



Estudios Económicos

ISSN: 0188-6916

jsempe@colmex.mx

El Colegio de México, A.C.

México

Caamal Olvera, Cinthya G.  
DECREASING RETURNS TO SCHOOLING IN MEXICO  
Estudios Económicos, vol. 32, núm. 1, enero-junio, 2017, pp. 27-63  
El Colegio de México, A.C.  
Distrito Federal, México

Available in: <http://www.redalyc.org/articulo.oa?id=59749085002>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative

## DECREASING RETURNS TO SCHOOLING IN MEXICO\*

Cinthya G. Caamal Olvera

*Universidad Autónoma de Nuevo León*

*Resumen:* Se estudia la relación entre el crecimiento de una fuerza laboral con mayor nivel educativo y la remuneración que obtiene en el mercado laboral en México. Se presentan coeficientes robustos de los rendimientos de la escolaridad; los sesgos de la habilidad y selectividad se subsanan al aplicar métodos paramétricos, semi-paramétricos y semi-no paramétricos. La estrategia empírica permite comparar los rendimientos de la escolaridad entre cuantiles para analizar la desigualdad en la distribución condicionada del salario a lo largo del tiempo. Los datos revelan una tendencia decreciente en el rendimiento derivado de la escolaridad al paso del tiempo. Los resultados indican que la desigualdad entre el cuantil más alto y el más bajo también se ha reducido.

*Abstract:* The aim of this paper is to study the relationship between a growing better educated labour force in Mexico and labour market reward. The paper provides robust estimates of the returns to education, dealing with the ability and selectivity biases using parametric, semi-parametric, and semi-nonparametric methods. The empirical strategy allows comparing the returns to education across quantiles to analyse the inequality variations within the conditional wage distribution over time. Data reveal a declining trend in the returns to education over time. Results also indicate that the inequality between the top and lower quantiles has been reduced.

*Clasificación JEL/JEL Classification:* I26, J21, C14

*Palabras clave/keywords:* returns to schooling, semi-nonparametric estimation, labour force, estimación semi-noparamétrica rendimientos de la educación, fuerza laboral

*Fecha de recepción:* 7 IV 2015

*Fecha de aceptación:* 27 X 2015

---

\* Facultad de Economía, cinthya.caamallv@uanl.edu.mx

## 1. Introduction

The objective of the present study is to analyse trends of a growing better educated labour force in Mexico and appraise how the labour market has valued the increasing human capital accumulation. This objective will be carried out through the estimation of the returns to education, specifically the effect on wages of a rising education level. The level of education chosen and the market rewards in the form of wages may have been affected by several structural changes that occurred in Mexico in recent decades, including an increase in compulsory education years and greater openness to trade and privatization of government-owned firms. The Mexican economy has faced sluggish economic growth in recent years; on average, its Gross Domestic Product (GDP) grew only about 2.35 percent per year between 2000 and 2013, which is low compared with other developing countries. To enhance growth, increase flexibility in the labour market, increase tax revenue, and encourage credit, several reforms were approved during 2012 and 2013. It is too soon, however, to study the effects of these reforms on the labour market.

The resources used to examine the relationship between wages and education were Mexico's National Urban Employment Survey (ENEU) for the period 1988-2004 and the National Survey of Occupation and Employment (ENOE) for the subsequent period 2005-2013. A standard measure to estimate the effect of education on workers income is applying the structural model of the returns to education proposed by Mincer (1958). In this study a nonparametric technique is used to deal with the unobservable characteristics across individuals because of the association between education and ability, career choice, or type of occupation over time. Quantile regression proposed by Koenker and Bassett (1978) will be used to account for this bias by separating the unobserved heterogeneity. To keep the comparability over time, the instrumental variables method is not recommended because of the issues related to finding a convenient instrument with the required properties.

A common problem in the household surveys is unreported wages by individuals, which can enlarge the bias in the estimates. To deal with the selectivity bias potentially present in the sample, several correction methods will be applied, such as the parametric method as developed by Heckman (1979) and a semi-parametric and semi-nonparametric method following Gallant and Nychka (1987), De Luca (2008) and De Luca and Perotti (2010), which are more flexible than the self-selection parametric correction method.

The contribution of the paper is to have found a robust declining trend over time on the returns to schooling in Mexico after 1997. Particularly if a declining trend on the returns to schooling is found in the upper quantiles but not in the lower quantiles, a reduction of inequality between the top and lower quantiles is predicted. These results are consistent with a phenomenon related to a reduction in income and wage inequality estimated after the North-American Free Trade Agreement (NAFTA) by Esquivel (2011) and Campos (2013b).

The paper is structured as follows: section 2 provides a brief literature review of the empirical evidence regarding the trend of the returns to education over time; section 3 presents a brief summary of the schooling and wage trends in Mexico and a descriptive statistic summary of the sample used; section 4 discusses the main issues regarding the estimation of the returns to schooling; section 5 describes the empirical strategy of the study; section 6 presents the results for male workers; and section 7 concludes. The appendix is presented at the end of the document.

## 2. Literature review

The estimation of returns to schooling is the main topic of Mincer (1958), Schultz (1961), Becker (1962), Becker (1964), Ben-Porath (1967), and Mincer (1974). In these studies, the authors asserted that productivity can be improved because schooling directly enhances wages. In contrast, Spence (1973) proposed that schooling is related to higher wages through a signalling effect of ability. Weiss (1995) claimed the main distinction between the two approaches is this: the first one assumes that education is the cause of workers' productivity differences, and the second one assumes that workers' differences exist prior to the education choice. Regan, Oaxaca, and Burghardt (2007) developed a neoclassical model of optimal schooling, arguing not only that the Mincerian schooling model overstates the returns to education due to the lack of an ability control variable but also that in the context of a simple schooling model, with a linear schooling specification, returns to education have identification problems and cannot be considered as an internal rate of return.

The empirical evidence regarding the returns to schooling has accumulated over time. Cross-sectional studies have attempted to disentangle the endogeneity of education and wages. Card (1999) reviewed literature including studies by Griliches (1977), Angrist and Krueger (1991), and Ashenfelter and Krueger (1994), among others,

using instrumental variables to account for this endogeneity. However, using instrumental variables methodology to estimate the returns to education over time may also present issues regarding the choice of a convenient instrument, as Heckman and Vytlačil (2005) pointed out, because different instruments define different parameters. Other studies have estimated returns to schooling by looking at the whole conditional earnings distribution, such as the quantile approach by Koenker and Bassett (1978). Its robustness properties in the presence of heterogeneity make quantile regression a suitable technique. Many researchers have also studied the evolution of the returns to education. Case country empirical evidence provides different patterns of the returns to schooling over time. Various studies reflect an increasing, stable, or declining trend; the direction of the trend depends on the characteristics of each country, time horizon considered, structural reforms, change in the political system, and the like. For example, Buchinsky (2001) estimated an increasing trend of the returns to schooling in the US between 1963 and 1980 at all quantiles of wage. Also, his results showed that returns to education and experience are different for any quantile, even if the change over time follows the same pattern for all the quantiles.

Machado and Mata (2001) provided evidence of increasing returns to schooling over the period 1982-1994 for Portugal, claiming that education is more valued for highly paid jobs because the impact of education at the tails of the distribution was distinct, the return at the 90th quantile increased by 3 percent, and the returns at the low quantiles decreased by 1.5 percent. Although returns to nine-year mandatory schooling decreased over the entire wage distribution, they argued this is because of the fall of the returns associated with the elementary education categories. The returns to having a college degree on average increased from 1982 to 1994 for the median and upper quantiles. They concluded that only after a certain degree, education pays off, and when this happens education is more valued for highly paid jobs.

An increasing trend of returns to education over time is also estimated for transition economies; Flabbi, Paternostro, and Tiongson (2007) used comparable data for eight countries<sup>1</sup> from the early transition period up to 2002. They suggested that the cause of this tendency is the institutional and structural factors that were present over this period of time.

---

<sup>1</sup> Bulgaria, Czech Republic, Hungary, Latvia, Poland, Russia, Slovak Republic, and Slovenia.

Some authors have compared different countries using similar controlled variables. Martins and Pereira (2004) presented quantile estimates for the returns to schooling for 16 European countries for the mid-1990s and compared trends over time across countries. They contrasted specific cross-country returns to schooling according to the accessibility of country data and accounted for differences in the data observations; diverse sources of information such as household, employee, or employer survey; and different wage measures, gross or net. They found a robust stylised fact that returns to schooling are larger for more skilled individuals, and thus returns to schooling increased across deciles for the set of countries analysed. Greece showed a rather decreasing trend across deciles, but they claimed this was because the available wage data were net of taxes, which makes the comparison to other countries difficult. Italy and Austria also reported net wages, and their trend across deciles was opposite to Greece.

Returns to education estimates for Germany have shown stable returns to schooling. Fitzenberger and Kurz (2003) provided an empirical analysis of the structure of earnings in West Germany across skill groups and industries for the period 1984-1994. They used panel data with a block bootstrap procedure to account for heterogeneity and autocorrelation in the error term. They found a uniform trend over time, as well as different effects of human capital and industry variables on earnings across quantiles.

Conversely, other studies have found declining trends over time. Naticchioni, Ricci, and Rustichelli (2010) compared public and private sector workers and found deeper decline in the private sector, which was the result of institutional factors such as stronger unions, higher wage compression, and other labour market conditions that did not affect the public sector in Italy. Unlike other countries where the trend of Educational Wage Premia (EWP) has been stable or slightly increasing for some groups, Austria has shown a decrease of the EWP for all educational attainments at all quantiles.

Fersterer and Winter-Ebmer (2003) estimated returns to education in Austria over the period 1981-1997 using cross-sectional data. They estimated that returns to secondary and tertiary education drop for all quantiles and the spread of returns is lower for women. In addition, they offered an explanation of this decline, which is consistent with a rise in the number of highly educated workers over the last two decades.

In general, they found higher estimated returns to education at higher quantiles, and all the coefficients are statistically different from each other. Furthermore, the fall in returns over time is relatively

similar across quantiles, although women's returns to schooling fell disproportionately in the lowest decile.

According to some studies for the case of Mexico, the tendency of the returns to schooling has not followed a linear development over 1987-2002. Rodríguez (2005) found that the highest rates of returns to education were present in 1991-1992 and the lowest, which coincide with the Mexican peso crisis, in 1994-1995. Moreover, the dispersion of the returns to education among regions increased, first because of the peso crisis and second because of the effects of the North American Free Trade Agreement (NAFTA). However, as he mentioned, it is striking to observe that returns to schooling have decreased in recent years. López (2006) found that returns to schooling for upper secondary rose sharply in the late 1980s and early 1990s and then fell after 1993. However, returns to tertiary education continued to rise until 1996 before falling to levels that remained superior to those observed in the early 1990s. Moreover, she explained this drop as a cyclical fall in education premium in recession times, which was observed in other Latin American countries.

After NAFTA, Robertson (2004) found that the relative price of skill-intensive goods fell, therefore the relative wage of skilled workers decreased. Esquivel (2011) also found an income inequality reduction, after NAFTA, explained by a reduction in labour income, a rise on remittances, and public transfers. Campos (2013b) found a decreasing wage inequality, after 1996, at the top of the wage distribution; he attributed this trend to a decrease in returns to education, slower demand growth, and an increase in supply of college educated.

### 3. Schooling and wage trends in Mexico

The information to carry out this study is obtained from Mexico's National Urban Employment Survey (ENEU) for the period 1988-2004 and the National Survey of Occupation and Employment (ENOE) for the subsequent period 2005-2013. The two surveys are equivalent in terms of the questions addressed in this research, even though the questionnaire changed. Another important change is the sampling strategy; nonetheless it is possible to make comparable inferences. ENEU-ENOE surveys present quarterly data and are a five-quarter round panel, which means they follow up with one family for five subsequent quarters. The estimation will consider only men who responded to the questionnaire in the third quarter of each year. The selection of the third quarter not only avoids the overrepresentation

of the individuals in the sample but also provides a constant measure of income that is not changing with additional payments for utility shares, vacations, and Christmas bonuses, among other items that are usually given in the first two and last quarters of the year. The analysis will be based on men aged between 20 and 55 years. The upper limit is a conservative bound because according to the Social Security Law<sup>2</sup> prevailing in Mexico during the studied period, workers could retire before they turned 60 years old if they had worked at least 1 250 weeks, among other requirements.<sup>3</sup>

The ENEU-ENOE total sample from 1988 to 2013 comprises 2'123 527 men, from which two main groups can be identified: 1) individuals who worked the week previous to the interview, who represented 84 percent, equivalent to 1'783 692 men; and 2) individuals out of the labour force or unemployed, who represented the remaining 16 percent of the total sample. To determine whether the individuals are regular workers, not seasonally employed, there is a query that asks precisely whether they have worked the whole year. The information from this question reveals that 98.32 percent of the workers were active the whole year; workers who report a positive wage represented 73.6 percent of the total sample. The majority of workers (72.69 percent) are employees, 25.41 percent are self-employed, and the remaining 1.90 percent of the workers does not receive a payment, see table A1 in the appendix.

A common problem in household surveys is unreported wages. Individuals, especially those with high wages, intentionally may not report wages or may deliberately report lower-than-actual wages. People with high wages may be cautious because of the high insecurity levels perceived in Mexico in recent years. For example, the rate of extortion per 100 000 inhabitants grew from 1.16 percent to 6.79 percent from 2000 to 2013. This crime level increased about five times more than any other type according to official accusation reports from the police offices of the 32 Mexican states.<sup>4</sup> At the same time, people earning low wages might report even lower wages so they can be considered beneficiaries in a social program. The percentage of workers who report wages is 87.5 percent, while a slightly lower percentage, 84 percent, of the whole-year workers reported wages. It seems a small percentage of workers are not reporting wages: 12.5

---

<sup>2</sup> *Ley del Seguro Social*, IMSS, 1997.

<sup>3</sup> For example, if a worker started his labour career at the age of 20, this person can retire even before turning 50 years old.

<sup>4</sup> [www.secretariadoejecutivo.gob.mx/es/SecretariadoEjecutivo/090820132](http://www.secretariadoejecutivo.gob.mx/es/SecretariadoEjecutivo/090820132).



percent out of the workers and 16.3 percent out of the whole-year workers. The relevant issue is to analyse the trend of the unreported wages. Using quarterly information from the ENOE survey, Campos (2013a) found an increasing proportion of workers with missing wages; in fact, he showed that individuals with the largest proportion of unobserved wages have higher levels of education, such as high school and university.

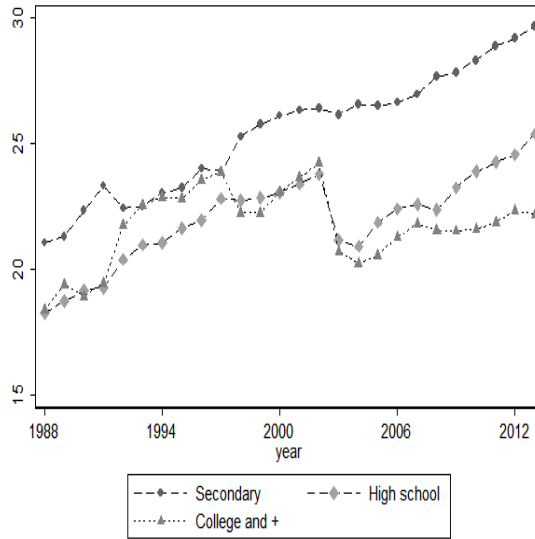
A simple regression between missing wages and schooling years indicates a coefficient of 0.005 and is statistically significant at 99 percent of confidence level; thus there is a small but positive correlation between unreported wages and higher levels of education. This reflects that the sample is still random but there could be a potential endogeneity issue related to the estimation of the returns to education if unreported wages are ignored.

The proportion of individuals acquiring education has been rising. The total men's sample shows that 53 percent of the population holds basic education primary and secondary education. People with upper levels of education -high school and college and more education- represent 44 percent of the total. Education levels follow a sustainable increasing trend over the period, which is consistent with the reduction of the labour force that has no approved schooling years, representing only 3 percent.

Figure 1 shows the percentages of the population with secondary, high school, and college and more levels of education over time; the percentages shown are calculated with respect to each year's observations. The secondary level refers to the population that attained between 7 and 9 schooling years, individuals with high school obtained between 10 and 12 years of education, and individuals with college or more attained at least 13 schooling years.

The figure also shows that the proportion of the population with secondary education increasingly grew and dominated the other education levels over time; in 1988 it was around 21 percent, and it increased to 30 percent in 2013. As shown in the figure, the proportion of the population with college or more education reached the same level as the population with secondary education only during 1991-1997. It can be seen that after 1997 the percentage of individuals with college education or more was reduced and about the same as the high school education percentages. The percentage of the population with college or higher education dropped from 23.87 percent in 1997 to 22.26 percent in 1998, and for the individuals with high school education the reduction was smaller, from 22.81 to 22.76 percent respectively.

**Figure 1**  
*Stock of human capital over time*

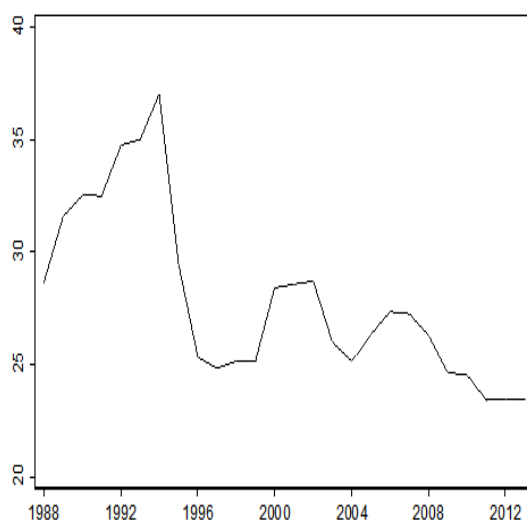


Source: Own calculations from ENEU 1988-2004 and ENOE 2005-2013. Proportions are obtained from the total observations per year.

Even though the evolution of the education achievement has not been linear and the education choice is influenced by the labour market results, there is a clear increase in the education levels over the last 26 years. The lower levels of education dropped by 43 percent; the secondary and high school levels increased by about the same amount, 40 percent; and college and more education rose 20.5 percent from 1988 to 2013. The resulting trend indicates that the Mexican population is increasing its schooling years when entering the labour market. Because data are a repeated cross-section, the evolution provides evidence of the presence of more educated workers, which is relevant to quantify how the returns to education will be affected given this change in the composition of the labour force.

Figure 2 shows real hourly wage evolution through the analysed period of 1988-2013; the largest reduction on the average real hourly wage coincides with the economic crisis 1994-1995, other drops are calculated after 2001 and in 2008. The implementation of NAFTA in 1994 was an important change that reduced wage inequality, as stated by Robertson (2004) wage inequality increased between 1988 and 1994 and after NAFTA the wage inequality fell.

**Figure 2**  
*Average real hourly wage*



Source: Own calculations from ENEU 1988-2004 and ENOE 2005-2013. Note: Average real hourly wage is expressed in Mexican Pesos at constant prices of 2013 and calculated for the objective sample of workers aged 20 and 55 years old who reported a positive wage.

Figure 2 shows an estimated reduction of 18 percent on the average real hourly wage comparing 2013 and 1988. Although the trend captures the evolution of the sample analysed, it can give hints about the association of an increasing labour force better educated and a reduction on real hourly wages.

#### 4. Estimation of returns to schooling

The human capital earnings function proposed by Mincer (1974) has commonly been used to estimate returns to schooling. This model focuses on the life-cycle dynamics of earnings, both observed and potential earning and human capital investments. It relies on assumptions in the functional form, such as linearity in schooling,<sup>5</sup> quadratic specification for experience, and independence in schooling and experience.

<sup>5</sup> See Regan, Oaxaca, and Burghardt (2007) for a detailed discussion regarding the identification of the returns to education coefficient.

Mincer's model approximates an equation that can be estimated linearly, where returns to schooling estimates are the same for any level of education.

The estimation of the returns to education has issues related to the bias present because of the correlation between wages, education, occupation choice, participation in the labour market, and latent unobservable characteristics. The following presents two of the most problematic biases ability and selectivity, and also the strategy to deal with these issues.

#### 4.1. *Ability bias*

In general, the studies that estimate the returns to education have acknowledged the correlation between earnings, education, and occupation choice. Researchers have been cautious with the inferences of potential causal effects because it is complex to identify whether higher income is due to higher education, or whether individuals with higher earnings have chosen to attain more schooling or a certain type of occupation. The causality could be established using instrumental variables; Card (1999) presents a review of papers that have tried to measure the causal effect of education on earnings by using instruments on the supply side of the education system as determinants of education outcomes. Other papers, such as Griliches (1977), Angrist and Krueger (1991), and Ashenfelter and Krueger (1994), have found that returns to schooling using instrumental variables are as big as or even greater than the corresponding OLS estimates, and they claim this as evidence of an ability bias in the OLS estimates.

The empirical strategy is not using the instrumental variables method to deal with the endogeneity caused by the ability bias; because the instrument may also change over time affecting the dynamic comparison of the estimates on the wage distribution (Imbens and Angrist, 1994). Instead it is applying quantile regression which employs information near to the specified section without reliance on any distributional assumptions. Therefore, quantile regression provides an association between wages and schooling that correspond to particular sections of the distribution, offering a more complete view of the relationship between the variables under study. However, the endogeneity issue may be still present.

The specification and interpretation of quantile regression are similar to that of ordinary least squares. On the one hand, quantile estimation minimises the sum of absolute residuals, which can be seen

as an optimal point estimator for asymmetric loss, if the symmetric case is the median, while OLS minimises the squared sum of the residuals. On the other hand, the interpretation of each quantile coefficient of the effect of schooling upon wages depends on the particular section of the distribution considered, while the OLS estimates represent the average effect. Unlike the minimisation problem of OLS, quantile equation is not differentiable, thus quantile estimators cannot be obtained directly. Therefore, nonparametric techniques can be applied to obtain the variance-covariance matrix of the estimation, such as bootstrap to provide standard errors to make inferences regarding the significance of estimated quantile coefficients.

Koenker and Bassett (1978) exposed formally the quantile regression estimation. This approach will estimate the local effect that education has upon wages at any point of the distribution, thus accounting for unobserved heterogeneity existent but not affecting the temporal comparison of returns to schooling (Naticchioni, Ricchi, and Rustichelli, 2010). Unlike the standard approach, OLS estimates the mean regression of the distribution, while quantile regression provides a more complete picture of the returns to schooling because it computes several regressions for different points within the wage distribution. For the case of quantile 50 or median, the symmetry implies that the minimisation of the sum of absolute residuals must equate the number of positive and negative residuals, then assuring that the number of positive and negative residuals is the same as the number of observations above and below the median. Other quantiles are obtained by minimizing asymmetrically weighted absolute residuals. Formally, the quantile estimator is obtained by solving the optimizing problem:

$$\min_b \frac{1}{n} \{ \sum_{i: y_i \geq x'_i b} \theta |y_i - x'_i b| + \sum_{i: y_i < x'_i b} (1 - \theta) |y_i - x'_i b| \} \quad (1)$$

The definition stands for a minimum point for which its density probability given  $x$  is larger or equal to the  $\theta$ -th quantile. The resulting  $\beta(\hat{\theta})$  is the estimator of the quantile  $\theta$ ; it is also called “quantile estimated coefficient”. Quantile coefficients arise from an optimization problem, where sample observations are sorted by a bootstrap procedure. To find the  $\theta$  - th sample quantile requires ranking the observations to split the residuals vector into its positive and negative parts. More specifically, the quantile estimator minimises the sum of absolute residuals to characterise the entire distribution; the median is a type of a Least-Absolute Deviation (LAD) method.

#### 4.2. *Selectivity bias*

The estimation of the returns to education can be biased if only those who report a positive wage are considered. According to Heckman (1979), there is a specification error or omitted variables problem resulting in biased estimated coefficients of the returns to education. The bias arises because individuals may not report a wage either because they are not working or because they choose not to report their wage. Gronau (1974) also established the presence of selectivity bias because of different job search strategies, which in turn affects the participation decision and the distribution of wage offers. Keane and Wolpin (1997) argued that school attendance is a choice that depends on endowments and financing constraints; if individuals have different characteristics then the returns to education may be mixed with these effects. Therefore, self-selection issues may also be present because schooling, work, and occupational choice are correlated.

Let  $w$  define the natural log of the real hourly wage, the entire wage distribution. Define  $w_1$  as missing wages from the surveys, which are unobserved not only because of lack of information regarding the reservation wage but also because some workers decide not to report a wage. Define  $w_2$  as positive wages that workers report in the surveys. Heckman (1979) proposed a parametric method to correct the estimation using a selection equation to account for factors that influence the decision to participate in the labour market, such as number of children, education, age, marital status, partners wage, and seasonal work, among others; these factors are included in the vector of characteristics,  $x_1$ . The coefficient that measures the relationship between the observed attributes and the decision to participate is  $\alpha_0$ . The equation (2) will be referred to as the participation equation:

$$w_1 = x_1' \alpha_0 + v \quad (2)$$

The idiosyncratic error term  $v$  represents unobservable characteristics that affect the decision to participate in the labour market, for example preferences for work, family, schooling, ability, and occupational choice, among others. By assumption,  $E(v|x_1) = 0$ .

The wage equation relates the observed wage with socioeconomic characteristics, such as education, age, marital status, and social security coverage. Equation (3) defines the relationship between observed wage and personal characteristics; the estimated coefficients are in the vector  $\beta_0$ . The wage equation is named also as the outcome equation:

$$w_2 = x_2' \beta_0 + u \quad (3)$$

However, the mean distribution of the disturbances in the equation (3) is not zero,  $E(u|x_2) \neq 0$ , because of the unobserved variables that are playing a role in the wage determination and that are related to education, participation decision, occupational choice, and ability.

The empirical strategy is to account for the ability and selectivity biases. Even though, the ability bias can be still present, quantile regression allows separating the endogenous relationship between wages and education within the wage distribution. Following Buchinsky (1998), equation (3) in the quantile context is:

$$w_2 = x_2' \beta_\theta + u_\theta \quad (4)$$

where  $\theta$  represents the selected quantiles: 10, 25, 50, 75, and 90.

Let  $Q_\theta(w|x_2)$  represent the quantile estimation of the relationship between wages conditional to the characteristics defined in  $x_2$  for each quantile  $\theta$ :

$$Q_\theta(w|x_2) = x_2' \beta_\theta + u_\theta \quad (5)$$

where  $u_\theta \equiv x_2'(\beta_0 - \beta_\theta) + u$ .

Also, the presence of the self-selection or selectivity bias may affect the decision to participate in the labour force, therefore making biased and inconsistent the relationship between wages and education. Because wage is observed only for those who are working or available for those who obtained a wage higher than their reservation wage, therefore:

$$Q_\theta(w|x_2) = x_2' \beta_\theta + Q_\theta(u_\theta|x_2, w_2 > w_1) \quad (6)$$

In the presence of the selectivity bias, the disturbances in the quantile regression form:

$$Q_\theta(u_\theta|x_2, w_2 > w_1) \neq 0 \quad (7)$$

This means that the disturbances causing the selectivity bias in the outcome equation have a relationship with the disturbances in

the participation equation, and therefore both disturbances  $u$  and  $v$  are related. Buchinsky (1998) used a semi-parametric technique to estimate a function known as index  $g$  to relate the disturbances  $u$  and  $v$  with a common error.

The estimation of the outcome wage equation in the quantile context, accounting for the selectivity bias, is:

$$w = x_2' \beta_\theta + h_\theta(g) + \varepsilon_\theta \quad (8)$$

Buchinsky (1998) defined  $h_\theta(g) \equiv Q_\theta(u_\theta | x_1, w_2 > w_1)$ , implying that the relevant variables for the estimation of the index  $g$  depend on the attributes from the reservation wage equation (2), which states the probability to work:

$$Pr(w_2 \geq w_1 | x_1) = Pr(\varepsilon \leq g(x_1, \alpha_0, \beta_0 | x_1)) \quad (9)$$

The empirical strategy to deal with the selectivity bias focuses on how to estimate  $h_\theta(g)$ . Heckman (1979) was the first to suggest a correction using a parametric specification, which assumes a bivariate normal density between the disturbances to obtain the inverse Mill's ratio. The strategy is obtaining a proxy of the unobserved omitted variables to include in the outcome equation. Klein and Spady (1993) have pointed out the risks of parametric misspecification of the density because it can bias even more the estimated coefficients. In fact, Heckman (1979) proposed to estimate the outcome equation in the second step using OLS which, as he claimed, provide unbiased but inefficient estimates.

The estimated coefficients are inconsistent when the assumption of the disturbances distribution is incorrect (Andrews and Schafgans, 1998). Newey (1988), Ahn and Powell (1993), and others used Heckman's approach to develop semiparametric estimators relaxing the normality assumption or any other functional form of error distribution.

Buchinsky (1998) estimated the selectivity bias using a two-step procedure similar to Heckman (1979), Newey and Powell (1990), and Klein and Spady (1993). In the first step,  $\alpha_0$  and  $\beta_0$  are estimated to get the function  $g$ , using the Semiparametric Least Squares (SLS) suggested by Ichimura (1993). In the second step, the estimated bias



is used to obtain a nonparametric correction term, which is approximated to a polynomial  $h(g)$ , to be included in the quantile regression with the observable characteristics defined in  $x_2$ , as it is expressed in the equation (8).

## 5. Empirical estimation

The motivation to study the returns to schooling is because of the rising better educated labour supply in Mexico and a real wage decrease in recent years. However, there are issues related to the estimation of the returns to schooling. Since quantile regression allows separating unobserved characteristics, such as ability, within the conditional distribution of wages and it is a semi-nonparametric estimation, the preferred method to deal with the selectivity bias should be one that does not assume a disturbance distribution *a priori*. The approach to correct for selectivity bias proposed by Heckman (1979) is a parametric estimation that assumes normality on the disturbance distribution. Semi-nonparametric methods are alternative methods to deal with the selectivity bias with more flexibility because they do not assume any particular functional form of the disturbance distribution.

Assuming that endogeneity and selectivity bias do not change over time, the quantile coefficients without any correction are still valid because they will describe the trend over time, even though the levels may be underestimated. In other words, the long-run trend is more important than the level return values for this study. Quantile regression can estimate robust returns to schooling not only for the average but also for every wage percentile in the conditional distribution, making possible the comparison of the inter-quantile estimates over time without making any inference regarding causality, but finding an association between wages and schooling.

The empirical strategy in this paper is estimating the returns to education in two steps, following Buchinsky (1998). In the first step a measure of the selectivity bias is estimated, and in the second step the quantile regression will be estimated after accounting for the selectivity bias obtained from the first step.

In the first step, the selectivity correction will consider a parametric, semi-parametric, and semi-nonparametric correction technique. The parametric correction is estimated using Heckman's (1979) approach to obtain the inverse Mill's ratio. The selection equation in the Heckman model is:

$$worker_{it} = \gamma_t way_{it} + \mu_t child_{it} + \delta_t X_{it} + v_{it} \quad (10)$$

The binary variable  $worker_{it}$  takes the value of 1 if the individual worked last week;  $way_{it}$  is a binary variable that indicates whether an individual worked the whole year. Otherwise, the worker is a seasonal worker. The variable  $child_{it}$  refers to the number of children. A singular fact from these surveys is that the question related to the number of children is asked only to the women in the household. This would create a measurement error in the sample because only men are considered for the estimation of the returns to education; and wife's number of children in the household will be matched to her husband to account for each man's number of children. This issue is controlled by considering information gathered from nuclear families, which are composed of a father, mother, and children; however, the measurement error would be large in households with more than one family, if the children are not correctly attributed to their corresponding fathers. The expression  $X_{it}$  represents the set of variables such as one that indicates whether the person works the whole year, approved schooling years variable, binary variables indicating whether they have or have not a medical service; a polynomial specification on age; and binary variables indicating the marital status. For a summary statistics of these variables for the total sample see table A1 in the appendix.

Following Buchinsky (1998), after the estimation of the Heckman selection equation the inverse Mills ratio is obtained and expanded in a quadratic polynomial; this will be the estimated bias,  $h(g)$ , to be included in the outcome equation.

An alternative method is to estimate the unknown disturbances distribution as a joint bivariate distribution using semi-parametric and semi-nonparametric methods. Following De Luca (2008), it is more efficient to estimate the equations (8) and (10) jointly by maximizing the log-likelihood function. For this purpose, the dependent variables are set to be binary: the selectivity equation would be equal to 1 if the individual is working and 0 otherwise. In the outcome equation, the dependent variable is a binary variable that takes the value of 1 when the individual reports a positive wage and 0 otherwise.

The bivariate binary-choice model estimates the predicted joint probability of the event that an individual reports wages and works:

$$Pr(wage = 1, worker = 1) = \pi(v, u) \quad (11)$$

The differences between semi-parametric and semi-nonparametric techniques are the method to approximate the unknown distribution of the disturbances.<sup>6</sup> Neither approach makes any parametric assumption related to the disturbances distribution, as opposite to the parametric approach. Buchinsky (1998), Martins (2001), and others have used Kernel functions to estimate the unknown disturbance distribution, which is a semi-parametric approach because implies choosing a parametric specification of the index function. If the semi-parametric method to correct selectivity bias assumes the disturbances to follow a bivariate Gaussian distribution, the model is the Seemingly Unrelated Bivariate Probit (Biprobit). The probability will allow estimating an index  $g$ , which in turn will help to approximate the selectivity bias using a quadratic polynomial. Analogous to the Heckman case, the same variables are used to estimate the selectivity equation.

If the Gaussian distribution is relaxed, then the methodology would be a semi-nonparametric correction (De Luca, 2008). The semi-nonparametric approach allows a more flexible functional form of the unknown disturbances distribution, because it uses a joint bivariate density function of the disturbances approximated by a Hermite Polynomial Expansion (HPE). According to Gallant and Nychka (1987), the HPE form is computationally easier than the traditional multivariate normal. In the first stage the unknown joint density between the disturbances, defined in equation (11), is approximated by a Hermite polynomial expansion following De Luca (2008) and De Luca and Perotti (2010), which in turn is based on the work of Gallant and Nychka (1987). The difference between these studies is that De Luca (2008) allows the Hermite polynomial order to differ. The empirical estimation will consider four different specifications of the degree of the HPE. A linear specification in both participation equation (2) and wage reporting equation (3) is denoted as (r1r1); a linear assumption in the participation and a quadratic specification in the wage reporting equation is denoted as (r1r2) and vice versa (r2r1); and quadratic specification in both equations is (r2r2). To calculate the probability in the participation equation only the schooling variable and the binary variables of the medical services are included, otherwise, there is collinearity in the second stage.

In the second step, the returns to education are estimated including the estimated selectivity bias from the first stage as a regressor in

---

<sup>6</sup> A nonparametric approach is not recommended because the coefficients are not identified and because of the “curse of dimensionality” as follows Gerfin (1996).

the quantile regression,  $\hat{h}_t(g)$ , which was obtained from parametric, semi-parametric, and semi-nonparametric techniques. Quantile regression coefficients represent robust measures of the location, which are not sensitive to outlier observations. The quantile partitions of the wage distribution are applied to separate by unobserved ability; this technique allows grouping unobserved heterogeneity among workers, such as ability. Quantile regression is a suitable approach to account for heterogeneity of the returns to education within the conditional distribution, and it can reflect the distribution of unobserved ability and then represent proxies for ability (Arias, Hallock, and Sosa, 2001).

The model follows a simple Mincer's specification of the earnings equation applying OLS and quantile regression considering the information of every male worker ( $i$ ) at any time ( $t$ ) over the period 1988-2013. The outcome equation is defined as:

$$\ln(wage)_{it} = \delta_t + \beta_t educ_{it} + \lambda_t form_{it} + \varphi_t marital_{it} + f(age)_{it} + \hat{h}_t(g) + u_{it} \quad (12)$$

Where  $\ln(wage)_{it}$  is the natural logarithm of the hourly real wage at current prices of 2013. Following Buchinsky (1998), the set of variables  $X_{it}$  included in the selection equation is also included in the outcome equation. In the set of variables are included a continuous variable of attained education,  $educ_{it}$ , where the relevant variable to this study,  $\beta_t$  is the estimated return to education at any year. A dummy variable that can serve as a proxy of the formality,  $form_{it}$ , takes the value of 1 if the individual has medical services provided by the Mexican Social Security Institute (IMSS); 39.86 percent of workers have access to this service. The marital status,  $marital_{it}$ , is a dummy variable that takes the value of 1 if the individual is married and 0 otherwise. From the dataset, tenure or any other variable that provides current experience<sup>7</sup> in the labour force is not available. Thus, to avoid specification error due to potential experience proposed by Mincer (1958), in this study, age will be used as a control variable, where  $f(age)_{it}$  is an age fourth-order<sup>8</sup> polynomial:

---

<sup>7</sup> It is a function of age, education, and a constant.

<sup>8</sup> Murphy and Welch (1990) argued that Mincers estimation could be improved by adding a fourth-order polynomial.

$$f(age)_{it} = \alpha_{1t}age_{it} + \alpha_{2t}(age_{it})^2 + \alpha_{3t}(age_{it})^3 + \alpha_{4t}(age_{it})^4 \quad (13)$$

Regan and Oaxaca (2009) proposed a measure of experience based on the actual hours of work over several years using panel data information.<sup>9</sup> Card (1999) mentioned that the estimates could be lower in comparison with those with the experience variable explicitly included. The estimated selectivity bias obtained from step one is represented by  $\hat{h}_\theta(g)$ , which can be obtained by parametric, semi-parametric, and semi-nonparametric correction methods. And finally,  $u_{it}$  is the error term.

## 6. Results

The dataset allows capturing the labour market conditions over time. It is a repeated cross section data set, which provides the information of different workers at different points in time. The estimation will quantify the effect of education for men over the period 1988-2013. Quantile estimation provides a closer look at the overall wage distribution; results are calculated for each year and  $\theta$  quantile 10th, 25th, 50th, 75th, and 90th. The results are graphically presented because it is easier to compare trends on the different quantile returns to education coefficients including different estimated selectivity bias from parametric, semi-parametric, and semi-nonparametric correction methods. The standard errors are obtained from a variance-covariance matrix via Bootstrap method which also includes quantile blocks in-between allowing coefficient comparisons of different quantiles.

The results are shown in two graphs. Figure 3 shows the schooling coefficient including the selectivity bias using a parametric (Heckman) and a semi-parametric (Biprobit) correction. Figure 4 shows the schooling coefficients with selectivity bias using different specifications of the Hermite Polynomial Expansion (HPE) approximations of the semi-nonparametric (SNP) technique.

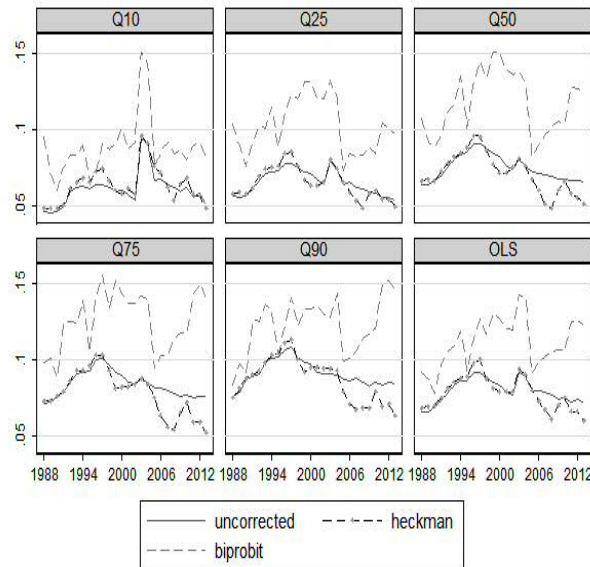
There are 156 quantile returns to education coefficients estimated for each of the selectivity bias correction specifications: five coefficients for each quantile over the period 1988–2013 and OLS coefficients

---

<sup>9</sup> Unfortunately, the data available are not a panel.

for each year during the same period of time. Quantile regression is a location measure of the effect of education on wages. In general, the estimated quantile coefficients are not crossing among them. Tests of the quantiles crossing are rejected for quantiles 50, 75, and 90 in any year under any of the selectivity bias correction model. However, for the lowest quantiles (10 and 25) for some years, the null hypothesis of quantile crossing is not rejected.

**Figure 3**  
*Evolution of the return to education:  
parametric and semi-parametric*



Source: Own calculations from ENEU 1988-2000, ENE 2000-2005, and ENOE 2005-2013. The estimated coefficients are shown in the appendix, table A2.

The estimates of the uncorrected model<sup>10</sup> of the lowest quantile (Q10) are in the range of 4.6 percent in 1988, 6.4 percent in 1997, and 5.1 percent in 2013. The Heckman and Biprobit selectivity correction

<sup>10</sup> Standard deviations are estimated to be small; therefore the estimated coefficients are highly significant.

models provided higher estimates than the uncorrected specification. The Heckman models estimates are in the range of 0.048 and 0.113, and Biprobit estimates are larger in the range of 0.06 and 0.146.

Figure 3 shows the same quantile specification defined in equation (12); the difference between the estimates is the inclusion of the selectivity bias correction by Heckman and Biprobit methods. The selectivity correction specification makes the returns to education coefficient vary. Parametric and semi-parametric imply assuming either a normal disturbance distribution or a particular distribution of the index  $g$ . These assumptions can lead to a misspecification of the true model. The figure shows that the semi-parametric Biprobit method had abrupt jumps compared to the uncorrected returns to education and the parametric Heckman correction. In fact, the parametric selectivity bias correction approximates to the uncorrected estimates, although at the end of the period the Heckman method implies lower returns to education.

The currency crisis of 1994–1995 reduced wages dramatically; this also reduced the returns to education for the lowest quantile in 1995. For the upper quantiles, the effect of the crisis is captured only by the Biprobit model. However, later estimated drops in the returns to education do not coincide with the crisis of 2001 and 2008. The uncorrected and Heckman corrected returns to education are notably higher in the first part of the analysed period than at the end.

Figure 4 presents the quantile coefficients using semi-nonparametric selectivity correction with four different specification of the HPE order. The joint probability of reporting wage and participation in the labour market is obtained in the first stage. The semi-nonparametric selectivity correction models are around the uncorrected estimates. They appeared to be different shifts of the uncorrected model because even though they differ in magnitude they follow the same trend. As the Hermite Polynomial Expansion increases, the estimated coefficients of the returns to education tend to be lower. The linear specification (r1r1) provided the highest estimates in the range of 0.32 to 0.155. The linear and quadratic combined specifications (r1r2 and r2r1) seemed to be similar in magnitude. And the quadratic specification in both equations provided the smallest coefficients.

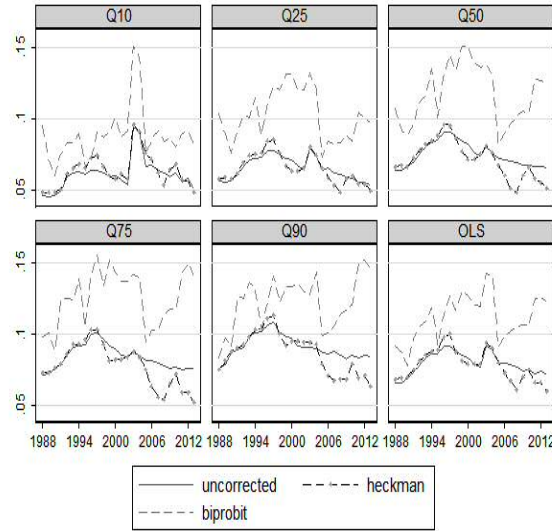
In order to choose which semi-nonparametric specification is preferred, De Luca (2008) and De Luca and Perotti (2010) suggested using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) tests to choose the polynomial degree in the specification of the HPE among the semi-nonparametric specifications.<sup>10</sup>

---

<sup>10</sup> See table A3 for the AIC and BIC for 1988 and 2013; for the rest of the

The AIC and BIC measures reveal that the r1r2 specification provides lower AIC and low BIC. The AIC and BIC measures became lower as the polynomial increases; the quadratic (r2r2) provided the lowest AIC and BIC. However, semi-nonparametric model (r2r2) may suffer from a misspecification because the returns to education estimates tended to zero. This issue is also stated by Gallant and Nychka (1987); even if the semi-nonparametric methods do not make any assumptions regarding the distribution of the disturbances, they do make assumptions regarding the polynomial orders that have to be chosen.

**Figure 4**  
*Evolution of the return to education: semi-nonparametric*



Source: Own calculations from ENEU 1988-2000, ENE 2000-2005, and ENOE 2005-2013. The estimated coefficients are shown in the appendix, table A2.

Results presented in the figures 3 and 4 show a declining trend over time, which is evident for most of the model specifications used. However, it is more appropriate to apply a formal test to determine

---

years the same pattern applies.



whether there is a linear or nonlinear trend of the returns to schooling over time, if any. A simple model is applied to each quantile regressing the estimated coefficients of the returns to education on a time trend for every selectivity bias correction model. The linear coefficients are shown in table A4 in the appendix. In general, the slope coefficients are negative, indicating a linear declining trend, except for the Biprobit selectivity bias specification in which coefficients are mostly positive and statistically significant only for the top quantile. The nonlinear trends (quadratic or cubic) are also shown but only for the semi-nonparametric specification (r1r2) in table A5 in the appendix. Quadratic coefficients were negative and statistically significant for most of the quantiles, indicating a concave trend. The nonlinear trend was rejected for the lowest quantile. The preferred correction estimation is the semi-nonparametric method because of the flexibility of not assuming a disturbance distribution prior the estimation; although there is still a choice in the degree of the polynomial specification in the first step, the problem can be solved by comparing the AIC and BIC measures.

After finding a concave declining trend over time, it would be useful to compare how the within-conditional distribution changed from 1988 to 2013. For comparison purposes, this exercise includes not only the semi-nonparametric specification preferred (r1r2) but the uncorrected estimated of the returns to education and the coefficients using the selectivity correction by Heckman and Biprobit.

The comparisons are based on the ratio of the top quantile estimate, Q90, with respect to lower quantiles, Q75, Q50, Q25, and Q10. This comparison provides a proxy of the inequality variation of the returns to schooling among the quantiles at two points in time, 1988 and 2013. Consistent with empirical evidence for other countries such as Buchinsky (1998), Martins and Pereira (2004), and Machado and Mata (2001), top quantiles obtained larger returns to education than the lower quantiles. Because the results showed that the estimated returns to education had been reduced over time, the question that would be address is how the inequality among the quantiles changed.

Table 1 shows that under no selectivity correction there is an estimated reduction in the ratio of the top and lowest quantile of 0.31 percent, while using a semi-nonparametric model this reduction is 39.91 percent and using a parametric model the inequality reduction is 16 percent; only the semi-parametric model showed an increase in the inequality between the top and lowest quantiles. A similar situation is obtained between the top and Q25, although in the uncorrected model an increase in the ratio is estimated, 16.78 percent.

**Table 1**  
*Inter-quantile ratio changes 1988–2013*

<i>Selectivity bias correction</i>	<i>Q90-Q10</i>	<i>Q90-Q25</i>	<i>Q90-Q50</i>	<i>Q90-Q75</i>
Heckman	-16.00	-2.29	8.71	17.92
Biprobit	106.31	86.78	51.79	22.26
Uncorrected	-0.31	16.78	8.83	3.25
Semi-nonparametric (r1r2)	-39.91	-21.07	-16.25	-8.46

Source: Own calculations from table A1. Note: Q90-Q10 means the ratio of the estimated coefficients Q90 over Q10.

The differences between the ratio coefficients of the upper quantiles, Q75 and Q50, between 1988 and 2013 are in the same direction. There is an increase in the inequality of the returns to education between the two ratios, Q90-Q50 and Q90-Q75; the Biprobit selectivity correction model estimated the largest rise of 51.79 and 22.26 percent respectively. The parametric selectivity correction and the uncorrected coefficients showed an increase from 1988 to 2013 in the range of 8.71 and 8.83 percent in the ratio of the top and median quantiles, Q90 and Q50. Completely different estimates of the inequality measurement are obtained between the top two quantiles.

In general, the preferred semi-nonparametric estimates show a reduction in the ratio of the top quantile relative to the lower quantiles, a difference that increases in absolute value moving downward in the conditional wage distribution. However, this means that the reduction in inequality between quantiles is not only because the education in the lowest quantile is more valued but also the opposite, education in the top quantile is less valued in 2013 than it was in 1988.

## 7. Conclusion

This study found the relationship between wages and education to be strongly positive. The labour market coordinates the supply of workers and the salaries that employers are willing to pay them according to their qualifications and therefore the result should be evident in the

workers wage premium depending on their education level. The Mexican labour market has been changing dramatically in recent decades because of several structural reforms. It is relevant to analyse how the market has valued additional human capital investments made not only at different parts of the wage distribution but also at different points in time.

The paper provided evidence of a declining trend on the quantile estimated coefficients over time. The empirical strategy tried to control for ability and selectivity biases. To deal with the ability bias, quantile regression allowed comparing returns to education in a particular section within wage distribution. Even though, the problem of endogeneity may still be present, the results reveal a positive association between education and wages. It is also found that upper quantiles obtained larger returns to education at any year considered under any model specification. To deal with the selectivity bias, parametric, semi-parametric, and semi-nonparametric corrections were applied. The Biprobit specification provided estimates that are out of the range of the rest of the estimates; the uncorrected and Heckman specification coefficients provided similar returns to education, but at the end, Heckman correction showed an even deeper declining trend. The semi-nonparametric specifications showed that the uncorrected estimates are shifted down as the Hermite Polynomial Expansion increased the order. The preferred method to deal with the selectivity bias is the semi-nonparametric because it is more flexible in the disturbance assumption, although, the order of the polynomial specification has to be still chosen. In this study the AIC and BIC statistics are used to select a linear specification in the worker equation and a quadratic specification in the wage equation, that is, the r1r2 specification.

The return to schooling increased moving from lower to upper quantiles, although the marginal increase across quantiles became smaller. The 1994 currency crisis deeply affected the reduction of wages; the effect is evident for workers in the lowest quantiles but most negative relative to workers in the upper quantiles. The estimated coefficients got smaller by the end of the period analysed. The inter-quantile analysis that allows estimating the inequality of the return to education between the top quantile and each of the lower quantiles, Q10, Q25, Q50, and Q75, reveals a reduction under the semi-nonparametric selectivity bias correction; the uncorrected estimates show a reduction in the gap between top and lowest quantile while the Heckman correction estimates a reduction in the inequality between the top quantile and the lower quantiles, Q10 and Q25. These

results imply a reduction in inequality between quantiles caused because education in the top quantile is less valued in 2013 than it was in 1988. The results obtained do not promote an optimistic view of future educational investment. If returns to education are indicators of the efficiency of education to raise wages, these results may discourage individuals from investing in their own human capital; then it should be a concern for policy makers to promote job creation to fulfil a growing better educated labour force.

## References

- Ahn, H. and J.L. Powell. 1993. Semiparametric estimation of censored selection models with a nonparametric selection mechanism, *Journal of Econometrics*, 58: 3-29.
- Andrews, D. and M.M.A. Schafgans. 1998. Semiparametric estimation of the intercept of a sample selection model, *Review of Economic Studies*, 65: 497-517.
- Angrist, J.D. and A.B. Krueger. 1991. Does compulsory school attendance affect schooling and earnings?, *Quarterly Journal of Economics*, 106: 979-1014.
- Arias, O., K.F. Hallock, and W. Sosa-Escudero. 2001. Individual heterogeneity in the return to schooling: Instrumental variables quantile regression using twins data, *Empirical Economics*, 26: 7-40.
- Ashenfelter, O. and A. Krueger. 1994. Estimates of the economic return to schooling from a new sample of twins, *American Economic Review*, 84(5): 1157-1173.
- Becker, G. 1962. Investment in human capital: A theoretic analysis, *Journal of Political Economy* LXX(5): 9-49.
- . 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, New York: Columbia University Press.
- Ben-Porath, Y. 1967. The production of human capital and the life cycle of earnings, *Journal of Political Economy*, 75: 352-365.
- Buchinsky, M. 1998. Recent advances in quantile regression models: A practical guideline for empirical research, *The Journal of Human Resources*, 33(1): 88-126.
- . 2001. Quantile regression with sample selection: Estimating women's return to education in the US, *Empirical Economics*, 26: 87-113.
- Campos-Vázquez, R.M. 2013a. Efectos de los ingresos no reportados en el nivel y tendencia de la pobreza laboral en México, *Ensayos Revista de Economía*, XXXII(2): 23-54.
- . 2013b. Why did wage inequality decrease in Mexico after NAFTA, *Economía Mexicana*, XXII(2): 245-278.

- Card, D. 1999. The causal effect of education on earnings, in O.C. Ashenfelter, and D. Card (eds.), *Handbook of Labor Economics*, vol. 3, Amsterdam: Elsevier Science.
- De Luca, G. 2008. SNP and SML estimation of univariate and bivariate binary-choice models, *The Stata Journal*, 8(2): 190-220.
- and V. Perotti. 2010. Estimation of ordered response models with sample selection, CEIS Tor Vergara, Research paper series, 8(3), 168.
- Esquivel, G. 2011. The dynamics of income inequality in Mexico since NAFTA, *Economía*, 12(1): 155-178.
- Fersterer, J. and R. Winter-Ebmer. 2003. Are Austrian returns to education falling over time?, *Labour Economics*, 10: 73-89.
- Fitzenberger, B. and C. Kurz. 2003. New insights on earnings trends across skill groups and industries in West Germany, *Empirical Economics*, 28: 479-514.
- Flabbi, L., S. Paternostro, and E.R. Tiongson. 2007. Returns to education in the economic transition: A systematic assessment using comparable data, World Bank Policy Research, Working Paper no. 4225.
- Gallant, R. and D.W. Nychka. 1987. Semi-nonparametric maximum likelihood estimation, *Econometrica*, 55(2): 363-390.
- Gerfin, M. 1996. Parametric and semi-parametric estimation of the binary response model of labour market participation, *Journal of Applied Econometrics*, 11(3): 321-339.
- Griliches, Z. 1977. Estimating the returns to schooling: Some econometric problems, *Econometrica*, 45(1): 1-22.
- Gronau, R. 1974. Wage comparisons. A selectivity bias, *Journal of Political Economy*, 82(6): 1119-1143.
- Heckman, J.J. 1979. Sample selection bias as a specification error, *Econometrica*, 47(1): 153-161.
- Heckman, J.J. and E. Vytlacil. 2000. Structural equations, treatment effects, and econometric policy evaluation, *Econometrica*, 73(3): 669-738.
- Ichimura, H. 1993. Semiparametric least squares (SLS) and weighted SLS estimation of single-index models, *Journal of Econometrics*, 58: 71-120.
- Imbens, G.W. and J.D. Angrist. 1994. Identification and estimation of local average treatment effects, *Econometrica*, 62: 467-476.
- Keane, M.P. and K.I. Wolpin. 1997. The career decisions of young men, *Journal of Political Economy*, 105(3): 473-522.
- Klein, R.W. and R.H. Spady. 1993. An efficient semiparametric estimator for binary response models, *Econometrica*, 61(2): 387-421.
- Koenker, R. and G. Bassett, Jr. 1978. Regression quantiles, *Econometrica*, 46(1): 33-50.
- Lopez-Acevedo, G. 2006. Mexico: Two decades of the evolution of education and inequality, World Bank Policy Research, WP no. 3919.
- Machado, J.A. and J. Mata. 2001. Earnings functions in Portugal 1982-1994: Evidence from quantile regressions, *Empirical Economics*, 26: 115-134.
- Martins, M.O.F. 2001. Parametric and semiparametric estimation of sample selection models: An empirical application to the female labour force in Portugal, *Journal of Applied Econometrics*, 16(1): 23-39.

- Martins, P.S. and P.T. Pereira. 2004. Does education reduce wage inequality? Quantile regression evidence from 16 countries, *Labour Economics*, 11: 355-371.
- Mincer, J. 1958. Investment in human capital and personal income distribution, *Journal of Political Economy*, LXVI(4): 281-302.
- . 1974. *Schooling, experience, and earnings*, National Bureau of Economic Research and Columbia University Press.
- Murphy, K.M. and F. Welch. 1990. Empirical age-earnings profiles, *Journal of Labor Economics*, 8(2): 209-229.
- Naticchioni, P., A. Ricci, and E. Rustichelli. 2010. Far away from a skill-biased change: Falling educational wage premia in Italy, *Applied Economics*, 42(26): 3383-3400.
- Newey, W.K. 1988. Two step series estimation of sample selection models, MIT, WP no. 90-04.
- and J. Powell. 1990. Efficient estimation of linear and type I censored regression models under conditional quantile restriction, *Econometric Theory*, 6(3): 295-317.
- Regan, T.L., R.L. Oaxaca, and G. Burghardt. 2007. A human capital model of the effects of ability and family background on optimal schooling levels, *Economic Inquiry*, 45(4): 721-738.
- and R.L. Oaxaca. 2009. Work experience as a source of specification error in earnings models: Implications for gender wage decompositions, *Journal of Population Economics*, 22: 463-499.
- Robertson, R. 2004. Relative prices and wage inequality: evidence from Mexico, *Journal of International Economics*, 64: 387-409.
- Rodriguez-Oreggia, E. 2005. Institutions, geography and the regional development of returns to schooling in Mexico, Research Institute for Sustainable Development and Social Equality (IIDSES), WP no. 8.
- Schultz, T.W. 1961. Investment in human capital, *American Economic Review*, 51(1): 1-17.
- Spence, A.M. 1973. Job market signaling, *Quarterly Journal of Economics*, 87: 355-374.
- Weiss, A. 1995. Human capital *vs.* signaling explanations of wages, *Journal of Economic Perspectives*, 9(4): 133-154.

## Appendix

**Table A1**  
*Variables summary*

<i>Variables</i>	<i>Observations</i>	<i>Mean value</i>	<i>Standard deviation</i>
Total sample size	2 123 527		
Employees	1 390 947	72.69%	
Self-employed	486 208	25.41%	
Workers without payment	36 313	1.90%	
Worker	1 783 692	84.00%	0.367
Medical service IMSS	744 634	39.86%	0.490
Medical service ISSSTE	122 292	6.55%	0.247
Do not have medical service	1 001 048	53.59%	0.499
Single	614 040	28.91%	0.453
Married or cohabitant	1 443 403	67.95%	0.467
Separated, divorced or widow	66 816	3.15%	0.175
Work all year	1 754 730	98.32%	0.128
Receive a positive wage	1 562 932	73.60%	0.441
Natural log of real hourly wage	1 562 932	3.33	0.764
Schooling years (approved)	2 115 224	9.55	4.254
Age	2 126 413	34.68	10.062
Number of children	2 017 538	3.24	3.001

Source: Own calculations from ENEU 1988-2000, ENE 2000-2005 and ENOE 2005-2013. Note: Natural logarithm of the real hourly wage is in constant Mexican pesos of 2013.

**Table A2**  
*Estimated returns to education 1988-2013*  
*Several model specifications*

<i>Q10</i>							
<i>Year</i>	<i>Heckman</i>	<i>Bi-probit</i>	<i>Uncorrected</i>	<i>snp_r1r1</i>	<i>snp_r1r2</i>	<i>snp_r2r1</i>	<i>snp_r2r2</i>
1988	.048	.095	.046	.04	.041	.039	.034
1989	.048	.07	.045	.047	.042	.042	.032
1990	.048	.06	.046	.051	.045	.047	.034
1991	.05	.075	.049	0.059	.055	.057	.033

**Table A2**  
(continued)

<i>Year</i>	<i>Heckman</i>	<i>Bi-probit</i>	<i>Uncorrected</i>	<i>snp_r1r1</i>	<i>snp_r1r2</i>	<i>snp_r2r1</i>	<i>snp_r2r2</i>
1992	.061	.083	.059	.067	.063	.065	.039
1993	.065	.083	.062	.07	.062	.066	.039
1994	.068	.089	.063	.074	.066	.07	.046
1995	.065	.066	.061	.062	.053	.057	.046
1996	.072	.072	.064	.062	.049	.053	.05
1997	.074	.091	.064	.063	.053	.056	.051
1998	.066	.087	.062	.068	.06	.062	.047
1999	.06	.091	.059	.069	.061	.062	.041
2000	.057	.101	.060	.071	.064	.066	.045
2001	.061	.087	.057	.065	.059	.061	.043
2002	.057	.092	.053	.065	.055	.06	.041
2003	.096	.151	.094	.032	.073	.06	.075
2004	.091	.143	.091	.048	.091	.082	.075
2005	.076	.075	.066	.044	.063	.055	.035
2006	.071	.085	.067	.053	.073	.07	.04
2007	.062	.092	.064	.039	.065	.064	.039
2008	.053	.084	.062	.065	.075	.074	.033
2009	.064	.086	.059	.066	.071	.071	.036
2010	.068	.08	.062	.061	.071	.07	.03
2011	.057	.089	.055	.061	.065	.067	.032
2012	.056	.091	.058	.068	.067	.069	.037
2013	.048	.081	.051	.065	.058	.063	.019
<i>Q25</i>							
1988	.057	.103	.056	.054	.052	.052	.04
1989	.059	.089	.055	.063	.059	.06	.038
1990	.057	.076	.056	.074	.066	.071	.041
1991	.061	.091	.061	.077	.069	.072	.05
1992	.069	.102	.066	.084	.076	.08	.046
1993	.074	.101	.071	.084	.072	.077	.046
1994	.075	.115	.072	.086	.076	.08	.056
1995	.076	.088	.073	.087	.073	.08	.054
1996	.084	.111	.078	.086	.071	.079	.054
1997	.085	.122	.078	.083	.069	.074	.055



**Table A2**  
(*continued*)

<i>Year</i>	<i>Heckman</i>	<i>Bi-probit</i>	<i>Uncorrected</i>	<i>snp_r1r1</i>	<i>snp_r1r2</i>	<i>snp_r2r1</i>	<i>snp_r2r2</i>
1998	.076	.12	.075	.087	.077	.08	.051
1999	.067	.131	.072	.091	.082	.084	.047
2000	.063	.131	.071	.087	.08	.083	.047
2001	.063	.121	.067	.085	.077	.079	.044
2002	.065	.12	.064	.083	.073	.077	.044
2003	.08	.132	.079	.072	.074	.073	.056
2004	.074	.121	.073	.069	.071	.07	.051
2005	.065	.073	.064	.074	.073	.073	.025
2006	.059	.084	.065	.07	.071	.071	.033
2007	.053	.082	.062	.066	.068	.068	.029
2008	.048	.083	.061	.068	.067	.068	.028
2009	.058	.089	.059	.068	.065	.066	.023
2010	.06	.084	.058	.067	.064	.065	.018
2011	.054	.104	.055	.059	.058	.059	.019
2012	.054	.100	.055	.057	.059	.059	.021
2013	.049	.097	.053	.058	.056	.057	.014
<i>Q50</i>							
1988	.066	.107	.064	.073	.067	.068	.053
1989	.067	.092	.064	.081	.077	.076	.051
1990	.066	.088	.066	.100	.09	.097	.054
1991	.072	.096	.069	.099	.084	.093	.061
1992	.078	.112	.076	.101	.094	.099	.058
1993	.082	.116	.080	.100	.085	.092	.056
1994	.084	.135	.083	.106	.094	.099	.063
1995	.088	.101	.085	.113	.096	.107	.065
1996	.096	.131	.091	.112	.093	.102	.064
1997	.095	.145	.091	.108	.091	.098	.066
1998	.086	.134	.088	.109	.099	.102	.061
1999	.077	.151	.084	.109	.100	.100	.059
2000	.071	.151	.082	.109	.102	.104	.053
2001	.071	.14	.076	.106	.095	.098	.053
2002	.075	.136	.074	.103	.091	.097	.051
2003	.081	.138	.08	.091	.079	.081	.052

**Table A2**  
(continued)

<i>Year</i>	<i>Heckman</i>	<i>Bi-probit</i>	<i>Uncorrected</i>	<i>snp_r1r1</i>	<i>snp_r1r2</i>	<i>snp_r2r1</i>	<i>snp_r2r2</i>
2004	.076	.131	.076	.086	.078	.078	.048
2005	.067	.082	.072	.100	.091	.094	.029
2006	.06	.089	.071	.096	.083	.087	.039
2007	.051	.096	.07	.085	.078	.078	.036
2008	.048	.102	.069	.083	.075	.076	.039
2009	.06	.105	.067	.082	.073	.075	.026
2010	.066	.104	.067	.084	.073	.075	.022
2011	.058	.128	.066	.069	.067	.067	.016
2012	.055	.127	.066	.066	.068	.068	.019
2013	.051	.124	.065	.071	.068	.069	.012
<i>Q75</i>							
1988	.073	.098	.071	.089	.084	.087	.068
1989	.073	.101	.072	.094	.09	.09	.059
1990	.076	.088	.076	.118	.117	.119	.07
1991	.079	.124	.079	.109	.100	.104	.07
1992	.086	.125	.085	.108	.104	.107	.069
1993	.093	.124	.091	.121	.109	.117	.07
1994	.093	.139	.092	.124	.114	.119	.077
1995	.096	.105	.093	.136	.124	.133	.08
1996	.103	.142	.100	.132	.121	.126	.078
1997	.103	.156	.101	.125	.112	.118	.084
1998	.093	.134	.096	.125	.12	.121	.079
1999	.081	.152	.092	.123	.119	.118	.072
2000	.082	.144	.09	.127	.124	.127	.068
2001	.082	.137	.085	.122	.114	.117	.067
2002	.084	.137	.084	.122	.111	.117	.065
2003	.088	.142	.087	.114	.093	.098	.061
2004	.084	.14	.085	.103	.09	.092	.059
2005	.075	.094	.081	.123	.114	.116	.045
2006	.063	.103	.081	.128	.109	.115	.052
2007	.056	.103	.08	.115	.100	.100	.052
2008	.054	.113	.078	.107	.089	.092	.053
2009	.065	.118	.076	.104	.083	.088	.032

**Table A2**  
(*continued*)

<i>Year</i>	<i>Heckman</i>	<i>Bi-probit</i>	<i>Uncorrected</i>	<i>snp_r1r1</i>	<i>snp_r1r2</i>	<i>snp_r2r1</i>	<i>snp_r2r2</i>
2010	.072	.118	.077	.113	.088	.094	.029
2011	.059	.143	.075	.084	.076	.077	.023
2012	.059	.15	.076	.086	.078	.079	.021
2013	.052	.141	.076	.092	.078	.084	.013
<i>Q90</i>							
1988	.075	.083	.076	.103	.100	.102	.084
1989	.081	.097	.079	.099	.100	.100	.075
1990	.088	.091	.087	.132	.134	.137	.087
1991	.09	.126	.089	.123	.115	.12	.09
1992	.093	.125	.091	.119	.119	.12	.082
1993	.098	.136	.098	.128	.12	.127	.083
1994	.103	.132	.101	.135	.129	.133	.092
1995	.104	.105	.102	.153	.148	.157	.1
1996	.111	.124	.107	.148	.147	.148	.099
1997	.113	.141	.108	.14	.129	.136	.103
1998	.100	.122	.102	.147	.15	.151	.097
1999	.092	.133	.098	.138	.135	.137	.09
2000	.095	.134	.097	.140	.140	.143	.087
2001	.095	.135	.092	.139	.134	.138	.085
2002	.094	.13	.091	.139	.135	.141	.084
2003	.094	.128	.091	.137	.108	.118	.076
2004	.093	.144	.090	.120	.098	.102	.075
2005	.079	.099	.087	.145	.138	.143	.066
2006	.071	.101	.086	.155	.138	.143	.067
2007	.067	.106	.088	.142	.121	.122	.073
2008	.068	.114	.085	.136	.101	.111	.073
2009	.068	.117	.082	.131	.095	.108	.041
2010	.079	.121	.085	.148	.100	.114	.044
2011	.069	.149	.083	.110	.085	.089	.036
2012	.071	.152	.085	.111	.088	.091	.034
2013	.063	.146	.084	.110	.085	.091	.024
<i>OLS</i>							
1988	.068	.092	.066	.074	.07	.071	.061

**Table A2**  
(*continued*)

<i>Year</i>	<i>Heckman</i>	<i>Bi-probit</i>	<i>Uncorrected</i>	<i>snp_r1r1</i>	<i>snp_r1r2</i>	<i>snp_r2r1</i>	<i>snp_r2r2</i>
1989	.069	.087	.066	.078	.072	.073	.055
1990	.07	.078	.069	.097	.09	.094	.063
1991	.075	.097	.073	.098	.087	.093	.067
1992	.081	.105	.079	.100	.093	.097	.064
1993	.085	.109	.083	.103	.09	.097	.064
1994	.088	.119	.086	.109	.098	.103	.072
1995	.089	.093	.086	.112	.097	.107	.073
1996	.098	.114	.092	.110	.093	.102	.074
1997	.100	.127	.092	.104	.089	.095	.076
1998	.088	.117	.088	.110	.102	.105	.071
1999	.081	.130	.085	.109	.100	.101	.065
2000	.079	.127	.083	.109	.102	.105	.063
2001	.079	.121	.079	.105	.096	.099	.062
2002	.079	.120	.077	.105	.093	.099	.060
2003	.094	.143	.092	.094	.090	.09	.068
2004	.09	.140	.089	.091	.090	.09	.065
2005	.08	.091	.08	.095	.091	.093	.045
2006	.073	.099	.08	.099	.093	.096	.050
2007	.067	.102	.079	.088	.086	.087	.050
2008	.061	.105	.077	.090	.085	.087	.048
2009	.07	.107	.074	.089	.081	.084	.036
2010	.075	.106	.075	.092	.083	.086	.033
2011	.066	.125	.072	.077	.075	.076	.030
2012	.066	.125	.074	.079	.077	.078	.031
2013	.060	.123	.071	.077	.074	.076	.021

Source: Own calculations using ENEU–ENOE 1988–2013 from equations 10–13. The selectivity correction methods are parametric (Heckman), semi-parametric (Biprobit), and semi-nonparametric with different Hermite Polynomial Expansions (r1r1, r1r2, r2r1, and r2r2). The standard errors in the quantile regression are obtained from a variance-covariance matrix via Bootstrap method.

**Table A3**  
*AIC and BIC results*

<i>Year</i>	<i>snp</i>	<i>AIC</i>	<i>BIC</i>
1988	r1r1	35653.44	35711.77
1988	r1r2	34618.41	34685.07
1988	r2r1	35168.38	35235.04
1988	r2r2	27743.24	27826.56
1988	r3r3	27127.67	27252.65
2013	r1r1	96655.93	96719.35
2013	r1r2	94084.41	94156.89
2013	r2r1	96052.26	96124.74
2013	r2r2	90919.30	91009.90
2013	r3r3	90720.93	90856.83

Source: Own estimates from the semi-nonparametric estimation with different HPE specifications.

**Table A4**  
*Linear trend coefficients*

<i>Quantile</i>	<i>Heckman</i>	<i>Bi-probit</i>	<i>Uncorrected</i>	<i>snp_r1r1</i>	<i>snp_r1r2</i>	<i>snp_r2r1</i>	<i>snp_r2r2</i>
Q10	.030	.063	.044	.008	.092	.080	-.017
Q25	-.057***	-.025	-.028	-.062**	-.019	-.036	-.124***
Q50	-.100**	.023	-.033	-.094**	-.068**	-.089**	-.179***
Q75	-.120***	.065	-.029	-.052	-.099	-.101**	-.209***
Q90	-.113***	.100**	-.031	.023	-.116*	-.095	-.221***
OLS	-.053*	.089**	-.002	-.053	-.026	-.040	-.154***

Source: Own calculations from table A2. Note: \*, \*\*, and \*\*\* significance level at 90, 95 and 99 percent respectively.

**Table A5**  
*Non-linear trend coefficients*

<i>Trend</i>	<i>Q10</i>	<i>Q25</i>	<i>Q50</i>	<i>Q75</i>	<i>Q90</i>	<i>OLS</i>
Linear	.0010 (.6602)	.0052*** (.00)	.0079*** (.00)	.0094*** (.0004)	.0104*** (.0069)	.0068*** (.00)
Quadratic	.0001 (.6358)	-.0003*** (.0009)	-.0006*** (.0001)	-.0006*** (.0034)	-.0006*** (.0435)	-.0004*** (.0002)
Cubic	-4.12E-06 (.4008)	4.11E-06*** (.04)	9.70E-06*** (.003)	1.047E-05*** (.0379)	8.73E-06 (.2423)	6.72E-06*** (.0081)
Constant	.0445 (.00)	.0508*** (.00)	.0628*** (.00)	.0773*** (.00)	.0891*** (.00)	.0654*** (.00)
<i>F-test</i>	8.36	40.90	35.95	25.21	16.84	38.43
<i>Adjusted r-squared</i>	.4690	.8272	.8075	.7439	.6552	.8179

Source: Source: Own calculations from table A2. Note: \* \*\*, and \*\*\* significance level at 90, 95 and 99 percent respectively. P-values are shown in parentheses.