firms over the surplus of the match, which is modeled using a standard Nash bargaining framework.

Heterogeneity of the labor force and schooling costs gives rise to self-selection in the educational system, a third key assumption in the model. Under this assumption agents face a binary choice, i.e. being educated or uneducated. Let $C>0$ denote the (ability-independent) schooling cost. At birth, each person compares his/her utility if educated (net of schooling costs) to that he/she would get if uneducated and invest accordingly. An individual endowed with characteristic a decides to get an education if and only if:

$$U_h(a) \geq U_l(a) - C \qquad (2)$$

The utility levels depend on expected wages and also on the job-finding rate $m(\theta_j)$ in each sector. As workers are charged a fixed cost for their schooling investment, only those with sufficiently high ability (a) choose to invest in education and are qualified to work on high–productivity jobs. The average abilities across the pools of unemployed in each sector are functions of an endogenous cutoff point, σ, and satisfy:

$$E_l(\sigma) = \int_0^\sigma \frac{\phi(a)}{\Phi(\sigma)} a \, da \text{ and } E_h(\sigma) = \int_0^\sigma \frac{\phi(a)}{1-\Phi(\sigma)} a \, da \qquad (3)$$

If educational attainment is positive correlated with labor market outcomes (wages and employment), self-selection gives rise to a composition effect. According to this effect, the mean ability among each education group decreases when the share of educated rises, i.e when the selection threshold, σ, decreases. As only the ablest get an education, the number of educated rises only if less-able individuals are drawn into education: the ablest among former uneducated now become the least able among educated. As a result, the average productivity across the two education groups falls. By differentiating equation (3) we see this effect of the threshold, σ, on the average abilities in each sector:

$$\frac{dE_l(\sigma)}{d\sigma} = \frac{\phi(\sigma)}{\Phi(\sigma)}\left[E_l(\sigma) - \sigma\right] > 0 \qquad (4)$$

$$\frac{dE_h(\sigma)}{d\sigma} = \frac{\phi(\sigma)}{1-\Phi(\sigma)}\left[\sigma - E_h(\sigma)\right] > 0 \qquad (5)$$

Self-selection creates an important externality: the composition of education groups affects firms' incentives to offer formal or informal jobs. A rise in the number of educated workers (fall in the threshold), hence a decline in the mean ability, deteriorates search prospects for firms in both sectors, which in turn advertise fewer vacancies. This leads to over-education, since the threshold individual considers her own earnings and employment prospects, which improve by schooling, but does not internalize the impact of her schooling decision on others' job opportunities, too many workers are willing to educate.

This model displays a negative structural relationship between tightness in each sector (define as the ratio vacancies to unemployment) and the share of workers with high education. A drop in labor market tightness increases the return to education, since it is relatively more important to raise one's job-finding rate when the market is loose. Workers acquire some excess education to improve their opportunities during the wage bargain and raising their chance of leaving unemployment.

Even though over-education will be a phenomenon that affects the economy as a whole, its incidence will be higher in that sector for which the market is looser and with a higher share of workers with high education. It is expected that in the formal sector the market is looser than in the informal sector. One possible explanation will be the existence of a minimum wage, a wage that is fixed above the market-clearing level, that only applies to the formal sector and it may lead to a significant difference between formal and informal sector in the wage pay for a work that requires the same level of education[3]. The above market-clearing wages paid in the formal sector rationed the jobs available and workers are forced to queue for the jobs available, like in Moen (1999) model. In such a setting, workers invest in education not only to get higher earnings, but also to improve their ranking in the job queue. In this competition, workers could end up applying to vacancies in which the required education is less than the acquired. Likewise, employers may use education as a criterion for filling the vacancies of the rationing jobs. Employers in the formal sector are in a position to 'skim the cream' by selecting the most proficient from the pool of available workers (Berry and Sabot, 1978).

3. Data and descriptive statistics

We use data from the Colombian Household Survey, a repeated cross-section conducted by the National Statistics Department (DANE), for 2008. The survey gathers information about employment conditions for population aged 12 or more including income, occupation and industry sector, in addition to the general population characteristics such as sex, age, marital status and educational attainment and covers the thirteen mayor metropolitan areas in Colombia.

The analysis was restricted to employed individuals that were not carrying formal studies aged between 15 and 60 years and who report working between 16 and 84 hours per week. Government employees, household employees, self-employed, bosses or employers, unpaid family workers, workers without pay in enterprises or other families business and journeyman or pawn are not included in the sample. The size of the sample is of 15675 observations.

---

[3] As Harberger (1971) points out "protected jobs can readily be identified because people want them. Companies paying wages higher than market levels for equivalent skills and working conditions tend to have very low labor turnover and long list of applicants for an opening to arise". Likewise, in the efficiency wage hypothesis, higher wages lead to higher productivity, and therefore the wage rate that minimizes total labor costs in the protected sector may not correspond to the one prevailing in the unprotected one.

There are four basic methods of measuring required education for a job and consequently over-education, two subjective measures and two objective measures. The first subjective measures use self-assessment to define the job's educational requirements and then comparing this with the worker's acquired education. The second and simply measure is asking directly the worker whether or not she or he is over-educated. Over-education can also be calculated objectively by using job analyst definition of the educational requirement for each occupation, as available in the United States (Dictionary of Titles), and comparing this with the educational level of the worker. The last objective measure is obtained by calculating the distribution of education for each occupation; employees who depart from the mean (Verdugo and Verdugo 1989) or mode (Mendes de Oliviera et al. 2000) by more than one standard deviation are classified as over-educated. Because the Colombian Household Survey only provides information of individual's two-digit occupations, in this study we measure over-education with the statistical method in its mean version[4]. The overall incidence of over-education for the sample is 15.44 (*see Table 1*).

Regarding informality, alternative definitions have been proposed in the literature and there is not an agreement about how to define and measure it, generally because data availability. According to DANE, informal workers are those who: (i) work in firms with five or fewer employees; (ii) are unpaid family workers and housekeepers; (iii) are self employed who work in places up to five persons other than independent professionals and technicians; or (iv) are business owners of firms with five or less employees. This definition, in line with the one proposed by the International Labor Office, has been criticized in the literature because it does not take into account the benefits associated with formal employment, such as social security system. The available data from the surveys permits us to determine whether workers are covered or not by the social security system, taking into account old-age insurance and health contributions. We classify workers as formal or informally employed according to social security system definition, under which, workers are classified as formal if they contributed to both health and old insurance. We find that 33.3% of our sample works in the informal sector[5].

Table 1 presents some basic summary statistics concerning the distribution of characteristics of our total sample and classified into those working in the formal and in informal sector. Accordingly to this statistics the incidence of over-education is 5.7% higher in the formal sector (17.3%) than in the informal sector (11.6%). This figure seems to support our first hypothesis; formal workers are more likely to be over-educated than informal ones. On the other hand, as can be seen, the percentage of female workers in the formal sector is higher than in the informa. A much larger proportion of the workforce in the formal sector is married. Informal workers have lower levels of education than formal workers, only 18% of informal workers have a higher education or more compare with 46% in the formal sector. There is little difference in the average hours of work in the two sectors. In contrast, there are some notable differences in the average of tenure between sectors; formal workers tend to accumulate much more tenure than informal workers, suggesting high stability of employment. In terms of the occupational structure, informal-sector workers are more likely to be in unskilled occupations compared to formal workers. Firms with less than 3 regular employees are substantially more likely to be part of the informal sector. In contrast, larger firms employ much of the formal-sector labor force with a workforce greater than one

---

[4] As Sanroma and Ramos (2004) explain, a 2 digit of occupational classifications is not an optimal level to apply the mode criterion.

[5] Self-employment in Latin America generally constitutes one of the principle sources of employment and a large of it operates in the informal sector, if the sample is not restricted the percentage of informal workers increased to 59% for 2008.

hundred. Nevertheless, while this exercise is informative it does not provide statistical significance of the relationship between over-education and labor market segmentation. To address this issue, we implement a multivariate analysis that we will present in the next section.

Table 1
Summary Statistics

	Total		Informal		Formal	
	Mean	SD	Mean	SD	Mean	SD
Over-education	0.15	0.36	0.12	0.32	0.17	0.38
Informal	0.33	0.47	1.00	0.00	0.00	0.00
Women	0.43	0.50	0.40	0.49	0.45	0.50
Married	0.50	0.50	0.46	0.50	0.53	0.50
By: Educational Attainment						
Non-education	0.01	0.08	0.01	0.12	0.00	0.05
Basic Primary	0.13	0.33	0.22	0.41	0.08	0.27
Basic secondary	0.14	0.35	0.22	0.41	0.10	0.30
Secondary	0.36	0.48	0.37	0.48	0.36	0.48
Higher education or more	0.37	0.48	0.18	0.39	0.46	0.50
Tenure (months)	48.19	67.33	27.39	47.64	58.60	73.08
Hours Work	50.48	10.89	50.89	13.24	50.27	9.49
Permanent contract type	0.48	0.50	0.12	0.33	0.65	0.48
By: Occupation						
Unskilled	0.32	0.47	0.42	0.49	0.27	0.45
Professionals and Technicians 1	0.07	0.25	0.02	0.14	0.09	0.29
Professionals and Technicians 2	0.05	0.22	0.04	0.20	0.05	0.22
Managers and Public Officials	0.03	0.18	0.02	0.13	0.04	0.20
Administrative Staff	0.20	0.40	0.14	0.34	0.24	0.42
Merchant and Vendor	0.17	0.37	0.18	0.39	0.16	0.37
Service Worker	0.16	0.36	0.18	0.39	0.15	0.35
By: Firm size						
1-3 workers	0.13	0.33	0.32	0.47	0.03	0.18
4-5 workers	0.10	0.29	0.20	0.40	0.04	0.20
6-10 workers	0.10	0.30	0.16	0.36	0.07	0.26
11-19 workers	0.08	0.26	0.07	0.26	0.08	0.27
20-30 workers	0.07	0.26	0.07	0.25	0.08	0.27
31-50 workers	0.07	0.25	0.04	0.20	0.08	0.27
51-100 workers	0.07	0.25	0.03	0.17	0.09	0.28
101 or more workers	0.39	0.49	0.11	0.31	0.53	0.50
By: Sector						
Agricultural, mining, electricity, gas and water	0.03	0.17	0.01	0.12	0.04	0.19
Industry	0.24	0.42	0.21	0.41	0.25	0.43
Construction	0.08	0.26	0.14	0.35	0.04	0.20
Sales, Hotels and Restaurants	0.29	0.45	0.37	0.48	0.25	0.43
Transportation	0.08	0.28	0.07	0.26	0.09	0.28
Financial Intermediation	0.12	0.33	0.07	0.26	0.14	0.35
Social Services	0.17	0.38	0.12	0.33	0.20	0.40
Observations	16043		5350		10693	

4. Empirical analysis

4.1 Probit model

To identify the determinants of educational mismatch we assume the following latent model that determines over-education:

$$Y_i^* = \beta X_i + \gamma S_i + \varepsilon_i \qquad (6)$$

7

The observed dichotomous realization Y_i of latent variable Y_i^* of whether the individual i is over-educated has the following form:

$Y_i = 1$ if the individual is over-educated ($Y_i^* \geq 0$)

$Y_i = 0$ otherwise

Thus, the determinants of over-education are estimated by using a univariate probit. X_i is a vector of individual characteristics such as age, gender, marital status, head of household, education, tenure, occupation, industry sector, contract type and the unemployment rate of the metropolitan area. Since the main aim of this study is to determine the effect of the sector in the probability of being over-educated it also includes a categorical, S_i, that takes the value of 1 if the individual is employed in the informal sector. ε_i is a normal distributed term that captures the unobservable component of over-education.

4.2 Probit with selection

The assumption that is often made in a simple univariate type of model is that the sector of employment, formal or informal, is exogenous to the probability of being over-educated. However, if the assignment of workers in each of the sectors is not random and some unobservable characteristics that influence the probability to choose a particular sector of employment could also affect the probability of being overeducated, then the estimation of a univariate probit suffers from selection bias[6]. One possible solution to the problem of sample selection from the process of classification is a two-stage Heckman estimation. We attempt to correct for selectivity bias by formulating a bivariate probit model with selection. This estimation consists of two equations, one for the classification of workers in each of the sectors and another for the propensity of being over-educated:[7]:

$$S_i^* = Z_i\gamma + \mu_i \tag{7}$$
$$Y_{ji}^* = X_i\beta_j + \varepsilon_{ji} \tag{8}$$

Where S_i^* is a latent variable that determines the sector in which individual i is employed and Y_{ji}^* is a latent variable that determines if the individual i in sector j (=formal, informal) is over-educated. Z_i is a vector of characteristics that influences the decision regarding the sector of employment and X_i is a set of individual characteristics that may influence the probability of being over-educated. To improve identification it is usually recommended that there must be a variable in the classification equation that is excluded from the outcome equation. Some family characteristics are believed to influence individual's choice of sector of employment but not affect over-education, such as the presence of children in the household. γ and β_j are vector of parameters. To allow for the possibility that unobserved

---

[6] We left aside another type of selectivity concerning the fact that an individual might not accept a job that does not match his or her level education and chooses instead to be unemployment or outside the labor force because in Colombia there is no unemployment benefit system and the family protection network against unemployment is low or exclusive for some group of individual with high income.

[7] The process in eq. (7) is in fact a classification process by which individuals are assigned to the formal or to the informal sector depending on their (observed and unobserved) characteristics. In contrast with the selection model, in such a process both outcomes allow observation of the main variable under analysis, that is if the individual is over-educated or not. A similar model with classification has been used in De la Rica (2004) in her study on the wage gap between workers with a permanent and with a fixed-term contract.

determinants of over-education and unobserved determinants of the sector of employment are correlated we assume that μ_i and ε_i are distributed bivariate normal, with $E[\mu_i] = E[\varepsilon_i] = 0$, $\text{var}[\varepsilon_i] = \text{var}[\mu_i] = 1$ and $\text{cov}[\varepsilon_i, \mu_i] = \rho_j$.

In the context of this classification model, our interest lies in the conditional probability of being over-educated given whether the worker is employed in the formal or informal sector. This conditional probability may be obtained through the joint probability of being over-educated and the probability of being employed in formal or informal sector:

$$prob\left(Y_i = 1 \mid S_i = j\right) = \frac{prob\left(Y_i = 1, S_i = j\right)}{prob(S_i = j)} = \frac{\Phi_2(X_{ij}\beta_j, Z_{ij}\gamma_j, \rho_j)}{\Phi\left(Z_{ij}\gamma_j\right)} \tag{9}$$

In addition, to complement the analysis in this case we also analyze the impact of the variables of interest on the unconditional probability of being over-educated, using the estimated parameters in the selection model:

$$prob\left(Y_i = 1\right) = \Phi\left(X_i\beta_j\right) \tag{10}$$

where $\Phi_2(\cdot)$ denotes the bivariate standard normal distribution and $\Phi(\cdot)$ the univariate standard normal distribution.

4.3 Decomposition analysis

The primary objective of the decomposition analysis is to consider the extent to which differences in the incidence of over-education between formal and informal employees reflect differences in characteristics (such as education), or differences in the response to those characteristics. We, therefore, perform a modified Oaxaca (1973) decomposition analysis in order to disentangle these two influences.

In the case of the probit model with selection, the decomposition based on the estimate parameters of the model can be written as:

$$\overline{prob\left(Y_i = 1 \mid S_i = formal\right)} - \overline{prob\left(Y_i = 1 \mid S_i = informal\right)} = \left[\frac{\overline{\Phi_2(X_{fo}\hat{\beta}_{fo}, Z_{fo}\hat{\gamma}_{for}, \hat{\rho}_{fo})}}{\Phi(Z_{fo}\hat{\gamma}_{fo})} - \frac{\overline{\Phi_2(X_{in}\hat{\beta}_{fo}, Z_{in}\hat{\gamma}_{fo}, \hat{\rho}_{fo})}}{\Phi(Z_{in}\hat{\gamma}_{fo})}\right]$$
$$+ \left[\frac{\overline{\Phi_2(X_{in}\hat{\beta}_{fo}, Z_{in}\hat{\gamma}_{fo}, \hat{\rho}_{fo})}}{\Phi(Z_{in}\hat{\gamma}_{fo})} - \frac{\overline{\Phi_2(X_{in}\hat{\beta}_{in}, Z_{in}\hat{\gamma}_{in}, \hat{\rho}_{in})}}{\Phi(Z_{in}\hat{\gamma}_{in})}\right] \tag{11}$$

The term on the left of the equality corresponds to the difference in the average probability of being over-educated between formal and informal workers. This difference may be explained by the sum of two terms. The first term on the right of equality, reflects the effect attributable to differences in characteristics between the groups, while the second term on the right of equality, quantifies the discrepancy caused by the difference in the effects of these characteristics.

5. Results

5. 1 Direct effect of informality on over-education

The coefficients and the marginal changes in probabilities (*dx/dy*) and their standard errors of estimating the univariate probit are displayed in Table 2. As is found in previews literature, the probability of being over-educated increases with the educational level (Alba-Ramirez, 1993; Kiker et al., 1997 and Quinn and Rubb, 2006). Overeducated workers may substitute education for the lack of previous job experience, taking jobs that required less education than they actually acquire in order to accumulate experience and improve their chances of finding a better job match (Rosen, 1972; Sicherman and Galor, 1990 and Mendes de Oliveira et al., 2000). Therefore we expect that over-educated workers will have lower experience. To test this hypothesis we use a variable that measures experience, particularly potential experience calculated as an individual's age minus five minus years of education. On the other hand, several studies have established that over-education may have a negative effect on job satisfaction (Tsang et al., 1991), if this is the case, then overeducated workers with more tenure in a firm are expected to be more prone to turnover. Consequently we hypothesize that over-educated workers will have less tenure. The results show that Experience has an inverted-U relationship with the probability of being over-educated; as experience grows first the probability also grows but then it stars to fall back. On the other hand, tenure does not affect the probability of being over-educated. These results gives some evidence of the substitutability between education and other forms of human capital postulated by the human capital theory, under which over-education might be seen as a transitory situation.

The results indicate that females are estimated to be less likely to be overeducated than males with similar characteristics and marital status does not have a statistically significant impact on the probability of being over-education for both genders. Females who are head of the household are more likely to be over-educated. By sector, workers in construction are more likely to be over-educated and as the size of the firm grows larger the probability of being over-educated is lower. Local labor market conditions seem to not be relevant, as the metropolitan unemployment rate is negative but not statistically significant. Having a permanent contract reduces the probability of being over-educated.

Contrary to what we were expecting, after controlling for observable characteristics, informal sector employees are found more likely to be overeducated rather than having adequate compare to formal workers. However, the coefficient is only statistical significant at 5% and its marginal effect is small, so there is no strong evidence of impact of informality over over-education once controlled for education and other observable characteristics. But it is possible that the effect of the sector of employment is subtler, there can be unobservable factors - ability among others - that affect both the probability of work in the formal or informal sector and the probability of being over-educated that a univariate probit analysis may not account for.

Table 2
Univariate Probit

Variable	Coefficient	S.E	dy/dx	S.E
schooling years	0.2433**	0.0056	0.0383***	0.0009
experience	0.0096+	0.0056	0.0015	0.0009
experience square	-0.0004*	0.0001	-0.0001*	0.0000
tenure	-0.0004	0.0006	-0.0000	0.0000
tenure square	0.0000	0.0000	0.0000	0.0000
gender (ref: men)				
women	-0.2389**	0.0419	-0.0368***	0.0064
marital status (ref: single)				
married	-0.0227	0.0493	-0.0036	0.0078
woman married	-0.0003	0.0662	-0.0002	0.0104
Household member (ref: others)				
head	-0.0474	0.0493	-0.0074	0.0077
women head	0.1750*	0.0713	0.0302*	0.0134
sector (ref: agricultural, mining, electricity, gas and water)				
Industry	0.0806	0.0846	0.0130	0.0141
Construction	0.5033**	0.0983	0.104***	0.0252
Hotel and Restaurants	-0.0966	0.0849	-0.0148	0.0126
Transportation	-0.3509**	0.0929	-0.0450***	0.0095
Financial Intermediation	-0.4072**	0.0897	-0.0517***	0.0090
Social Services	-0.4962**	0.0881	-0.0625***	0.0086
Firm size (ref: 1 - 3 workers)				
4-5 workers	-0.1144+	0.0616	-0.0169*	0.0085
6-10 workers	-0.1393*	0.0614	-0.0203*	0.0083
11-19 workers	-0.0971	0.0683	-0.0144	0.0096
20-30 workers	-0.1095	0.0701	-0.0161	0.0096
31-50 workers	-0.1125	0.0711	-0.0165	0.0097
51-100 workers	-0.2941**	0.0757	-0.0388***	0.0081
101 or more workers	-0.0758	0.0553	-0.0118	0.0085
contract type (ref: temporal term)				
Permanent	-0.0923**	0.0329	-0.0145**	0.0051
Informal	0.0982*	0.0412	0.0158*	0.0068
Metropolitanarea unemployment rate	0.0059	0.0078	0.000921	0.0012
Constant	-3.7439**	0.1571	-	-
N	15,675			

+ p<0.1, * p<0.05, ** p<0.01

5.2 Results with sample selection

As noted above, the univariate probit estimates of the probability of being over-educated could be biased if the allocation of workers into one sector or another is nonrandom. In order to obtain consistent estimates we implement the bivariate selection correction as presented before.

Table 3 the coefficients and the marginal effects, conditional and unconditional, of the bivariate selectivity-corrected over-educated equations for formal and informal workers. The central element of this discussion is on the probability of being over-educated, because the parameters in the formal and informal equations are relatively standard, and largely conform to results reported elsewhere (Magnac, 1991 and Pradhan and Soest, 1995). In general, the estimated coefficients of the over-education equation are roughly identical to those estimated with the bivariate probit for informal workers and we will not describe them again for brevity reasons. In contrast, we observe some interesting differences with respect to the univariate probit estimates for the case of formal workers, which appear to be worth analyzing in more detail. After we control for being selected as a formal worker, the size of the firm have a positive and significant effect of being over-educated. This result

might be interpreted as follows; large firms usually have better job opportunities, aside of higher wages, in this type of firms workers have more probability to be promoted and more training on the job. These characteristics make job's offers from large firm high valuable for job seekers who compete for this works. In this competition, workers could end up applying to vacancies in which the required education is less than the acquired. Likewise, employers of large firms in the formal sector are in a position to select the most skilled from the pool of available workers.

Regarding the correlation of the disturbance terms of the two-step equations, ρ, this parameter is positive and highly statistical significant for the formal equation. In contrast, ρ, is insignificant throughout for informal workers. This finding suggests that for formal workers non-observable characteristics, that exert a positive effect on the probability of being employed in the formal sector, have a positive impact on the probability of being over-educated. This could be interpreted as evidence that for formal workers in the sample; over-education is caused, at least in part by the desire to become part of the formal sector (better employment opportunities, social system protection, etc.). A worker with considerable formal education possibly will enter a job for which less education is required, because that job is protected, for example, by the minimum wage, while the occupation for which the worker is best suited pays less than the minimum in the informal sector. Nevertheless, if ability is considered a non-observable characteristic, and according to the theoretical framework presented before, this result seems somewhat contradictory. One will expect that workers with more ability, and hence with higher skills, end up working in the formal sector, and that this ability prevents them from being over-educated. In this case, r will be negative rather than positive. Our results imply, therefore, that this negative effect of ability is more than offset by the positive effect of the desire to become part of the formal sector, among other factors.

Table 3
Probit with selection

Variable	Informal			Formal		
	Coefficients	Uncondicional dy/dx	Condicional dy/dx	Coefficients	Uncondicional dy/dx	Condicional dy/dx
schooling years	0.2992**	0.0285**	0.0403**	0.2301**	0.0243**	0.0335**
experience	0.0036	0.0003	-0.0007	0.0173**	0.0018**	0.0039**
experience square	-0.0004	0.0000	0.0000	-0.0005**	-0.0001**	-0.0001**
tenure	-0.0002	0.0000	0.0001	0.0007	0.0001	-0.0002+
tenure square	-0.0000	0.0000	0.0000	-0.0000	0.0000	0.0000
gender (ref: men)						
women	-0.2688**	-0.0250*	-0.0350**	-0.1803**	-0.0188**	-0.0267**
marital status (ref: single)						
married	-0.0388	-0.0037	-0.0025	-0.0161	-0.0017	-0.0077
woman married	0.1673	0.0175	0.0205	-0.0477	-0.0049	0.0006
Household member (ref: others)						
head	-0.1021	-0.0096	-0.0141	0.0021	0.0002	0.0018
women head	0.3265*	0.0386	0.0514	0.0909	0.0102	0.0158
sector (ref: agricultural, mining, electricity, gas and water)						
Industry	0.0110	0.0011	-0.0021	0.0309	0.0033	0.0101
Construction	0.8842**	0.1517*	0.1775	-0.0378	-0.0039	0.0132
Hotel and Restaurants	-0.3719	-0.0312	-0.0462	-0.0441	-0.0046	-0.0037
Transportation	-0.5618*	-0.0363*	-0.0542	-0.3207**	-0.0273**	-0.0347**
Financial Intermediation	-0.8329**	-0.0482**	-0.0700	-0.3040**	-0.0267**	-0.0386**
Social Services	-0.7809**	-0.0500**	-0.0747	-0.4976**	-0.0407**	-0.0509**
Firm size (ref: 1 - 3 workers)						
4-5 workers	-0.1380	-0.0120+	-0.0105	0.1089	0.0124	0.0048
6-10 workers	-0.2252*	-0.0185**	-0.0133	0.2985**	0.0383**	0.0259+
11-19 workers	-0.1699	-0.0143+	0.0039	0.4438**	0.0632**	0.0457**
20-30 workers	-0.2134	-0.0174*	-0.0004	0.4752**	0.0692**	0.0513**
31-50 workers	-0.4555*	-0.0312**	-0.0243	0.5734**	0.0892**	0.0699**
51-100 workers	-0.6100**	-0.0376**	-0.0360	0.4231**	0.0597**	0.0383*
101 or more workers	-0.2753	-0.0251**	0.0019	0.6396**	0.0775**	0.0387**
contract type (ref: fixed term)						
Indetermine	-0.2648**	-0.0252**	-0.0355	-0.0662*	-0.0070*	-0.0096*
Metropolitanarea unemployment rate	-0.0020	-0.0002	-0.0007	0.0026	0.0003	0.0011
Constant	-4.0400**	_	_	-4.4515**	_	_
ρ	0.1887	_	_	1.6212**	_	_
Observations	15675			15675		

+ p<0.1, * p<0.05, ** p<0.01

5.3 Decomposition results

As described in Section 4, the predicted probability of over-education is decomposed using the estimates from the bivariate probit models with selection. Table 4 presents the decomposition results. The incidence of over-education for formal workers is 17% while the incidence for informal workers is 12%. Thus, the raw difference in the incidence of over-education between formal and informal workers is 5.6%. Of this difference 91% can be explain due to differences in the characteristics between the groups. In other words, if the distribution of observable characteristics of worker of the formal and the informal sector were the same, no significant differences in the propensity of over-education between the two sectors would be observed. This is so, because just 9 percentage

points of the total gap is attributable to differences across groups in the impact of the observable characteristics and to the effect of unobservables. These results are consistent with results of previews works that do not find a formal sector wage premium, after they controlled for individual and firm characteristics (Pratap and Quintin, 2006).

Table 4
Decomposition of the conditional probability

Incidence of overeducation	
Formal sector	17.3%
Informal sector	11.6%
Gap (Formal - Informal)	5.7%
Explain by characteristics	5.1%
Explain by parameters	0.5%

6. Conclusions

This study has attempted to extend the over-education literature for developing countries by examine the connection between labor market segmentation, a modern protected formal sector and a traditional-unprotected-informal sector, and over-education in Colombia. So far, studies about informality in developing countries have focused primarily on the size of the informal sector and on the effects of labor market rigidities on employment, wages and its distribution, and on the probability to become informal. However, no attention has been paid to the effects of a large informal sector in the way workers match their education with the one required to perform their job. A first descriptive analysis using recent data from Colombia indicates that the incidence of over-education for formal workers is higher than that presented for informal workers. However, once we control for individual and firm characteristics this relationship vanishes. Actually, we have provided empirical evidence that suggests that labor market segmentation, rather than having a direct effect, has a subtler impact on over-education. Some unobservable factors that affect the probability of working in the formal sector affect also the probability of being over-educated, whereas this mechanism does not exist in the case of the informal sector.

Using micro data from Colombia, we estimated two types of empirical models in order to test the relationship between over-education and informality: a simple univariate probit model for the probability of being over-educated that includes as an argument the sector in which the individual is employed, and a bivariate probit model with selection, that takes into account that the process of classification into the formal or the informal sector has an effect in the probability that an individual is over-educated. The results from the univariate probit estimation indicate that, apparently, informal workers are found more likely to be overeducated rather than having adequate compare to formal workers. However, we have also proved that the assignment of workers in formal or informal sector is not random and that some unobservable characteristics that influence the probability to choose a particular sector of employment also affect the probability of being overeducated. In such scenario, results from the estimation of a standard probit model do not provide consistent results due to sample selection.

Actually, the results obtained from the probit model for the probability of over-education controlling for the classification into a given sector, confirm that non-observable characteristics that exert a positive effect on the probability of being employed in the formal sector have a positive impact on the probability of being over-educated. This could be interpreted as evidence that for formal workers in the sample, over-education is caused at

least in part by the desire to become part of the formal sector (better employment opportunities, social system protection, etc.). A worker with considerable formal education possibly will enter a job for which less education is required, because that job is protected, for example, by the minimum wage, while the occupation for which the worker is best suited pays less than the minimum in the informal sector. Nevertheless, given that ability is among the non-observable characteristic, and according to the theoretical framework presented, this result seems somewhat contradictory. One will expect that workers with more ability, and hence with higher skills, end up working in the formal sector, and that this ability prevents them from being over-educated. Our results imply, therefore, that the negative effect of ability on over-education is more than offset by the positive effect of the desire to become part of the formal sector, among other factors. It is in our agenda, for future research, to explore more in detail the unobservable factors that affect the probability of being over-educated and become part of the formal sector.

In any case, the decomposition analysis suggests that the difference of the incidence of over-education between formal and informal workers is explained mostly by differences in observable characteristics, with only a small part that is attributed to differences in the effects and, potentially, to the impact of unobservable characteristics.

Appendix 1. Average Explanatory Variables

	Mean	Min	Max
schooling years	10,98526	0	26
experience	18,01435	0	55
experience square	458,0555	0	3025
tenure (*in months*)	48,63445	0	520
tenure square	6948,51	0	270,4
women	0,4328549	0	1
married	0,5094099	0	1
woman married	0,1719298	0	1
head	0,4213716	0	1
women head	0,1075598	0	1
Industry	0,236236	0	1
Construction	0,0749601	0	1
Hotel and Restaurants	0,2835088	0	1
Transportation	0,0828708	0	1
Financial Intermediation	0,1216587	0	1
Social Services	0,1723764	0	1
4-5 workers	0,0947368	0	1
6-10 workers	0,101244	0	1
11-19 workers	0,076236	0	1
20-30 workers	0,0729187	0	1
31-50 workers	0,0678788	0	1
51-100 workers	0,0688995	0	1
101 or more workers	0,3906858	0	1
Permanent	0,4786603	0	1
Informal	0,3303349	0	1
Metropolitan area unemployment rate	11,86843	9,089485	16,31971

Appendix 2. Probit with selection - Informal equation estimates

Variable	Coefficients	Standard error
age	-0.2145**	0.0194
age2	0.0026**	0.0002
Educational Attainment (Ref: none)		
Primary	-0.3982*	0.1642
Basic secondary	-0.4960**	0.1806
Secondary	-0.7442**	0.2000
Higher education or more	-0.9037**	0.2239
experience	0.0640**	0.0143
experience square	-0.0011**	0.0002
tenure	-0.0085**	0.0005
tenure square	0.0000**	0.0000
gender (ref: men)		
women	-0.0200	0.0417
marital status (ref: single)		
married	-0.1464**	0.0459
woman married	0.1864**	0.0609
Household member (ref: others)		
head	0.0294	0.0449
women head	0.0608	0.0640
Occupation (ref: Agricultural or forestry)		
Professionals and Technicians 1	-0.0258	0.0865
Professionals and Technicians 2	0.4597**	0.0847
Managers and Public Officials	-0.0723	0.0954
Administrative Staff	-0.1509**	0.0499
Merchant or Vendor	-0.0717	0.0490
Service Worker	-0.0383	0.0452
sector (ref: agricultural, mining, electricity, gas and water)		
Industry	0.2011*	0.0900
Construction	0.4582**	0.0973
Hotel and Restaurants	0.1173	0.0907
Transportation	0.1511	0.0974
Financial Intermediation	-0.0948	0.0950
Social Services	0.2399*	0.0944
Firm size (ref: 1 - 3 workers)		
4-5 workers	-0.4010**	0.0518
6-10 workers	-0.8321**	0.0493
11-19 workers	-1.3303**	0.0543
20-30 workers	-1.4036**	0.0555
31-50 workers	-1.6465**	0.0592
51-100 workers	-1.8328**	0.0618
101 or more workers	-2.0778**	0.0446
Metropolitan area unemployment rate	0.0256**	0.0069
Number of child 0 -2	0.0669*	0.0265
Number of child 3 -5	0.0187	0.0284
Number of unemployed	0.0502*	0.0252
Household education	-0.0308**	0.0063
Constant	4.9592**	0.3184
Observations	15675	

+ $p<0.1$, * $p<0.05$, ** $p<0.01$

Appendix 3. Probit with selection - Formal equation estimates

Variable	Coefficients	Standard error
age	0.1865**	0.0204
age2	-0.0025**	0.0002
Educational Attainment (Ref: none)		
Primary	0.4676**	0.1632
Basic secondary	0.6204**	0.1824
Secondary	0.8634**	0.2046
Higher education or more	0.9231**	0.2288
experience	-0.0409**	0.0152
experience square	0.0010**	0.0002
tenure	0.0083**	0.0005
tenure square	-0.0000**	0.0000
gender (ref: men)		
women	0.0287	0.0410
marital status (ref: single)		
married	0.1639**	0.0465
woman married	-0.2152**	0.0606
Household member (ref: others)		
head	-0.0439	0.0453
women head	-0.0458	0.0632
Occupation (ref: Agricultural or forestry)		
Professionals and Technicians 1	0.5160**	0.0758
Professionals and Technicians 2	0.0084	0.0757
Managers and Public Officials	0.4571**	0.0835
Administrative Staff	0.3214**	0.0449
Merchant or Vendor	0.1676**	0.0457
Service Worker	0.0324	0.0442
sector (ref: agricultural, mining, electricity, gas and water)		
Industry	-0.1535+	0.0906
Construction	-0.4472**	0.0975
Hotel and Restaurants	-0.0828	0.0913
Transportation	-0.1423	0.0984
Financial Intermediation	0.0912	0.0980
Social Services	-0.2695**	0.0953
Firm size (ref: 1 - 3 workers)		
4-5 workers	0.3993**	0.0505
6-10 workers	0.8203**	0.0487
11-19 workers	1.3181**	0.0541
20-30 workers	1.3950**	0.0555
31-50 workers	1.6336**	0.0589
51-100 workers	1.8259**	0.0622
101 or more workers	2.0775**	0.0444
Metropolitan area unemployment rate	-0.0229**	0.0069
Number of child 0 -2	-0.0819**	0.0253
Number of child 3 -5	-0.0148	0.0274
Number of unemployed	-0.0399+	0.0239
Household education	0.0336**	0.0062
Constant	-4.7926**	0.3245
Observations	15675	

+ p<0.1, * p<0.05, ** p<0.01

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.

Alba-Ramirez, A., & San Segundo, M. J. (1995). The returns to education in Spain. *Economics of Education Review, 14*(2), 155-166.

Alpin, C., Shackleton, J. R., & Walsh, S. (1998). Over- and undereducation in the UK graduate labour market. *Studies in Higher Education, 23*(1), 17-34.

Bauer, T. K. (2002). Educational mismatch and wages: a panel analysis. *Economics of Education Review, 21*(3), 221-229.

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

Berry, A., & Sabot, R. H. (1978). Labour market performance in developing countries: A survey. *World Development, 6*, 1199-1242.

Büchel, F., & Battu, H. (2003). The Theory of Differential Overqualification: Does it Work? *Scottish Journal of Political Economy, 50*(1), 1-16.

Castillo, M. (2007). Desajuste educativo por regiones en Colombia: ¿competencia por salarios o por puestos de trabajo? In: Cuadernos de Economía, v. XXVI, n. 46, (pp. 107-145). Bogotá.

Charlot, O., & Decreuse, B. (2005). Self-selection in education with matching frictions. *Labour Economics, 12*(2), 251-267.

Chevalier, A. (2003). Measuring Over-education. *Economica, 70*(279), 509-531.

De la Rica, S. (2004). Wage gaps between workers with indefinite and fixed-term contracts: The impact of firm and occupational segregation. *Moneda y Crédito*, *219, 43-69.*

Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the U.K. graduate labour market. *Economics of Education Review, 19*(2), 179-198.

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

Duryea, S., Galiani, S., Nopo, H., & Piras, C. C. (2007). The Educational Gender Gap in Latin America and the Caribbean. In: IDB Working Paper No. 502.

Flórez, C. E. (2002). The Function of the Urban Informal Sector in Employment: Evidence from Colombia 1984-2000. In: Documento CEDE No. 2002-04, Universidad de Los Andes. Bogotá D.C.

Gong, X., & Van Soest, A. (2001). Wage differentials and mobility in the urban labor market: A panel data analysis for Mexico. In: IZA, Bonn Discussion Paper No. 329.

Green, F., McIntosh, S., & Vignoles, A. (2002). The Utilization of Education and Skills: Evidence from Britain. *The Manchester School, 70*(6), 792-811.

Groot, W., & van den Brink, H. (2000). Skill mismatches in the Dutch labor market. *International Journal of Manpower 21*(8), 584–595.

Harberger, A. C. (1971). On Measuring the Social Opportunity Cost of Labour. *Int. Lab. Rev., 103*(6), 559-579.

Hartog, J., & Oosterbeek, H. (1988). Education, allocation and earnings in the Netherlands: 0verschooling? *Economics of Education Review, 7*(2), 185-194.

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.

Kugler, A., & Kugler, M. (2009). The Labor Market Effects of Payroll Taxes in a middle Income country: Evidence from Colombia. *Economic Development and Cultural Change*.

Magnac, T. (1991). Segmented or competitive labor markets. . *Econometrica, 59* 165–187.

Maloney, W. F. (2004). Informality Revisited. *World Development, 32*(7), 1159-1178.

Maloney, W. F., & Núñez, J. (2004). Measuring the Impact of Minimum wages: Evidence from Latin America. In *Law and Employment: Lessons from Latin America and the Caribbean,* (pp. 109-130). University of Chicago Press and National Bureau of Economic Research.

McGoldrick, K., & Robst, J. (1996). Gender differences in overeducation: a test of the theory of differential overqualification. *American Economic Review* (AEA Conference proceedings).

McGuinness, S. (2006). Overeducation in the Labour Market. *Journal of Economic Surveys, 20*(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*(6), 1334-1347.

Mendes de Oliveira, M., Santos, M. C., & Kiker, B. F. (2000). The role of human capital and technological change in overeducation. *Economics of Education Review, 19*(2), 199-206.

Moen, E. (1999). Education, ranking, and competition for jobs. *Journal of Labor Economics , 17,* 694 - 723.

Mondragón-Veléz, C., Peña, X., & Wills, D. (2010). Labor Market Rigidities and Informality in Colombia. In: Documento CEDE No. 2010-07, Universidad de Los Andes. Bogotá D.C.

Mora, J. J. (2005). Sobre educación en Cali (Colombia). ¿Desequilibrio temporal o permanente?: algunas ideas, 2000-2003. *Documentos Laborales y Ocupacionales, 2 SENA*.

Núñez, J. (2002). Empleo informal y evasión fiscal en Colombia. In *Archivos de Economía, 210, DNP*.

Pradhan, M., & van Soest, A. (1995). Formal and informal sector employment in urban areas of Bolivia. *Labour Economics, 2*(3), 275-297.

Pratap, S., & Quintin, E. (2006). Are labor markets segmented in developing countries? A semiparametric approach. *European Economic Review, 50*(7), 1817-1841.

Quinn, M. A., & Rubb, S. (2006). Mexico's labor market: The importance of education-occupation matching on wages and productivity in developing countries. *Economics of Education Review, 25*(2), 147-156.

Rosen, S. (1972). Learning and experience in the labor market. *Journal of Human Resources, 7*, 326–342.

Sicherman, N. (1991). Overeducation in the labor market. *Journal of Labor Economics, 9* (2), 101–122.

Sicherman, N., & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy, 98*(1), 169–192.

Tansel, A. (1999). Formal versus informal sector choice of wage earners and their wages in Turkey. In: Economic Research Forum Working Paper No. 9927.

Tenjo, J. (1990). Opportunities, Aspirations, and Urban Unemployment of Youth: The Case of Colombia. *Economic Development and Cultural Change, 38* (4), 733-761.

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

Tinbergen, J. (1956). On the theory of income distribution. *Weltwirtchaftliches Archiv, 77*, 156–175.

Tsang, M. C., Rumberger, R. W., & Levin, H. M. (1991). The Impact of Surplus Schooling on Worker Productivity. *Industrial Relations: A Journal of Economy and Society, 30*(2), 209-228.

Verdugo, R., & Verdugo, N. (1989). The impact of surplus schooling on earnings. *Journal of Human Resources, 24* (4), 629–643.