Borradores de ECONOMÍA

Heterogeneity in the Returns to Tertiary Education for the Disadvantage Youth: Quality vs. Quantity Analysis

By: Leonardo Fabio Morales
Christian Posso
Luz A. Flórez

No. 1150
2021

BANCO DE LA REPÚBLICA
COLOMBIA

Heterogeneidad en los retornos a la educación terciaria de los jóvenes de bajos ingresos: Análisis de calidad vs cantidad.

Leonardo Fabio Morales, Christian Posso, Luz A. Flórez [*]

Resumen

Este artículo utiliza registros administrativos de diferentes fuentes para construir un base datos de estudiantes de bajos ingresos en Colombia nacidos entre 1980 y 1990. Estos datos incluyen resultados de pruebas cognitivas, información socioeconómica en su último año de secundaria, e información sobre su trabajo, años después de la graduación de la secundaria. Evaluamos los rendimientos de la educación terciaria estimando el "efecto marginal de tratamiento" (MTE) de la inversión en educación terciaria. El MTE permite estimar un parámetro aleatorio para el rendimiento de la educación, que varía con la heterogeneidad no observada de los trabajadores. Encontramos una heterogeneidad considerable en los retornos, al punto de que, para una masa de la población, el retorno es cercano a cero. Utilizando los modelos estimados, simulamos dos tipos de políticas: una que aumenta la oferta de educación terciaria y otra que mejora la calidad de la educación secundaria. Se encuentra que una política que mejora la calidad de la educación secundaria da rendimientos similares a una política ambiciosa de incrementos en la oferta de educación terciaria.

Codigos JEL: J31, I21, I26, I28

Palabras claves: Retornos a la educación, MTE, expansión universitaria, Calidad de la educación, Salarios.

[*] Morales: *Banco de la República*, lmoralzu@banrep.gov.co (autor de contacto); Posso: *Banco de la República*, cpossosu@banrep.gov.co; Flórez: *Banco de la República*, lflorefl@banrep.gov.co. Las opiniones, declaraciones, conclusiones e interpretaciones expresadas en este documento son responsabilidad de los autores y no representan la posición del Banco de la República o su Junta Directiva. Como es usual, todos los errores y omisiones en este trabajo son nuestra responsabilidad.

Heterogeneity in the Returns to Tertiary Education for the Disadvantage Youth: Quality vs. Quantity Analysis.

By Leonardo Fabio Morales, Christian Posso, Luz A. Flórez [†]

The opinions contained in this document are the sole responsibility of the authors and do not commit Banco de la República or its Board of Directors

Abstract

This paper uses administrative records from different sources to construct a unique data set of low-income students in Colombia born from 1980 to 1990. This data includes cognitive test results, socio-economic information at their high school final year, and information on their labor market results, many years after high school graduation. We evaluate the returns by estimating the Marginal Treatment Effect (MTE) of the tertiary investment decision. The MTE allows estimating a random parameter for tertiary education return, which varies with unobserved heterogeneity across workers. We find sizeable heterogeneity in returns, as recent literature has also identified, to the extent that for a considerable mass of the population, the return is negligible. Using the estimated models, we simulate two types of policies: one that increases the supply of tertiary education and another that enhances secondary education quality. We find that a less costly policy that improves secondary education quality gives similar returns than a more ambitious policy that increases tertiary education supply.

JEL codes: J31, I21, I26, I28

Keywords: Returns to tertiary education, marginal treatment effect, university expansion, quality of education, education, wages.

[†] Morales: *Banco de la República*, lmoralzu@banrep.gov.co (corresponding author); Posso: *Banco de la República*, cpossosu@banrep.gov.co; Flórez: *Banco de la República*, lflorefl@banrep.gov.co. The opinions contained in this document are the sole responsibility of the authors and do not commit *Banco de la República de Colombia* nor its Board of Directors. All errors are authors' own responsibility.

1. Introduction

There is an implicit consensus over the positive effects of education in labor market outcomes, social mobility, and economic growth in social sciences. There is a long tradition in the literature on social mobility that identifies education as one of the primary channels of social mobility improvements in meritocratic societies (Blanden, 2013; Hertz et al., 2007; Becker and Tomes, 1986; Blanden and Macmillian, 2014; Garcia et al., 2015). Education is one of the silver bullets in poverty reduction efforts. There is evidence on the high cross-country correlations between educational achievement and GDP per-capita and estimations of the return of investments in education using the traditional wage regressions (Card, 2001; Glaeser, 2004). Education is widely understood as the most successful strategy to overcome poverty, to the extent that the relationship between development and the accumulation of knowledge in a society is almost tautological (Topel, 1999).

Using different administrative data sources, we study the returns to tertiary education in Colombia of a census cohort of underprivileged students who were born from 1980 to 1990. Focusing on this representative sample, we can test if the promise of enhancing socio-economic conditions as a result of educational investments is fulfilled for low-income students. We evaluate the returns of investment in tertiary education of this cohort by estimating the Marginal Treatment Effect (MTE) (Bj rklund and Moffitt,1987; Heckman and Vytlacil, 1999). The MTE allows estimating a random parameter for tertiary education return, which varies with unobserved heterogeneity across workers. Recent literature in return to schooling has shown with different data, for other countries, that indeed the returns are heterogeneous across individuals; furthermore, individuals seem to know their idiosyncratic return (Heckman and Vytlacil 2001, 2005, 2007; Carneiro et al. 2010; Belskaya et al., 2020).

One of the strongest advantages of the MTE estimation is that it allows the estimation of the returns of investment in tertiary education for the marginal individual, who is induced to invest in tertiary education due to a specific policy (Heckman and Vytlacil 2001). Therefore, this methodology allows measuring the returns of particular policies. In this paper, we assess in the Colombian context a policy of expansion in the supply of tertiary education,

which is a policy simulation that had shown important positive effects in the U.S., China, and Russia cases (Carneiro et al. 2010; Heckman and Li, 2004; Wang et al., 2014; Belskaya et al., 2020; Kyui, 2016).[3] In a similar study, Carranza and Ferreyra (2019) study Colombia's Higher Education expansion in the early 2000s using student- and program-level data.[4]

Additionally, we assess the effect of an alternative policy that increase the quality of public schools that provide secondary education. A recurring debate in developing countries is whether investing public resources in expanding tertiary education or improving the quality of basic and secondary education. In the case of Colombia, for instance, the coverage of secondary education is close to universal (in 2014, the gross and net coverage rates were 101% and 72%, respectively); nevertheless, the quality is poor from international comparison from standardized tests (Busso et al., 2013).

Using the MTE framework, we assess the impact of marginal increases in the average quality of secondary education on the probability of enrolling in university, and by this way, we find the return of increasing the quality of secondary level education in the margin. There are many reasons to believe that quality enhancements at the secondary level will translate into tertiary education returns. The subsidized college education is based on meritocratic systems, where high school test or public universities admission test are the essential screaming mechanism for determining admissions. Furthermore, even in private universities, the mechanisms for accessing government-subsidized credit are also meritocratic and partially determined by high school exit tests. Therefore, improving the secondary level's quality can increase the probability of enrolling in tertiary education (and access to government subsidies), especially for students who belong to low-income families.

[3] See also Dickson & Harmon, (2011) and Gunderson & Oreopolous (2020) for a summary of the trends in the literature on economic returns to education and the econometric issues in developed economies and Patrinos & Psaharopoulos (2020) for the case of developing countries.

[4] Carranza and Ferreyra (2019) study Colombia's Higher Education expansion in early 2000s using student- and program-level data. This study aims to decompose changes on access to higher education associated with demand factors and unobserved factors. Nonetheless, they do not focus on the individual's returns to higher education. They find the decision to open new programs was driven by the economic incentives facing the institutions. Also, Barrera & Bayona (2019) study the impact of attendance of a prestigious Colombian private university in the individuals labour market outcomes; while Gamboa & Rodriguez (2018) study the earnings expectations of a representative sample of students who finish their high school in Colombia.

In this paper, we evaluate how an increase in universities and tertiary programs increases the probability of enrolling in university; in this way, we find the return of increasing the supply in the margin. In the same way, we take advantage of the expansion in secondary education to assess the impact of marginal increases in the average quality of secondary education on the probability of enrolling into university; by this way, we find the return of increasing the quality of secondary level education in the margin.

We contribute to clarify a classical debate in developing countries' education policy: ¿is it better to invest in education coverage or invest in standards' improvements in the supply already provided? We provide evidence that the returns of an ambitious expansion policy of coverage in the tertiary segment are similar to the returns of enhancing the quality in the secondary component. We find that an increase in the tertiary education supply produces an average annualized return of 7% in wages (adding one university in each city). Regarding the policy that simulates an increase in the quality of secondary education, we find that a rise in 0.5 standard deviations in the high school's average exit exam score would produce an annualized return in wages of nearly 7% in wages as well.

As some previous literature has shown, we find that tertiary education returns are heterogeneous; according to our parametric estimation, for the percentile 95 the tertiary education is approximately 3% per year of education, while for the percentile 10 this return is around 12% per year of education. The average return of an additional year of tertiary education is 5.5% per year. We find positive sorting into university; therefore, due to unobserved heterogeneity, more skilled individuals have more probability of getting into university and, at the same time, earning a better wage in the formal labor market.

Finally, we identify how returns to university change with worker's tenure in the labor market; we estimate heterogeneous ATE of an additional tertiary education year, across different samples according to years after graduation. Our results indicate that returns to education increase with labor market experience. For workers with at least two years of experience in the labor market, the ATE of an additional year of tertiary education is 5.5%. It increases to around 7.5% for workers with at least five years of experience. We observe

the same behavior with all other evaluation parameters. tertiary education investments provide higher returns when individuals accumulate more experience in their professions. As a consequence of this pattern in returns, income gaps would increase between college-educated workers and non-college-educated ones.

This paper is organized into seven sections, including this introduction. In the second section, we briefly comment on some of the fundamental studies in the literature of MTE and literature studying the effects of the expansion in the tertiary education level. In the third section, we offer a brief context of the higher education market in Colombia. In the fourth section, we describe our data, and in the fifth section, we describe the MTE methodology. In the sixth section, we describe our results; and finally, in the seventh section, we offer some conclusion and policy recommendations derived from our main empirical findings.

2. Supply and Quality of Secondary Education and Tertiary levels of education.

The expansion of tertiary education is a phenomenon that has taken place in all developing world. This expansion is, in part, a response to rapid growth in the secondary level of education. For instance, in the context of Latin America, during the period 1991-2010 high school enrollment and graduation rates increase substantially (Galiani, et al., 2017; Manocorda, Sanchez & Schady, 2010). On the one hand, unconditional enrollment on-time rate in secondary school increased from 44 percent to 58 percent; on the other hand, the average unconditional graduation rate in secondary schooling increased 15 percentage points (Bassi et al., 2015). To catch up with this increase in secondary graduation, the supply of tertiary has grown as well; Camacho et al. (2017) report the significant increases in new universities in countries as Mexico, Brazil, and Peru, where the growth rate of the total numbers of tertiary education supply from 2005 to 2010 was 14%, 6.8%, and 6.5%, respectively.

According to the Ministry of education's official information, in Colombia in 2000, 277 institutions offer tertiary education (not only universities), and in 2010 this number increased to 349; this is an astonishing 26% growth rate. In terms of tertiary education programs, this increment was even more prominent; the number of programs in institutions

offering some degree of tertiary education increased more than 70% from 2000 to 2010.[5] As a result of this rapid expansion, reductions in the institutions' and average graduates' quality might happen. Recent research offers evidence that this rapid expansion in higher education supply has increased the heterogeneity of returns of tertiary education across programs. Camacho et al. (2017) show that there is a 15% unconditional difference in wages between workers who attend programs that exist before 2000 and the ones who attended new programs. This evidence suggests that lower returns from new higher education programs reflect the lower quality of the marginal student who enters these new programs and the fact that the new programs are created in areas traditionally associated with lower returns. Summarizing, the supply of secondary and tertiary levels of education has increased substantially in the context of the Latin American region, and in general, in the developing world (Carnoy et al., 2013). Some of the literature mentioned before have raised concern about the consequences of this expansion on tertiary education quality. Nevertheless, studies that relate the returns of tertiary education as consequences of quality changes in secondary education are missing links in the literature.

The connection between quality on the secondary level and the tertiary education returns has been studied previously in the literature (see for instance, Card & Krueger, 1990). Graduates from bad quality secondary level institutions have lower chances to access to University education, and therefore, low returns to schooling. This positive relationship between secondary education quality and returns might be exacerbated for students from un-advantaged backgrounds. In Colombia, as in many other developing countries, the access to subsidized education (public universities) is based on meritocratic systems. High school standardized tests and public universities admission test are the essential screaming mechanisms for determining admissions.

Furthermore, even in private universities, the mechanisms for accessing government-subsidized credit are also meritocratic and partially determined by high school exit tests. There are more reasons to argue that high school quality increases individuals' probability of enrolling in university. In general, quality secondary schools do a better job motivating

[5] Similar facts are reported in Carranza and Ferreyra (2019).

students to pursue ambitious endeavors, which is crucial in the case of poor students from families with probably no precedents of college graduate family members. In other words, low-income students can find in good quality high schools the social references, good peer relationships, role models, and motivation they cannot find at home (Leslie, Johnson & Carlson, 1977; King, 1996).

Previous literature has described the evolution of secondary education in the Latin American region as the half-empty glass because the supply and coverage of secondary education have grown fast, but at expenses of inferior quality in terms of international comparisons. For example, the Latin American students rank in the lowest percentile of the international standardized test as PISA (Bassi et al., 2015). There is ample evidence that the Latin American Region scores substantially below the average of OECD countries. Colombia, in particular in 2009, ranked 59 and 53 among 66 countries in math and reading, respectively. There is evidence as well indicating that poor students score worse than wealthy students (Bassi et al., 2015).

This paper's primary purpose is to evaluate how the increase in universities and programs increases the probability of enrolling in university. In this way, we find the return of increasing the supply in the margin; to do this, we take advantage of the expansion of tertiary education, which took place in the first decade of the century. Analogous to the expansion in tertiary education, there has been an expansion in secondary education; as discussed before, such an expansion may imply a reduction of secondary-education quality. We take advantage of the expansion in secondary education to assess the impact of marginal increases in the average quality of secondary education on the probability of enrolling in university; in this way, we find the return of increasing the quality of secondary level education in the margin.

3. Data

In this paper, we construct a comprehensive data set from several sources of administrative records in Colombia. The estimation sample consists of a cohort of individuals who graduate from high school from 1994 to 2006, the year in which they took the official high school exit standardized test named SABER 11. In many cases, public and private universities use this high school exit exam as a screening method for determining admissions. All individuals in the sample had a characteristic in common; at some point in their youth, their families were registered in the System of Identification of Potential Beneficiaries of Social Programs (SISBEN), the low-income census households in Colombian. Knowing that the individual´s family was interviewed in the SISBEN[6] is how we identify that the individual comes from a low-income household.

To characterize this cohort of students' labor market results, years after they graduate from high school, we use another administrative source called Integrated Record of Contributions to Social Security (PILA). The PILA is a system that registers payments of labor taxes for all formal employees in all formal firms. PILA is a unique source of information containing wages, some firms, and workers' characteristics.[7] We link our estimation sample at the individual level using the PILA in 2010. That is, we measured the individuals' earnings in 2010.

In addition, we match our estimation sample at the individual level with two additional information sources provided by the Colombian Institute for the Evaluation of Higher Education (ICFES). The first source of information is the high school exit exam score in Colombia, and some socio-demographic information collected just before the student takes the test (known as SABER 11). The second source comes from the ICFES; it is the database

[6] SISBEN is a targeting mechanism to select beneficiaries of social programs as subsidized health insurance, conditional cash transfers, elderly subsidies, among others (Bottia, Cardona and Medina, 2012). It was introduced in Colombia in 1995 followed by two revisions in 2005 and in 2009/10 (version II and III, respectively).

[7] PILA has been used recently in the literature for the study of different aspects of the Colombian labor market, for instance in the study of labor dynamics (Morales and Medina, 2016; Flórez et al., 2020), in the study of the consequences on labor demand of trade liberalization policies (Mejia et. al, 2017), in the study of consequences of changes in payroll taxes policies (Kugler et al., 2017; Bernal et al, 2017; Morales and Medina, 2017).

of all individuals enrolled in tertiary education who took the official exit exam of tertiary education in Colombia, widely known as SABER-pro or ECAES. This ECAES test was not mandatory in the 2004-2008 period. Nevertheless, the exam takers in those years were very similar to the total number of people who graduated from the universities (Guarín et al., 2016).

Finally, from ICFES, we know the municipality of individuals' residences in our sample just before they graduated from high school. We characterize the cities of residence of the individuals just before completing high school in terms of the unemployment rate and the Index of Unsatisfied Basic Needs (NBI), a poverty index popular in developing countries, which can be easily computed Census data. Finally, we can characterize the individual's high school institution in terms of its quality, using average ICFES test scores and additional characteristics.

A table of summary statistics is presented in Appendix B. Our final sample includes more than 542K individuals. The mean age of 28.01 years, 48% of the workers are female, only 10.8% have a parent with post-secondary education, and around 63% of them came from families with total incomes smaller than two minimum wages. Finally, the share of individuals that in the cohort invest in tertiary education was nearly 10%.

4. Empirical Strategy: Estimating the Marginal Treatment Effect of Tertiary Education

4.1 The MTE Framework

The MTE framework is based on a generalized Roy Model, a choice-theoretic model very appealing for policy analysis (Carneiro, Heckman, and Vytlacil, 2011). In this section, we follow the model originally proposed in Heckman and Vytlacil (2005). Let us represent wages for individuals with tertiary education W_1 and individuals who only complete high school by W_0, where wages for individuals with and without tertiary education are modeled as $W_1 = X\beta_1 + U_1$ and $W_0 = X\beta_0 + U_0$, respectively. Therefore, the potential wage is defined as:

$$W = S(X\beta_1 + U_1) + (1 - S)(X\beta_0 + U_0) \quad (1)$$

S is a binary variable equal to one for individuals with tertiary education and zero in other cases. In addition, The decision of investing in tertiary education is modeled as a standard discrete choice model, where the latent variable is the individual's net benefit of enrolling into college (S^*) (Carneiro, Heckman, and Vitlacil, 2011). This benefit of enrollment is modeled as a linear function, which depends upon observed variables (Z) and unobserved perturbation U_s. Therefore, the decision of enrolling in university can be explained as:

$$S = 1\{S^* \geq 0\} = 1\{Z\gamma + U_S \geq 0\} \qquad (2)$$

From this previous equation, we know that the probability of attending university is given by $P(Z) \equiv P(S = 1) = P(U_s \leq Z\gamma) \equiv F_{U_s}(Z\gamma)$, where F_{U_s} is the cumulative probability distribution of the random variable U_s. We can define the variable $V \equiv F_{U_s}(U_S)$, which by construction is uniformly distributed; this previous transformation is inconsequential because U_S can have any density (Carneiro et al., 2017). Note that expression (1), after simple algebraic manipulations, can be expressed as:

$$W = X\beta_0 + S(X\beta_1 - X\beta_0) + S(U_1 - U_0) + U_0 \qquad (3)$$

The Marginal Treatment Effect can be defined as the derivative of the conditional expectation of wages -equation (3) with respect to the probability of attending tertiary education; this is:

$$MTE = \frac{\partial E[W|X = x, P(Z) = p]}{\partial p} = \frac{\partial E[W|X = x, V]}{\partial p} \qquad (4)$$

Where the last equality comes from the definition of V in the previous paragraph. As the reader can notice from equation (4), the MTE measures changes in wages due to marginal increments in the probability of tertiary education. MTE is a simple way of characterizing the heterogeneity of wages given the heterogeneity in the probability of attending university. Let us now focus in the conditional expectation of wages in equation (3):

$$E[W|X = x, V = v] = X\beta_0 + (X\beta_1 - X\beta_0)p + \int_0^p E\big[(U_1 - U_0)|X = x, V = v\big]\, dv \quad (4)$$

$$= X\beta_0 + \int_0^p \big((X\beta_1 - X\beta_0) + E[(U_1 - U_0)|X = x, V = v]\big)\, dv \quad (5)$$

From the definition of MTE represented in equation (4), and from equation (5) we can obtain a useful expression, which is the basis for the computation of the MTE. If we derive expression (5) with respect to p, as it is indicated in the definition of equation (4) we get:

$$MTE(x, v) = X(\beta_1 - \beta_0) + E[(U_1 - U_0)|X = x, V = v] \qquad (6)$$

As equation (6) reveals, the MTE characterizes the heterogeneity of the returns to tertiary education. These returns may vary because of differences in individuals' observed characteristics as their socio-economic background or their results in high school final standardized tests. Nevertheless, individuals may be different on un-observables as well, and this component of the heterogeneity is captured by the second component on the right-hand side of the equation (6). Equation (6) offers a more comprehensive way of framing heterogeneity of returns to schooling because individuals may be identical from the econometrician's point of view, but they can be different in terms of unobserved skills. A standard interpretation of the variable V in the literature on MTE is as the psychic cost of education, or in other words, a negative ability; therefore, individuals with higher values of V have less probability of enrolling in university (Carneiro et al., 2017)[8].

A great feature of the MTE is that traditional parameters of interest in the literature of impact evaluation can be expressed as a function of the MTE (Heckman and Vytlacil, 2005). The expectation of the MTE over all the support of V is the traditional Average Treatment Effect (ATE), which make sense because ignoring the heterogeneity of the return and taking the unconditional expectation, would result in the "average effect" of tertiary schooling. In the same way, conditional expectation of the MTE over the distribution V for those who attended to university and those who do not would result in the average treatment on the treated (ATT) and the average treatment on the un-treated (ATU), respectively.

Arguably, the strongest attribute of the MTE is that it allows computing the returns for the marginal individual, since the return of tertiary education can be computed in all support of V. The marginal individual is the one who is indifferent between investing or not

[8] Heckman, et al. (2018) estimate the returns to education using a dynamic model that synthesizes the two approaches used in the literature: the structural dynamic discrete choice and the reduced-form treatment effect.

in tertiary education; therefore, it is the one that in the margin may be induced to attending university as a result of a policy. Heckman and Vytlacil (2001) introduce a new parameter called the Policy Relevant Treatment Effect (PRTE), which measures the average return to schooling for those induced to enroll in university as a response to a specific policy. This PRTE has an extremely appealing interpretation since it may be used to simulate the returns to policies that are expected to increase the chances of tertiary education enrollment.

All the parameters mentioned in the previous paragraph can be expressed as a function of the MTE using the following formulas developed by Heckman and Vytlacil (2001). In the case of the PRTE, a continuous exclusion restriction (Z) is needed, and a policy that generates a change in this variable from $Z = z$ to $Z = z'$.

$$ATE(X = x) = \int MTE(X = x, V = v) f_{V|X}(v|x) dv \quad (7)$$

$$ATT(X = x) = \int MTE(X = x, V = v) f_{V|X}(v|x, S = 1) dv \quad (8)$$

$$ATU(X = x) = \int MTE(X = x, V = v) f_{V|X}(v|x, S = 0) dv \quad (9)$$

$$PRTE(X = x) = \int MTE(X = x, V = v) f_{V|X}(v|x, S(Z = z) = 0, S(Z = z') = 1) dv \quad (10)$$

4.2 Estimation of Average, Marginal returns and Policy Relevant Treatment Effects

We follow two approaches to estimate the MTE. First, we estimate a parametric Roy model assuming that errors U_0, U_0, U_S are jointly normal distributed (Appendix A presents the assumptions of this parametric estimation). The second alternative is a semiparametric estimation, in which no assumptions on the distribution of U_0, U_0, U_S is made. In this paper, we estimate the MTE using both alternatives, but we focus on the estimation's parametric version, which allows us to estimate policy simulations. In a further section on robustness checks, we show that both estimations are statistically equivalent in the common support of the probability of attending college.

Considering tertiary education as a treatment, the average treatment effect (ATE) describes the effect of tertiary education in wages for the average individual; this is the difference in wages between those enrolled in university and those who do not. As mentioned before, other parameters of interest are the effect on the treated (ATT) and the effects on the untreated (ATU). These parameters describe the impact of tertiary education just for the ones who indeed went to university and what effect tertiary education would have had for those who did not enroll, respectively. The most interesting parameter, from a policy evaluation perspective, is the average return of tertiary education for the individuals who are at the margin between investing in tertiary education or not (MTE); These are the "marginal individuals," who are induced to enroll to university as a result of a policy. In this paper, we compute the MTE, and we use formulas 7, 8, 9 to calculate ATE, ATT, and ATU.

Literature has defined several marginal policies (Carneiro Heckman and Vytlacil, 2010; 2011). In this paper, we estimate two types of policy-relevant treatment effects (PRTE). We compute the policy-relevant treatment effect (see equation 10) of a policy that increases a single component of vector Z, holding everything else constant; therefore, as a result of the policy, a new probability of enrolling into college can be represented as $P_{Z_\delta} = P(Z_\delta)$, where Z_δ is a vector such that ${Z^k}_\delta = Z^k + \delta$ and ${Z^j}_\delta = Z^j \; \forall j \neq k$. Using this framework, we simulate two policies: *i)* a policy that expands the tertiary education supply by increasing one university at the municipality where the individual lived when he took the high school exit exams. In addition, policy *ii)* simulates an increase in the quality of secondary education, as the increases in half standard deviation of the average score in the high school exit exam where the individual got her high school degree.

5. Results

In this section, we present the MTE estimation results, the relevance of the main instruments, and show some evidence of the heterogeneity in the returns. We begin by discussing the results from the estimation of our parametric model given in equations A2-A3 from Appendix, as the reader may remember in this parametric estimation assumed that $U_0, U_1, U_S \sim N(0, \Omega)$. Since individuals endogenously choose tertiary education, we use several exclusion restrictions in the selection equation (enrollment in university). On the first hand, as a measure of the tertiary education supply, at the municipality level, we use the number of tertiary education institutions and the number of total students enrolled in universities when the individual graduated from high school.[9] On the other hand, we use the average score in the secondary institution's high school exit test, where individuals got their graduation. We control for a battery of socio-economic background characteristics in the selection and wages equation. In addition, we control for the individual score in the high school exit test, in the selection and the wage equations as well.

Results for the selection and wages equation are presented in Appendix C. In the selection equation, we control some characteristics of the labor market when individuals graduated from high school, for instance, the unemployment rate and a poverty index at the municipal level. We focus our attention on two instruments: the first as the main instruments for the tertiary education supply, in terms of the number of tertiary-level institutions, and the second as the quality of the secondary level education in terms of the high school average score in the exit exam. Our first instrumental variable is close to the commonly used binary variable: the proximity of a university campus in the county of residence during teenage years (Card, 1995; Cameron and Taber, 2004); this instrument has been used before in the MTE literature (Belaska et. al, 2020; Carneiro, et al., 2010). We test the relevance of all of our instruments in a joint significance test. We robustly reject the hypothesis of no joint significance; furthermore, all instruments are individually significant, as shown from Appendix C (column 3).

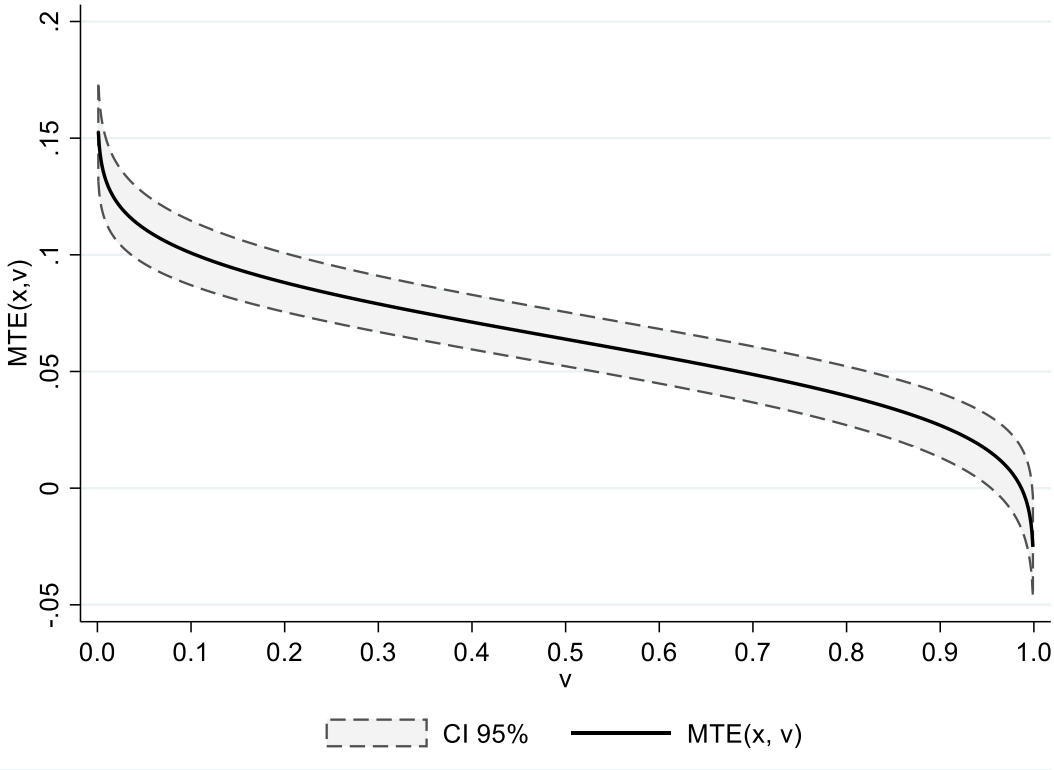
[9] This instrument is an extension over a commonly used binary instrumental variable "the presence of a college" (Card, 1995; Cameron and Taber, 2004).

Returns to college may vary because of differences in observed and unobserved characteristics of individuals. The component of the unobserved heterogeneity is captured by the second element on the right-hand side of the MTE definition (equation 6). In Graph 1, we present our estimated MTE. The heterogeneity in the MTE across individuals is substantial. On the one hand, per year, the return to tertiary education for the percentile 95 is approximately 3%; these are individuals with lower levels of unobserved ability and lower probability of enrolling into college.

On the other hand, for the percentile 10 is around 12%, they have higher unobserved ability and a higher probability of enrollment. As shown in Graph 1, the MTE decreases in v, which constitutes evidence that individuals select themselves into tertiary education based on expected (unobserved) returns. This phenomenon usually is denominated "selection on gains." The highest returns are observed for individuals with the highest probability of enrolling in university (with lows v) (Carneiro et al., 2011). A formal test for selection on gains relies on the estimated covariance terms $\sigma_{1,U_S}, \sigma_{0,U_S}$ presented at the end of the Table in Appendix C (see T test (r0-r1)=0). The null hypothesis that the slope of the MTE is zero, i.e., H_0: $\sigma_{1,U_S} - \sigma_{0,U_S} = 0$, is rejected at the 1 percent level (t=3.02). This pattern of the heterogeneity in returns to tertiary education has been found before for different countries as the U.S., Russia, Indonesia, and Chile[10] (Carneiro et al., 2010; Belskaya et al., 2020; Carneiro et al., 2017; Rodríguez, Urzúa & Reyes, 2016). The pattern is consistent with the intuition that individuals self-select into higher education levels based on their comparative advantage (Carneiro et al., 2017).

[10] Peet, et al. (2015) find evidence of heterogenous returns to education in 25 developing countries between 1985 and 2012. Even though their results should not be interpreted as the causal effect of one year of schooling, rather than a conditional association; the authors found important heterogeneity in the returns within individuals.

Graph 1: Estimated Marginal Treatment Effect



Notes: These figures plots the values of MTE for each value of a uniformly distributed variable v in equation (11). The MTE is generated as indicated in equation (A1) where all the parameters involved in this estimation came from a switching regression model regression of the system of equations A2, A3, A4 estimated by Maximum Likelihood. The confidence intervals are constructed for a 95% of confidence, and the MTE standard error is computed analytically using the Delta Method. The returns are annualized by dividing the total estimated return by 4, the mean duration of a tertiary education program in Colombia.

Based on the MTE, we compute several other interest parameters as the ATE, ATT, and ATU (see Table 1). These parameters are simulated using the equations 7, 8, and 9.[11] The average return of education (ATE) on our sample was 5.5% per year of university. This result is comparable to the findings by Carneiro et al. (2011) for the case of the U.S. The return of tertiary education in the treated population is greater; the most skilled individuals select themselves into tertiary education; therefore, their return per year in university is almost 12%, substantially higher than the average. On the contrary, since there is sorting on returns, one can expect that the untreated returns would be lower; indeed, the untreated

[11] In all cases, standard errors are estimated using bootstrap methods for 250 replications.

population return would have been 5%, which are lower than the average, but not excessively. Therefore, given the distribution of skills in the economy, there is still a relative-to-media large return of attending university for non-treated; the coverage rates of tertiary education in Colombia are still low; this evidence is not surprising.

5.1 Quality Vs. Quantity

Given our estimation sample, young workers belonging to a low-income family at some point in their past, we focus on two promising policies. The first policy, which has been documented previously in the literature of MTE, is a global expansion of tertiary education. More universities and technical institutions mean more opportunities for admission in public universities or affordable private ones (in general, such an expansion may reduce the average tuition in private institutions). The second policy, which, to the best of our knowledge, has not been explored in the literature, enhances the quality of secondary institutions where these individuals got their high school degrees.

As explained before, subsidized tertiary education is based on a meritocratic system, where high school tests or public universities admission test are the essential screening mechanism for determining admissions. Therefore, high-quality secondary education constitutes the best of the preparations for admission tests. In most cases, the admission is based on the aspirant student's high school exit exam. The mechanisms for accessing government subsidized credit for private tertiary education investments are also meritocratic and partially determined by high school exit tests.

Based on the MTE, we compute the PRTE for the two policies describes in the previous paragraph. In each simulation, a new probability of enrolling in university is generated, which can be represented as $P_{Z_\delta} = P(Z_\delta)$, where Z_δ is a vector such that ${Z^k}_\delta = Z^k + \delta$ and ${Z^j}_\delta = Z^j \; \forall j \neq k$. The PRTE of these two policies is presented as well in Table 1. The effect of expansion (a new university at each municipality) is an average annualized return of 7.07% in wages compared to workers with no university; the effect is statistically significant. An enhancement in the secondary school's quality, increasing in 0.5 standard deviations the high school's average score at which individual got her high school degree (16

additional points in total score), generates an annualized return of 7.06%; this effect is statistically significant as well.

Our policy experiments contribute to a classical debate in developing countries' education policy: ¿is it better to invest in education coverage or invest in improvements in the supply already provided? The previous question is difficult to answer because the researcher requires a computation of the returns of any policy for the marginal individual, the one who is induced to invest in more education as a direct result of the policy. Colombia experienced the highest increase in secondary graduation rates in the Latin American and Caribbean regions; graduation of individuals at schooling ages increases from 20% in 1990 to 47% nearly ten years later (Camacho et al., 2017). Nevertheless, there have been no substantial improvements in the quality of secondary and elementary education levels; this is the international standardized test's conclusion. Colombia, in 2009 ranked 59 and 53 among 66 countries in math and reading, respectively.

Table 1: Treatment and Policy Parameters

Treatment Parameter	Estimated Parameter
Average Treatment Effect (ATE)	0.054731
	(0.0000517)
Average Treatment on the Treated Effect (ATT)	0.1170813
	(0.0000609)
Average Treatment on the Un-treated Effect (ATU)	0.0492478
	(0.0000409)
Policy Relevant Treatment Effect (PRTE)1	0.0707349
	(0.0001756)
Policy Relevant Treatment Effect (PRTE)2	0.0706937
	(0.0001741)

Notes: All the treatment parameters and policy parameters are estimated as indicated in equations 7, 8, 9 and 10 using simulation methods. Standard errors are estimated using bootstrap methods for 250 replications. The returns are annualized by dividing the total estimated return by 4, which is the mean duration of a tertiary education program in Colombia.

We provide evidence that the returns of a very ambitious expansion policy of coverage in the tertiary segment are similar to the returns of a policy of enhancing the quality in the secondary education segment. This evidence is very intuitive considering our

estimation sample; for low-income students, the only two options for enrolling in tertiary education are public universities or subsidized educative loans supplied by the government. For both of these options, the potential demand exceeds the government's capabilities to provide these educational services; therefore, meritocratic systems have been implemented. These systems grant the subsidies recipients based on the high school exit exam or some other admission exam.

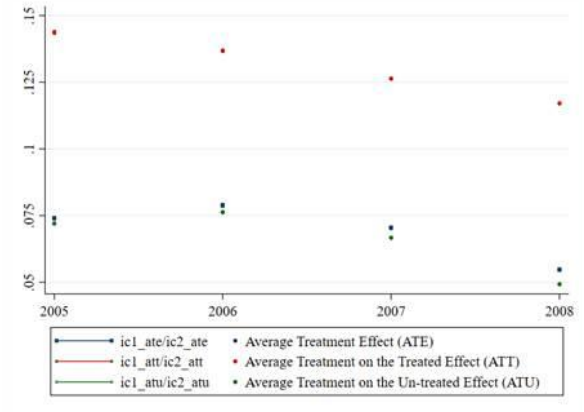
5.2 Returns to Education and Labor Market Tenure

In this last subsection, we explore the heterogeneity, across different labor market tenure levels, of our main evaluation parameters: ATE, ATT, ATE, and policy simulations. We estimate these parameters for four samples of individuals according to the time we observed them in our sample. In the 2008 sample, we observe individuals with at least two years of experience because their wages were observed in 2010; likewise, in the 2007 sample, we observe individuals with at least three years in the labor market, and so on. Heterogeneous estimations for the ATE, ATT, ATU are presented in Graph 2, and estimations for policy simulations are presented in Graph 3. In all cases, the treatment parameters and policy parameters are estimated as indicated in equations 7, 8, 9, and 10 using simulation methods. Standard errors are estimated using bootstrap methods for 250 replications. Tables with estimated evaluation parameters in all cases are presented in Appendix E.

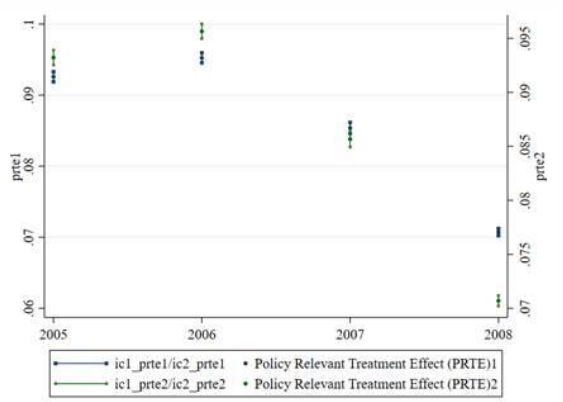
Our results indicate that returns to education increase with labor market experience. For workers with at least two years of experience in the labor market, the ATE of an additional year of tertiary education is 5.5%. It increases to around 7.5% for workers with at least five years of experience. We observe the same pattern with all other evaluation parameters. For instance, let us consider the ATT, an additional year of college, has an ATT for workers with at least two years of experience of nearly 12%, and it increases to more than 14% for workers with five years of experience. Finally, our findings indicate that the returns of both policies we simulate, expansion of tertiary education supply, and enhancing quality in the secondary segment, increase with experience in the labor market. Therefore, we provide evidence that tertiary education investments provide higher returns when individuals accumulate more experience in their professions. As a consequence of this pattern in returns,

income gaps would increase between college-educated workers and non-college-educated ones. We can only observe a span of a few years; nevertheless, this seems to be the case.

Graph 2: Heterogeneous ATE, ATT, ATU, across different experience in the Labor Market



Graph 3: Heterogeneous PMRE, across different experience in the Labor Market



Notes: All the treatment parameters and policy parameters are estimated as indicated in equations 7, 8, 9, and 10 using simulation methods. Standard errors are estimated using bootstrap methods for 250 replications. The returns are annualized by dividing the total estimated return by 4, the mean duration of a tertiary education program in Colombia. The parameters estimated in 2008's regression use a sample of workers with at least one year of labor market experience. The 2007's parameters were estimated only with workers with at least two years of experience in the labor market. The 2006's parameters were estimated only with workers with at least three years of experience in the labor market. Finally, the 2005's parameters were estimated only with workers with at least four years of experience in the labor market.

5.3 Robustness check: Semi-parametric Estimation of the MTE:

Imposing the distribution of U_1, U_0, U_S could be a strong assumption; using the methodology of Local Instrumental Variables, which does not impose distributional restrictions on U_0, U_0, U_S (Carneiro et al., 2017; Heckman and Vytlacil, 2000; Belskaya et al., 2020), we can retrieve a more flexible estimation of the MTE. Local Instrumental Variables is a semiparametric estimation method, which consists of two steps procedure. In the first step of the procedure, we are interested in estimating the following version of equation 4 (see equation 11), and in the second stage, we solve for $K(p)$[12] and their derivatives with respect

[12] Note that we have replaced $E\left[(U_1 - U_0)|X = x, V = v\right] \cdot p$ in (4) by a flexible function of p, $K(p)$. For the estimation of equation (11) we follow estimate a semiparametric model using a partially linear regression estimator (Robinson, 1988). In the first stage we estimate nonparametric regressions of W, X and Xp on p, which is the variables entering non-

to p. The MTE estimation using this semiparametric methodology can only be identified for the common support of $P(Z)$.

$$E[W|X=x,V=v] = X\beta_0 + (X\beta_1 - X\beta_0)p + E\left[(U_1 - U_0)|X=x,V=v\right]\cdot p$$
$$= X\beta_0 + Xp(\beta_1 - \beta_0) + K(p) \qquad (11)$$

The graph in Appendix D presents the results of this MTE semiparametric estimation; the parametric estimation is contained within the 95% confidence interval of the parametric estimation. As we can observe, the results are not significantly different from those estimated with the parametric analysis. In this case, we also find an important heterogeneity in the returns: per year the return to tertiary education for the percentile 85 is negative, while for the percentile 10 is approximately 11%. Moreover, as before, we find that MTE decreases with v, confirming that individuals select themselves into tertiary education[13].

parametrically in equation (8) (Carneiro et al., 2017). In a second stage, we obtain coefficients $\beta_0, (\beta_1 - \beta_0)$ by running an OLS regressions of the residuals $W - E[W|p]$ on residuals $X - E[X|p]$ and $pX - E[Xp|p]$. In the second stage of the local instrumental variable procedure, we solve for $K(p)$ in equation (11), and we get $K(p) = W - X\hat{\beta}_0 + Xp(\widehat{\beta_1 - \beta_0})$. Therefore, $K(p)$ and it first derivative are estimated using a no parametric regression of $\left(W - X\hat{\beta}_0 + Xp(\widehat{\beta_1 - \beta_0})\right)$ on p. Finally, applying the definition of the MTE in equation (6) is the derivative of equation (8) with respect to p, this is represented in the following equation: $\frac{\partial E[W|X=x,V=v]}{\partial p} = X(\beta_1 - \beta_0) + K'(p)$. The parameter $(\beta_1 - \beta_0)$ is estimated in the first stage of the procedure and the second one, $K'(p)$, is estimated in the second stage. Since the parameter in the second stage is the first derivative, in the second stage we estimate a local polynomial regression of order 2, as suggested by Fan and Gijbels (1996) the best order of the polynomial to estimate is $n+1$ if the researcher is interested in the derivative of order n (Carneiro et al., 2017).

[13] Authors such as Heckman (2008) and Henderson et al. (2011) also use nonparametric methods to investigate heterogeneity in the returns to education

6. **Conclusions**

The expansion of the supply of educational services in the developing world has been outstanding in the last three decades, which is an expected result given the overwhelming evidence of the desirable consequences of education and the long tradition of models that remark human capital as an engine for economic growth (Card 2001, Lucas, 1988; Romer, 1990; Topel, 1999; Becker and Tomes, 1986). For instance, in the last 20 years, in the Latin American and Caribbean regions, enrollment and graduation rates in the secondary education segment have shown unprecedented growth (Bassi et al., 2015; Gasparini et al., 2011). This expansion in the secondary level has created pressures for supply increases at the tertiary level as well. For Colombia, in particular, the growth in the tertiary level is of such extent that the number of tertiary education degree programs more than double in the first decade of this century (Camacho et al., 2017). Despite the desirable consequences of these expansions on both education levels, they have raised concerns about their effects on quality standards. In the case of the secondary level expansion, the coverage growth has not been coupled with enhancements in quality; Latin American students rank in the lowest percentile of the international standardized test as PISA (Bassi et al., 2015).

In this paper, we compute the MTE to return to tertiary education for a cohort of low-income students who took the exit high school test between 1998 and 2006, which are observed in the formal labor market in 2010. Even though many developing countries are experiencing educational expansion, the estimations of heterogeneous returns to university in developing economies are very uncommon. To the best of our knowledge, there are three exceptions, two studies for China (Heckman and Li, 2004; Wang et al., 2014) and one study for Indonesia (Carneiro et al., 2017).

Our findings suggest a large degree of heterogeneity in annualized returns to tertiary education. We find that the return to tertiary education for the percentile 90 is approximately 3% per year, while for the percentile 10 this return is around 12%. Moreover, the average return for an additional year of a college education is 5.5%. Then MTE decreases in v, where the highest returns are observed for individuals with the highest probability of enrolling in university (with lows v) (Carneiro et al., 2011). These results are relevant when we are

focusing on individuals that belonged to low-income families. Policies that increase the probability of enrolling in university students' are the most successful strategy to overcome poverty.

The estimation of the MTE allows us to simulate the effects of two policies that are usually in policymakers' toolkit in poverty reduction efforts. The first policy is a global expansion of tertiary education, which implies an increase in public universities' admission opportunities or affordable private ones. The other policy is enhancing the quality of the secondary level of education. There are at least two channels through which this second policy can increase the probability of attending tertiary education and increasing wages. On the one hand, a subsidized tertiary education is based on a meritocratic system. On the other hand, the mechanisms for accessing government-subsidized credit for investments in private tertiary education are meritocratic as well.

The simulation of an increase in the tertiary education supply (in one college) at the municipality where the individual lived when he took the high school exit exam gives an average annualized return of 7% in wages. Regarding the second policy, which simulates an increase in the quality of secondary education, as the increases in 0.5 standard deviations in the high school's average score at which individuals got her high school degree, it generates an annualized return of 7% as well. These policy experiments contribute to clarify a classical debate in developing countries' education policy: ¿is it better to invest in education coverage or to invest in standards' improvements in the supply already provided? We provide evidence that the returns of a very ambitious expansion policy of coverage in the tertiary segment are similar to the returns of a policy of enhancing the secondary segment's quality.

An additional finding is that returns to education increase with labor market experience. We observe an increasing pattern of all our evaluation parameters with years of experience in the labor market. The same is true with the return of our simulated policies. Therefore, tertiary education investments provide higher returns when individuals accumulate more experience in their professions. As a consequence of this pattern in returns,

income gaps would increase between college-educated workers and non-college-educated ones.

This evidence is plausible, considering our estimation sample. For low-income students, the only two options for enrolling in college are public universities or subsidized educative loans granted by the government. In the Latin America and Caribbean region, education policies have favor increasing the supply of education services instead of rising standards of the services already provided. To some extent, this could be a regressive policy given that meritocratic systems only provide subsidized tertiary education for the most qualified students. In this paper, we provide evidence that enhancing quality in the secondary segment could a very successful policy for low-income students, given that this policy increases the chances of enrolling in university for these students as much as an extensive expansion plan on the supply side.

References

Bassi, M., Busso, M., & Muñoz, J. (2015). Enrollment, Graduation, and Dropout Rates in Latin America: Is the Glass Half Empty or Half Full?. *Economía*, 113-156.

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

Blanden, J. (2013). Cross-country Rankings in Intergenerational Mobility: A Comparison of Approaches From Economics and Sociology. *Journal of Economic Surveys*, 27(1), 38-73. https://doi.org/10.1111/j.1467-6419.2011.00690.

Blanden, J & Macmillan, L (2014). Education and Intergenerational Mobility: Help or Hindrance? *Centre for Analysis of Social Exclusion, Working Paper No. 8.* The London School of Economics and Political Science.

Becker, G., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics,* 4(3, Part 2), S1-S39.

Belskaya, O., Sabirianova Peter, K., & Posso, C. (2020). Heterogeneity in the Effect of College Expansion Policy on Wages: Evidence From the Russian Market. *Journal of Human Capital*, Vol. 14-1, 2020

Bernal, R., Eslava, M., Meléndez, M., & Pinzón, A. (2017). Switching From Payroll Taxes to Corporate Income Taxes: Firms' Employment and Wages after the 2012 Colombian Tax Reform. *Economia*, *18*(1), 41-74. DOI: 10.2307/90017435

Bj rklund, A. and Moffitt, R. (1987). The Estimation of Wage Gains and Welfare Gains in Self-selection Models. *Review of Economics and Statistics*, 69(1):42–49. DOI: 10.2307/1937899.

Bottia, M., Cardona, L., & Medina, C. (2012). El SISBEN Como Mecanismo de Focalización Individual del Régimen Subsidiado en Salud en Colombia: Ventajas y Limitaciones. *Revista de Economía del Rosario*, 15(2), 137-177.

Busso, M., Bassi, M., & Muñoz, J. S. (2013). Is the Glass Half Empty or Half Full? School Enrollment, Graduation, and Dropout Rates in Latin America. *IDB Working Paper No. IDB-WP-462*. http://dx.doi.org/10.2139/ssrn.2367706

Camacho, A., Messina, J., & Uribe Barrera, J. P. (2017). The Expansion of Higher Education in Colombia: Bad Students or Bad Programs?. *Documento* CEDE, No. 2017-13. https://dx.doi.org/10.2139/ssrn.2921965

Cameron, S. V., & Taber, C. (2004). Estimation of Educational Borrowing Constraints Using Returns to Schooling. *Journal of political Economy,* 112(1), 132-182.

Card, D. (1995) Using Geographic Variation in College Proximity to Estimate the Return to Schooling. *National Bureau of Economic Research*, No. W4483. DOI: 10.3386/w4483.

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, 69(5), 1127-1160. https://doi.org/10.1111/1468-0262.00237.

Card, D., & Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. Journal of political Economy, 100(1), 1-40.

Carneiro, P., Heckman, J., & Vytlacil, E. (2010). Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin. *Econometrica*, 78(1),377-394. https://doi.org/10.3982/ECTA7089

Carneiro, P., Heckman, J., & Vytlacil, E. (2011). Estimating Marginal Returns to Education. *American Economic Review*, *101*(6), 2754-81. DOI: 10.1257/aer.101.6.2754

Carneiro, P., Lokshin, M., & Umapathi, N. (2017). Average and Marginal Returns to Upper Secondary Schooling in Indonesia. *Journal of Applied Econometrics,* 32(1), 16-36. https://doi.org/10.1002/jae.2523.

Carnoy, M., Loyalka, P ., Dobryakova, M., Dossani, R., Froumin, I., Kuhns, K., ... & Wang, R. (2013). *University expansion in a changing global economy: Triumph of the BRICs?*. Stanford University Press.

Carranza, J. E., & Ferreyra, M. M. (2019). Increasing Higher Education Access Supply, Sorting, and Outcomes in Colombia. *Journal of Human Capital,* 13(1), 95–136. doi:10.1086/701435

Dickson, M., & Harmon, C. (2011). Economic Returns to Education: What we Know, What we don't Know, and Where We Are Going—Some Brief Pointers. *Economics of Education Review*, 30(6), 1118-1122. https://doi.org/10.1016/j.econedurev.2011.08.003

Flórez, L., Morales, L., Medina, D., & Lobo, J. (2020). Labor Flows Across Firm Size, Age, and Economic Sector in Colombia vs. the United States. *Small Business Economics,* 1-32. https://doi.org/10.1007/s11187-020-00362-8.

Galiani, S., Cruces, G., Acosta, P., & Gasparini, L. (2017). Educational Upgrading and Returns to Skills in Latin America: Evidence From a Supply-Demand Framework. *National Bureau of Economic Research*, No. w24015 . DOI: 10.3386/w24015

Gamboa, L., & Rodríguez, P. (2018). Ingresos subjetivos y expectativas académicas de la educación terciaria en Colombia. *Ensayos Sobre Politica Economica*, *36*(86), 159-177. https://doi.org/10.32468/espe.8601

García, S., Rodríguez, C., Sánchez, F., & Bedoya, J. (2015). La lotería de la cuna: La movilidad social a través de la educación en los municipios de Colombia . *Universidad de los Andes-Documento CEDE. No.* 013816.

Gasparini, L., Galiani, S., Cruces, G., & Acosta, P. (2011). Educational Upgrading and Returns to Skills in Latin America: Evidence from a Supply-Demand Framework, 1990-2010. *The World Bank*. https://doi.org/10.1596/1813-9450-5921

Glaeser, E., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2004). Do Institutions Cause Growth?. *Journal of Economic Growth*, *9*(3), 271-303. https://doi.org/10.1023/B:JOEG.0000038933.16398.ed

Guarín, A., Londoño, S., Medina, C., Parra, J., Posso, C., & Vélez, C. (2016). Estimating the Effect of Attending a Public versus a Private University in Colombia on Academic Achievement. *Borradores de Economía; No. 968*. Banco de la República. https://doi.org/10.32468/be.968.

Gunderson, M., & Oreopolous, P. (2020). Returns to Education in Developed Countries. *The Economics of Education* (pp. 39-51). Academic Press. https://doi.org/10.1016/B978-0-12-815391-8.00003-3

Heckman, J., & Li, X. (2004). Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000. *Pacific Economic Review*, 9(3), 155-171. https://doi.org/10.1111/j.1468-0106.2004.00242.x

Heckman, J., Lochner, L., & Todd, P. (2008). Earnings Functions and Rates of Return. *Journal of Human Capital*, 2, 1–31.

Heckman, J., Humphries, J., & Veramendi, G. (2018). Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking. *Journal of Political Economy*, *126*(S1), S197-S246. https://doi.org/10.1086/698760

Heckman, J., Tobias, J., & Vytlacil, E. (2000). Simple Estimators for Treatment Parameters in a Latent Variable Framework with an Application to Estimating the Returns to Schooling. *National Bureau of Economic* No. w7950. https://doi.org/10.1162/003465303322369867

Heckman, J. J. and Vytlacil, E. (1999). Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects. *Proceedings of the National Academy of Sciences*, 96(8):4730–4734. DOI: 10.1073/pnas.96.8.4730.

Heckman, J., & Vytlacil, E. (2000). The Relationship Between Treatment Parameters within a Latent Variable Framework. *Economics letters*, *66*(1), 33-39. https://doi.org/10.1016/S0165-1765(99)00181-0

Heckman, J., & Vytlacil, E. (2001). Policy-relevant Treatment Effects. *American Economic Review*, 91(2), 107-111. DOI: 10.1257//aer.91.2.107

Heckman, J., & Vytlacil, E. (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica*, *73*(3), 669-738. https://doi.org/10.1111/j.1468-0262.2005.00594.x

Heckman, J., & Vytlacil, E. (2007). Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments. *Handbook of econometrics*, *6*, 4875-5143. https://doi.org/10.1016/S1573-4412(07)06071-0

Henderson, D. J., Polachek, S. W., & Wang, L. (2011). Heterogeneity in Schooling Rates of Return. *Economics of Education Review,* 30(6), 1202-1214. https://doi.org/10.1016/j.econedurev.2011.05.002

Hertz, T., & Jayasundera, T. (2007). School Construction and Intergenerational Mobility in Indonesia. *American University, Washington DC Department of Economics Working Paper Series*, No. 2007-18.

King, J. E. (1996). The Decision To Go to College: Attitudes and Experiences Associated with College Attendance Among Low-Income Students.Washington, D.C.: The College Board.

Kugler, A., Kugler, M., & Herrera, L.(2017). Do Payroll Tax Breaks Stimulate Formality? Evidence from Colombia's Reform. *National Bureau of Economic Research,* No. 23308.. DOI: 10.3386/w23308

Kyui, N. (2016). Expansion of Higher Education, Employment and Wages: Evidence From the Russian Transition. *Labour Economics*, 39:68–87. DOI: 10.1016/j.labeco.2016.01.001.

Leslie, L., Johnson, G., & Carlson, J. (1977). The Impact Of Need-Based Student Aid Upon The College Attendance Decision. *Journal of Education Finance, 2*(3), 269-285.

Lucas Jr, R. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, *22*(1), 3-42. https://doi.org/10.1016/0304-3932(88)90168-7

Manacorda, M., Sánchez, C., & Schady, N. (2010). Changes in Returns to Education in Latin America: The Role of Demand and Supply of Skills. *ILR Review*, *63*(2), 307-326. https://doi.org/10.1177/001979391006300207

Mejía, J., Morales, L., & Medina, D. (2018). Trade Liberalization and its Effects on Labor Fluidity: Evidence From Colombia. *The International Trade Journal*, *32*(1), 43-75. https://doi.org/10.1080/08853908.2017.1389324

Morales, L., & Medina, D. (2016). Labor Fluidity and Performance of Labor Outcomes in Colombia: Evidence From Employer-Employee Linked Panel. *Borradores de Economía, No. 926*. https://doi.org/10.32468/be.926

Morales, L., & Medina, C. (2017). Assessing the Effect of Payroll Taxes on Formal Employment: The Case of the 2012 Tax Reform in Colombia. *Economía*, *18*(1), 75-124.

Patrinos, H., & Psacharopoulos, G. (2020). Returns to Education in Developing Countries. *The Economics of Education* (pp. 53-64). Academic Press. https://doi.org/10.1016/B978-0-12-815391-8.00004-5

Peet, E. D., Fink, G., & Fawzi, W. (2015). Returns to Education in Developing Countries: Evidence from the Living standards and Measurement Study Surveys. *Economics of Education Review*, 49, 69-90. https://doi.org/10.1016/j.econedurev.2015.08.002

Robinson, P.(1988). Root-N-Consistent Semiparametric Regression. *Econometrica*, 56, 931–954. [391]. DOI: 10.2307/1912705

Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5, Part 2), S71-S102.

Rodríguez, J., Urzúa, S., & Reyes, L. (2016). Heterogeneous Economic Returns to Post-secondary Degrees: Evidence From Chile. *Journal of Human Resources*, 51(2), 416-460. doi:10.3368/jhr.51.2.0213-5474R1

Topel, R. (1999). Labor Markets and Economic Growth, in *Handbook of Labor Economics*, Vol. 3C, 2943-2984. Ed. by Orley Ashenfelter and David Card. Amsterdam and New York: North Holland. https://doi.org/10.1016/S1573-4463(99)30035-3

Wang, X., Fleisher, B. M., Li, H., and Li, S. (2014). Access to College and Heterogeneous Returns to Education in China. *Economics of Education Review,* 42:78–92. DOI:10.1016/j.econedurev.2014.05.006.

Appendix

Appendix A
Parametric Estimation of the MTE

The parametric estimation requires assuming that $U_0, U_1, U_S \sim N(0, \Omega)$, and that this un-observed component are independent of X and Z. Note that Ω represents a variance covariance matrix such as:

$$\Omega = \begin{bmatrix} \sigma_0^2 & \sigma_{0,1} & \sigma_{0,U_s} \\ \sigma_{0,1} & \sigma_1^2 & \sigma_{1,U_s} \\ \sigma_{0,U_s} & \sigma_{1,U_s} & \sigma_{U_s}^2 \end{bmatrix}$$

where $\sigma_0^2, \sigma_1^2, \sigma_{U_s}^2$ represent variances of U_0, U_1, U_S, respectively, and $\sigma_{1,U_s}, \sigma_{0,U_s}$ represent covariance terms. This jointly normality assumption allows writing equation (6) as:

$$MTE(x, v) = X(\beta_1 - \beta_0) - \left(\frac{\sigma_{1,U_s}}{\sigma_{U_s}} - \frac{\sigma_{0,U_s}}{\sigma_{U_s}}\right)\Phi^{-1}(v) \qquad (A1)$$ [14]

All parameters $\beta_1, \beta_0, \sigma_{1,U_s}, \sigma_{U_s} = 1, \sigma_{0,U_s}$ in equation (11) come from the jointly estimation of the following set of equations in a switching regression model:

$$W_1 = X\beta_1 + U_1 \quad \text{(A2)}$$

$$W_0 = X\beta_0 + U_0 \quad (A3)$$

$$P(S = 1) = F_{U_s}(Z\gamma) \quad (A4)$$

[14] This result comes from the following property for conditional expiation of jointly normal distributed random variables.

Appendix B: Summary Statistics

Table B1

Variable	Obs	Mean	Std. Dev.
Ln(Wage)	542,313	13.31207	0.7031794
Some college	542,313	0.0905602	0.286983
high school exit exam score	542,313	230.9019	31.44234
Household Income at the time when student took high school exit exams.			
between 1 and under 2 minimum wages	542,313	0.4480383	0.4972931
between 2 and under 3 minimum wages	542,313	0.1916347	0.3935875
between 3 and under 5 minimum wages	542,313	0.0949599	0.2931598
5 or more minimum wages	542,313	0.0376111	0.190254
Parent's education at the time when student took high school exit exams.			
Elementary	542,313	0.5258532	0.4993316
Secondary	542,313	0.3068302	0.4611787
Technical and technological	542,313	0.0372147	0.1892876
Higher	542,313	0.0760871	0.2651376
Age	542,313	28.01142	2.673656
Age squared	542,313	791.7882	150.9258
Gender	542,313	0.5233398	0.4994554
Unemployment rate at exit exam	542,313	11.2655	1.471908
UBN at exit exam	542,313	38.54664	15.53801
Supply of tertiary education in the municipality and at the time student took their high school exit exams.			
Number of institutions of tertiary education	542,313	26.54091	38.21171
Number of students	542,313	0.6113429	0.817438
Number of students squared	542,313	1.041944	1.92109
Log high school average score in the exit exam	542,313	229.8173	17.84238
Average UBN	542,313	23.29038	14.84891
Average Unemployment rate	542,313	17.06429	2.487913

Notes: Authors' calculations. The estimation sample consists of a cohort of individuals who graduate from high school during 1994 to 2006. Information of households at the time of exit exams comes from administrative records from the Colombian Ministry Education. Information of wages comes from administrative records from the Social Security (PILA). Unemployment rate and the Unsatisfied Necessities Index (UBN) comes from census information

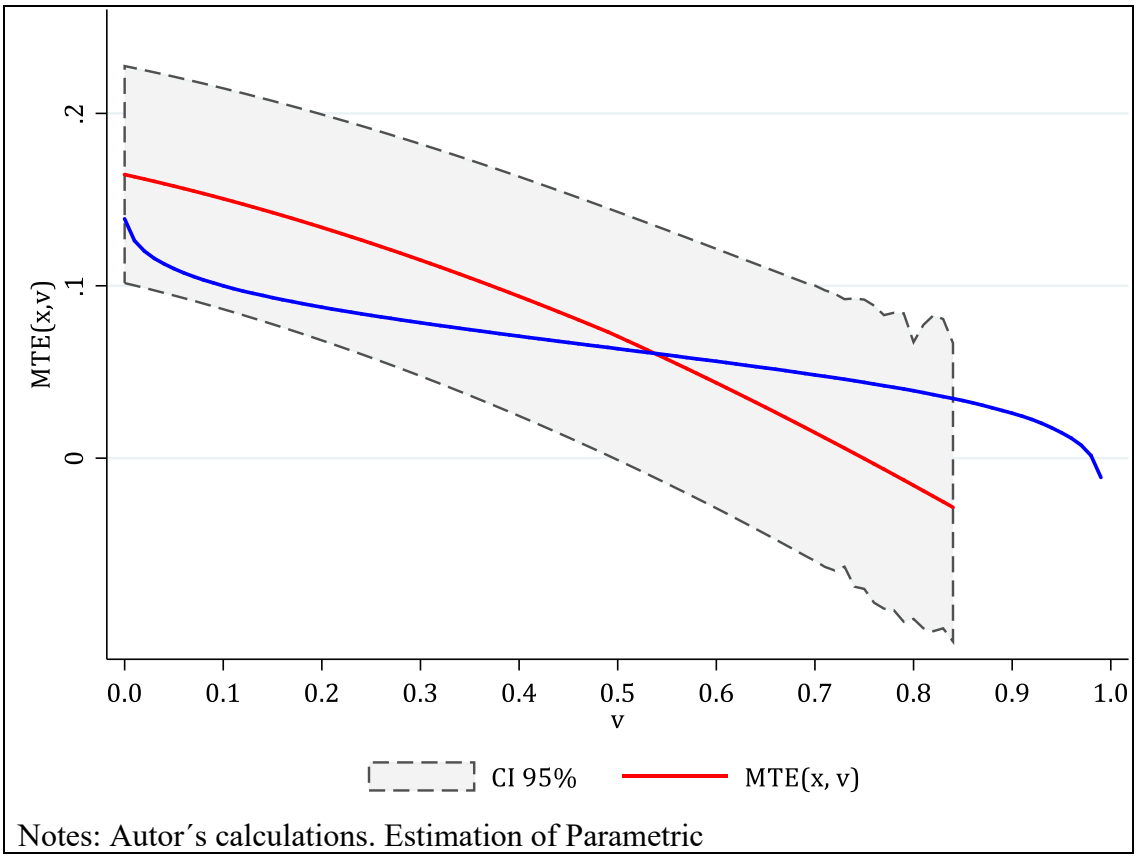
Appendix C: Parametric estimation

	(1) Ln(wage) with college	(2) Ln(wage) no college	(3) Selection Equation	Variance-Covariance
High school exit exam score	0.0030***	0.0039***	0.0110***	
	(0.0000)	(0.0001)	(0.0001)	
Household Income at the time when student took high school				
Between 1 and under 2 minimum wages	0.0395***	0.0750***	0.2508***	
	(0.0024)	(0.0119)	(0.0082)	
Between 2 and under 3 minimum wages	0.0726***	0.1282***	0.4342***	
	(0.0031)	(0.0130)	(0.0091)	
Between 3 and under 5 minimum wages	0.0751***	0.1828***	0.5346***	
	(0.0042)	(0.0143)	(0.0105)	
5 or more minimum wages	0.0552***	0.1824***	0.5386***	
	(0.0066)	(0.0167)	(0.0137)	
Parent's education at the time when student took high school				
Elementary	0.0271***	0.0205	0.1601***	
	(0.0042)	(0.0215)	(0.0152)	
Secondary	0.0503***	0.0731***	0.3487***	
	(0.0045)	(0.0219)	(0.0154)	
Technical and technological	0.1009***	0.1445***	0.5673***	
	(0.0072)	(0.0242)	(0.0184)	
Higher	0.0609***	0.1206***	0.5567***	
	(0.0058)	(0.0233)	(0.0168)	
Age	0.1034***	0.2456***		
	(0.0068)	(0.0274)		
Age squared	-0.0017***	-0.0043***		
	(0.0001)	(0.0005)		
Gender	0.0323***	0.0202***	-0.3393***	
	(0.0019)	(0.0073)	(0.0053)	
Supply of tertiary education in the municipality and at the time student took their high school exit exams.				
Number of institutions of tertiary education			0.0054***	
			(0.0004)	
Number of students			-0.5324***	
			(0.0265)	
Number of students squared			0.1188***	
			(0.0060)	
Log high school average score in the exit exam			0.0023***	
			(0.0002)	
UBN at exit exam			0.0034***	
			(0.0002)	
Unemployment rate at exit exam			0.0678***	
			(0.0011)	
Lagged school average score	0.0019***	0.0019***		

	(0.0001)	(0.0001)		
Average Unemployment rate	-0.0127***	-0.0333***		
	(0.0007)	(0.0024)		
Average UBN	-0.0005***	-0.0026***		
	(0.0001)	(0.0002)		
Constant	10.6187***	9.0003***	-6.0994***	
	(0.0942)	(0.3861)	(0.0440)	
r0				-0.0487***
				(0.0032)
r1				0.1296***
				(0.0195)
T test (r0-r1)=0				8.144
Observations	542,313	542,313	542,313	542,313

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
Column (1) presents estimation of equation one (see equation A2, appendix A) for individuals who enrolled into college (S=1). Column (2) presents estimation of equation one (see equation A3, appendix A) for individuals who did not enroll into college (S=0). Column (3) presents the estimation of equation 2 (see equation A4, appendix A), the selection equation. The entire system is estimated by maximum likelihood assuming a distribution of errors as described in Appendix A. We use heteroscedasticity-consistent standard errors.

Appendix D: Semi-Parametric and Parametric MTE Estimation



Notes: Autor´s calculations. Estimation of Parametric

Appendix E: Heterogeneous evaluation parameters by levels of experience in the labor market.

Evaluation parameters 2007 Sample

Treatment Parameter	Estimated Parameter
Average Treatment Effect (ATE)	0.070451
	(0.0000485)
Average Treatment on the Treated Effect (ATT)	0.126354
	(0.0000851)
Average Treatment on the Un-treated Effect (ATU)	0.066688
	(0.0000429)
Policy Relevant Treatment Effect (PRTE)1	0.085384
	(0.0002746)
Policy Relevant Treatment Effect (PRTE)2	0.085678
	(0.000266)

Notes: All the treatment parameters and policy parameters are estimated as indicated in equations 7, 8, 9, and 10 using simulation methods. Standard errors are estimated using bootstrap methods for 250 replications. The returns are annualized by dividing the total estimated return by 4, the mean duration of a tertiary education program in Colombia. The 2007's parameters were estimated only with workers with at least two years of experience in the labor market. The 2006's parameters were estimated only with workers with at least three years of experience in the labor market.

Evaluation parameters 2006 Sample

Treatment Parameter	Estimated Parameter
Average Treatment Effect (ATE)	0.0788319
	(0.0000669)
Average Treatment on the Treated Effect (ATT)	0.1367783
	(0.0001193)
Average Treatment on the Un-treated Effect (ATU)	0.0762745
	(0.0000664)
Policy Relevant Treatment Effect (PRTE)1	0.095248
	(0.000251)
Policy Relevant Treatment Effect (PRTE)2	0.0956575
	(0.0002411)

Note: All the treatment parameters and policy parameters are estimated as indicated in equations 7, 8, 9, and 10 using simulation methods. Standard errors are estimated using bootstrap methods for 250 replications. The returns are annualized by dividing the total estimated return by 4, the mean duration of a tertiary education program in Colombia. The 2006's parameters were estimated only with workers with at least three years of experience in the labor market.

Evaluation parameters 2005 Sample

Treatment Parameter	Estimated Parameter
Average Treatment Effect (ATE)	0.0740044
	(0.0000457)
Average Treatment on the Treated Effect (ATT)	0.1436674
	(0.0001471)
Average Treatment on the Un-treated Effect (ATU)	0.0720276
	(0.0000479)
Policy Relevant Treatment Effect (PRTE)1	0.0925832
	(0.0002527)
Policy Relevant Treatment Effect (PRTE)2	0.0932288
	(0.0002482)

Note: All the treatment parameters and policy parameters are estimated as indicated in equations 7, 8, 9, and 10 using simulation methods. Standard errors are estimated using bootstrap methods for 250 replications. The returns are annualized by dividing the total estimated return by 4, the mean duration of a tertiary education program in Colombia. The 2005's parameters were estimated only with workers with at least four years of experience in the labor market.


BANCO DE LA REPÚBLICA
COLOMBIA