Economics Bulletin

Volume 45, Issue 1

Women's returns to education in Mexico: a revisit 30 years after

Alfonso Miranda
Applied Economics Division, Center for Economic Research and Teaching (CIDE), Mexico

Maria Benitez-Banegas
Independent consultant, Mexico.

Yahaira Rodriguez-Martinez
Applied Economics Division, Center for Economic Research and Teaching (CIDE), Mexico

Abstract

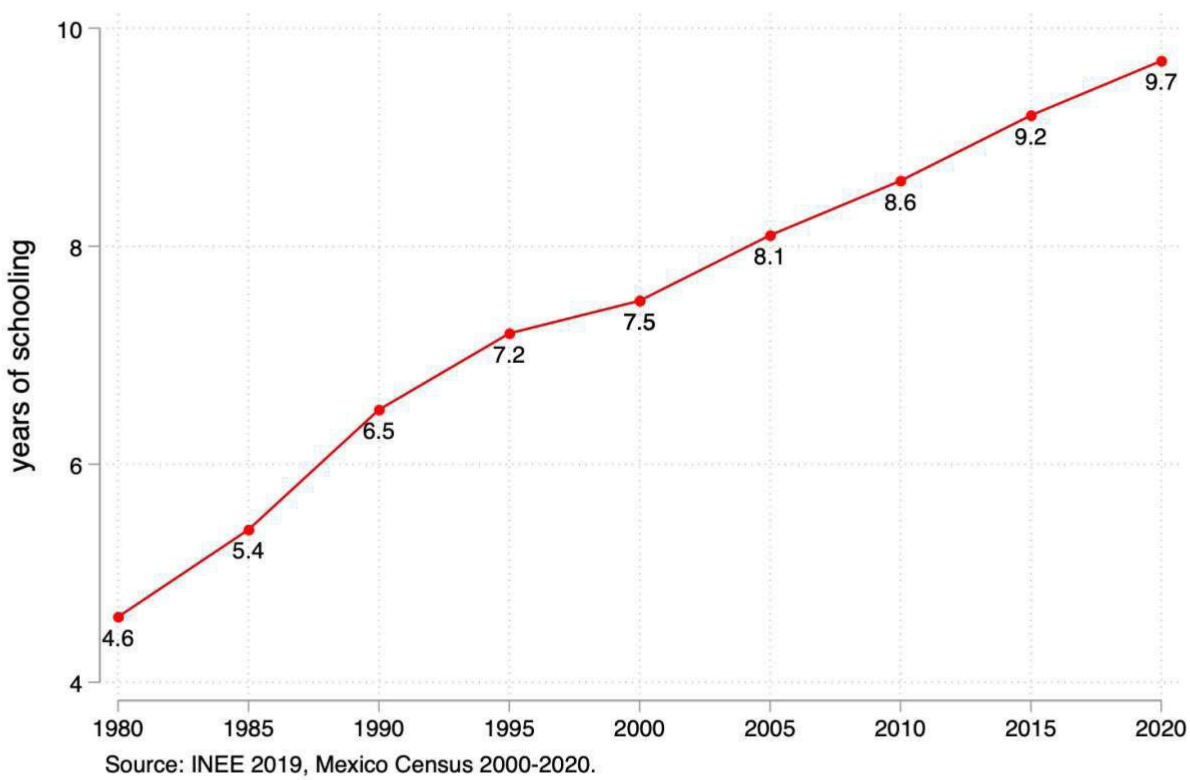
Correcting for sample selection and endogeneity, we estimate the rate of return to education for women in Aguascalientes, Mexico. Our point estimate of 12.9% is similar to figures reported for Mexico 30 years ago, suggesting that investment in education remains as profitable now as it was three decades ago regardless the increment on average education from 6.5 to 9.6 schooling years that occurred since then.

Corresponding author: Alfonso Miranda (alfonso.miranda@cide.edu)(https://orcid.org/0000-0002-8176-0946). Applied Economics Division, Center for Economic Research and Teaching (CIDE). Circuito Tecnopolo Norte #117, Col. Tecnopolo Pocitos II, 20313 Aguascalientes, Mexico. Declarations of interest: None. Use of AI for manuscript preparation: None. Funding: None. Data availability: Data and Stata do-files are available from the authors upon request.
Citation: Alfonso Miranda and Maria Benitez-Banegas and Yahaira Rodriguez-Martinez, (2025) "Women's returns to education in Mexico: a revisit 30 years after", *Economics Bulletin*, Volume 45, Issue 1, pages 11-17
Contact: Alfonso Miranda - alfonso.miranda@cide.edu, Maria Benitez-Banegas - ferma28.banegas@gmail.com, Yahaira Rodriguez-Martinez - yahaira.rodriguez@cide.edu.
Submitted: April 25, 2024. **Published:** March 30, 2025.

1. Introduction

Human capital has been the most dynamic production factor in the last 30 years and the only available form of investment for most people around the world. Demographic changes and technological advances demand a highly skilled workforce. In Mexico, these forces stimulated a large education expansion, taking average schooling from 6.5 to 9.7 years in the last 30 years (INEGI 2000-2020, INEE 2019). Academics, policy makers, and commentators, however, have argued recently that the Mexican economy has been unable to create the jobs to leverage on the supply of this human capital. Limiting, or even reducing, its returns over time (Campos-Vazquez et al. 2016, Levy and López-Calva 2020, Rios 2023). Against this backdrop it is worth asking: How returns to education have changed over time in Mexico? How do they compare with returns gained in high income countries and/or other Latin American countries?

Figure 1. Average years of education in Mexico 1980-2020.



To approach an answer to these questions in the present paper we estimate the returns to education of women in Aguascalientes, Mexico, in 2018. A three-stage estimator proposed by Wooldridge (2010) is used to simultaneously correct for sample selection and endogeneity bias in a Mincerian equation. Despite the large education expansion experienced ever since, we show evidence that returns to education in Mexico are nearly identical nowadays than 30 years ago.

Patrinos and Psacharopoulos (2020) estimate that returns to education in developing countries are around 11% on average, while in developed countries returns are around 8%. Despite substantial discussion over the years, globally, there is no evidence that returns to education have decreased over time even though school attendance has increased (Patrinos and Psacharopoulos 2018). Most studies on returns to education around the 80's and 90's do not correct for the endogeneity due to omitted variables in the Mincer equation, which hinders the analysis (Psacharopoulos and Patrinos 2004). For the Mexican case, returns-to-education studies are limited.

Table 1. Selected estimates of the rate of return to education in Mexico.

Authors	Year	Men	Women
Carnoy	1967	12.0	NA
Bracho and Zamudio	1994	11.9	11.6
Garro, Gómez and Melendez	1997	10.7	11.35
Psacharopoulos and Patrinos	2004	13.2	14.7
Ordaz	2007	10.2	6.4
Urciaga and Almendarez	2008	10.3	9.85
Montenegro and Patrinos	2013	10.53	12.5

As it is seen in Table 1, returns to education in Mexico have not changed significantly since 1967. However, available studies have not simultaneously addressed issues of sample selection and endogeneity of education, probably due to a lack of high-quality data. The use of data with detailed controls and a clean identification strategy to correct for the endogeneity of education as well as potential sample selection bias—caused by individuals self-selecting into education and work—is an important contribution to this research area. It is there where we hope to contribute. We use mother's years of schooling as instrument for education. Similarly, we use the presence of children younger than one at the household as instrument for women's labour market participation. We argue we have good quality instruments and a sound identification strategy.

We use the case of Aguascalientes because of the quality of the data available with the baseline of the Longitudinal Study of the Development of Children in Aguascalientes (EDNA), which are better than existent data for any other state, or the country, on issues regarding returns to education. While performing the study for Aguascalientes could be seen as a limitation, the reality is that Aguascalientes is not particularly atypical for a Mexican state. In fact, in terms of income and education, this state is near the national average. Women in Aguascalientes complete on average 10.3 years of schooling, while the mean for women in Mexico is 9.64 years. For the first trimester in 2020, nationwide, the monthly wage for men was on average $5,220.00 MXN, while that for women was $4,100.00 MXN. On the other hand, for Aguascalientes, the average male wage was $4,800.00 MXN, while the average female wage was $3,900.00 MXN. (Data Mexico 2024).

2. Data

We use the baseline of EDNA, a prospective, multitopic and multidisciplinary study that analyses the main variables that impact the physical, intellectual, and emotional development of children (Miranda et al., 2020). EDNA has data for adult women but not for adult men. The analytical sample is composed by women aged between 21 and 45 who are mothers of the study's children. While working with women who have already entered motherhood may induce selection bias, the National Survey of Demographic Dynamics 2018 shows that 50% of women enter motherhood by age 22 and 87% by age 28 (INEGI 2018). Hence, if present, we do not expect a 'motherhood' selection bias to be large for estimating the conditional mean.[1] On average, mothers in Aguascalientes complete 9.4 years of schooling. While their monthly wage is $4,036.92 MXN with a significant variance in the data.

Table 2. Summary statistics of analytical sample.

Variables	Observations	Media
Women's monthly wage	356	4036.92
Years of schooling	947	9.407
Age	956	31.983
Work experience	843	15.231

Source: EDNA database.

3. Econometric Methods

We use the method suggested by Wooldridge (2010, p. 809) to simultaneously correct for sample selection bias and endogeneity bias of education in a Mincerian equation. The structural system is,

$$y_i^* = \log(wage_i^*) = \beta_1 + \beta_2 \text{exp}_i + \beta_3 \text{exp}_i^2 + \beta_4 \text{IQ}_i + \beta_5 \text{strat}_i + \beta_6 \text{yrsedu}_i + u_i, \quad (1)$$
$$yrsedu_i = \theta_1 + \theta_2 \text{exp}_i + \theta_3 \text{exp}_i^2 + \theta_4 \text{IQ}_i + \theta_5 \text{strat}_i + \theta_6 \text{motheredu}_i + v_i, \quad (2)$$
$$s_i^* = \gamma_1 + \gamma_2 \text{exp}_i + \gamma_3 \text{exp}_i^2 + \gamma_4 \text{IQ}_i + \gamma_4 \text{strat}_i + \gamma_5 \text{underone}_i + \epsilon_i \quad (3)$$
$$s_i = 1(s_i^* > 0), \quad (4)$$
$$y_i = s_i \times y_i^*, \quad (5)$$

where strat_i represents the survey stratum to which the $i-th$ woman belongs, $y_i^* = \log(wage_i^*)$ is her potential monthly log-wage (an unobserved or latent variable), yrsedu_i is her years of education, exp_i is her labour experience (current age minus age at first job), IQ_i is a proxy of her cognitive ability as measured by a 12 item Raven score (see Raven 1936), motheredu_i is her

[1] As a robustness check we fit regressions restricting the analytical sample to women aged 28 and over. Fitting regressions with this restricted sample we find similar results to those reported in table 3. This suggest that if a 'motherhood' selection bias exist, the size of the bias is small and do not pose a major inconsistency problem for our estimators. These results are available from the authors upon request.

mother's years education, and underone_i is a dummy variable indicating whether she has children aged one or less. A woman works when her utility of working s_i^* (a latent variable) is larger than a given threshold $s_i = 1(s_i^* > 0)$; normalised to zero without loos of generality. s_i is the labour participation or selection rule: An always observed binary variable. Wage y_i is observed when $s_i = 1$ and unobserved when $s_i = 0$.

The parameter of interest is women's rate of return to education β_6 (i.e. the causal effect of education on log-wage).

We suspect that $E(u|v) \neq 0$ and $E(u|\epsilon) \neq 0$, while $Cov(v, \epsilon) \neq 0$ can be accommodated without much consequence for estimation. $E(u|v) \neq 0$ makes years of education an endogenous variable in the wage equation (1). This will cause the OLS estimator to suffer from endogeneity bias. On top, ignoring the endogeneity of yrsedu, we cannot really estimate equation (1) because we do not observe y_i^*. What we can estimate is,

$$y_i = \beta_1 + \beta_2 \text{exp}_i + \beta_3 \text{exp}_i^2 + \beta_4 \text{IQ}_i + \beta_5 \text{strat}_i + \beta_6 \text{yrsedu}_i + u_i, \tag{6}$$

on the selected sample $s = 1$. But $E(u|\epsilon) \neq 0$ implies that fitting (6) by OLS delivers an inconsistent estimator because there will be sample selection bias given that $E(u|\boldsymbol{w}, s_i = 1) \neq 0$; $\boldsymbol{w} = (\text{exp}, \text{exp}^2, \text{IQ}, \text{strat}, \text{yrsedu})$. In summary: We have endogeneity bias compounded with sample selection bias.

There are various possible approaches to estimate the structural parameter of interest β_6. One option is to impose distributional assumptions and estimate the whole structural system (1)-(5) by maximum likelihood (ML). Assuming multivariate normality $(u, v, \epsilon) \sim MN(\mathbf{0}, \boldsymbol{\Sigma})$ is a popular choice. Small deviations from multivariate normality, however, can result in a seriously biased ML estimator (Vella 1998). Another option will estimate structural equation (1), while taking a reduced form approach for estimating equations (2) and (3) simply projecting yrsedu and s into the space spanned by the exogenous control variables and instruments in the system. This is the spirit of Wooldridge's (2010, p. 809) three-stage estimator (3SE), which suggest addressing the endogeneity of education using a two-stage least squares estimator in the second stage of a Heckman two-stage sample selection model (Heckman 1979).

This is a three-stage estimator (3SE) with stages: (1) Fit a probit regression of women's labour market participation status on a dummy indicator of the presence of children under one at home and controlling for all exogenous variables in the system—the instrument for selection is the presence of children under one at home—and get the predicted inverse Mills ratio estimator $\hat{\lambda}$; (2) Fit a OLS regression of years of education on schooling of the women's mother—used as instruments for women's education—, and other exogenous variables in the system, to get a predicted value of education $\widehat{edu}$; (3) Fit a OLS regression of log(wages) on predicted education $\widehat{edu}$ from the second stage and the predicted inverse Mills ratio from the first stage, controlling for other exogenous variables in the system. Control variables include experience, women's IQ proxy as measured by a 12 item Raven score (see Raven 1936), and the survey strata. Notice that the second and third stages can be fit by 2SLS of log(wages) on education using years of schooling of the mother as instrument for education and controlling for the inverse Mills ration from the first stage $\hat{\lambda}$ as well as all other exogenous variables in the system.

This is, in fact, a control function approach that does not require full blown multivariate normality. All that is needed is the expected value of u given ϵ be a linear function of ϵ—i.e. $E(u|\epsilon) = \kappa\epsilon$, with $\kappa \in \mathbb{R}$ (Vella 1998). To correct for parameter variation in the first two stages, standard errors are Bootstrapped.

Join modelling s and $yrsedu$ is unnecessary because we take a 'reduced form' approach to model the selection mechanism. To see why, notice that: (1) $\hat{\lambda}_i$ is only a function of the instruments in the system $\boldsymbol{z}$ and hence it is exogenous in (6); (2) conditioning on $\hat{\lambda}_i$ in (6) ensures that the resulting residual $\varepsilon = u - E(u|\boldsymbol{z}, s = 1)$ has, by construction, zero conditional mean $E(\varepsilon|\boldsymbol{z}, s = 1) = E[u - E(u|\boldsymbol{z}, s = 1)|\ \boldsymbol{z}, s = 1] = 0$ in the selected sample. Hence, sample selection bias has been corrected whether $Cov(v, \epsilon) = 0$ or $Cov(v, \epsilon) \neq 0$ is true. All that remains to be dealt with is the endogeneity of education. This is why fitting,

$$y_i = \beta_1 + \beta_2 \text{exp}_i + \beta_3 \text{exp}_i^2 + \beta_4 \text{IQ}_i + \beta_5 \text{strat}_i + \beta_6 \text{yrsedu}_i + \beta_7 \hat{\lambda}_i + \varepsilon_i, \qquad (7)$$

by 2SLS is needed to obtain a consistent estimator. Notice that we have a just identified system, so no overidentification restrictions are available.

Regarding identification, our selected instruments have been used widely in the returns to education literature and we argue they are valid and of good quality. The ability to condition for IQ throughout the system reduces concerns that unobservable cognitive skills could affect at the same time wage, education, and selection into work. Therefore, we believe our study has a sound identification strategy.

4. Results

Results in column 1 from table 3 show that an OLS estimate of the effect of education on women's monthly wage is about 9.6%. Once potential sample selection bias is accounted for by a Heckman model in column 2, the estimate remains similar at 9.4%. There is evidence of sample selection bias in the Heckman model, as a Wald test for $H_0\!: \rho = 0$ is rejected at 1% with a $\chi^2_{(1)} = 19.4$ $(p = 0.00)$. However, the role of sample selectivity is minor as the OLS and Heckman estimates are very similar. In fact, the OLS point estimates falls within the 95% confidence interval of the Heckman point estimate. Accounting for the potential endogeneity of years of education but not for sample selection, a 2SLS estimator gives an estimate of the returns to education of about 12.6%. A Montiel Olea-Pflueger (2013) statistic $\text{F}^{\text{eff}} = 60.5$ rejects the null of a weak instrument at 5%. Hence, our instrument for education (i.e. mother's education) does not suffers from a weak instrument problem. Further, an endogeneity test fails to reject the null that years of education is exogenous at 1%. In the third column we report our Wooldridge 3SE. Standard errors are bootstrapped to account for parameter variation in the 1st and 2nd stage. Once both sample selection and endogeneity of education are corrected for, the estimate of the returns to education goes up to 12.9%. Again, diagnostic tests show that neither the instrument for education nor the instrument for selection suffer from a weak instrument problem. Hence, we are confident that our 3SE estimates are well identified. In line with what we found for 2SLS, a diagnostic test for the endogeneity of years of education fails to reject the null of exogeneity. At this point is important to notice that the OLS point estimate falls well within the 95% confidence interval of the 3SE point estimate. Hence, neither endogeneity nor selection bias are large enough to play a substantial role.

Table 3. Results for log(wage) for females in Aguascalientes, Mexico in 2018. Robust standard errors clustered at school level are reported in parentheses and bootstrapped for the 3SE. Heckman sample selection model is fitted by Maximum Likelihood. Montiel Olea-Pflueger (2013) effective first-stage F statistic reported. Exogenous variables include experience, experience square, women's IQ, and the survey strata. Mother's years of education is used as instrument for women's years of education (eduyrs). The presence of children younger than one is used as instrument for selection (labour marker participation).***p<0.01, ** p<0.05, * p<0.1.

	OLS	**Heckman**	**2SLS**	**3SE**
eduyrs	0.096***	0.094***	0.126***	0.129***
	(0.012)	(0.011)	(0.032)	(0.045)
exp	-0.022	-0.030	-0.020	-0.020
	(0.020)	(0.020)	(0.021)	(0.020)
expsq	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
IQ	-0.001	-0.003	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
IMR				0.047
				(0.297)
Intercept	7.295***	8.339***	6.989***	6.911***
	(0.186)	(0.225)	(0.306)	(0.650)
No. obs. selection		787		713
No. obs. wage	351	333	333	333
No. clusters	126	135	126	126
Diagnostic tests				
$\rho = 0$ test $\chi^2(1)$		19.4		
$\mathrm{P} > \chi^2(1)$		0.000		
Ex. instr. sel. $\chi^2(1)$		10.7		21.2
$\mathrm{P} > \chi^2$		0.001		0.000
Eff. first-stage F			60.5	46.4
Endog. test F			0.8	0.7
$\mathrm{P} > F$			0.359	0.418

5. Conclusions

Using Mincerian regressions that correct for endogenous education and potential sample selection bias, we find an estimate of women's returns to education of about 12.9% per schooling year—which implies an increase in monthly income by $404.58 MXN. This estimate is very similar to those reported 30 years ago by Zamudio and Bracho (1994) and Montenegro and Patrinos (2014). Moreover, compared to other regions, we find that Mexican women's investment in education remains competitive as the average return is about 11% in Latin America and about 8% in high-income countries (Patrinos and Psacharopoulos 2020). Overwhelming evidence suggest that OLS estimates do not suffer from substantial endogeneity or selection bias. Our findings are consistent with those of Patrinos and Psacharopoulos (2018), suggesting that there is no evidence of a declining trend in the rate of return to education across the globe in the last three decades despite the fact the supply of human capital has increased substantially since then.

6. References

Bracho, T., Zamudio Carrillo, A. (1994) "Los rendimientos económicos de la escolaridad en México, 1989" *Economía Mexicana Nueva Época* III (2), 345-377.

Campos-Vazquez, R.M., Lopez-Calva, L.F., Lustig, N. (2016) "Declining wages for college-educated workers in Mexico: are younger or older cohorts hurt the most?" *Revista de Economía Mundial* 43, 93-111.

Carnoy, M. (1967) "Earnings and schooling in Mexico" *Economic Development and Cultural Change*, *15*(4), 408-419.

Data México. (2024) *"Aguascalientes - Datos económicos y laborales" SE.* Retrieved March 20th 2024.

Garro, N.; Gómez M. y Meléndez, J. (1997) "Situación ocupacional y niveles de ingreso de los trabajadores en relación con su educación y ocupación." STPS, Reporte de investigación 212.

Heckman, J. J. (1979) "Sample selection bias as a specification error" *Econometrica* 7(1), 153-161.

INEE. (2019). "Panorama educativo de México. Indicadores del Sistema Educativo Nacional 2018. Educación básica y media superior", INEE.

INEGI (2000-2020) *Census of Population and Housing* 2000, 2010, and 2020.

INEGI (2018). *National Survey of Demographic Dynamics* 2018.

Levy, S., López-Calva, LF. (2020) "Persistent misallocation and the returns to education in Mexico" *The World Bank Economic Review*, *34*(2), 284-311.

Miranda, A., Osiel, D.G., Aguilar-Rodriguez, A., Antonio, A.A., Zizumbo-Colunga, D., Rodríguez-Martínez, Y., Sainz-Santamaría, J. (2020) "The Aguascalientes Longitudinal Study of Child Development: Baseline and first results" *Longitudinal and Life Course Studies* 11(3), 409-423.

Montenegro, C. E., Patrinos, H. A. (2013) "Returns to Schooling around the World" Background paper for the world development report, 8258024-132095074719.

Olea, J. L.M., Pflueger, C. (2013) "A robust test for weak instruments" *Journal of Business & Economic Statistics*, *31*(3), 358-369.

Ordaz, J. L. (2007) "México: Capital humano e ingresos. *Retornos a la educación, 1994-2005"* CEPAL.

Patrinos, H. A., Psacharopoulos, G. (2020) "Returns to education in developing countries" in *The Economics of Education* by Bradley and C. Green, Eds., Academic Press.

Psacharopoulos, G., Patrinos, H. A. (2004) "Returns to investment in education: A further update" *Education Economics*, *12*(2), 111-134.

Psacharopoulos, G., Patrinos, H. A. (2018) "Returns to investment in education: A decennial review of the global literature" *Education Economics* 26(5): 445–458.

Raven, J. C. "Mental tests used in genetic studies: The performance of related individuals on tests mainly educative and mainly reproductive" MSc diss., University of London, 1936.

Rios, V. "El mito de que la educación nos sacará de pobres" *Milenio*, 03.07.2023.

Urciaga García, J., Almendarez Hernández, M.A. (2008) "Salarios, educación y sus rendimientos privados en la frontera norte de México: Un estudio de capital humano" *Región y sociedad*, 20(41), 33-56.

Vella, F. (1998) "Estimating models with sample selection bias: a survey" *Journal of Human Resources* 33(1),127-169.

Wooldridge, J. M. (2010) *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.