ECONOMICS
& Sociology

Mora, J. J., & Herrera, D. Y., Alvarez, J. F.., & Arroyo, J. S. (2023). Returns to human capital in a developing country: A pseudo-panel approach for Colombia. *Economics and Sociology*, *16*(1), 57-70. doi:10.14254/2071-789X.2023/16-1/4

RETURNS TO HUMAN CAPITAL IN A DEVELOPING COUNTRY: A PSEUDO-PANEL APPROACH FOR COLOMBIA

Jhon J. Mora*
Universidad Icesi, Profesor Titular Departamento de Economía, Cali, Colombia
E-mail: jjmora@icesi.edu.co
ORCID 0000-0002-9974-1874
**Corresponding author*

Diana Y. Herrera
Universidad Icesi, Departamento de Economía. Cali, Colombia
E-mail: diana.herrera1@u.icesi.edu.co
ORCID 0000-0001-5110-0683

Juan F. Álvarez
Universidad Icesi, Departamento de Economía. Cali, Colombia
E-mail: jfalvarec@icloud.com
ORCID 0000-0002-4290-023X

Jose S. Arroyo
Universidad del Valle, Cali, Colombia
E-mail: santiago.arroyo@correounivalle.edu.co
ORCID 0000-0002-2581-6298

ABSTRACT. This article analyzes the recent returns to human capital in Colombia. Using a pseudo-panel approach, the results show a human capital return of 9.7% from 2016 to 2020. Comparisons with previous data show a reduction of approximately 5 p.p. in return on human capital. Our results show the importance of investing in Colombia's human capital due to the positive externalities compared to other forms of investment in Colombia. Also, we find statistically significant differences between men and women; Around 2 percentage points (p.p.). A more effective public policy is needed to correct these disparities, despite Colombia's public policies such as equal pay for equal work or advances in equal parental leave.

Received: January, 2022
1st Revision: August, 2022
Accepted: February, 2023

DOI: 10.14254/2071-789X.2023/16-1/4

***JEL Classification*:** I26, J24, C21, J16.

Keywords: returns to education, human capital, repeated cross-sectional models, selection bias, wages differences by sex.

Introduction

Over 60 years have passed since Mincer (1958) established the relationship between income and education, which calculates the rate of return to education. The results at the global level show how the returns to a year of schooling are positive in most countries, fluctuating around 2%–20% (Montenegro & Patrinos, 2014).

In Colombia, empirical evidence confirms that investing in education is profitable, however, the return has gradually decreased in recent years. According to Psacharopoulos and Patrinos (2018), in the mid-1970s, the country's rate of return to education reached levels close to 13%. Later, in the early 1980s, the rate decreased to 9%, stabilizing until the mid-1990s. After getting maximum returns in 2000 (14%), annual estimates have shown a rate of around 10% per academic year. These results are reasonable considering that average schooling increased from less than four years in the 1970s to approximately 12 years in 2014.

However, the data on the returns to education in Colombia are based on cross-sectional analyses, which allow us to make estimates for a particular year but not obtain conclusions over several years, since it is impossible to correctly estimate the returns over several periods if the mean of the individual heterogeneity of the income equation varies substantially over time. Moreover, the latter leads to inconsistent estimations when all the information is aggregated, since the Mincer equation is estimated from a pooled model.

Thus, this article’s main contribution is estimating the returns to education for Colombia using pseudo-panel data and correcting for selection bias, allowing for consistent estimates over the analyzed time. Our results show that the return on an additional year of education from 2016 to 2020 was 9.7%, which indicates a drop in the returns of almost 4 percentage points (p.p.) compared to 1996–2000 (Mora & Muro, 2014). In turn, we find that women receive lower education returns than men, with a gap of 2.6 p.p.

1. Literature review

There is vast literature analyzing the relationship between human capital and wages (Folloni & Vittadini, 2010; Wößmann, 2003). Most research relates to returns in developed countries (Ashworth et al., 2021; Carneiro & Sokbae, 2011; Gunderson & Oreopolous, 2020; Valleta, 2019), but recently, researchers have considered countries that have exhibited significant economic development, e.g., Asian countries (Fang et. al., 2012; Gao & Smyth, 2015; Ma & Iwasaki, 2021). Manacorda et al. (2010), Popli (2011), and Murakami and Nomura (2020) have contributed significantly to such studies. Finally, Patrinos and Psacharopoulos (2020) found that the average rate of return to human capital is 18.8% among 139 countries.

Like international estimates, the research on income and human capital in Colombia has focused on analyzing wage returns derived from increases in education levels. For example, Tenjo (1992, 1993) investigated the evolution of the returns to education for 1976 and 1989 using the methodology of Mincer equations—with and without selectivity correction—combined with spline models to obtain different estimates of the returns to various levels of education. The results show a decrease in average returns.

Chávez and Arias (2002), using a spline model with selection correction for 1990–1995 and 1999–2000, found that returns were higher for women than men in both periods and that, probably because of the economic recession of 2000, a more considerable drop was observed in the second period. Similarly, Prada (2006) analyzed the returns to education in Colombia and its evolution across 1985–2000, considering a wage equation with linear splines, finding that the returns are progressively increasing by educational level.

Mora (2003) examined the effect of obtaining a diploma on the distribution of income, finding that in the Colombian labor market, alongside the years of education accumulated and academic degrees obtained, characteristics of the educational institution and choice of the area of study are differential factors in quantifying education’s returns. Analogous to which, Hernández (2010) provided evidence that corroborated the increasing positive relationship between the degree level (technical, professional, postgraduate, et al.) and income. However, the higher income of highly qualified employees can be linked with essential professional burnout as it is typical, for instance, for educational workers in the public administration sphere (Kryshtanovych et al., 2022).

Zárate (2003) analyzed the changes in the returns to education and experience at different points of the wage distribution for the Colombian labor market using the semi-parametric quantile regression technique between 1991 and 2000. Its results show dissimilar returns in magnitude and variance and an increase in wage inequality toward the end of the period. Similarly, Castillo et al. (2017) used this approach to analyze the wage returns of youth and adults. Finally, Galvis (2010) analyzed wage differences by gender, comparing the influence of education and experience in each group.

Tenjo et al. (2004), in a comparative study of six Latin American countries, estimated Mincerian income equations with selection correction for men and women from 1980 to 1998. For Colombia, the results show higher human capital returns for women than men, both with and without selectivity correction.

At Colombia’s municipal level, Castellar and Uribe (2004) found that the return to education for the Cali metropolitan area averaged 12.7% from 1988–2000. For Bogotá D.C, Forero and Gamboa (2007) estimated, correcting for selection bias, a rate of return to education of 15.9% and 13.7% for 1997 and 2003, respectively, which equals a 15% reduction in returns between the two.

Pave and Blom (2005) analyzed the returns to formal education for different age groups in 1995–2000, using the internal rate of return (IRR) methodology. The results show that the rate of return to education in Colombia remains profitable at all levels of education up to about 40 years old.

Vargas (2013) used the Mincer equation to calculate Colombia's returns to education, considering the areas (urban or rural) where individuals completed their educational cycle. His results are consistent with the hypothesis of returns to investment in education. Additionally, he found that the returns to education of individuals educated in rural areas and working in urban areas are like those of urban workers educated in the city but higher than those working in rural areas.

Garcia et al. (2009) used Heckman, Lochner, and Todd’s methodology (2006, 2008) to estimate the value of the IRR of higher education in Colombia for 2001–2005. The results showed that the IRR for higher education is between 0.074 and 0.128.

Herrera et al. (2015) examined the returns to education in Colombia’s formal and informal employment. Their results showed that formal workers earned about twice as much return to education than informal workers. Additionally, returns to education increased across the wage distribution for formal workers, while informal workers had no comparable pattern.

Generally, there is a consensus on the positive effect of years of education on wages (Gastón & Tenjo, 1992). However, the conclusions are still under discussion regarding the mechanisms through which education leads to higher wages. The most generalized and well-known effect is explained by the positive impact of education on performance as a source of remuneration (Samoliuk et al., 2021). Arteaga (2018) used the University of The Andes data to estimate the effects of reducing the length of economics and finance majors. Using a difference-in-differences approach, his results argued that human capital is essential in determining wages and rejected a pure signaling-based model.

Barrera and Bayona (2019), using data from the admissions process of a prestigious university applied to a regression discontinuity model, found that an increase in graduate salaries was more from a university signaling effect than higher human capital. These results complement the findings of Saavedra (2009) and MacLeod et al. (2017), who, using similar data, concluded that university reputation is positively correlated with graduate earnings. These findings are important, especially in light of the growing discussion on quality of the higher education (Draskovic et al., 2020).

Fuentes et al. (2020) analyzed the impact of education on the wages of household heads in Colombia in 2019. Their results complement the Mincerian models estimated for the country, where the rate of return to education increases as the level of schooling increases.

Among the research that has used the pseudo-panel approach to estimate the returns to education are Warunsiri and McNown (2010), who, applied to the Thai case, examined these returns for workers born between 1946 and 1967 considering selectivity bias. Their findings confirm the reliability of the pseudo-panel approach, finding an overall return between 14% and 16%. In addition, their results indicate that women and urban workers had higher returns than men and rural workers, respectively. Similarly, Kemelbayeva (2019) found relatively high returns to education in Kazakhstan, between 7% and 13% for 2002–2016, while documenting higher returns for women (between 10% and 13% compared to 8% and 12% for men), despite having lower earnings than men. These patterns are not similar for all developing countries. Particularly, it was confirmed a lesser return for females in Azerbaijan, however, the rule of "to earn more, learn more" is typical for this country too (Ismayilov et al., 2022). On a microeconomic level some disparities in gender differentiation of incomes can be successfully mitigated implementing the social responsibility principles (Oliinyk, 2020).

For Latin America, Sapelli (2009) used cohorts based on information from the Occupancy Surveys of Greater Santiago for 1957–2000 and the National Socio-Economic Characterization Survey for 1990–2006 applied to the Chilean case. His results revealed a much higher level of returns for all levels of education than those obtained through Mincer's methodology, which shows how estimates through cross-sectional studies underestimate the accuracy of returns of the Chilean educational process.

Finally, Sánchez and Núñez (2003) examined changes in household structure, human capital, and returns to education across cohorts for Colombia. The results showed a generalized decline in the returns to education for men across cohorts. In contrast, the returns remained relatively constant for women for cohorts born between the 1910s and 1950s and then fell slightly for the youngest cohorts.

2. Methodological approach

Estimation of the human capital returns using pool regressions suffers an errors-in-variables problem due to changes in individual heterogeneity in the time and inconsistence if selection bias is present. Suppose the following mincer equation (Mincer, 1958,1970),

$$Y_{i(t)} = \beta_1 X_{i(t)} + \rho\, \lambda_{i(t)} + f_i + \epsilon_{i(t)} \qquad (1)$$

Where $Y_{i(t)}$ represents the income, $X_{i(t)}$ are variables that explain the human capital theory (education and potential experience), $\lambda_{i(t)}$ is the selection bias, and f_i is the individual heterogeneity. The "$i(t)$" subscripts indicate that the observations come from representative and independent cross-sections where individuals are only available for a single period.

Deaton (1985) showed the inconsistency (1) when individuals are different along time and discussed the strategy to estimate (1) using cohorts in a pseudo-panel approach. Mora and Muro (2014) showed the strategy when selection bias was present in (1), which we follow.

Using the generalized method of moments corrected (GMMC) system, the moment's equation associated with (1) is,

$$E[(Y_{it} - X_{it}'\beta_1 - Z_{0i}'\delta - \rho\lambda_{ct})h(Z_{0i}, Z_{1it})] = B\beta + b \tag{2}$$

Where Z_0 is a dummy cohort indicators matrix, Z_{1it} are time-varying instrumental variables but exclude Z_0; $h(\cdot)$ is a known function usually a set of time and cohort-time interactions although any other time-varying variable is not discarded; $\beta = (\beta_1'\delta'\rho)'$; B, b depends on the covariance matrix of the measurement errors.

For the selection process, we use a probit to model the selection rule,

$$E[(s_{it} - Z_{1it}'\gamma_t)A_t] = 0 \tag{3}$$

Where s_{it} is the selection process, and A is a cohort-means operator, $(Z_0'Z_0)^{-1}Z_0'$. From below, the cohort expression for the moment's equations (2) and (3) can be expressed as,

$$E[s_{ct} - Z_{1ct}'\gamma_t] = 0; t = 1, \dots, T, c = 1, \dots C. \tag{4}$$

$$E[(\Delta Y_{ct} - \Delta X_{ct}'\beta_1 - \rho\Delta\lambda_{ct})\Delta W_{ct}] = B\beta + b \tag{5}$$

Where $\Delta W_{ct} = (\Delta X_{ct}', \Delta\lambda_{ct})'$. Equation (4) is a system of T cross-section linear regressions. In equation (5), we have used the first differences of the synthetic panel (Deaton, 1985). Substituting $\widehat{\gamma_{ct}}$ in (5), we get,

$$E\left[\left(\Delta Y_{ct} - \Delta X_{ct}'\beta_1 - \rho\Delta\,\widehat{\lambda_{ct}}\right)\Delta X_{ct}\right] = B\beta + b \tag{6}$$

Finally, The GMMC estimator is

$$\hat{\beta} = \left[\sum_{c=1}^{C}(\Delta W_c'\Delta W_c + B')D_c\sum_{c=1}^{C}(\Delta W_c'\Delta W_c + B)\right]^{-1}\left[\sum_{c=1}^{C}(\Delta W_c'\Delta W_c + B')D_c\sum_{c=1}^{C}(\Delta W_c'\Delta Y_c - b)\right] \tag{7}$$

Where $\Delta W_c = (\Delta W_{c2}, \Delta W_{c3}, \dots, \Delta W_{cT})'$, $\Delta Y_c = (\Delta Y_{c2}, \Delta Y_{c3}, \dots, \Delta Y_{cT})'$. The optimal choice of D_c, Hansen (1982), is any consistent estimator of the inverse covariance matrix of $\Delta W_c'\Delta W_c$. The asymptotic distribution of the GMMC estimator, for B, b, ΔW_c known, can be derived using standard assumptions and GMM theory (See Mora & Muro, 2014).

Following Deaton (1985), Newey and McFadden (2005), and Mora and Muro (2014), a convenient expression for an upper bound of the covariance matrix V_β is:

$$V_\beta = [M_{WW} - \Sigma]^{-1}\left[\Sigma_{WW}\left(\sigma_\mu^2 + \sigma_{00} + \theta'\Sigma\theta - 2\sigma'\theta\right) + (\sigma - \Sigma\theta)(\sigma - \Sigma\theta)'\right][M_{WW} - \Sigma]^{-1} + \Pi'\hat{V}\Pi \tag{8}$$

The first additive term in equation (8) is the covariance matrix for a pseudo-panel data model (Deaton, 1985, p. 118). The second term is the correction matrix (for selectivity bias) required for using in the estimation of the pseudo-panel data model an estimated regressor instead of the "true" regressor in the second step of the two-step-GMMC estimation procedure

(Mora & Muro, 2014). Finally, we correct by bias in the estimates covariance matrix using Newey and McFadden (2005).

Data

The data comes from the Great Integrated Household Survey (GEIH for its acronym is spanish), developed by the National Administrative Department of Statistics. Given that there is no longitudinal survey for Colombia that allows for following the individual over time, a pseudo-panel consisting of a time series of independent and representative cross-sections between 2016 and 2020 was constructed from information from GEIH (Due to a change in methodology in GEIH since 2016, previous periods were omitted). Since the observations are independent cross-sectional data for each period, eight five-year cohorts of individuals aged between 12 and 51 years were defined.

The sample had 60,411 individuals, where 54.67% (33.027) were women and 45.33% (27,384) were men. *Table 1* presents the distribution by cohort. Each cohort had more than 1,700 individuals. The cohort with the youngest individuals (cohort 1) presented an average of 2,210 individuals per year, while the average number of individuals in the oldest cohort (cohort 8) was 1,113.

Table 1. Number of individuals by cohort

Cohort/ Year	2016	2017	2018	2019	2020	Total
1	2,378	2,253	2,094	2,216	2,107	11,048
2	2,342	1,962	1,850	1,909	1,586	9,649
3	2,151	1,668	1,591	1,784	1,447	8,641
4	1,964	1,381	1,460	1,497	1,341	7,643
5	1,895	1,245	1,198	1,294	1,134	6,766
6	1,576	1,072	1,003	1,114	1,049	5,814
7	1,631	898	895	952	911	5,287
8	1,722	982	876	1,021	962	5,563
Total	15,659	11,461	10,967	11,787	10,537	60,411

Source: *own compilation*

Table 2 shows that between 2016 and 2020, Colombia's average years of education of the different cohorts increased by at least 10%. The highest growth was in the first cohorts due to greater access to education and greater intergenerational mobility (Mora & González, 2019). Overall, cohort 3 presented the highest level of human capital with an average of 12.9 years of education.

Table 2. Average years of education by cohort and year

Cohort/ Year	2016	2017	2018	2019	2020	Mean (2016–2020)
1	6.81	7.95	7.95	7.89	7.94	7.71
2	10.68	12.62	12.45	12.47	12.41	12.13
3	11.74	13.12	13.10	13.25	13.37	12.92
4	11.63	12.78	12.75	13.07	12.98	12.65
5	11.18	12.49	12.37	12.81	13.15	12.40
6	10.57	11.89	11.70	12.26	12.38	11.76
7	9.70	11.22	11.01	11.27	11.64	10.97
8	9.29	10.45	10.19	10.30	10.79	10.21

Source: *own compilation*

Concerning salaries, the data indicated a higher real salary level as the cohort progress, with averages of 2,483 COP (0.654 USD) and 6,372 COP (1,679 USD) in cohorts one and eight (1 USD = 3,795 COP). Additionally, there was a marked salary difference in favor of men starting in cohort 5, fluctuating between 382 COP (0.101 USD) and 782 COP (0.206 USD). Appendix *Table A1* shows the average real hourly wage by cohort.

Finally, the distribution of the other variables is in Appendix *Table A2*. For 2016–2020, on average, labor participation was around 56.4%, years of education at 11.2, about 36% of individuals were married, they were head of the households in 26.2%, and number of the other employed individuals in the household was 1.92.

3. Results

To estimate returns to education, we used the standard mincer equation (Mincer, 1958,1970) to extend to pseudo panel notation:

$$lwhr_{i(t),t} = \alpha_{i(t)} + \beta_0' S_{i(t),t} + \beta_1' Exp_{i(t),t} + \beta_2' Exp_{i(t),t}^2 + \lambda Sel_{i(t),t} + \mu_{i(t),t}; \quad t = 1,\dots,T;\ i = 1,\dots,N \tag{9}$$

$$Sel_{i(t),t} = Married_{i(t),t} + Head\_household_{i(t),t} + NempHH_{i(t),t} + \eta_{i(t),t} \tag{10}$$

where $lwhr$ represents the logarithm of the real hourly wage; $S_{i(t),t}$ is the number of years of education; $Exp_{i(t),t}$ is the potential experience ($Age_{i(t),t} - S_{i(t),t} - 6$) and $Exp_{i(t),t}^2$ is the potential experience squared; while $\alpha_{i(t)}$ represents unobserved individual heterogeneity; and $\mu_{i(t),t}$ is the error term. In equation (9), β_0 is the returns to human capital, β_1, and β_2 explain the returns to potential experience, and a positive β_1 and negative β_2 explain the diminishing returns to the potential experience.

λ is the inverse Mills ratio, which is included in the wage equation since only the income of employed people is observed, creating a selection bias. Equations (9) and (10) are estimated using the GMMC for selection bias (equation 7).

In contrast, in the selection equation, $Sel_{i(t),t}$ is a labor participation dummy variable, which takes the value of 1 if the individual participates in the labor market (employee or unemployed) and 0 otherwise. The following were used as covariables of the selection process: $Married_{i(t),t}$, is a dummy variable that takes the value of 1 if the individual is married and 0 otherwise; $Head\_household_{i(t),t}$, is a dummy variable equaling 1 if the individual is the household head and 0 otherwise; and $NempHH_{i(t),t}$ is the number of the other employed individuals in the household.

First, estimates were made for each year, and the results were *(Table 3):*

Table 3. Mincer equation in Colombia 2016–2020

	HCS2016	HCS2017	HCS2018	HCS2019	HCS2020	HCS2016-2020
S	0.113201***	0.095818***	0.091013***	0.096734***	0.097726***	0.1034022***
	(0.00198974)	(0.00164527)	(0.00165147)	(0.00162308)	(0.00175911)	(0.0008128)
Potential Experience	0.016843***	0.022397***	0.020254***	0.020656***	0.020623***	0.020819***
	(0.00229982)	(0.00195790)	(0.00199168)	(0.00194012)	(0.00218323)	(0.000982)
Squared Potential Experience	-0.000144*	-0.000236***	-0.000203***	-0.000191***	-0.000179**	-0.0002118***
	(0.00006035)	(0.00005424)	(0.00005456)	(0.00005322)	(0.00005983)	(0.0000266)
			Selection Equation			
Married	0.574500***	0.474825***	0.423370***	0.413993***	0.345781***	0.4497018***
	(0.02591498)	(0.02851280)	(0.02906190)	(0.02797976)	(0.02889838)	(0.0124538)
Head Household	1.599269***	1.475800***	1.509555***	1.471020***	1.394537***	1.492413***
	(0.03437850)	(0.03659469)	(0.03673250)	(0.03496955)	(0.03573460)	(0.0158278)
NempHH	0.533212***	0.543591***	0.508948***	0.497436***	0.487763***	0.5201848***
	(0.01193353)	(0.01350234)	(0.01299619)	(0.01259725)	(0.01365796)	(0.0057286)
Inverse Mills Ratio	-0.179646***	-0.091274***	-0.100527***	-0.086435***	-0.062018***	-0.0989515***
	(0.02156922)	(0.01757718)	(0.01739296)	(0.01743929)	(0.01822654)	(0.0087438)
LL	-18,212.2	-10,564.5	-9,900.9	-10,843.1	-9,381.3	-60,470.1
Lr(Rho)	86.019	29.738	35.652	25.464	11.582	148.556
N	15,659	11,461	10,967	11,787	10,537	60,411

Note: Standard errors in parentheses; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.
Source: *own compilation*

In *Table 3*, all variables are statistically significant. *Table 3* also shows a statistically significant selection bias, and Lr test of Rho rejects the hypothesis of zero correlation between the main equation and the selection equation.

From *Table 3*, the returns to education fell during the first three years of study, going from 11% (2016) to 9.1% (2018), and then increased by one p.p. during the following two years, standing at 9.7% (2020)[1]. Additionally, the returns after an additional year of experience went from 1.6% (2016) to 2% (2020).

Next, we estimated the returns to education using GMMC (equations 7 and 8), and the results were (*Table 4*):

[1] Pool estimation shows a 10.3% of return to education. However, section 3 shows that the pool estimation parameters are inconsistent.

Table 4. Pseudo-panel regression

Variables	2016–2020
S	0.097304***
	(0.0000262)
Potential Experience	0.0094258***
	(0.0000467)
Potential Experience Squared	-0.00000757***
	(0.00000114)
Inverse Mills ratio	-6.299479***
	(0.0006428)
N	60,411

Note: Standard errors in parentheses; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.
Source: *own compilation*

Table 4 shows the results of the pseudo-panel model; thus, the rate of return to human capital for 2016–2020 is 9.7%. However, a decrease is observed when the return to human capital is compared with pooling (1996–2000), which was 10.3%.

This decrease could be explained by the increase in the levels of human capital that the Colombian labor force has experienced in recent decades. Also, during the analyzed period, it is observed that the return to an additional year of experience is 9.4%, and decreasing returns to the experience are observed[2].

Finally, we analyzed differences by sex (*Table 5*). Garcia et al. (2009) found around two p.p. of differences in the returns to human capital.

Table 5. Pseudo-panel regression by sex

Variables	2016–2020	
	Men	Women
S	0.1121699***	0.0905967***
	(0.0000696)	(0.0000423)
Potencial Experience	0.0010935***	0.0176901***
	(0.0001092)	(0.0000848)
Potencial Experience Squared	0.0002242***	-0.0001374***
	(0.00000274)	(0.00000203)
Inverse Mills ratio	-6.117336***	-6.365147***
	(0.0016157)	(0.0011117)
N	33,027	27,384

Note: Standard errors in parentheses; *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.
Source: *own compilation*

Table 5 shows differences in the returns of human capital between men and women of around 2.1 p.p. between 2016 and 2020; these differences are statistically significant[3]. Note that both men and women have experienced a drop in the returns to their education in recent decades[4]. As Tenjo et al. (2017) show, women in the last decade experimented with higher returns to human capital compared with men without selection bias. However, when they're correct, by selectivity bias the results are opposite. Our results also show the same results with

[2] García et al. (2009) found a rate of around 7% using pool data for 2001–2005.
[3] Following Clogg et al. (1995) and Paternoster et al. (1998), the z statistic for the difference between men and women is 2.37, which is statistically significant at 1%.
[4] The estimate by sex for 1996–2000 showed that men received higher returns to education by 2.3 p.p.

two p.p. of difference with selectivity bias. However, the average return to experience for women was 1.7%, while that for men was 1%.

Conclusion

This article provides new evidence on the return to human capital in Colombia using the pseudo-panel methodology and correcting selection bias. Our results indicate that the rate of return to education for 2016–2020 is lower than the rate of return to human capital for 1996–2000. This decrease is explained by the increase in the supply of skilled labor, mainly derived from sustained growth in schooling levels.

In the analysis by sex, the results show statistically significant differences in the returns to human capital, which are 2.6 p.p. for 2016–2020, increasing by 13% from 1996–2000.

Regarding public policy, it is profitable to invest in education because of its positive externalities compared to other forms of investment in Colombia. Therefore, it is necessary to continue promoting Colombia's economic development through more and better human capital. However, the results by sex show that there are significant differences in the potential experience, which is why a more effective public policy is needed to correct these disparities, despite Colombia's public policies such as equal pay for equal work (Law 1482 of 2011) or advances in equal parental leave (Law 2114 of 2021).

Acknowledgement

The authors are thankful to the Internal Grant Agency of Universidad Icesi No.: COL0014387-782 “DETPLMIGRAVEN” for financial support to carry out this research.

References

Arteaga, C. (2018). The effect of human capital on earnings: Evidence from a reform at Colombia's top university. *Journal of Public Economics*, 157, 212–225. doi: 10.1016/j.jpubeco.2017.10.007

Ashworth, J., Hotz, V.J., Maurel, A., & Ransom, T. (2021). Changes across cohorts in wage returns to schooling and early work experiences. *Journal of Labor Economics*, 39(4), 1-80. doi: 10.1086/711851

Barrera, F., & Bayona, H. (2019). Signaling or better human capital: evidence from Colombia. *Economics of Education Review*, 70, 20-34. doi: 10.1016/j.econedurev.2019.02.006

Carneiro, P., & Sokbae L. (2011). Trends in quality-adjusted skill premia in the United States, 1960-2000. *American Economic Review*, 101(6), 2309-49. doi: 10.1257/aer.101.6.2309

Castellar, C., & Uribe, I. (2004). Capital humano y señalización: evidencia para el área metropolitana de Cali 1988-2000. *Revista Sociedad y Economía*, 6, 51-79.

Castillo, C.D., Da Silva, J., & Pérez, M. (2017). Retornos salariales para Colombia: un análisis cuantílico. *Apuntes del CENES*, 36, 211-246. doi: 10.19053/01203053.v36.n63.2017.5830

Chávez, A.H., & Arias, H.Y. (2002). Cálculo de la tasa interna de retorno de la educación en Colombia. *Documento de Trabajo UEC No. 002077*, Universidad Externado de Colombia.

Clogg, C.C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, 100(5), 1261–1293.
doi: 10.1086/230638

Deaton, A. (1985). Panel Data from Time Series of Cross-Sections. *Journal of Econometrics*, 30(1-2), 109-126. doi: 10.1016/0304-4076(85)90134-4

Draskovic, V., Jovovic, R., & Rychlik, J. (2020). Perceptions of the declining quality of higher education in the selected SEE countries. *Journal of International Studies, 13*(4), 286-294. doi: 10.14254/2071-8330.2020/13-4/20

Fang, H., Eggleston, K., Rizzo, J., Rozelle, S., & Zeckhauser, R.J. (2012). The returns to schooling: evidence from the 1986 compulsory education law. *NBER Working Paper No 18189.* doi: 10.3386/w18189

Folloni, G., & Vittadini, G. (2010). Human capital measurement: a survey. *Journal of Economic Surveys*, 24(2), 248-279. doi: 10.1111/j.1467-6419.2009.00614.x

Forero, N., & Gamboa, L.F. (2007). Cambios en los retornos de la educación en Bogotá entre 1997 y 2003. *Lecturas de Economía*, 66, 225-250.

Fuentes, H.J., Contreras, L.E., & González, K. (2020). Educación de la fuerza laboral en Colombia y sus efectos en los salarios. *Revista Venezolana de Gerencia*, 25(4), 364-376.

Galvis, L.A. (2010). Diferenciales salariales por género y región en Colombia: una aproximación con regresión por cuantiles. *Documentos de Trabajo sobre Economía Regional y Urbana No. 131***.**

Gao, W., & Smyth, R. (2015). Education expansion and returns to schooling in urban China, 2001–2010: evidence from three waves of the China Urban Labor Survey. *Journal of the Asia Pacific Economy*, 20(2), 178-201. doi: 10.1080/13547860.2014.970607

García, A.F., Guataqui, J.C., Guerra, J.A., & Maldonado, D. (2009). Beyond the Mincer Equation: The Internal Rate of Return to Higher Education in Colombia. *Education Economics*, 22(3), 328-344. doi: 10.1080/09645292.2011.595579

Gastón, N., & Tenjo, J. (1992). Educational attainment and earnings determination in Colombia. *Economic Development and Cultural Change*, 41, 125-139.

Gunderson, M., & Oreopolous, P. (2020). Returns to education in developed countries. In A. Bradley & C. Green. (Eds.), *The Economics of Education: A Comprehensive Overview* (pp. 39-51). Academic Press. doi: doi:10.1016/B978-0-12-815391-8.00003-3

Hansen, L. P. (1982). Large Sample Properties of Generalized Methods of Moments Estimators. *Econometrica*, 50(4), 1029-1054. doi: 10.2307/1912775

Heckman, J., Lochner, L., & Todd, P. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In E. Hanushek & F. Welch (Eds.), *Handbook of Economics Education* (pp. 307-458). Elsevier.

Heckman, J., Lochner, L., & Todd, P. (2008). Earnings functions and rates of return. *Journal of Human Capital*, 2(1), 1-31. doi: 10.1086/587037

Hernández, G. (2010). ¿Cuán rentable es la educación superior en Colombia? *Lecturas de Economía*, 73, 181-214.

Herrera, P., López, E., & Motellón, E. (2015). Double penalty in returns to education: informality and educational mismatch in the Colombian labor market. *The Journal of Development Studies*, 51(12), 1683-1701. doi: 10.1080/00220388.2015.1041516

Ismayilov, I., Aliyev, K., & Bakirova, N. (2022). Return to education in Azerbaijan. Does gender matter?. *Economics and Sociology, 15*(3), 11-27. doi: 10.14254/2071-789X.2022/15-3/1

Kemelbayeva, S. (2020) Returns to schooling in Kazakhstan: an update using a pseudo-panel approach. *Eurasian Economic Review volume*, 10, 437-487. doi: 10.1007/s40822-020-00148-z

Kryshtanovych, M., Akimova, L., Akimov, O., Parkhomenko-Kutsevil, O., & Omarov, A. (2022). Features of creative burnout among educational workers in public administration system. *Creativity Studies*, *15*(1), 116-129.

Ma, X., & Iwasaki, I. (2021). Return to schooling in China: a large meta-analysis. *Education Economics,* 29(4), 379-410. doi: 10.1080/09645292.2021.1900791

MacLeod, W.B., Riehl, E., Saavedra, J.E., & Urquiola, M. (2017). The big sort: college reputation and labor market outcomes. *American Economic Journal: Applied Economics*, 9(3), 223-261. doi: 10.1257/app.20160126

Manacorda, M., Sánchez-Páramo, C., & Schady, N. (2010). Changes in returns to education in Latin America: the role of demand and supply of skills. *Industrial and Labor Relations Review*, 63(2), 307-326. doi: 10.1177/001979391006300207

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4), 281-302. doi: 10.1086/258055

Mincer, J, (1970). The distribution of labor incomes: A survey. *Journal of Economics Literature*, 8(1). 1-26.

Montenegro, C.E., & Patrinos, H.A. (2014). Comparable estimates of returns to schooling around the world. *World Bank Policy Research Working Paper No. 7020*.

Mora, J.J. (2003). Sheepskin effects and screening in Colombia. *Colombian Economic Journal*, 1, 95-108.

Mora, J.J and González, A.M. (2019). Intergenerational educational progress: an analysis for Colombia. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 29, 143-159.

Mora, J.J., & Muro, J. (2014). Consistent estimation in pseudo panels in the presence of selection bias. *Economics,* 8, 1-25. doi: 10.5018/economics-ejournal.ja.2014-43

Murakami, Y., & Nomura, T. (2020). Expanding higher education and wage inequality in Chile. *Journal of Economic Studies*, 47(4), 877-889. doi: 10.1108/JES-12-2018-0445

Newey, W.K., & McFadden, D. (2005) Large sample estimation and hypothesis testing. In R. Engle & D. McFadden (Eds.), *Handbook of Econometrics* (pp. 2111–2245). North-Holland.
doi: 10.1016/S1573-4412(05)80005-4

Oliinyk, O. O. (2020). Social Responsibility Assessment in the Field of Employment (Case Study of Manufacturing). *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, 3, 131–136.

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the Correct Statistical Test for the Equality of Regression Coefficients. *Criminology*, 36(4), 859-866.
doi: 10.1111/j.1745-9125.1998.tb01268.x

Patrinos, H., & Psacharopoulos, G. (2020). Returns to education in developed countries. In A. Bradley & C. Green (Eds.), *The Economics of Education: A Comprehensive Overview* (pp. 53-64). Academic Press. doi: 10.1016/B978-0-12-815391-8.00004-5

Pave, T., & Blom, A. (2005). Is formal lifelong learning a profitable investment for all of life? How age, education level, and flexibility of provision affect rates of return to adult education in Colombia. *World Bank policy res. work. pap. No. 3800*.

Prada, C.F. (206). ¿Es rentable la decisión de estudiar en Colombia? *Ensayos Sobre Política Económica*, 24(51), 226-323. doi: 10.32468/Espe.5105

Psacharopoulos, G., & Patrinos, H.A. (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics*, 26(5), 445-458.
doi: 10.1080/09645292.2018.1484426

Popli, G.K. (2011). Changes in human capital and wage inequality in Mexico. *Oxford Development Studies*, 39(3), 369-387. doi: 10.1080/13600818.2011.596276

Saavedra, J.E. (2009). The learning and early labor market effects of college quality: a regression discontinuity analysis. *Investigaciones del ICFES*, 1-61.

Samoliuk, N., Bilan, Y., & Mishchuk, H. (2021). Vocational training costs and economic benefits: exploring the interactions. *Journal of Business Economics and Management. 22*(6), 1476-1491. doi: 10.3846/jbem.2021.15571

Sánchez, F., & Núñez, J. A. (2003). Dynamic analysis of human capital, female work-force participation, returns to education and changes in household structure in urban Colombia, 1976-1998. *Colombian Economic Journal*, 1, 109-149.

Sapelli, C. (2009). Los retornos a la educación en Chile: estimaciones por corte transversal y por cohortes. *Documentos de Trabajo No. 349*. Pontificia Universidad Católica de Chile.

Tenjo, J. (1992). Labor markets, the wage gap, and gender discrimination: the case of Colombia. In G. Psacharopoulos & Z. Tzannatos (Eds.), *Case Studies on Women Employment and Pay in Latin America.* The World Bank.

Tenjo, J. (1993). Evolución de los retornos a la inversión en educación 1976-1989. *Revista de Planeación y Desarrollo*, 24, 85-102.

Tenjo, J., Medina, R.R., & Bernat, L.F. (2004). Evolution of salary differences between men and women in six Latin American countries. In C. Piras (Ed.), *Women at Work*: *Challenges for Latin America* (pp. 139-185). Inter-American Developmnet Bank.

Tenjo, J., Alvarez, O., Gaviria, A., & Jimenz, M. (2017). "Evolution of Returns to Education in Colombia (1976-2014)," Coyuntura Económica, Fedesarrollo, vol. 47(1 y 2), 15-48.

Valletta, R.G. (2019). Recent flattening in the higher education wage premium polarization, skill downgrading, or both? In C. Hulten & V. Ramey (Eds.), *Education, Skills, and Technical Change: Implications for Future US GDP Growth* (pp. 313-354). University of Chicago Press.

doi: 10.7208/9780226567945-010

Vargas, B. (2013). Retornos a la educación y migración rural-urbana en Colombia. *Revista Desarrollo y Sociedad*, 72, 205-223. doi: 10.13043/dys.72.5

Warunsiri, S., & McNown, R. (2010). The returns to education in Thailand: a pseudo-panel approach. *World Development*, 38(11), 1616-1625. doi: 10.1016/j.worlddev.2010.03.002

Wößmann, L. (2003). Specifying human capital. *Journal of Economic Surveys*, 17(3), 239-270. doi: 10.1111/1467-6419.00195

Zarate, H. (2003). Cambios en la estructura salarial: una historia desde la regresión cuantílica. *Borradores de Economía*, 2-45.

Appendix A

Table A1. Average real hourly wage by cohort

Cohort	Men		Women		Total	
	COP	USD	COP	USD	COP	USD
1	2,430.21	0.640	2,605.28	0.687	2,483.23	0.654
2	3,462.85	0.912	3,444.12	0.908	3,454.54	0.910
3	4,365.86	1.150	4,441.48	1.170	4,401.28	1.160
4	5,235.81	1.380	5,553.41	1.463	5,388.69	1.420
5	6,141.22	1.618	5,662.77	1.492	5,909.65	1.557
6	6,153.12	1.621	5,771.59	1.521	5,957.57	1.570
7	6,480.78	1.708	5,727.94	1.509	6,096.86	1.607
8	6,764.76	1.783	5,983.26	1.577	6,371.91	1.679

Source: *own compilation*

Table A2. Average of the variables of the selection equation by year

Year	S	Potential Experience	Potential Experience Squared	Selection	Married	NempHH	Head household
2016	10.15	13.76	330.2	0.66	0.38	2.09	0.29
2017	11.42	11.67	269.31	0.55	0.35	1.96	0.25
2018	11.35	11.74	269.10	0.55	0.36	1.94	0.26
2019	11.54	11.77	270.09	0.56	0.35	1.91	0.26
2020	11.58	11.97	275.95	0.50	0.36	1.73	0.25

Source: *own compilation*