Gender Differences in the Returns to Education in Mexico

Cinthya G. Caamal-Olvera[*]

The aim of this paper is to examine the trend over time of the rewards to education granted by the Mexican labour market accounting for gender differences. The empirical strategy is to make time comparisons across quantiles to relate the trend with a possible workers' overeducation effect. The paper provides robust estimates of the returns to education across the conditional wage distribution using quantile regression methodology. The estimated coefficients reveal a robust declining trend stronger for males than for females. The estimates of the returns to education are larger for females than for males. Data show that in 1988 around 17 per cent of the upper educated workers were employed in unskilled jobs; this percentage rose to 44 per cent in 2011. Given this stark increase, it is important to quantify and compare how the labour market is rewarding additional years of education to this rising better-educated labour force over time. Particularly females are an interesting case to analyse because they increased their level of education 22 per cent more than males during the same time period.

JEL Codes: C21, J01, and J16

1. Introduction

The purpose of this paper is to analyse the trend over time of the returns to education in Mexico with a gender differentiation. On average, the gender wage gap has been larger during the last 20 years in the lower part of the unconditional wage distribution and smaller at the top, 16.6 per cent and 10 per cent respectively. It is also important to consider the increase in the schooling years that females have attained because the proportion of women in Mexico with college education increased at least five-fold during this period of time.

Comparisons between males and females are complicated because the conditions to choose whether to participate in the labour market, the schooling years and the occupations pursued present gender differences. Compared to female workers, a higher proportion of male workers with higher education are performing jobs requiring low qualifications. However, there was a smaller proportion of women in 1988 performing unskilled jobs, and by 2011 this proportion was about the same as males; the larger proportion of college educated workers performing unskilled jobs could be related to an overeducation effect. The analysis is not only concentrated on the mean but will consider different parts of the conditional wage distribution over time. This study will provide robust estimations regarding the reward that the labour market gives to males and females across the conditional wage distribution

---

[*] Dr. Cinthya G. Caamal-Olvera, Facultad de Economía, Universidad Autónoma de Nuevo León, México. Email: cinthya.caamallv@uanl.edu.mx

estimated by the quantile regression. This methodology will make it possible to separate individuals by unobserved characteristics, such as ability, accounting for the endogeneity bias.

Results indicate that returns to education for females have been larger than for males. Also interesting is the fact that the declining trend over time is deeper for males than for females. The limitation of the study, however, is the assumption of the constant sample selectivity bias over time.

The paper is organised as follows: section 2 provides the literature related to the returns to education by gender for the international evidence and also for the Mexican case; section 3 presents a descriptive analysis of the data used for the estimation; section 4 explains the methodology used; section 5 presents the results; and section 6 concludes.

2. Literature review

Papers related to the study of the returns to education for males and females have increased during the past decades. The model proposed by Mincer (1958, 1974) is commonly used for the structural equation. However, other studies have raised theoretical concerns related to the identification of the structural equation, per Oaxaca and Burghardt (2007). Three important biases will be taken into consideration. The empirical estimation using the Ordinary Least Squares (OLS) is biased because of the endogenous relationship between education and wages, since it is hard to disentangle whether the more educated workers earn higher wages because they are more able or whether they acquire more education because they are more able. This will create an ability bias in the estimation that may overestimate the returns to education. The second bias is related to the measurement issues related to education and their relationship with the error term; this measurement bias can underestimate the returns to education. Griliches (1977) found that ability and measurement bias compensate each other; Angrist and Krueger (1991), Ashenfelter and Krueger (1994), and Card (1999) found evidence that the ability bias is small.

The third relevant bias when comparing gender coefficients is the selectivity bias, which could be stronger for females than males. The estimations obtained for females that report a positive wage may impose a selection bias to females with a reservation wage larger than the wage offered by the market, according to Heckman (1979). However the direction of the bias is not clear since there is selectivity observed and unobserved. Mulligan and Rubinstein (2008) found that for the United States, the selection of females into the labour market changed sign, from negative in the 1970s to positive in the 1980s. They argued that the reduction in the gender wage gap was the result of the changing selectivity and investment biases rather than females´ wage increases.

Among studies that have corrected for selectivity bias in the context of quantile regression are Nicodemo (2009) and Picchio and Mussida (2010), who found a better fit correcting for selectivity bias studying the gender wage gap in

Mediterranean and Italian families respectively. Bosio (2009) also estimated the quantile gender wage gap; however, after correcting for the endogenous self-selection, he did not find changing patterns with respect to the standard quantile regression. Coelho, Soares and Veszteg (2008) estimated returns to education for Brazil using quantile regression and correcting for selectivity using a probit specification; they found that the selectivity correction overstates the returns to education.

Evidence has shown that returns to education are larger for females than males. Psacharopoulos and Patrinos (2004) presented a comparison of 42 countries over time. They found that the returns to education are diminishing with the development level. The average coefficients are significantly larger for females (9.8 per cent) and for males (8.7 per cent).

In the Mexican context, there is some evidence that females´ returns to education are larger than males. Garro, Gómez and Meléndez (1997) estimated 11.35 per cent for females and 10.75 per cent for males after correcting for sample selectivity; Zamudio (1995) estimated coefficients for workers with college education and obtained coefficients for females of 29.7 per cent and for males 24 per cent. However, other studies have found that males´ returns to education are larger than females'; Bracho and Zamudio (1994) and Urciaga and Almendares (2008) found in the range of 11.86 to 11.6 and 10.3 to 9.85 per cent respectively.

3. Data description

The database used is the National Survey of Urban Employment (ENEU, for its acronym in Spanish) for the period 1988-2004 and the National Survey of Occupations and Employment (ENOE) for 2005-2011. These surveys provide information regarding the Mexican labour market. The basic questionnaire is comparable over time even though there were major changes in 1994 (third quarter) and 2005, when the name changed as well. The design is a quarterly rotating panel; however, to avoid a seasonal component, it is selected only in the third quarter ignoring increases in transitory wages due to Christmas bonuses, utilities, and other.

The strategy is to select nuclear families, which are composed of a mother, father and children. In many families in Mexico, parents or other relatives, such as brothers or sisters of the head of the family, are present. To avoid household compositional effect, nuclear families are selected to keep comparable the family decisions made mainly by the head of the family and the spouse. The sample targets males and females who report either to be heads of families (and their spouses) or to be between 20 and 55 years old. Although the upper limit could be too low, the justification for this choice is that in the Social Security Law workers can retire before turning 60 years old if they have worked 1,250 weeks. The sample excludes students because the objective is to study those who can participate actively on the labour market in the short run.

Data from non-workers implies selectivity bias in the estimation because the reservation wage that non-workers would accept to participate in the labour market is not observed. Table 1 presents a comparison between non-workers and workers in terms of their socio-demographic characteristics. The purpose is to examine differences among the relevant variables to determine whether to participate or not in the labour market, especially for women.

Table 1: Socio-Demographic Characteristics

Characteristics	Females		Males	
	Do not work	Work	Do not work	Work
Head of family	0.08	0.32	0.97	0.98
Age	37.21	37.33	44.14	37.64
Children	2.12	1.87	2.64	2.00
Schooling	7.23	9.23	7.83	9.25
Married	0.95	0.74	0.86	0.94
Total observations	552,196	374,387	23,878	872,378
Gender total	926,583		896,256	

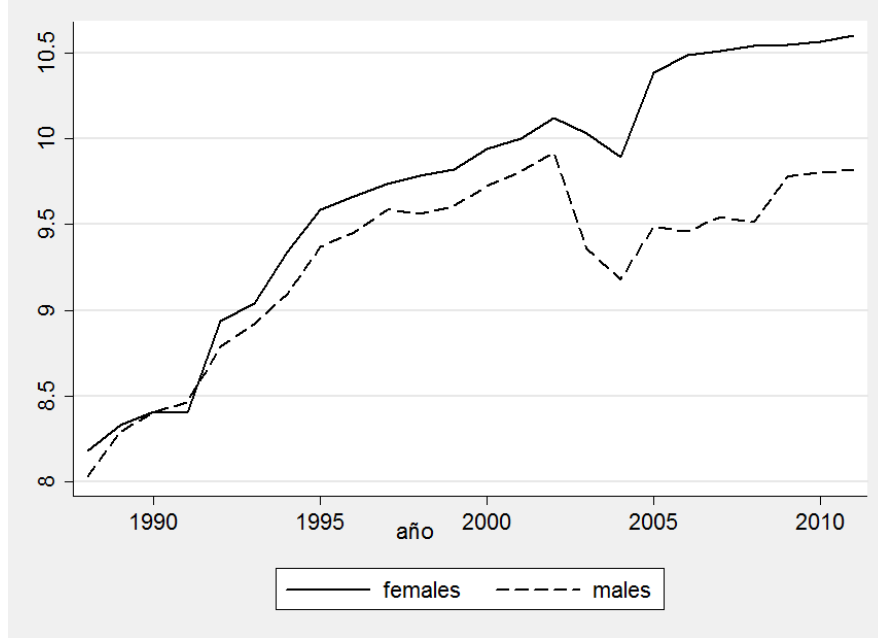Source: Own calculations from ENEU-ENOE 1988-2011.

Data shows that males are mainly heads of the families; females who work are the head of the family only 32 per cent of the time. The majority of females who do not work are spouses of the family heads. Males who do not work are older than any other group, and females who do not work are younger than any other group. The question related to number of children is asked only to females; therefore, to include this information for males, an inference was made directly from the wife in each household; however, if there is no female in the house, this information is unfortunately lost. It is clear that women who work have fewer children than any other group. The average completed schooling is about the same for females and males who work: 9.23 and 9.25 respectively. For those not working, the schooling is substantially lower: 7.23 and 7.83 for females and males respectively.

To have a more stable group and also to have a proxy of the formal sector of the economy, the paper focuses on employees. The sample particularly chooses only employees who have worked at least 20 hours and not more than 72 hours per week and also declared to be employed for the whole year. This definition refers to male and female Full-Time Full-Year (FTFY) workers. Total sample size for females is 280,318 and for males is 618,041. The sample is reduced due to missing values on education, wages and total hours worked per week; it also excludes those who declared to have worked more than 72 hours per week. There are no suppositions on the missing values to avoid any additional bias.

From Table 1, it is evident that on average during the period analysed, the schooling years for female and male workers were alike. It is of interest, however, to analyse the pattern of education over time and not only the average of the period. Figure 1 shows the gender schooling gap evolution. The figure shows that

females' average schooling rose in a larger proportion after 1992, the year of a reform that established nine years of schooling as compulsory basic education.

Figure 1: Evolution of the Average Schooling of FTFY Workers

Source: Own calculations from ENEU-ENOE 1988-2011.

Table 2 shows the proportion of unskilled jobs performed by FTFY workers given their education level. There are two types of jobs defined in this study: skilled and unskilled. The skilled jobs are those that require specific training to perform them. In this group are professionals, technicians, teachers, and directors in the private and public sectors. The unskilled jobs do not require specific training to be performed. Agriculture workers, craftsmen, drivers, administrative assistants, in-store employees, traveling sales people, domestic workers and security workers are considered unskilled. Results show that in 1988, unskilled jobs were mainly performed by low-educated workers, but by the end of the period percentages rose at every level of education (except for females with no approved education).

Table 2: Proportion of Unskilled Jobs

	FEMALES		MALES	
Education	1988	2011	1988	2011
No approved	98.43	97.72	71.73	89.04
Basic	70.70	83.24	62.00	71.17
High-school	12.53	72.23	39.36	58.62
College	3.19	23.4	13.77	25.02

Source: Own calculations from ENEU-ENOE 1988-2011.

Opposite to the predictions of the human capital theory of Becker (1962), it is shown that larger increases in unskilled jobs are estimated for the more educated workers; particularly a larger proportion of college workers are performing jobs with

low-qualification requirements. This effect is stronger for females, and it could indicate that the large increase in female workers with high education level was not assimilated into the labour market. This result implies that having high education does not guarantee holding a skilled job.

4. Methodology

From the previous section, it is evident that as of 2011, a quarter of unskilled jobs in Mexico are performed by workers with college education. The empirical strategy is to estimate the returns to education following a structural equation proposed by Mincer (1958, 1974). In this equation, human capital — in the form of education and experience — is considered the main factor to explain wages:

$$ln(wage)_i = \alpha_t + \beta \, educ_i + \mu \, exp_i + \delta \, exp_i^2 + \varepsilon_i \tag{1}$$

Where $ln(wage)_i$ is the natural logarithm of wage, $educ_i$ is the schooling years and exp_i is the experience. The equation also has a quadratic term on experience, exp_i^2, and a random error term, ε_i.

Equation (1) assumes a linear structure on education, and experience has a quadratic form. However, the experience measurement can be another source of bias because it is a linear combination of age minus education minus six and also ignores exits and entries in the labour market or workers who could not have worked right after school. The following equation will be used to estimate the returns to education for males and females separately:

$$ln(wage)_{it} = \delta_t + g(age)_{it} + \beta_t educ_{it} + \varphi_t marital_{it} + \lambda_t ins_{it} + e_{it} \tag{2}$$

Equation (2) is similar to equation (1), and it adds variables that can exogenously affect the worker's wage, such as marital status, $marital_{it}$. It includes a dummy variable for married and single status, and a variable that can provide a proxy of the formal labour market and participation, ins_{it}, which represents the proportion of workers that has medical insurance provided either by private firms or government. ENEU database does not provide the real experience of workers, although ENOE provides an estimate of tenure. Murphy and Welch (1990) have suggested a quartic specification on age to reduce the measurement bias on the experience. The function $g(age)$ represents a quartic function on age:

$$g(age)_{it} = \alpha_{1t} age_{it} + \alpha_{2t} (age_{it})^2 + \alpha_{3t} (age_{it})^3 + \alpha_{4t} (age_{it})^4 \tag{3}$$

Card (1999) mentions that using age instead of potential experience as a control variable may implicate smaller coefficients.

The empirical strategy is to deal with the endogenous relationship between wages and education to make comparisons over time. Equations (2) and (3) are estimated

applying quantile regression in order to make inter-temporal comparisons in each section of the conditional wage distribution. OLS provides biased coefficients because of the unobserved omitted variables and is the mean estimator. The coefficient of interest is β_t, which represents the return to education for each year; quantile regression will allow having a coefficient β_t for every section of the conditional wage distribution:

$$min_b \frac{1}{n}\left\{\sum_{i:y_i \geq x'_i b} \tau |y_i - x'_i b| + \sum_{i:y_i < x'_i b}(1-\tau)|y_i - x'_i b|\right\} \qquad (4)$$

Following Koenker and Bassett (1978) the quantile regression will allow identifying the effect of education on wage on each section of the wage distribution controlling for observable variables that are relevant in the wage determination. The quantile coefficients can be a measure to identify unobserved ability, per Arias, Hallock and Sosa-Escudero (2001). The interpretation of the quantile regression is similar to OLS, but the technique to estimate the standard errors proceeds from a non-parametric process, bootstrap. Even so, the coefficient comparison among quantile coefficients will provide a wider picture of the distribution of the wages for males and females over time.

5. Results and Findings

The dataset is composed of full-time full-year workers because it is a way to represent the formal labour market in Mexico. In the informal economy, education may not be valued in the same way as in the formal economy; therefore, even if it is interesting to analyse the informal economy because it is a large sector, doing so is outside the scope of this paper. Wages are calculated as the natural logarithm of the real hourly wages with 2002 as the base year. Education is measured as schooling years completed, so the variable is continuous. From previous sections, it was shown that females had larger increases in their education.

To have a wider picture of the trend over time, and given space limitations, the results of the OLS and quantile regressions for males and females are shown in Figure 3. There are six returns to education coefficients for each year (OLS, Q10, Q25, Q50, Q75 and Q90), the time period is 24 years (1988-2011) and estimations are for males and females. The total estimated coefficients are 288.

Figure 3 shows four interesting facts. First, the returns to females´ education are higher than males´ in any part of the wage distribution; this result is consistent with other papers that found the same gender effect, such as Bracho and Zamudio (1994) and Garro, Gomez and Melendez (1997). Second, the concave trend over time is also evident for both males and females; it is also shown that during the first part of the period, 1988-1997, the returns to education are increasing and after 1997 the trend is declining, which is consistent with the results of Robertson (2004) and Hazarika and Otero (2008). Third, the growth rate of the females´ return to education is higher during the first part of the period — before the North America

Free Trade Agreement (NAFTA) — and lower after. It is also evident that drops in the returns to females´ education are not as deep as males´; this result is observed in every part of the wage distribution. And fourth, there is a gender returns to education gap with a structural change in 1997, because in that year the returns are about the same magnitude.

Figure 3: Returns to Education Coefficients over Time
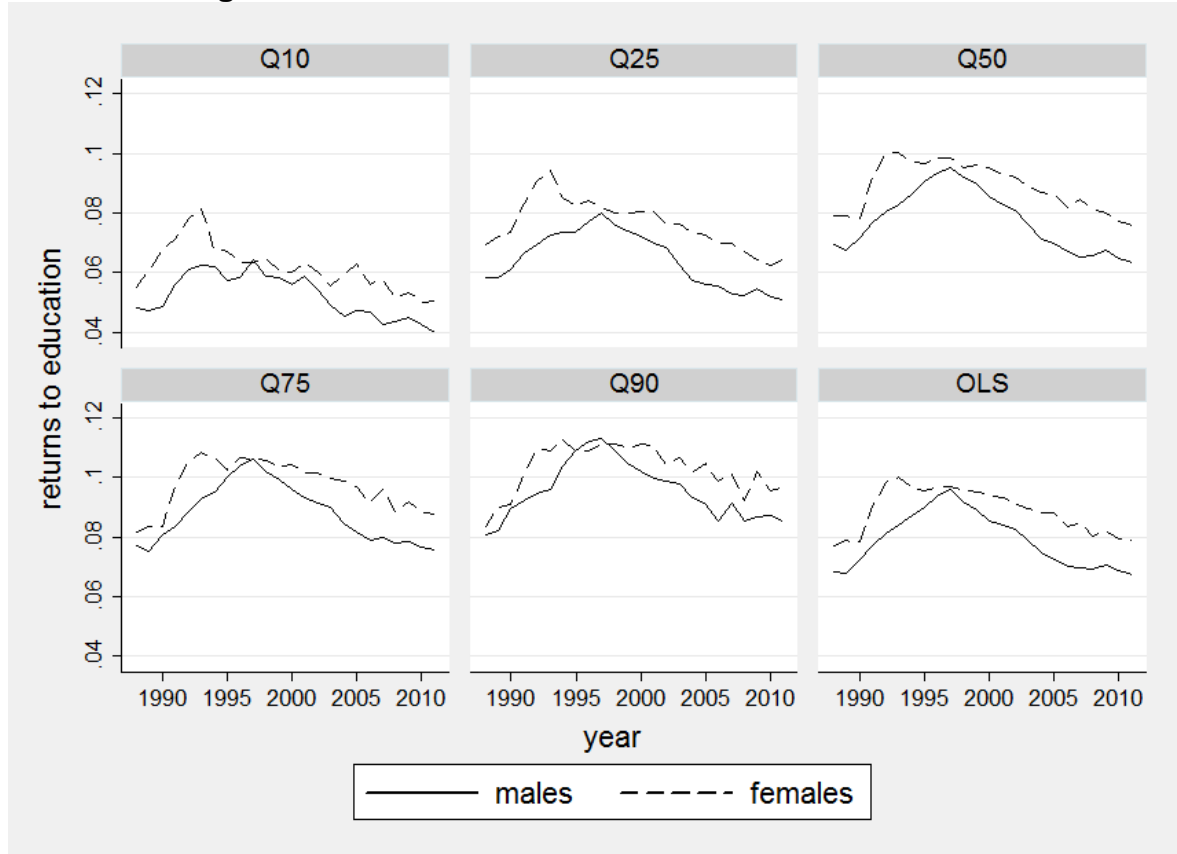
Source: Own estimations from equations (2) and (3) from ENEU-ENOE 1988-2011.
Coefficients are significant at 95% confidence level.

The financial crisis that took place in 1994 reduced wages deeply; however, the effect on the returns to education is evident only in the lowest quantile, Q10. Trade liberalisation could compensate for the negative effect on wages for workers with more schooling, those in the upper quantiles. The males´ coefficients of the lowest quantile, Q10, from 1988-2011 are around 4.8 per cent to 3.98 per cent respectively, and the peak was reached in 1997 at 6.44 per cent. For females, the coefficients during the same period were 5.5 per cent to 5.05 per cent, and the maximum was 8.1 per cent also estimated for 1997. Females´ coefficients are evidently larger than males´ returns to education coefficients.

A similar trend is estimated over the rest of the quantiles. The males´ returns to education are lower than the females´ returns for every quantile in any year. In the median quantile, Q50, the return to education estimated was 6.92 per cent for

males and 7.89 per cent for females in 1988. The females´ maximum value was 10.04 per cent, which was estimated in 1993 — a year before NAFTA took place; on the contrary, for males the peak was reached after NAFTA in 1997, and the value was 9.52 per cent — significantly smaller than females´ return. At the end of the analysis period, there is an evident downgrading in the value of education assigned by the market since it was 6.32 and 7.59 per cent for males and females respectively.

In the top quantile, the same pattern is observed. Females obtained larger returns to education than males, and the reduction of the returns is smaller for females than males. The estimated coefficients for this quantile are larger than for any other quantile, consistent with the fact that more able individuals will obtain higher education and higher wages, so then the unobserved ability is taken into account. The females´ returns to education were in the range of 8.34 to 9.65 per cent over the period 1988-2011; the maximum value 11.24 per cent was reached in 1994, and several other years saw values larger than 11 per cent as well: 1997, 1998, 2000 and 2001. The maximum males' returns to education was reached in 1997, and only in this quantile was males' returns to education larger than females, 11.31 per cent. The estimated coefficient in 1988 was 8.09 per cent, and different from the other quantiles, the lowest was not estimated in 2011; however, the trend over time is declining, as is the case with the rest of the quantiles.

The inter-quantile analysis over time is relevant to compare the coefficients and observe how the reductions differ within the conditional wage distribution; see Table 3.

Table 3: Inter-Temporal Changes 1988-2011

Quantile	Males	Females
Q10	-17.04	-8.30
Q25	-12.39	-7.67
Q50	-8.80	-3.86
Q75	-1.57	7.48
Q90	5.31	15.74
OLS	-1.32	2.33

Source: Own estimations from equations (2) and (3).

It is evident that in the lower quantiles there are deeper and negative changes compared to upper quantiles, and in absolute value the negative drops are smaller for females than males. The positive increase in the returns to education is observed for males only on the top quantile, while for females it is evident there is an increase in their returns to education among the top 25 and top 10 per cent of the conditional wage distribution.

The inferences obtained by comparing the OLS coefficients for males and females reveal a differentiated education effect on every part of the wage distribution. OLS males' coefficients indicate that on average the return to education was reduced by

1.32 per cent and for females increased by 2.33 per cent. However, when comparing the median, Q50, the effects are negative and evidently deeper: -8.8 and -3.86 per cent for males and females respectively. Therefore, the OLS coefficients that provide only the average effect are not enough to analyse the education effect on the wage determination.

Table 3 confirms that education has reduced the effect to enhance wages over time; however, the decline is not as deep for females. In the lowest quantile, the reduction of the females´ returns to education was half of the males´ reduction. In the females´ upper quantiles, Q75 and Q90, the increases were 7.48 and 15.74 per cent respectively. Data have shown that females had increased not only their participation in the labour market but also their education level, so the outcome is that females' education is more valued than males' in the labour market, even though females received lower wages than males by about 6 to 10 per cent, particularly in the lower part of the unconditional wage distribution.

6. Conclusion

The evidence presented in this study showed that females´ returns to education are higher than for males along the conditional wage distribution. The results also showed that returns have been reduced for the bottom part of the distribution for both but in a larger magnitude for males. Female returns to education rose in the top 25 per cent of the conditional wage distribution, while male returns increased only at the top ten for the period 1988-2011. Larger increases in the return to education were estimated for females. However, even though females obtained larger increases in their returns to education, their average real hourly wages are lower by about 6 to 10 per cent than males, and females have on average more schooling years than males. The fact that the returns to education for females do not decline as much as for males implies an interesting phenomenon that deserves to be examined in more detail. A possible explanation for this result could be a slow assimilation of highly educated females mainly in occupations that pay higher; the existence of an overeducation effect is not clear. An extension to this paper is to account for changing selectivity over time in the quantile context.

References

Angrist, JD & Krueger, AB 1991, 'Does compulsory School Attendance Affect Schooling and Earnings?', *Quarterly Journal of Economics*, vol. 106, pp. 979-1014.

Arias, O, Hallock, KF, & Sosa-Escudero, W 2001, 'Individual heterogeneity in the re- turn to schooling: instrumental variables quantile regression using twins data', *Empirical Economics*, vol. 26, pp. 7-40.

Ashenfelter, O, & Krueger, A 1994, 'Estimates of the Economic Return to Schooling from a New Sample of Twins', *American Economic Review*, vol. 84 no. 5, pp. 1157-1173.

Becker, G 1962, 'Investment in Human Capital: A Theoretic Analysis', *Journal of Political Economy*, vol. LXX, no. 5, pp. 9-49.

Bosio, G 2009, 'Temporary employment and wage gap with permanent jobs: evidence from quantile regression', MPRA Paper No. 16055.

Bracho, T & Zamudio, A 1994, 'Economic Returns to education II: Estimations for the Mexican case , 1989'. Working Paper, 31, CIDE. (In Spanish)

Card, D 1999, 'The Causal Effect of Education on Earnings', in O.C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics,* Vol. 3 (pp. 1801-1863), Amsterdam: Elsevier Science, B.V.

Coelho, D, Soares, F & Veszteg, R 2008, 'Quantile Regression with Sample Selection: Estimating Married Women´s Return of Education and Racial Wage Differential in Brazil'. Mimeo, Instituto de Pesquisa Econômica Aplicada, Brazil.

Garro, N, Gómez, M & Meléndez, J 1997, 'Workers´ occupational situation and income level according to their education and occupation', Mexico: Secretariat of Labor and Social Insurance. Research Report 212. (In Spanish)

Griliches, Z 1977, 'Estimating the Returns to Schooling: Some Econometric Problems', *Econometrica*, vol. 45, no. 1, pp. 1-22.

Hazarika, G & Otero, R 2008, 'North-South Trade Liberalization and Returns to Skill in the South: The Case of Mexico', *The Journal of International Trade and Economic Development,* vol. 20 (April), no. 4, pp. 449-465.

Heckman, J 1979, 'Sample Selection Bias as a Specification Error', *Econometrica*, vol. 47, no. 1, pp. 153-162.

Koenker, R & Bassett Jr., G 1978, 'Regression Quantiles', *Econometrica*, vol. 46, no. 1, pp. 33-50.

Mincer, J 1958, 'Investment in Human Capital and Personal Income Distribution', *Journal of Political Economy*, vol. LXVI, no. 4, pp. 281-302.

Mincer, J 1974, 'Schooling, Experience, and Earnings, Technical Report', *National Bureau of Economic Research*, NBER, New York.

Mulligan, CB & Rubinstein, Y 2008, 'Selection, investment, and women´s relative wages over time', *The Quarterly Journal of Economics*, vol. 123, no. 3, pp. 1061-1110.

Murphy, KM & Welch, F 1990, 'Empirical Age-Earnings Profiles', *Journal of Labor Economics*, vol. 8 (April), no. 2, pp. 202-229.

Nicodemo, C 2009, 'Gender Pay Gap and Quantile Regression in European Families', IZA Discussion Paper No. 3978.

Oaxaca, RL & Burghardt, G 2007, 'A Human Capital Model of the Effects of Ability and Family Background on Optimal Schooling Levels', *Economic Inquiry*, vol. 45, no. 4, pp. 721-738.

Picchio, M & Mussida, C 2010, 'Gender Wage Gap: A Semi-Parametric Approach with Sample Selection Correction', IZA Discussion Paper No. 4783.

Psacharopoulos, G & Patrinos, HA 2004, 'Returns to Investment in Education: A Further Update', *Education Economics,* vol. 12, no. 4, pp. 111-134.

Robertson, R 2004, 'Relative prices and wage inequality: evidence from Mexico', *Journal of International Economics*, vol. 64 (December), no. 2, pp. 387-409.

Urciaga García, J & Almendarez Hernandez, MA 2008, 'Wages, education and their private returns in the North-Frontier of Mexico. A Human Capital study', Region and Society, vol. XX, no. 41, pp. 33-56.

Zamudio Carrillo, A 1995, 'Rendimientos a la educación superior en México: Ajuste por sesgo utilizando máxima verosimilitud', *Economía Mexicana, Nueva Época*, vol. IV, no. 1, pp. 69-91.