Another Brick in the Wall: *

The Economic Consequences of Setting Foot in a College in Colombia

Author:

Luis Omar Herrera-Prada†

Ibero-America Institute for Economic Research

Universität Göttingen

International Monetary Fund

February 1, 2021

Keywords: Returns to education, Sheepskin effect, Higher Education, Education, Graduation, Gradiation on time

JEL Classification Numbers: I23, I24, I26, J24, J31

*I am thankful to Inmaculada Martínez-Zarzoso, Sarah Belford, Juan Esteban Saavedra, Andrés Córdoba, Pablo Druck, Óscar Bernal, Paula Triana, Diana Durán, Luz Emilse Rincón, Clara Pava, Pelin Berkmen, Dmitry Gershenson, Julia Faltermeier, Sidonia Mckenzie, Hannah Winterberg, SPADIES team at Universidad de los Andes (2005-2011), Ministry of Education of Colombia, Ministry of Health and Social Protection of Colombia, and all the very useful comments received in the Development Economics Seminar - Universität Göttingen; WHD Seminar - International Monetary Fund., Public Policy Seminar - Universidad de los Andes (Bogota-Colombia), Jornada NIP 2020-Universidad de la Republica (Montevideo-Uruguay). An online appendix with more detailed charts for labor supply and school premium using other characteristics is available in `https://herreraprada.com` . All views and errors are mine.

†Corresponding Author: l.herreraprada@stud.uni-goettingen.de

Abstract

Using administrative data, I track the path of all the secondary school graduates in Colombia from 2002 to 2012 that enter higher education and/or the formal labor market (5.4 million graduates). I compare graduates within the same secondary school and cohorts to estimate the premium of higher education. I estimate the sheepskin effect by exploiting the phenomenum of students who enrolled in the labor market after finished 90% or more of the college course-work but did not graduate and comparing them against workers that did earn a bachelors degree. Using a modified Mincer equation, I find that the Colombian labor market values a college graduate at the time of graduation the same as a secondary school graduate with five years of formal labor market experience. I also find high positive correlations between the quality of higher education institutions and students' skills and earnings, and between on-time graduation and earnings. High-quality higher education institutions boost the entry-level salary for their graduates, but this boost fades over time as others gain experience and the graduates' skills as workers are revealed. I find evidence that higher education is slowly reducing the gender income gap and improving income distribution in Colombia. Finally, the sheepskin effect is about 12.6% on average and the returns for bachelors, diplomas, and masters are 15.1%, 33.6%, and 53.2%, respectively.

1 Introduction

Every semester, secondary school seniors face what is perhaps the most crucial decision of their lives: to step or not to step into higher education. Several studies analyze the advantages and disadvantages of this decision, its consequences, the future income, and working life of these students; but, no study so far has been able to longitudinally analyze the total population of secondary graduates in a country over a long enough period to examine the impact of higher education on their lives.

In Colombia, several authors have analyzed the returns of education and the sheepskin effect[1] on the labor market(Rodríguez, 1981; Psacharopoulos, 1985, 1994; Tenjo, 1993; Mesa and Gutiérrez, 1996; Núñez and Sánchez, 1998; Cárdenas and Bernal, 1999; Santamaría, 2001; Arias and Chávez, 2002; Mora, 2003; Zárate, 2005; Arango et al., 2005; Prada, 2006; Farné, 2006; Posso, 2008; Forero and Ramírez, 2008; Hernández, 2010; Herrera-Prada and Caballero, 2013; Galarza et al., 2015; González-Velosa et al., 2015; Busso et al., 2020). The analysis advanced in the country at the pace that the sophistication of the data allowed. Before 2006, the authors based their studies on household surveys with inaccurate questions, and after 2006 with a new impressive information system created by the Ministry of Education. The Ministry's new data solved problems that household surveys could not address, but papers with that data used samples.

This document uses a set of administrative databases to generate a novel database that follows the lives of 5.4 million students who finished secondary school from 2002 to 2012 during their college life (if they had one), training in non-formal education (if they had one), and then their entry into the formal Colombian labor market.

Perhaps the greatest advance of this research is the ability to measure the impact of enrollment in higher education among the closest peers in high school. In the case of the impact of higher education, it compares the paths of schoolmates in the same generation. In the case of the higher education degree, the comparison is made with students who after 2 semesters of complete more than 90% of all their studies course-work did not receive the degree; even comparing within the same higher education institution.

[1]The sheepskin effect is a phenomenon when people with an academic degree earn a higher income than those with an equivalent schooling level but without the credential. It was analized for first time by Hungerford and Solon, 1987.

The main results show that if the student graduates from higher education, each year of study is valued by the labor market as one year of experience, confirming the results obtained by Jaeger and Page (1996). The household income is less determinant than skills for future earnings. These results are consistent with the relevance of the skills found by Farber and Gibbons (1996), but contrary to their findings, this paper presents evidence of relevance between experience and wages. The unknown information about the student's real skills starts to be revealed once the worker is hired. This revelation explains the growth in salaries over time for those who achieve a higher education degree and experience.

The student's skills and the quality of the higher education institution are the more relevant determinants of a student's future income in my study. As in Epple et al. (2006), an equilibrium between individual skills and a high-quality college generates a very high probability of success in terms of the future earning potential. My results show a positive correlation between reputation and earnings growth, suggesting that reputation of college matters, according to MacLeod et al. (2017). Graduating from a certified higher education institution sends a signal to the market which is reflected in an increase in the labor market entry-level salary. This increase fades over time, once the student's skills are known as in Farber and Gibbons (1996).

For women and low-income students, premiums are given by the mere fact of being enrolled in higher education. The gender gap in returns is 5.3% in favor of men, although women have a higher premium than women who did not enroll in higher education. Female college graduates earn higher premiums and faster than male college graduates. These are signals of slow but certain improvements in the reduction of the gender gap. In some cases of socioeconomically disadvantaged groups, the mere fact that they have pursued a higher education degree represents an improvement over their peers in terms of income from 5 to 10 years after secondary school. This result provides evidence of redistributive improvements due to the enrollment in higher education. Students with high incomes or high skills who fail to complete higher education are punished by wage differentials with their peers who did not go on to higher education.

In general terms, the returns to education and the sheepskin effect are very close, if not equal, since entering higher education and not graduating does not have significantly different results from those who never went to higher education but work in the formal labor market. What is clear is that graduating from higher education, and even better, doing it on time, boosts the income of

those who accomplish it regardless of their socioeconomic characteristics.

In many cases, institutions or society impose requirements on students that keep them away from the degree, putting another brick in the wall and preventing them from having a better future. The degree and better if it is on-time will improve the income and open doors to new education levels for graduates. These new levels will also improve productivity and the human capital of society.

The following section presents the Colombian context; Section 3 describes the data; Section 4 discusses the identification strategy; Section 5 presents the results; section 6 provides conclusions and policy implications.

2 Context

Beginning in the 1990s, a number of papers have examined the performance of workers in the Colombian labor market according to their educational level. These papers explored the demand side of the labor market from 1976 to 2000 by disaggregating various characteristics, including relative wages, gender, level of education, and other salary determinants(Tenjo, 1993; Mesa and Gutiérrez, 1996; Núñez and Sánchez, 1998; Cárdenas and Bernal, 1999; Santamaría, 2001; Arias and Chávez, 2002; Arango et al., 2005). Another set of papers studied the labor supply side of the Colombian labor market. These papers estimated the returns of education, including the sheepskin effect by sector, level, the field of knowledge (Rodríguez, 1981; Psacharopoulos, 1985, 1994; Mora and Muro, 2008; Zárate, 2005; Prada, 2006; Posso, 2008; Herrera-Prada and Caballero, 2013; González-Velosa et al., 2015; Busso et al., 2020). All these studies used household surveys as their primary data source. In summary, these authors found that wages for Colombian workers increased primarily due to an improvement in relative demand. However, they could not determine what level of qualification triggered better salaries. These papers faced the same problem: the poor capacity of the household surveys to provide rich indicators (Farné, 2006).

Unbeknownst to them, at the same time, the Ministry of Education in Colombia was starting a program that would solve this problem. Beginning in late 2004, the Ministry of Education promoted a massive program to collect individual data on students in higher education. Colombia´s higher education database (SNIES) is now one of the biggest education data information systems in the

world. The database has 4 components: SNIES (the main program, focused on information of the higher education system), SPADIES (the anti-dropout program), OLE (the program that tracks graduates into the labor market), and SACES (the program for system quality improvement). This novel database will become the main source of abundant literature on higher education in Colombia.

One of the first papers to use the OLE database was (Forero and Ramírez, 2008). They estimated the returns of education using follow up surveys from OLE. However, the main variable used for the returns to education was a poor measure. The survey's question on household income captured only a range of income, not the specific level. At that point in time, the OLE survey's sample was small and likely biased. The Ministry of Education continued to develop the database. By 2010, the OLE database included individual social security records for all graduates from higher education institutions (from the Ministry of Labor). This was a huge improvement in the analysis of returns to higher education and it was exploited by Hernández (2010); Herrera-Prada and Caballero (2013); González-Velosa et al. (2015); Busso et al. (2020). However, the OLE data still had several structural problems, some of them noted by Hernández (2010), which still exist today:

1. OLE does not include the self-employed. Self-employed represents between 23% of colombian formal labor force[2]. As it was not mandatory to contribute to the Social Security system in Colombia before 2008, information on the self-employed was not reported in the database used by Hernández (2010). The self-employed are a crucial pillar of the labor market; they were a dynamic force driving labor market trends from 2002 until the labor reform in 2012. Not only self-employed but also the informal workers are missing from the OLE data. My analysis includes the self-employed. Note that to be consistent with historical data, the Ministry of Labor does not report data on the self-employed to OLE.

2. OLE can only track higher education' graduates; it compares programs, levels, and areas of knowledge. This document has a database capable of doing the same thing as OLE and more. This paper tracks all students who completed secondary school since 2002, whether or not they went on to higher education, and who enrolled in the formal labor market. I can measure the real impact of higher education by comparing secondary school students who had different

[2]This number depends on the definition used. According to (OECD 2020), the figure is 52.1%. However, both the household survey and the social security data register 23% (see A1 for some stats from the household survey used in this study).

paths. In addition, I can estimate a more accurate sheepskin effect by comparing secondary school graduates with those who met all the requirements but left post-secondary education before obtaining a degree.

3. OLE may have a biased report in the 2006 dataset. From 2004 to 2007, data management at the personal level was archaic in almost all Colombian higher education institutions. Therefore, data collection for the first waves of any of the 4 information systems of the Ministry of Education (mentioned above), at that time, it was focused on those institutions with computerized data management; usually the large traditional public institutions of higher education and the high-income private ones. For example, the first wave for the SPADIES program in 2006 had only 33 higher education institutions Herrera-Prada (2020). Surprisingly, several higher education institutions had all their data on paper and could not be included in the systems until late 2008 or early 2009. In 2008, SPADIES became the source of data for OLE and part of SNIES. In 2010, all new non-SPADIES data were centralized in SNIES. The systems worked in parallel to audit the reports from higher education institutions and complement each other's dataset. This document contains the universe of students who were enrolled in higher education in the 296 higher education institutions in the system for 20 years (1998-2017); SNIES began collecting the student data at the individual level in 2007.

3 Data

In 2018, the Colombian higher education system had 2.4 million students: 93% as undergraduate, 4% in diplomas, 3% in master's degrees; less than 0.2% in PhDs. They are 51% in the public sector, and 53% are women. The system has 298 higher education institutions, 10,990 programs; an attendance rate of 52%; 52 higher education institutions with high-quality certification; 162,209 professors -only 8.5% with doctoral degrees- (observatorio de la universidad colombiana, 2020). In March 2020 -pre COVID-, the labor market had 20.5 million workers, an employment rate of 51.7%, and an unemployment rate of 12.6%.

This study uses a collection of administrative datasets that allow me to track secondary school graduates into higher education and/or the formal labor market sector. I use the same algorithm

that the Ministry of Education of Colombia uses to merge the databases with SPADIES[3]. The secondary exit exam test (ICFES dataset) provides the information of all students that graduated from secondary school between 2002 and 2012 in Colombia (5,425,850 million students). I also use information from the students' anti-dropout system database (SPADIES), which collects the information of all those that were enrolled between 1998 and 2017 (7.2 million students). I found 2,764,503 (50.95% -similar to the attendance rate of 52% mentioned above-) students from ICFES in the SPADIES dataset. Finally, Then, I merged the ICFES database with the PILA using the individual's ID number. I use Social Security records (PILA) from 2008 to 2014 for the wage and labor supply (16.8 million workers). I found that 464,871 (8.57%) students from ICFES were in the PILA database, and from those 156,759 (2.89%) students were found also in SPADIES dataset (Table 1). This mean than just one third of those reported in the formal labor sector were enrolled at least one period in the higher education system. This small percentage could be easily explained by the age and the high amount that are still studying. This also restrict the analysis for the informal sector, as the imputation of data from other sources cause disbalances in time. This document will explore the results for those in the formal sector and analize an upper bound with the informal sector imputation.

[Table 1 about here.]

3.1 ICFES

Colombian Institute for the Promotion of Higher Education (ICFES - by its acronym in Spanish) administers the Saber 11 exam, which is required to graduate at the end of secondary education and it is presented every semester for students that are finishing the high school. The ICFES database provided information on student gender, school location (department and county), secondary school area (urban or rural), household income, school sector (public or private), and the Saber 11 exam score (total, and by math and verbal components). Complementary information about the secondary

[3]The algorithm receives from each source (main and using databases) two key variables: i) a sequence of alphabetic characters and ii) a number (usually the date of birth). In both databases, it decomposes the strings (from i and ii) into all possible combinations of these characters and compare them among databases. If the result of the comparison reaches a given level "triger", the observation is considered matched. The algorithm is conservative, in case of more than one potential matching option it does not execute the matching. The triger value used in this document is 98%, the same value used by the Ministry of Education in the SPADIES-ICFES match. The variables used for the process are full names and date of birth.

schools were added from the MEN's Census of schools 2016 using the ICFES's school id code to merge the data (15,000 schools). All schools were merged.

Because ICFES used different score range over time, I standardized it by assigning each student their percentile on the Saber 11 exam in the period they took the exam. This is the same process that the Ministry of Education uses in its systems. The Ministry of Education, however, does not standardize subscores such as mathematics or verbal. To capture this information, I created the percentile version based on subscores. To create the verbal variable, I used the reading and/or verbal scores on each test; if necessary, I added them and then estimated the percentile. To create the maths results score variable, I added the scores of the mathematical components (ICFES used to have two questions: one related to knowledge and one related to skills; if the test contained only one question, only that one was used for this estimation) and estimated the percentile with the result of this addition. For Total, Math, and Verbal, I created a dummy that takes the value of 1 if a student has a high score (above 66%) and 0 otherwise. The variable used as control in the regressions is each student's decile in the Total score (Ministry considers high score over 90 percent).

As household income was not collected for some periods. I imputed it by using the household income mode in the same school in other periods when it was collected. In case there are two or more options as a mode, the imputation took out the lower value. The income level is standardized by the ICFES into 9 ascending categories[4]. The average and the median are in the range between 2 and 3 minimum wages (category 2). Therefore, I created a dummy that thakes the value of 1 if a student is considered to have a high income (above category 2) and 0 otherwise. Finally, the gender is the sex reported by the student at the time of taking the Saber 11 exam.

3.1.1 Ministry of Education– Census of schools

For this paper, a secondary school is a business on a specific shift, not a building. If a secondary school is not operated by the government, it is considered a private school. A public sector building can be used as a secondary school by the government in the morning and rented to a private organization for use in the afternoon or evening. Public schools that are controlled by a private entity under a contract with the government are considered private. For the school zone, if the

[4]0 "[0.1] minimum wages" 1 "[1.2] minimum wages" 2 "[2.3] minimum wages" 3 "[3.5] minimum wages" 4 "[5.7] minimum wages" 5 "[7.9] minimum wages" 6 "[9.11] minimum wages" 7 "[11.13] minimum wages" 8 "[13.15] minimum wages" 9 "[15] minimum wages"

school zone is not explicitly defined as urban, it is considered rural. All reported mixed urban-rural or rural-urban categories are coded as rural.

3.2 Ministry of Education - SPADIES

The SPADIES database provided information about the higher education where the student enrolled, their identification information, the area of study, level of the program, timeframe of studies, and the status of the student in the system (e.g. dropout, graduated, or active). Definitions of all status categories are explained in the identification strategy chapter.

Students in math, natural science, and engineering schools are coded as STEM using a dummy that takes the value of 1 if the students belong to this programs.

The program level is University level if the student is pursuing a professional degree or technical degree in another manner. Some colleges offer technical programs that with extra coursework become a professional degree. Note that professional programs usually last 4-5 years and technology programs 2-3 years. To advance from technician to professional, the student must graduate from the technical program before beginning the extra coursework. For the purpose of this document, those students are considered in a technical program until they obtain a professional degree.

3.2.1 Ministry of Education - SNIES

SNIES database provides the information for the characteristics of the programs, higher education institutions (HEI), and educational levels. The term, level, and area of knowledge for each program were merged using the program identification code provided from SPADIES and SNIES. The characteristics of each HEI as the location, level, and sector were merged using the SNIES id code. The Ministry of Education uses a figure called "High Quality Accreditation" which qualifies each HEI in different subjects, after a long examination process, and if the requirements are met, the Ministry of Education grants the higher education institution a quality certificate. Those higher education institutions with the high quality certificate of the Ministry of Education until 2017 are named as certified higher education institutions. I created a dummy that takes the value of 1 if the student is enrolled in a certified HEI.

SPADIES only accounts undergrad students, but using the SNIES's information from 2007 to 2013, I complemented the historical data from SPADIES with the lastest educational level reached

by each student, it means that I was able to know if a student finished a higher level of education as diploma, masters, or phd. The matching between the ICFES and SNIES data was done using the same algorithm used in the case of the ICFES-SPADIES discussed above. With this process, I organized the graduates in the following categories (Descriptive statistics of SPADIES-SNIES in Table 2):

1. Bachelor's degree. This is the group of students who graduated from higher education but did not continue their studies in postgrade.

2. Diploma is a post undergraduate degree, but less than a master's degree; it is more flexible and has a shorter duration. It has not international recognition. At the local level, it is known as a specialization.

3. Master's degree.

4. Ph.D. Regarding this category, it is essential to clarify that very few students succeed in reaching this level of education and receive salaries in the time studied. Reaching the Ph.D. level implies at least 5 years in undergraduate studies and another 5 years in the doctoral program. Consequently, the students with a Ph.D. level that we find in the database have low salaries, and it is assumed that these salaries correspond to a partial job during their academic career. This helps to explain the results of the regression for this category. Time is not the only barrier to reach the level; the brain drain is high. A substantial number of students who are able to attain postgraduate levels of education continue their studies outside the country. In the case of master's degrees, many students return, and this situation is somewhat different for doctoral students.

[Table 2 about here.]

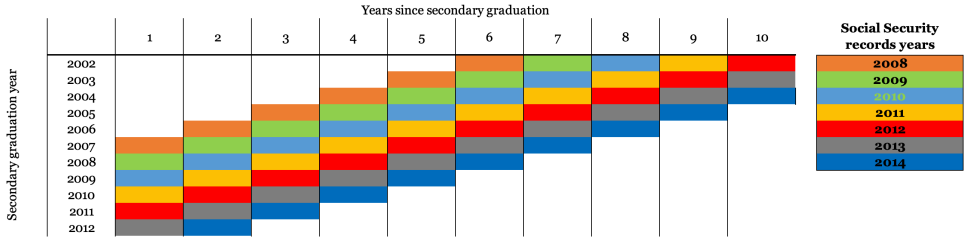
3.3 PILA

The PILA database contains the records of all payments to Social Security for every person in the formal sector, annual days worked, type of emplyment (including self employed), and type of employeer. This information is collected for the contribution of all formal Colombian workers to pension and health funds. The raw dataset is the full list of payments in a given year. Each payment

reported includes information on current salary, labor time reported for that payment, location of the payment, organization or individual making the payment, and economic activity of the worker. I calculated the salaries from the contribution to the health fund in each payment. Annual income is the sum of all contributions reported in the health care system during a calendar year. Salaries are deflated to 2013 values in colombian peso and converted to US dollars, then transformed to their logarithmic version. The days worked are reported at the time of contribution to Social Security. All payment reports are added, so it is possible to have people with more than 360 working days per year.

Since it is impossible to know the tenure for cohorts pre-2006. This paper uses the difference between the year of reporting on the PILA and the year of graduation from secondary school, which is determined from the year the student took the Saber 11 test, to estimate the years after secondary. So, each ICFES cohort contributes in different periods to the normalized variable (See Box 1).

Box 1. Cohorts distribution



Each column shows which cohorts are considered in the timing variable depending on the year of Social Security records.

Self-employed workers are identified using their type of contributor in the social security records. Self-employed workers, self-employed workers in association, self-employed workers without regulations to contribute, and other codes for transition from employee to self-employed (code 42 and code 49) were marked as self-employed. Public employees are identified by the code 3 of the variable for the type of employer. It is important to take into account that the State is a large employer, but not all its workers are registered in the PILA [5] (Descriptive statistics of PILA in Table 3).

[Table 3 about here.]

[5]Teachers and military are examples of groups not reported in the social security system

3.3.1 SENA

The SENA is the National Apprenticeship System (by its acronym in Spanish), it was founded in 1957 to promote job training at no cost to employers and reduce moral risk in the training of the workforce Medina and Saavedra (2011). However, although it was not considered formal education, the trainings became so tailored that they were adopted as a qualification for employment, in some cases structured as a formal education program, since many apprentices used the knowledge learned in the training as a source of income. SENA sends the apprentices who are completing the training to companies as interns. Their employers are required to pay for the social security of these apprentices and report them in the PILA report as SENA apprentices. The variable was created using this report. At this point, it is important to make two observations: The students who are marked as SENA are the few who arrive in the formal labor market, however they are the comparable ones because they go to the labor market under the same conditions as higher education students. We only know about those who are in the final stages of their training, and we do not have any information about the status of any other apprentice. The jobs of the apprentices have in many cases a market in the informal sector.

4 Methods

Conceptual Framework

There is abundant literature on the relationship between human capital formation and the benefits of education. The basic definition of human capital refers to intangible assets - such as education - that improve earnings, good habits, and health. Investment in these assets is crucial since these assets cannot be separated from their owner as a tangible asset can (Becker, 1962). According to this theory, schooling will increase human capital and, consequently, earnings based on the years of education received. This will happen until the decreasing marginal utility of education is outweighed by the opportunity cost of obtaining that education. This whole process is reflected in the labor market equation from the seminal contribution of (Mincer, 1974).

However, the literature quickly questioned the relationship between years of schooling and labor market earnings, as education is claimed to be a sign or even a filter. Employers are often in a

scenario with asymmetric information that makes it difficult to choose the right employee (Phelps, 1972; Arrow, 1973; Spence, 1973).

So, the degrees work as a quality signal, creating a shortcut in the hiring process. The signals are mutual, employees showing their credentials in response to companies' requests for better-prepared people to reduce uncertainty. In fact, employees try to show their skills to employers by getting an education. High salaries are paid to increase the opportunity cost of continuing schooling, particularly for those with high skills. The sense is that it is less expensive for highly skilled workers to prepare, so they move up the education ladder faster and further. The less skilled will soon realize that they cannot compete with the skilled, and then the opportunity cost of staying in the system is higher, dropping out or stopping at a certain level of education (Wood, 2009).

Under this scenario, the theory of credentials rose. It is argued that two individuals with exactly the same education but only one with the degree (credential) should receive the same salary. This theory predicts that salaries will rise faster if the individual has more education and holds the degree. This phenomenon is known as the sheepskin effect. The development of this theory can be divided into three parts: 1) from the late 1960s to the early 1980s, when researchers had access to individuals' basic background, years of education, and earnings; 2) in the 1980s and 1990s, databases included general information on individuals' background, years of education, earnings, and degree completion; and 3) after 2000, the new databases included more sophisticated information on individuals' background (including cognitive skills), more details on earnings and degrees, and more robust data.

The innovative work of this theory is (Hungerford and Solon, 1987) that uses a kind of sophisticated Mincer equation with discontinuity in the years they found significantly higher gains (8, 12, and 16 years of schooling) compared to the previous year. Later, the literature on the sheepskin effect had a great boom. A number of relevant papers found that the labor market incorporated years of schooling as experience, but the returns for each additional year of schooling were small compared to the significant difference in earnings between individuals with and without diplomas (Shabbir, 1991; Belman and Heywood, 1991, 1997; Jaeger and Page, 1996; Kane et al., 1999; Gibson, 2000; Ferrer and Riddell, 2002; Mora, 2003; Schady, 2003; Bauer et al., 2005; Crespo and Reis, 2009; Hernández, 2010; Bilkic et al., 2012; Son, 2013; Heckman et al., 2016; Yunus, 2017; Olfindo, 2018).

Status assignation and model

SPADIES brought a lot of advances to the Higher Education System in Colombia, the most important of which was the consolidation of the definitions for the status of students in the system. I made some adjustments to the status categories created for SPADIES so that my dataset would be more specific for the purpose of this document:

1. Graduated: SPADIES defines this status as a person who finished the coursework and got the degree from a higher education program. This paper divides this category into two groups: i) Graduated on time (those who graduated within one year of the expected time of graduation), and ii) Graduated late (those who graduated more than 1 year after their expected time of graduation). Graduated late and Graduated on time encompass all individuals that graduated from higher education. Expected graduation year was estimated as five years for professional programs and 3 years for tech.

2. Dropout: SPADIES defined dropout as a student that has not been reported in the system for 2 or more consecutive semesters. For the purpose of this paper, I made a small adjustment to this definition. I created a new group for those students who finished more than 90 percent of their coursework but are reported as Dropouts by SPADIES because they did not graduate after 2 or more consecutive periods without been part of the system. This sub-group is called "egresado no graduado" in the Colombian technical language, in this paper I will refer to them as "Candidates". The Candidates are the best counterfactual with which to the group of individuals who did graduate from higher education in Colombia. The remaining Dropouts (who completed less than 90 percent of their coursework)are named "incomplete". So, incomplete plus candidates are the total of dropouts.

3. Active: I retain the SPADIES definition for Active, which is any student still active and enrolled in the system as of 2017.

4. I took the apprenticeships from the SENA.

While an individual's status does change over time in the system, for the purpose of my study each individual is identified by their status in 2017 (latest available). For the regression model in its general form for the following equation, the status indicator takes the value of one (1) if the

individual is in the treated group and zero (0) if not. Outputs are the log of wages or the work supply measured as the days worked. Equation 1 is based in (Hungerford and Solon, 1987).

$$ln\left(Y\right)_{it} \quad = \quad \beta_0 + \beta_1 Status_i + \beta_2 Exp_{it} + \beta_3 Exp_{it}^2 + \phi_{it} + \vartheta_i + \varepsilon_{it}$$

Where, Status is a vector of dummies for each one of the statuses mentioned before as treatments. Experience is dynamic in time and it is measured as the difference between the year of apperance in the Social Security records and the year of graduation from secondary. Sex is a dummy for females. Score is the standarized score. Finally, ϑ comprends: a dummy that takes the value of 1 if the secondary school is located in an urban area, a dummy that takes the value of 1 if the secondary schools is a public school, and a dummy that takes the value of 1 if the student reported a household income higher than 3 minimum monthly wages. ϕ are the controls that change over time as self employment or public servant. Other non shown fixed effects are the year of graduation from secondary, and School, Department, or higher education institutions (depending on regression).

An extension for this model tests if it is just the degree is relevant or if the results of the sheepskin effect are due skills and experience. To do so, I added two new variables with the interaction of the score and the dummy of graduate, and the interaction of the years of experience and the dummy of graduate. In this case, as in the exercise of Farber and Gibbons (1996), the dependent variable is the level an not the logarithm.

$$Y_{it} \quad = \quad \beta_0 + \beta_1 Status_i + \beta_2 Exp_{it} + \beta_3 Exp_{it}^2 + GradXScore + GradXExp + \phi_{it} + \vartheta_i + \varepsilon_{it}$$

Finally, I analize the returns by year since graduation "j" using differents groups to compare results. Figures show the relevant coeficients β_{1j} for eq 3. Where, ϑ comprends controls by sex, percentile of secondary test score, and household income;. ϕ are the controls that change over time as self employment or public servant, regression controls also by high-school and year of graduation fixed effects.

$$\left\{ ln\left(Y\right)_{ij} = \beta_0 + \beta_{1j} Status_i + \phi_{ij} + \vartheta_i + \varepsilon_{ij} \right\} \quad \forall \quad j = 1...10$$

[Table 4 about here.]

5 Results

The main results for all students who finished secondary school in Colombia between 2002 and 2012 and went to the formal labor market are shown in Table 5. I find that the experience is marginally decreasing and for each year since secondary the salary increases by 20.6%; the gender gap is greater for men by 5%; private secondary school students have a premium of about 6% compared to public secondary school students; the annual income increases as the household income increases; for each score percentile on the secondary school test, the annual income increases by 0.2%. Finally, on one hand, students who have no get the degree have no premium, and those that are still active have negative returns compared to those who never make it to higher education. On the other, candidates have a 2% premium, but those who did earn a degree showed an 18.4% premium. This means that the effect of sheepskin in Colombia is about 16.5% (Table 5).

Let's assume the above results as an upper limit, now let me define the lower limit as the results of the same exercise but now comparing with the closest peer for each student, their classmates in high school. Column 2 reports the coefficients for the regression using high-school fixed effects. There are no major changes in family income, skills, gender, and experience. However, public schools are now positive and the effect of being in an urban area is stronger. Students whom drop-out of school continue to have negative results, while the graduated students showed no difference with those who never went to college. Finally, the graduate adjusted to 15.6%. Under these results, the effect of sheepskin in Colombia for higher education is about 12.6% (Table 5). We also learn that the returns for a Bachelor are 15.1%, a student with a Diploma degree are 33.6%, a student with a masters degree are 53.2%. Despite, I included the Ph.D. graduates in the regression, I do not consider Phds in the analysis as very few reach this level. Their results are not shown (Table 6).

[Table 5 about here.]

[Table 6 about here.]

5.1 GRAPHICAL ANALISYS

Now, to analyze the dynamics of the higher education premium per year since high school graduation, I will concentrate on the values to the right of the value 5 on the "X" axis, I will use this value as an approximation to the completion of the academic program. The values reported during college are lower than those observed in the control group basically because the students do not work full time (Figure 1). This situation also explains the delay in graduating, since they are students who study and work.

The analysis is made for three types of variables: i) those collected during the test such as: gender, score, household income, school area, and school sector; ii) those collected during reporting to social security ; iii) those collected during tertiary education such as: program level, higher education institutions quality, program area. For the first and second group, the comparison is against all the same groups that were not enrolled in higher education. For the third group, the comparison includes all remaining students. Thus, for example, women are compared to women who were not enrolled in higher education, but STEM is compared to all non-STEMs including those who never enrolled in higher education.

5.1.1 General results

During college, working students reported about 10% less income and 20% fewer days worked (Figure 1) than those who never went to college. I also found evidence of a couple of relevant facts:

1. The sheepskin effect begins to grow since the graduation from college, reaching about 12.5% for those with high skills. After 10 years from secondary school graduation, the sheepskin effect reaches 34% for those with high income; maybe because they were more able to reach a higher level of education after the bachelor degree. In all cases after 10 years from secondary, the sheepskin effect reaches values over 20%, except by the self-employees. (Table 9).

2. Only for those who graduated on time, four years in college equals four years of work without education, confirming the results obtained by Jaeger and Page (1996) (Figure 1). For late graduates, dropouts, or incomplete, it takes 7 years to reach the same income level as those

who did not enroll. For the second group, after year 7 only the graduates show a premium. While the labor supply is the same, there is no evidence of a significant difference in annual earnings between being enrolled in higher education and becoming a dropout, graduate, or active, and never having been enrolled in higher education.

3. Delayed graduate students may be working at the same time as their college, which explains the delay in obtaining a degree. Although they achieve the same income as those who never went to college in Year 7, the slope of their earnings is similar to that of on-time graduates, increasing each year the gap against students who never enrolled in or completed their college (Figure 1).

4. Graduates on time are having more income also because they are more likely to enroll in postgraduate programs.

5.1.2 Results of variables from ICFES test exam

Gender As for the gender, despite the fact that men's income is higher than women's, the premium for women is a little higher premium than men's, but not significantly different. In line with this, the sheepskin effect for women is smaller than for men. This is explained because women's incomes are higher than men's, even for those who never finished college (Figure 2).

There is no difference between women and men in the labor supply for those who completed higher education. For those who did not finish, men work less and women work the same as their colleagues who did not go to university. Women tend to report more days of work during college than men; this explains why women report more compensation than men in comparison to their control group (Online Appendix 1).

Skills In terms of skills, 10 years after finish secondary school, students with high scores reach a premium of 42%. In the case of students with a low score, the premium is 25.1% after 10 years from the secondary. The sheepskin effect at year 5 is 12.4% for highly skilled students and 1.9% for lowly skilled students. In year 10, these figures are 25.1% in students with a high score and around 24.8% in the case of students with a low score. Results are similar to the results of the performance only in math and only in verbal (Appendix 1 and Appendix 2). Note that returns for candidates

with high skills are the highest among all the categories for candidates.

Household Income Income is less important than skills for future premiums. However, it shows a particular redistributive characteristic. Students with low incomes but without a degree still have premiums, even though the sheepskin effect remains around 24.4%. A totally opposite picture is faced by students with high income but without the degree, all undergraduates have a penalty compared to their peers who did not enroll in higher education, at best there is no difference after 10 years between those who did higher education but without the degree and those who never went to tertiary school.

School sector Similar to what I found in the analysis of returns to education according to household income, in the case of the school sector, private school students have a penalty if they do not obtain a degree. The mere fact of having the degree represents about 33.3% of the returns and 32.8% for sheepskin effect for them in the year 10. In the case of public schools, obtaining the degree represents a higher income, but it is not very different from other statuses that also receive premiums in the long term.

School area The comparison between urban and rural areas reflects the fact that most schools are in urban areas. This explains why the results are similar to the analyses presented above for all students. However, the results for students in rural areas are similar to those for late graduates in urban areas, active students in rural areas have a higher reward than in urban areas, and the effect of sheepskin is less in urban than in rural areas.

5.1.3 Results of variables from PILA

Self-Employed One of the most important contributions of this paper is the ability to analyze the self-employed. In the case of this type of worker the results are frustrating. Self-employed graduates who graduate on time does not have any difference in their income after they receive the college degree compared to those who never went to tertiary education. Their situation in year 10 is not much different from those students who graduated late, their premium is 7.8% for graduates late and 11.5% for graduated on time, the lowest of all categories. This result is novel, but should be interpreted with caution. Although several studies show that the potential for development and

growth for the self-employed is complex, this category can also include many students who have an informal part-time job with an under-reporting of income in the formal sector.

5.1.4 Results of variables from SPADIES data

Quality of HEI The returns to education and the sheepskin effect are greater for those students who enrolled in the higher education institutions certified high quality. One particular behavior is that students working during their college years in high-quality institutions reported premiums compared to their peers. Being enrolled in these colleges sends a signal of quality to the marketplace that fades over time when students do not reach the grade. However, students at high-quality institutions only show penalties during the time of delay for active or late graduates, for the rest of the time it is a premium. This is particularly special, since no other feature shows premiums over long periods, even during college time.

Level of program One of the most interesting results emerged from the chart by levels of the program, University level students show considerably higher returns and sheepskin effects. The large difference compared to technology students easily explains why Colombia concentrates 70% of its higher education population at the professional level. By having a higher future salary, career choice at the time of entry into the system is mainly concentrated on future income. At the same time, the graph shows that there is no difference between those with technological degrees and students who did not enroll in higher education, particularly after finishing college. Also, the penalty of not obtaining the degree and becoming a candidate is the greatest among the other comparisons and analyses that are being presented. At this point, it is easy to make assumptions, but I would like to be more careful with this as it could be explained by a subreport in the Social Security records due to two main reasons: i) the main activity with a higher income is in the informal sector, which is not reported to minimize the payroll cost. ii) The brain drain to other countries; the Ministry mentioned in an international conference in 2008 that Colombia exports its technological graduates to countries like Canada and France.

Type of program An interesting behavior is found in STEM analyses. Students in the STEM programs have a boost after obtaining the degree, only one year after graduation they obtained

a premium of about 30% of the results with respect to their peers. However, this improvement remains at the same level the rest of the time of the analysis. Getting the degree on time is crucial if the student is enrolled in STEM programs. Unlike other careers, if the degree is delayed there is not much difference between states before the eighth grade.

Finally, let me summarize in the following table the main results of years 5 and 10 on graduate performance and the sheepskin effect.

[Table 7 about here.]

6 Conclusions

In this document, I show that the decision to go to college drastically improves an individual's life path, particularly for low-income or low-skilled students. Alternatives to formal education like *Apprenticeships* show similar outcomes or worse outcomes than those who never go to college. This can be explained because these students can be qualified labor in the informal sector. Also, the quality of the HEI is very relevant. Graduated or even being enrolled in certified colleges sends a signal of quality to the labor market. However, the boost received fades over time when students do not reach the degree or when the labor market learns the real skills of the student. Only for those who graduated on time, four years in college equals four years of work without education. For those who graduated late, it takes seven years to reach the same income level as those who did not enroll.

During college, working students reported about 10% less income and 20% fewer days worked than those who never went to college. Delayed graduate students may be working at the same time as their college education, which explains the delay in obtaining a degree. These support the results of Herrera-Prada (2013) that show the delay in graduation is explained by an increase in vulnerable population in the system and a trade-off between dropouts and lag in graduation. The opportunity cost of working and studying simultaneously was not correctly estimated in the literature. I find that these students earn less than the one who never went to higher education, but it is explained because they are working less time. However, the delay exacerbates the opportunity cost of continuing studying as the investment in education continues growing, and the future earnings after the graduation will grow slowly.

College education brings income redistributive improvements over their peers for the socially

disadvantaged who went to college and a reduction of the gender gap for the women who graduated from college. Low-income students or students from the public secondary schools that went to college have positive returns just by setting foot in a college, but high-income students or students from private secondary schools who went to college but did not graduate have no return or even they receive punishment in salary terms versus their peers. In gender terms, women have a higher premium compared to the women that did not enroll in higher education, and female college graduates earn higher premiums, and they earn them faster than male college graduates. There is no difference between women and men in the labor supply for those who have completed higher education. These are signals of slow but certain improvements in the reduction of the gender and social gap.

The greater the secondary school test score the greatest return in the future, but at the same time the greatest punishment. Students with high skills but no degree have the greatest penalty after high-income graduates. After 10 years, graduating from higher education represents a 33.2% bonus for late graduates and 42.1% for on-time graduates. The sheepskin effect in the high skilled students is the highest in year 5 for all the categories analyzed. This is explained by the penalty for no graduates.

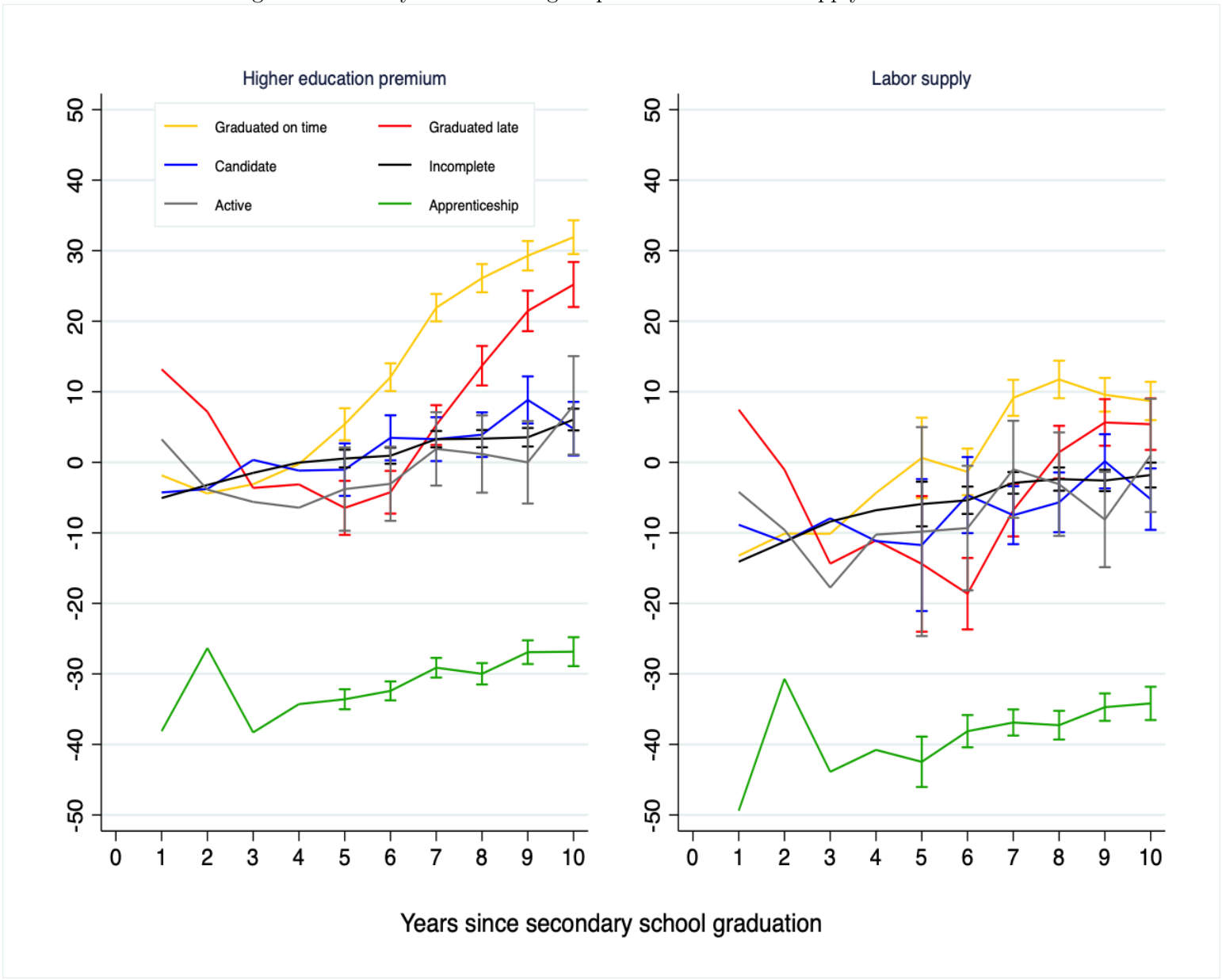
The quality of the higher education institutions is reflected in the long-term returns; no certified institution's students show higher returns in year 5. This can be explained by different events at the same time. No certified higher education institution has more in the transition to work program than certified higher education institutions; non-certified higher education institutions depend less on their prestige and need to promote their students in the labor market; students start to work and never come back to finish the program. In fact, this is one of the reasons that non-certified higher education institutions use to explain their dropout rate. If graduates are already working, the opportunity cost to complete the degree is higher, since at that time students do not realize that they will need the degree in the future and leave school and continue working. Arriving at a high-quality institution guarantees some bonus, even if the student does not get the degree. ¿signal or selection?. Well, this result shows a positive correlation between reputation and earnings growth, suggesting that reputation of college matters, according to MacLeod et al. (2017).

This document shows the benefits of pursuing a career in higher education in Colombia, mainly at the university level. It also highlights the income redistributive improvements over their peers

for the socially disadvantaged who attended college, the stagnation of the self-employed, and the great premiums for people with high cognitive skills. However, the paper cannot show whether the large differential generated in returns to education after a few years of the degree is due to a progression in the level of education after the degree, this issue is interesting for future research. What is clear is that graduating from higher education, and even better, doing it on time, boosts the income of those who accomplish it regardless of their socioeconomic characteristics. In some cases of socioeconomically disadvantaged groups, the mere fact that they have pursued a career in higher education represents an improvement over their peers in terms of income for the next 5 years. Colombia needs urgently to change/improve the understanding, the quality, and the articulation with the labor market for the community college and *Apprenticeship* programs. Colombia also needs to increase the college graduation rate at all levels, especially the on-time graduation rate, which would boost income for those who accomplish it regardless of their socioeconomic characteristics.

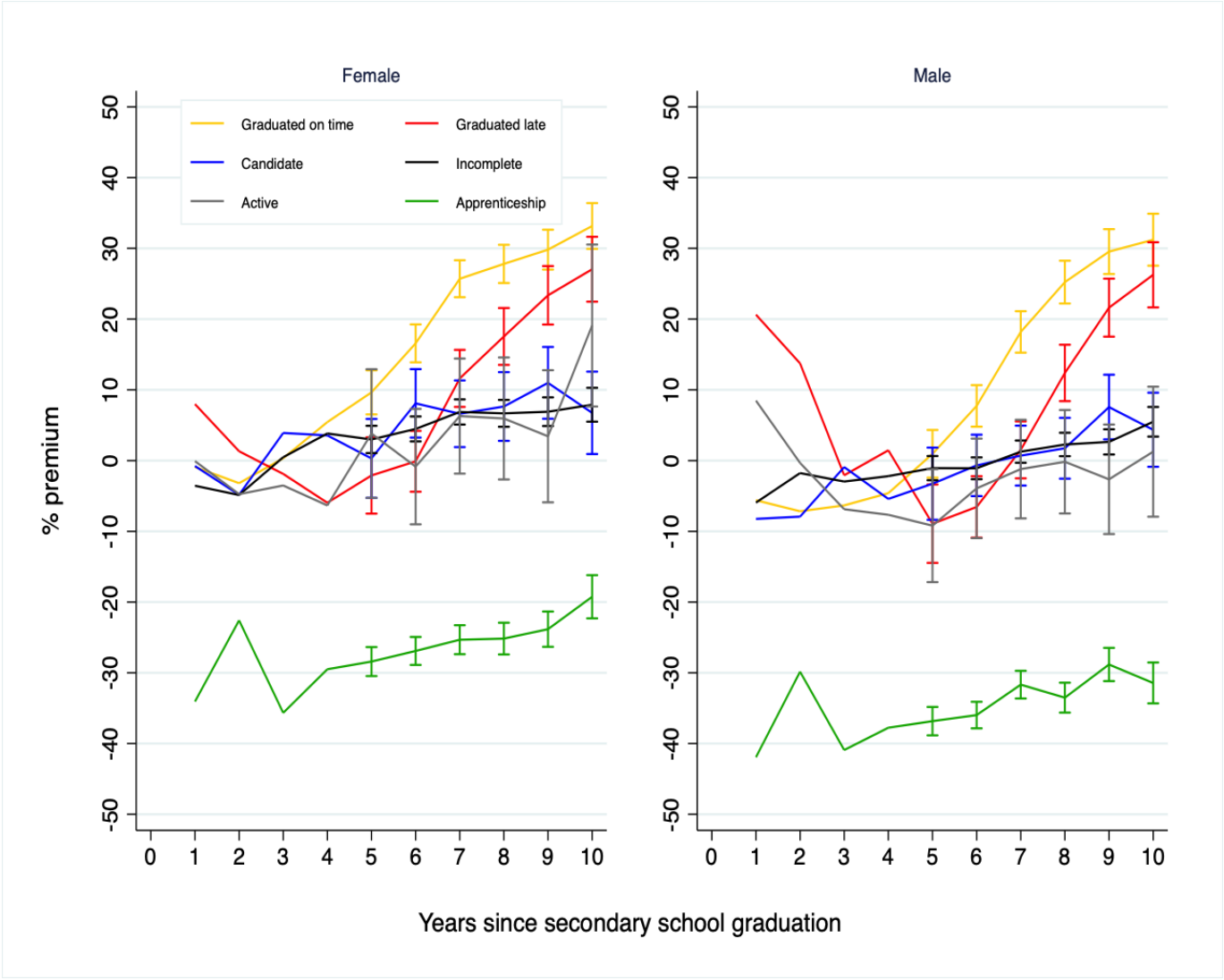
Finally, Colombia needs to reduce restrictions for graduation. In many cases, institutions or society impose requirements on students that keep them away from the degree, putting another brick in the wall and preventing them from having a better future. The country not only needs to increase the graduation rate to improve the quality of the workforce, but also to reduce dropout (Herrera-Prada, 2013), and also to increase student income (as shown in this paper). This is not only a problem for the Ministry, the higher education institutions can also evaluate if their degree requirements are affecting their students. The degree not only sends signals to the market or increases the knowledge and training of students, but it also opens the door to new levels of education that will further empower the student.

Figure 1: Tertiary education degree premium and labor supply in Colombia



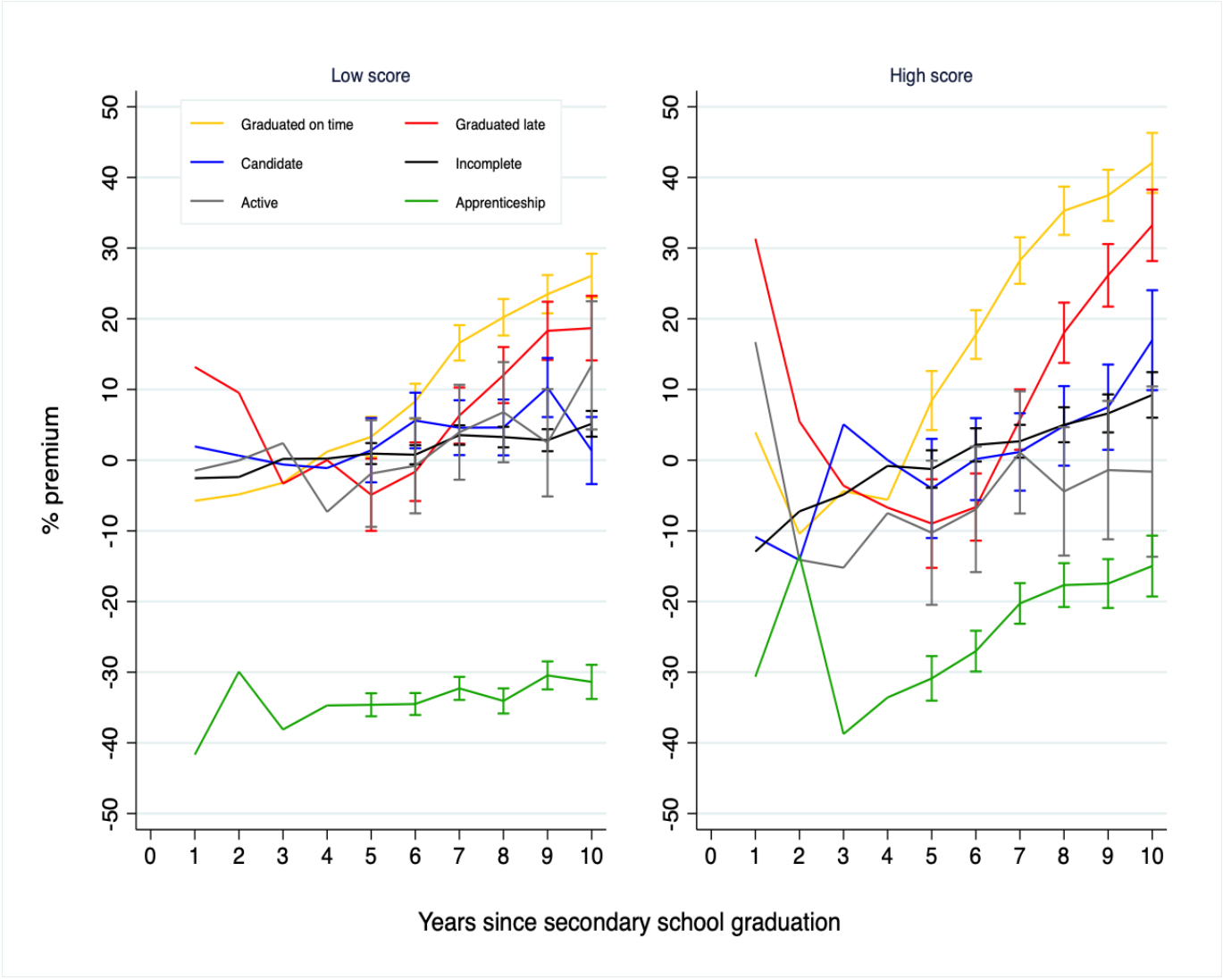
Notes: The figure shows the Mincer regression coefficients for returns to education by higher education status for late graduates, on-time graduates, incomplete, candidates and active students vs. secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Students who drop-out of school, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation from college is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of the year of graduation from secondary school and secondary school (Table 1 for descriptions).

Figure 2: Tertiary education degree premium in the formal sector in Colombia by gender



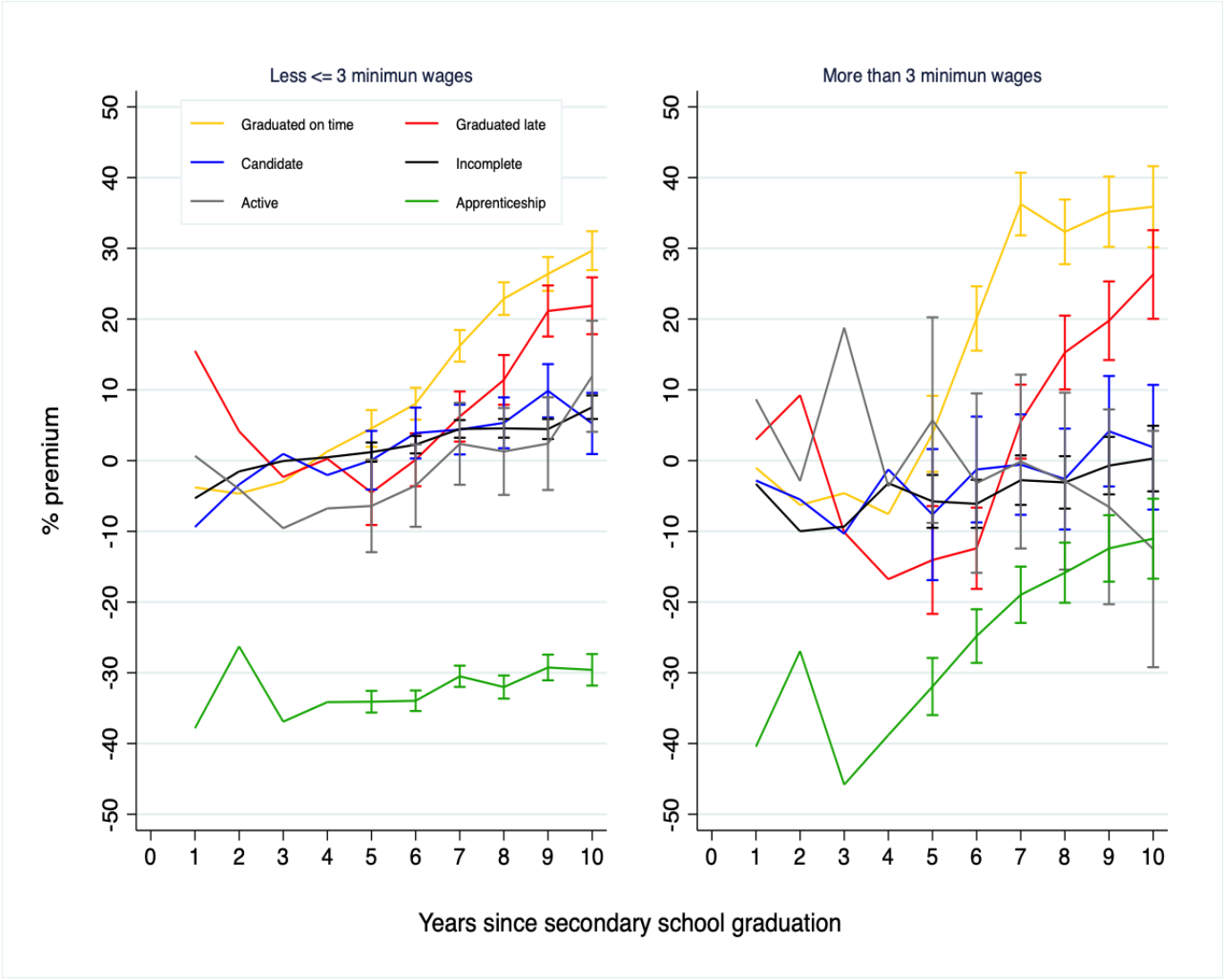
Notes: The figure shows the Mincer regression coefficients for returns to education by gender, according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions).

Figure 3: Tertiary education degree premium in the formal sector in Colombia by secondary test score



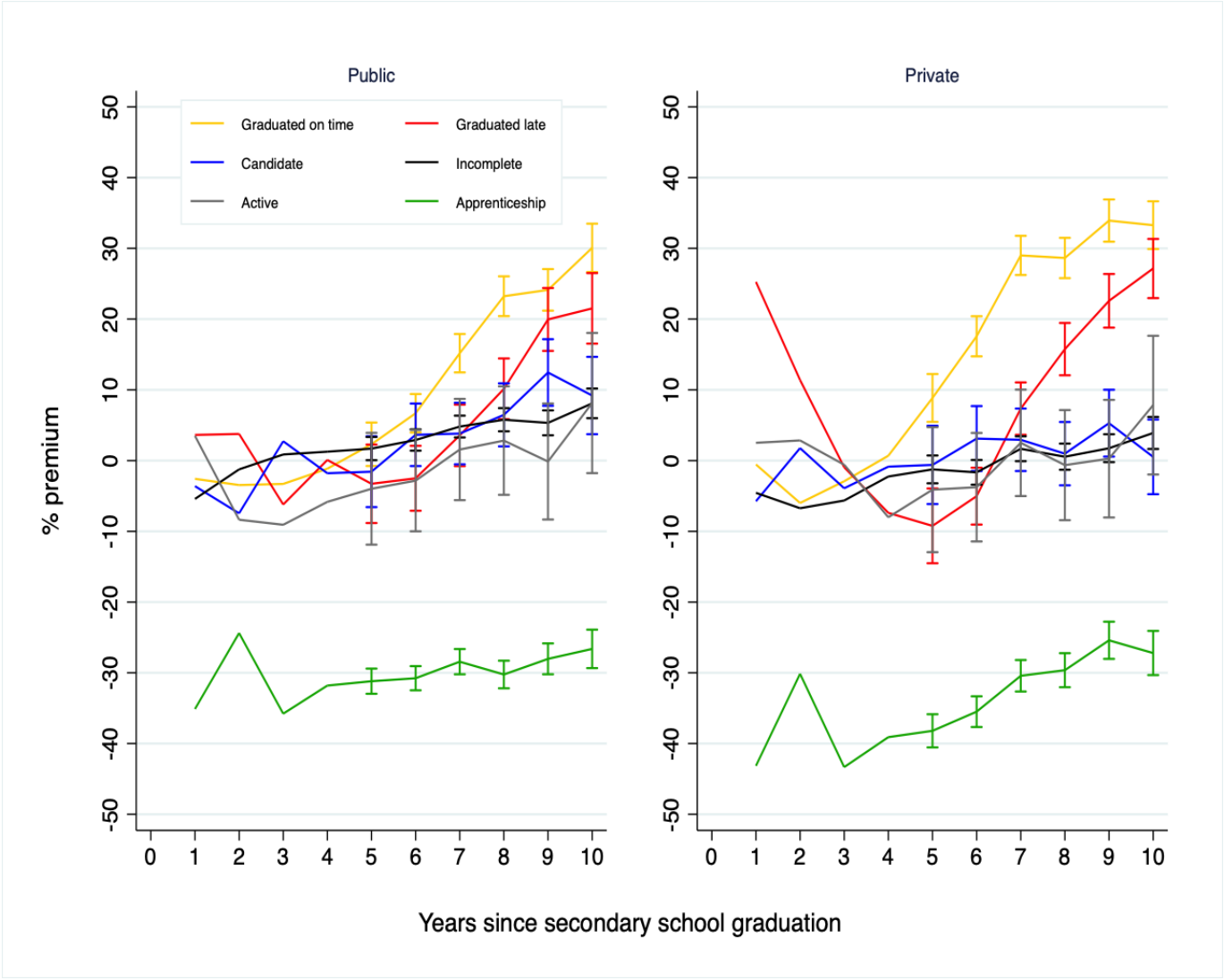
Notes: The figure shows the Mincer regression coefficients for returns to education by score test (cognitive skills), according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions). Charts for skills in verbal or math are found in the appendix.

Figure 4: Tertiary education degree premium in the formal sector in Colombia by household income



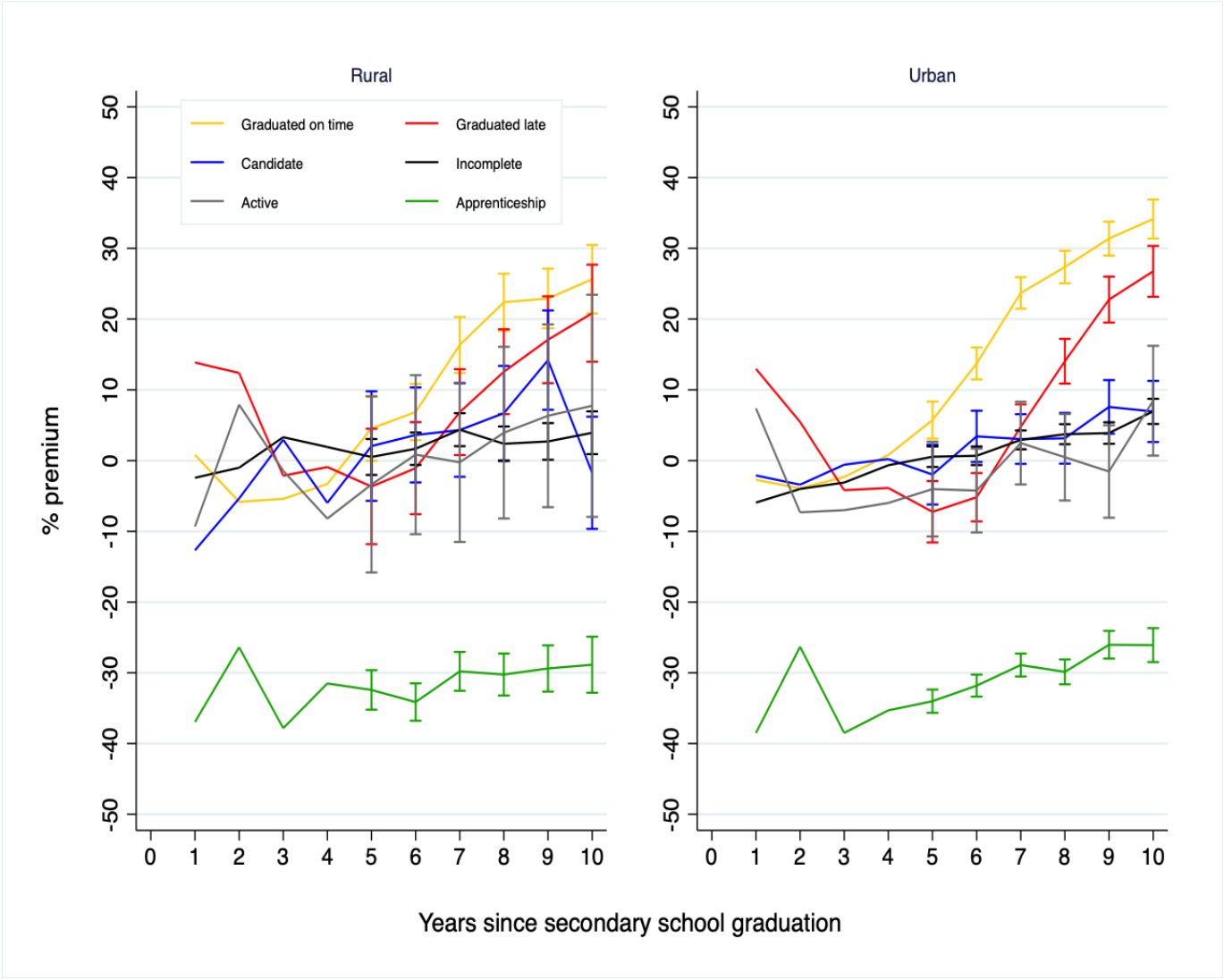
Notes: The figure shows the Mincer regression coefficients for returns to education by household income, according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions).

Figure 5: Tertiary education degree premium in the formal sector in Colombia by school sector



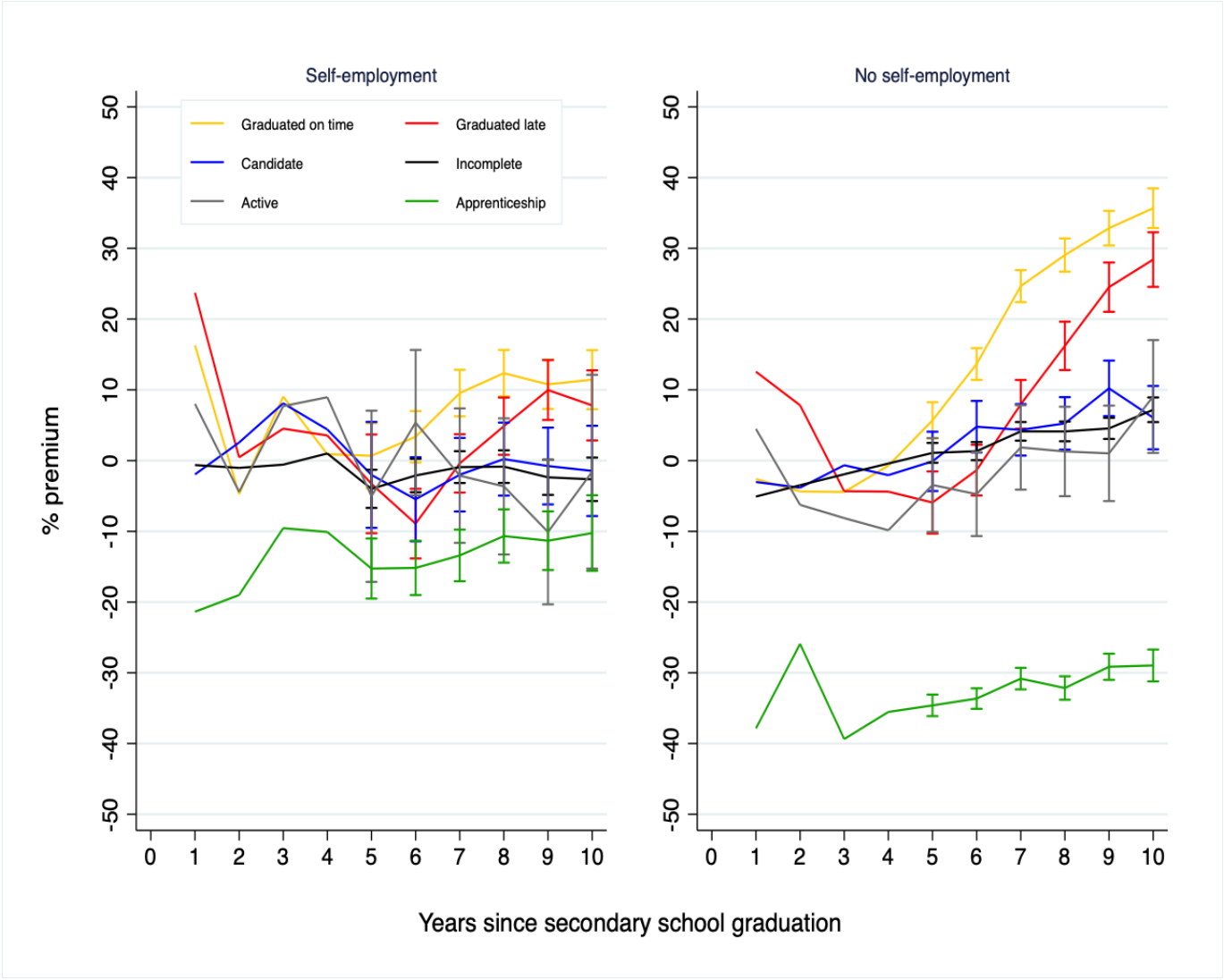
Notes: The figure shows the Mincer regression coefficients for returns to education by school sector, according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions).

Figure 6: Tertiary education degree premium in the formal sector in Colombia by school area



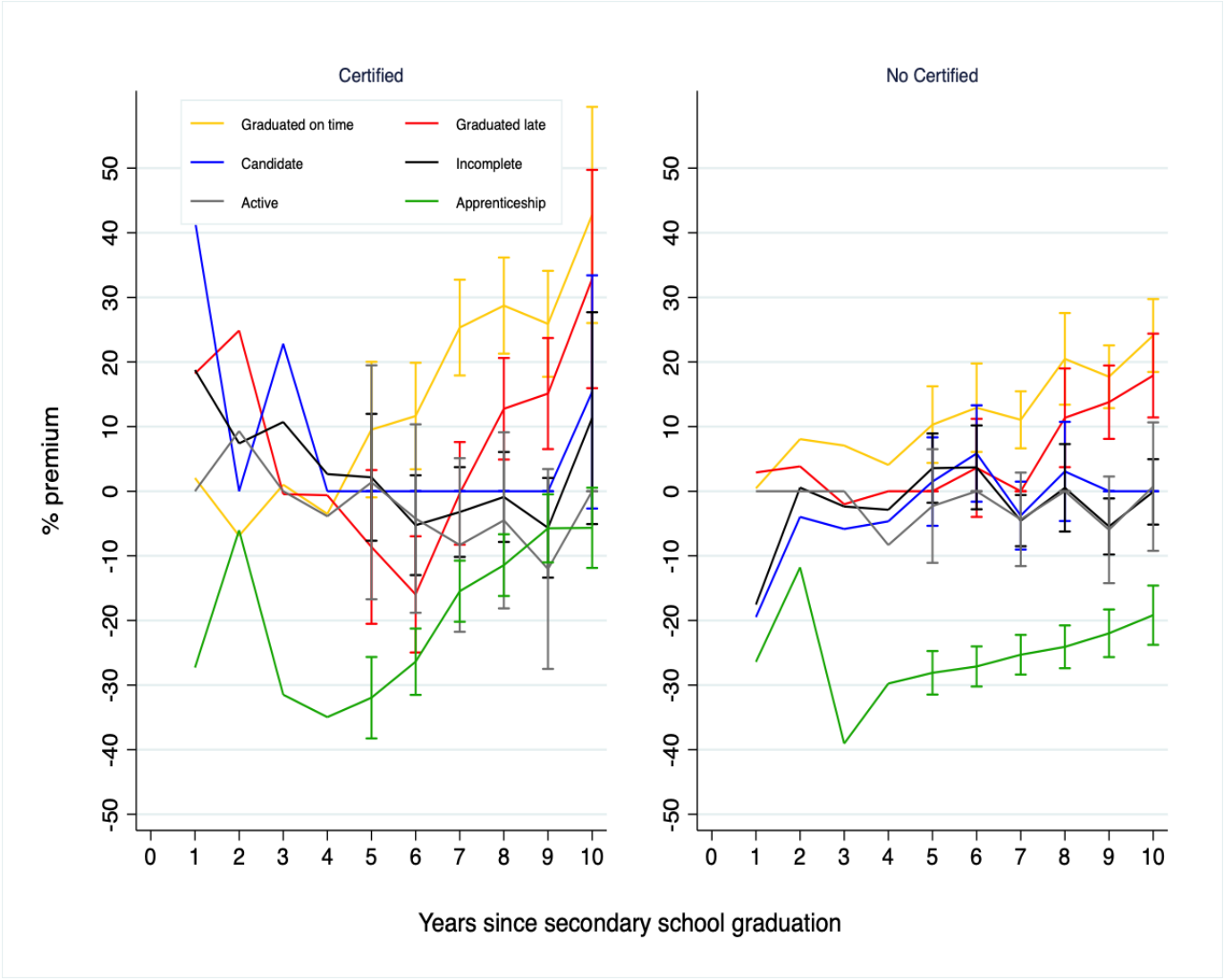
Notes: The figure shows the Mincer regression coefficients for returns to education by school area, according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions).

Figure 7: Tertiary education degree premium in the formal sector in Colombia by self-employment



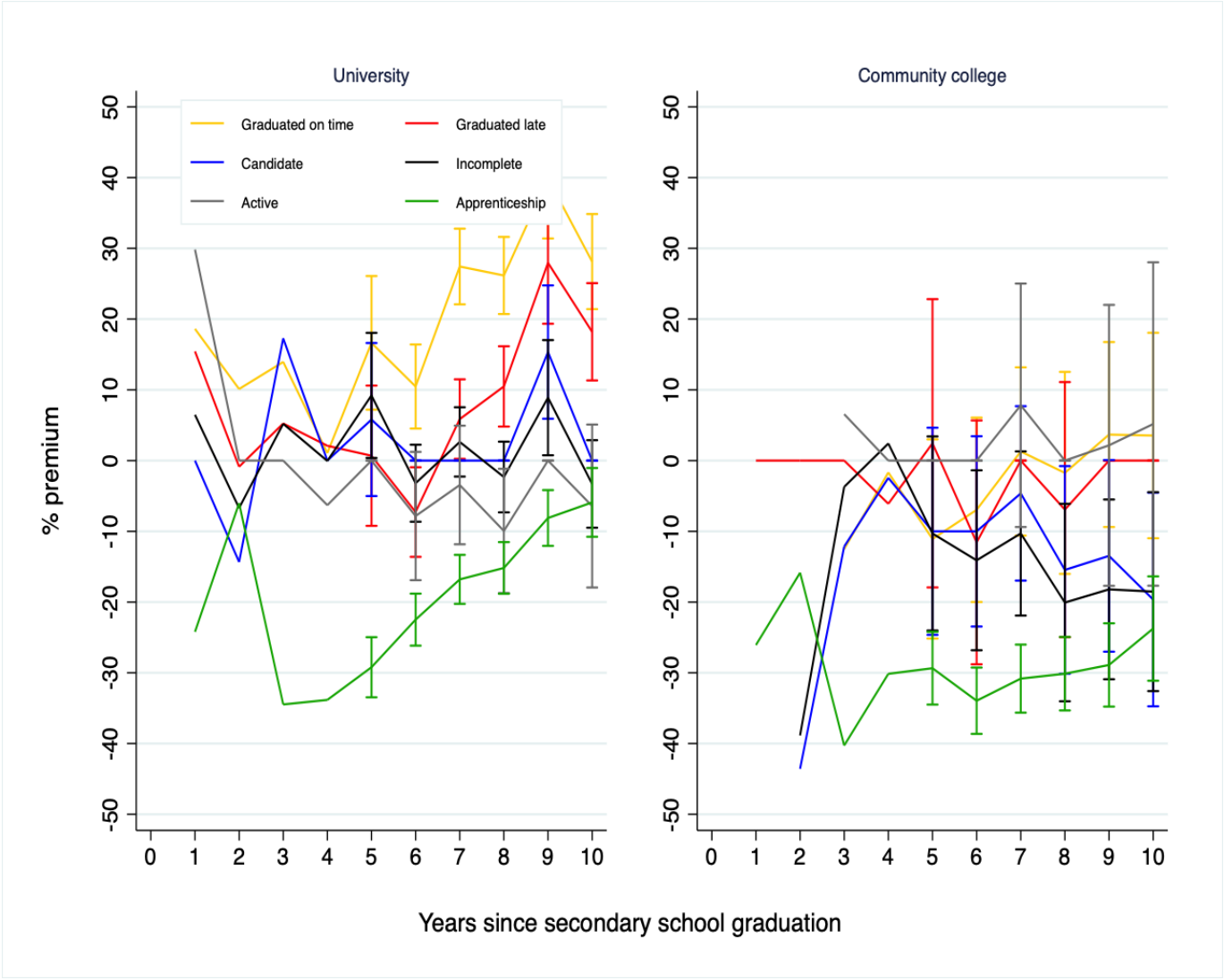
Notes: The figure shows the Mincer regression coefficients for returns to education by self-employment, according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions).

Figure 8: Tertiary education degree premium in the formal sector in Colombia by quality of HEI



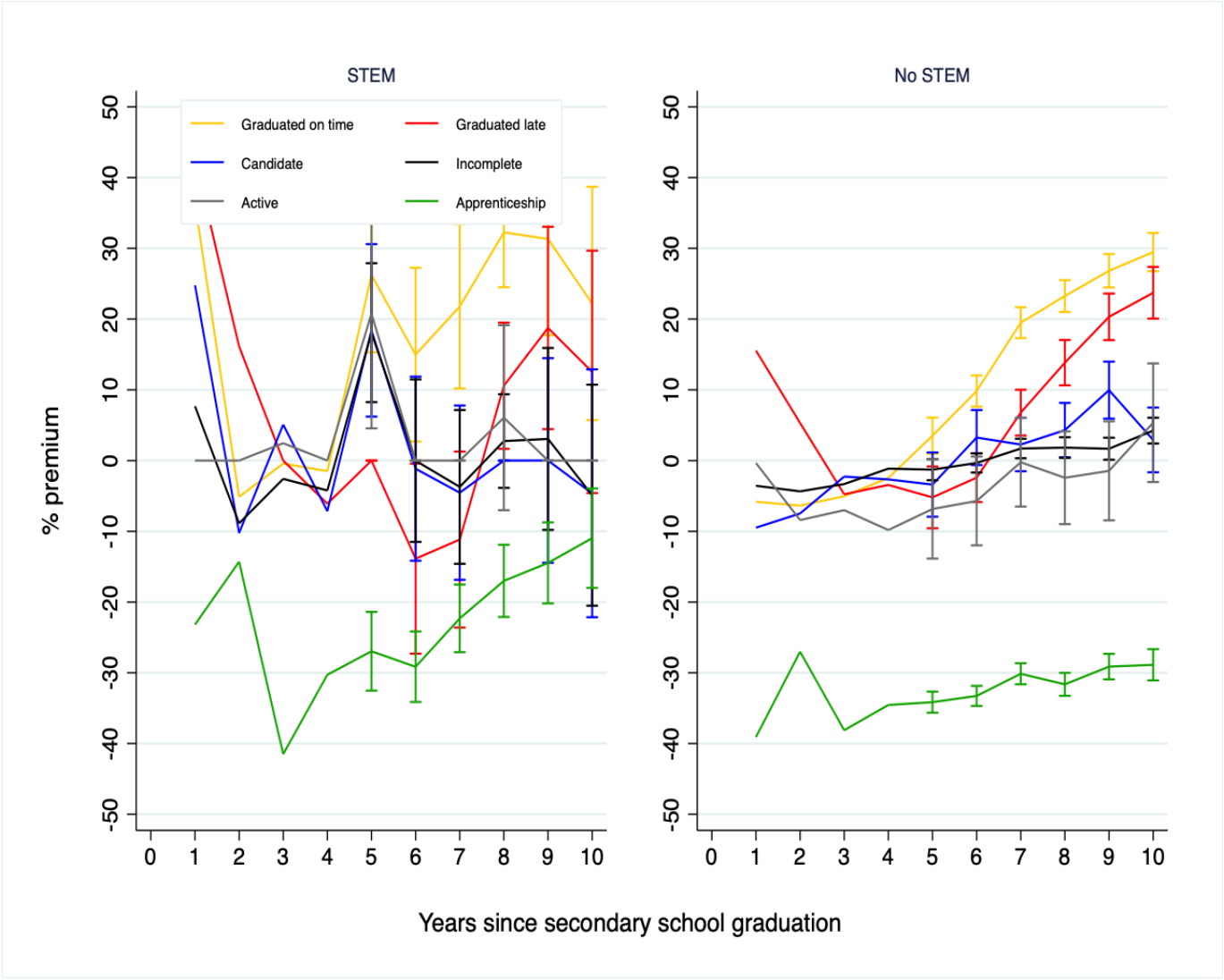
Notes: The figure shows the Mincer regression coefficients for returns to education by higher education institutions certification, according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions).

Figure 9: Tertiary education degree premium in the formal sector in Colombia by level of the program



Notes: The figure shows the Mincer regression coefficients for returns to education by level of the program, according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions).

Figure 10: Tertiary education degree premium in the formal sector in Colombia by type of program



Notes: The figure shows the Mincer regression coefficients for returns to education by type of the program, according to the situation of higher education for late graduates, on-time graduates, incomplete, candidate and active students versus secondary school students who never enrolled in higher education. Days worked are estimated using the same equation as returns. Whiskers are 95 percent confidence intervals. X-axis years from secondary is estimated as the difference between the report on social security records and the year of secondary school graduation. Incomplete means that the student completed more than 90 percent of their college program but did not obtain a degree. Dropouts, according to Colombian authorities, are incomplete plus candidates. As a reference, the year of graduation in higher education is expected to be 5 years after secondary school graduation for the professional degree and 3 years for the technical programs. The premiums come from a regression model that controls for skills, household income, gender, program level, program type, higher education institutions quality, self-employment, employment sector, and fixed effects of year of graduation from secondary school and secondary school (Table 1 for descriptions).

References

Arango, Luis Eduardo, Carlos Esteban Posada, and José Darío Uribe, "Cambios en la estructura de los salarios urbanos en Colombia, 1984-2000," *Lecturas de Economia*, 2005, *63* (63), 7–39.

Arias, Helmut and Álvaro Chávez, "Cálculo de la tasa interna de retorno de la educación en Colombia," 2002.

Arrow, Kenneth, "Higher Education as filter," *Journal of Public Economics*, 1973, *2*, 193–216.

Barro, Robert J, "Human Capital and Growth," *American Economic Review*, 2001, *91* (2), 12–17.

Bauer, Thomas K., Patrick J. Dross, and John P. Haisken-DeNew, "Sheepskin effects in Japan," *International Journal of Manpower*, 2005.

Becker, Gary, "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 1962, *70* (5), 9–49.

Belman, Dale and John S. Heywood, "Sheepskin Effects in the Returns to Education: An Examination of Women and Minorities," *The Review of Economics and Statistics*, 1991.

_ and _ , "Sheepskin effects by cohort: Implications of job matching in a signaling model," *Oxford Economic Papers*, 1997.

Bilkic, N., T. Gries, and M. Pilichowski, "Stay in school or start working? - The human capital investment decision under uncertainty and irreversibility," *Labour Economics*, 2012.

Bound, John and Sarah Turner, "Cohort Crowding: How Resources Affect Collegiate Attainment.," 2006, *12424*.

_ , M. Lovenheim, and Sarah Turner, "Understanding the decrease in college completion rates and the increased time to the baccalaureate degree," *Ann Arbor, Mich.: Population Studies Center, University of Michigan Institute for Social Research*, 2007, pp. 1–91.

Busso, Matías, Juan Sebastián Muñoz, and Sebastián Montaño, "Unbundling Returns to Skills: Evidence from Postsecondary Education in Colombia," 2020.

Cárdenas, Mauricio and Raquel Bernal, "Changes in the distribution of income and the new economic model in Colombia," *Cepal Naciones Unidas*, 1999, (002116).

Crespo, Anna and Mauricio Cortez Reis, "Sheepskin effects and the relationship between earnings and education: Analyzing their evolution over time in Brazil," *Revista Brasileira de Economia*, 2009.

de Roux, Nicolás and Evan Riehl, "Disrupted academic careers: the returns to time off after high school," 2019.

Epple, Dennis and Richard E. Romano, "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects," *American Economic Review*, 1998.

_ , Richard Romano, and Holger Sieg, "Admission, tuition, and financial aid policies in the market for higher education," 2006.

Farber, Henry S and Robert Gibbons, "Learning and the wage dynamics," *Quaterly Journal of Economics*, 1996, (November).

Farné, Stefano, "El Mercado de Trabajo de los Profesionales Colombianos," *Boletín del Observatorio del Mercado de Trabajo y la Seguridad Social*, 2006, *9*.

Ferrer, Ana and W. Craig Riddell, "The role of credentials in the Canadian labour market," *Canadian Journal of Economics*, 2002.

Flores-Lagunes, Alfonso and Audrey Light, "Interpreting degree effects in the returns to education," *Journal of Human Resources*, 2010, *45* (2), 439–467.

Forero, Nohora and Manuel Ramírez, "Determinantes de los ingresos laborales de los graduados universitarios durante el período 2001-2004," *Documentos De Trabajo*, 2008, (31).

Galarza, J, O Vos, A Gaviria, and M Jiménez, "Evolution of returns to education in Colombia (1976-2014) Programa de Economía Documentos de Trabajo," 2015, p. 46.

Gibson, John, "Sheepskin effects and the returns to education in New Zealand: Do they differ by ethnic groups?," *New Zealand Economic Papers*, 2000, *34* (2), 201–220.

González-Velosa, Carolina, Graciana Rucci, Miguel Sarzosa, and Sergio Urzúa, "Returns to higher education in Chile and Colombia," 2015.

Guataquí, Juan Carlos, Andrés Felipe García, and Mauricio Rodríguez, "Estimaciones De Los Determinantes De Los Ingresos Laborales En Colombia Con Consideraciones Iferenciales Para Asalariados Y Cuenta Propia," *Universidad del Rosario-Facultad de Economia*, 2011, *16* (8), 929–935.

Hanushek, Eric and Ludger Woessmann, "Schooling, educational achievement, and the Latin American growth puzzle," *Journal of Development Economics*, 2012, *99* (2), 497–512.

Harberger, Arnold C and Sylvia Guillermo-Peón, "Estimating Private Returns To Education in Mexico," *Latin american journal of economics*, 2012, *49* (1), 1–35.

Heckman, James J, John Eric Humphries, and Gregory Veramendi, "Returns to Education: The Causal Effects of Education on Earnings," 2016.

Hernández, Gustavo, "Cuán rentable es la educación superior en Colombia?," *Lecturas de Economia*, 2010, *73* (julio-diciembre), 181–214.

Herrera-Prada, Luis Omar, "Determinantes de la tasa de graduación y graduación a tiempo en la educación superior de Colombia 1998-2010," *Coyuntura Económica*, 2013, *XLIII* (1), 143.177.

_ , "Ending the Musical Chairs Game in Higher Education: Impacts of SPADIES Software on Educational Outcomes in Colombia," 2020.

_ **and Carlos Caballero**, "El financiamiento de la educación superior en Colombia," in "La educación superior:Retos y perspectivas" 2013, chapter 3, pp. 121–175.

Hungerford, Thomas and Gary Solon, "Sheepskin Effects in the Returns to Education," *The Review of Economics and Statistics*, 1987.

Jaeger, David A. and Marianne E. Page, "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education," *The Review of Economics and Statistics*, 1996.

Kane, By Thomas J and Cecilia Elena Rouse, "Labor-Market Returns to Two- and Four-Year College," *he American Economic Review*, 1995, *85* (3), 600–614.

Kane, Thomas J and Douglas O Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," *National Bureau of Economic Research*, 2008, (14607), 51.

_ , Cecilia Elena Rouse, and Douglas Staiger, "Estimating returns to schooling when schooling is misreported," *NBER working paper series*, 1999, (July), 1–43.

Kharbanda, Varun, "The Role of Endogeneity in the Relationship Between Education and Earnings : A Semiparametric Approach," 2013, pp. 1–37.

Kugler, Adriana, Maurice Kugler, and Luis Omar Herrera-Prada, "Do Payroll Tax Breaks Stimulate Formality ?," *Economia - Journal of the Latin America and the Caribbean Economic Association*, 2017, (Fall), 3–40.

MacLeod, W. Bentle, Evan Riehl, Juan E. Saavedra, and Miguel Urquiola, "The big sort: College reputation and labor market outcomes," *American Economic Journal: Applied Economics*, 2017, *9* (3), 223–261.

Manacorda, Marco, Carolina Sánchez-Páramo, and Norbert Schady, "Changes in Returns to Education in Latin America : The Role of Demand and Supply of Skills Changes in Returns to Education in Latin America : The Role of Demand," *Industrial and Labor Relations Review*, 2007, *63* (2).

Mesa, Fernando and Javier Gutiérrez, "Efectos de la apertura económica en el mercado laboral industrial," *Revista de Planeación y Desarrollo*, 1996.

Mincer, Jacob, "Schooling, Experience, and Earnings," 1974.

Mora, Jhon James, "Sheepskin effects and screening in Colombia," *Colombian Economic Journal*, 2003.

_ and Juan Muro, "Sheepskin effects by cohorts in Colombia," *International Journal of Manpower*, 2008.

Morales-Ramos, Eduardo, "Los Rendimientos de la Educación en México," *Banco de México Working Papers*, 2011.

Núñez, Jairo and Fabio Sánchez, "Educación y Salarios relativos en Colombia, 1976-1995. Determinantes, Evolución e Implicaciones para la Distribución del Ingreso," 1998.

observatorio de la universidad colombiana, El, "El observatorio de la universidad colombiana," 2020.

OECD, "Self-employment rate (indicator)," 2020.

Olfindo, Rosechin, "Diploma as signal? Estimating sheepskin effects in the Philippines," *International Journal of Educational Development*, 2018.

Orozco Silva, Luis Enrique, Alberto Roa Valero, and Luis Carlos Castillo Gómez, "La Educación Superior en Colombia," 2011, pp. 1–83.

Park, Jim, "Estimation of Sheepskin Effects Using the Old and the New Measures of Educational Attainment in the Current Population Survey.," *Economics Letters*, 1999, *62* (2), 237–240.

Phelps, Edmund S, "The Statistical theory of Racism and Sexism," *American Economic Review*, 1972.

Posso, Christian, "Desigualdad salarial en Colombia 1984-2005: cambios en la composición del mercado laboral y retornos a la educación post-secundaria," *Borradores de Economía*, 2008, (529), 1–31.

Prada, Carlos Felipe, "Is the decision of study in Colombia profitable?," 2006, (July 2005), 12–69.

Psacharopoulos, G. and Patrinos, H.A, "Returns to Investments in Education: A Further Update. Education Economics," *Education Economics*, 2004, *12* (2).

Psacharopoulos, George, "Returns to Education: A Further International Update and Implications," *Journal of Human Resources*, 1985, *20* (4), 583–604.

_ , "Returns to Investment in Education. A Global Update," *World development*, 1994, *22* (9), 1325–1343.

Purnastuti, Losina, Paul W. Miller, and Ruhul Salim, "Declining rates of return to education: Evidence for Indonesia," *Bulletin of Indonesian Economic Studies*, 2013, *49* (2), 213–236.

Rodríguez, Eduardo, "Rentabilidad y crecimiento de la educación superior en Colombia 1971-1978," 1981.

Saavedra, Juan Esteban and Carlos Medina, "Formación para el Trabajo en Colombia *," 2012.

_ , Lucrecia Santibáñez, Darío Maldonado, and Luis Omar Herrera-Prada, "Premium or penalty ? Labor market returns to novice public sector teachers," 2016.

Sánchez, Fabio, Margarita Quirós, Carlos Reverón, and Alberto Rodríguez, "Equidad Social En El Acceso Y Permanencia En La Universidad Pública Determinantes Y Factores Asociados," 2002, *7191*, 1–48.

Santamaría, Mauricio, *External Trade, Skill Technology and the Recent Increase of Income Inequality in Colomnbia* 2001.

Sapelli, Claudio, "Tasa De Retorno De Educacion En Chile," 2003, (0716-7334), 1990–1998.

Schady, Norbert R., "Convexity and sheepskin effects in the human capital earnings function: Recent evidence for Filipino men," 2003.

Shabbir, T., "Sheepskin effects in the returns to education in a developing country," *Pakistan Development Review*, 1991.

Son, Hyelim, "Human Capital Investment When Sheepskin Effects Matter : Evidence from Income Shocks in Indonesia," 2013, pp. 1–48.

SPADIES, "Reporte Modelo SPADIES al MEN," Technical Report, Universidad de los Andes, Bogotá 2008.

Spence, Michael, "Job market signaling," *The Quarterly Journal of Economics*, 1973, *LXXXVII*, 355–379.

Tenjo, Jaime., "Evolución de los retornos a la inversión en educación 1976- 1989." PhD dissertation 1993.

Wood, Tom, "The Sheepskin Effect," *National Association of Scholars*, 2009, (July).

Yunus, Norhanishah Mohamad, "Sheepskin effects in the returns to higher education: New evidence for Malaysia," *Asian Academy of Management Journal*, 2017, *22* (1), 151–182.

Zárate, Héctor Manuel, "Cambios en la estructura salarial: una historia desde la regresión cuanfílica," *Monetaria*, 2005, *XXVIII* (4), 339–364.

Table 1: Sort of variables from the ICFES test and SPADIES

Student in...	PILA		Total
SPADIES	NO	YES	
NO	2,680,704	319,128	2,999,832
	59.7%	7.1%	66.8%
YES	1,390,552	99,571	1,490,123
	31.0%	2.2%	33.2%
Total	4,071,256	418,699	4,489,955
	90.7%	9.3%	100.0%

Source:ICFES, SPADIES and PILA.

Table 2: Descriptive statistics of using variables

	All Students		Males		Females	
	Mean	Sd	Mean	Sd	Mean	Sd
	(1)	(2)	(3)	(4)	(5)	(6)
Female	.479	.5	0	0	1	0
Income over 3 mmw	.188	.391	.185	.388	.192	.394
Student from a public secondary school	.665	.472	.67	.47	.66	.474
Student from an urban secondary school	.715	.451	.703	.457	.729	.445
Student with high test score	.297	.457	.328	.469	.263	.44
Student with high test score (verbal)	.286	.452	.276	.447	.296	.457
Student with high test score (maths)	.283	.451	.289	.453	.277	.447
Student in a professional program	.21	.407	.186	.389	.235	.424
Student in a STEM program	.111	.314	.142	.349	.077	.266
Student enrolled in a high quality HEIs	.116	.321	.118	.323	.114	.318
Student reported as self employeed ‡	.022	.145	.02	.14	.023	.15
Student reported as public servant ‡	0	.02	0	.019	0	.02
Observations	4489955.0		2340821.0		2149134.0	

Source:ICFES, PILA. ‡ as percent of all students

Table 3: Descriptive statistics PILA

	All Students		Males		Females	
	Mean	Sd	Mean	Sd	Mean	Sd
	(1)	(2)	(3)	(4)	(5)	(6)
Annual income in 2008	3.0	5.2	3.2	5.0	2.8	5.5
Days worked in 2008	227.0	404.0	228.0	363.0	225.0	460.0
Annual income in 2009	3.2	5.9	3.4	5.5	3.1	6.4
Days worked in 2009	248.0	532.0	245.0	480.0	254.0	600.0
Annual income in 2010	2.9	3.2	3.1	3.3	2.7	3.0
Days worked in 2010	220.0	178.0	220.0	171.0	219.0	186.0
Annual income in 2011	3.6	3.6	3.8	3.7	3.4	3.4
Days worked in 2011	255.0	183.0	253.0	172.0	258.0	199.0
Annual income in 2012	4.1	4.0	4.3	4.1	3.8	3.9
Days worked in 2012	266.0	198.0	264.0	183.0	268.0	219.0
Annual income in 2013	4.7	4.3	4.9	4.5	4.3	3.9
Days worked in 2013	287.0	125.0	286.0	125.0	289.0	125.0
Annual income in 2014	4.9	4.7	5.2	4.8	4.5	4.3
Days worked in 2014	285.0	115.0	285.0	114.0	284.0	117.0

Source:PILA. Income in thousands of 2013 USD (1 USD = 1869.1 colombian pesos. Workers can have more than one job, which explains the high values in the report of days worked.

Table 4: Descriptive statistics of model variables

	All Students		Males		Females	
	Mean	Sd	Mean	Sd	Mean	Sd
	(1)	(2)	(3)	(4)	(5)	(6)
Apprenticeship	.008	.09	.008	.089	.008	.092
Enrolled in Higher Ed.	.332	.471	.316	.465	.349	.477
Retired	.186	.389	.189	.392	.181	.385
Active	.009	.094	.009	.096	.008	.091
Candidate	.018	.134	.018	.131	.019	.136
Graduated	.119	.324	.1	.3	.14	.347
Graduated on time	.084	.277	.069	.253	.1	.3
Graduated late	.035	.185	.031	.174	.04	.197
Bachelor	.117	.321	.098	.297	.137	.344
Postgraduate	.003	.051	.002	.044	.003	.058
Diploma	.003	.052	.002	.043	.004	.06
Master	0	.013	0	.012	0	.014
PhD	0	.009	0	.01	0	.008

Source:ICFES, DANE, and SPADIES.

Table 5: Main results with formal workers

	Log Income (1)	Log Income (2)	Log Income (3)	Income (4)	Income (5)	Income (6)	Income (7)	Income (8)
Graduated	0.184***	0.156***	0.204***	-0.110***	-1.782***	-2.584***	-2.069***	-1.837***
	(0.003)	(0.003)	(0.010)	(0.026)	(0.032)	(0.038)	(0.061)	(0.062)
Candidate	0.019***	0.030***	0.032***	0.225***	0.226***	0.254***	0.153***	0.156***
	(0.006)	(0.006)	(0.012)	(0.024)	(0.024)	(0.024)	(0.047)	(0.048)
Incomplete	0.003	0.016***	0.023**	0.135***	0.125***	0.147***	0.033	0.045
	(0.002)	(0.002)	(0.010)	(0.008)	(0.008)	(0.008)	(0.041)	(0.042)
Active	-0.028***	-0.010		0.120***	0.104***	0.133***		
	(0.010)	(0.010)		(0.039)	(0.039)	(0.039)		
Apprenticeship	-0.334***	-0.326***	-0.260***	-0.788***	-0.780***	-0.782***	-0.766***	-0.748***
	(0.002)	(0.002)	(0.004)	(0.010)	(0.010)	(0.010)	(0.018)	(0.019)
Female	-0.053***	-0.053***	-0.014***	-0.425***	-0.429***	-0.428***	-0.256***	-0.314***
	(0.002)	(0.002)	(0.003)	(0.006)	(0.006)	(0.006)	(0.013)	(0.014)
Years after secondary	0.206***	0.205***	0.209***	0.342***	0.342***	0.342***	0.298***	0.290***
	(0.001)	(0.001)	(0.003)	(0.005)	(0.005)	(0.005)	(0.011)	(0.011)
Years after secondary^2	-0.006***	-0.006***	-0.005***	0.004***	0.001***	0.001***	0.013***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Secondary test score	0.002***	0.002***	0.003***	0.006***	0.007***	0.006***	0.008***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban school	0.018***	0.038**	0.015***	0.100	0.092	0.096	0.042***	-0.678***
	(0.002)	(0.019)	(0.004)	(0.072)	(0.072)	(0.072)	(0.014)	(0.153)
Public school	-0.059***	0.042*	-0.044***	0.237***	0.235***	0.228***	-0.249***	0.873***
	(0.002)	(0.022)	(0.003)	(0.084)	(0.084)	(0.084)	(0.014)	(0.185)
More than 3mmw	0.055***	0.021***	0.038***	0.208***	0.221***	0.219***	0.467***	0.220***
	(0.002)	(0.003)	(0.004)	(0.012)	(0.012)	(0.012)	(0.018)	(0.022)
Self-employeed	-0.083***	-0.087***	-0.181***	-0.962***	-0.977***	-0.976***	-1.732***	-1.705***
	(0.002)	(0.002)	(0.004)	(0.009)	(0.009)	(0.009)	(0.017)	(0.017)
Public servant	0.208***	0.183***	0.155***	-0.444***	-0.488***	-0.486***	-1.405***	-1.261***
	(0.009)	(0.009)	(0.016)	(0.028)	(0.028)	(0.028)	(0.052)	(0.053)
Graduate X Score				0.019***		0.017***	0.019***	0.013***
				(0.000)		(0.000)	(0.000)	(0.001)
Graduate X Years after secondary					0.397***	0.384***	0.297***	0.298***
					(0.004)	(0.004)	(0.005)	(0.005)
FE School	No	Yes	No	Yes	Yes	Yes	No	Yes
FE HEI	No	No	Yes	No	No	No	Yes	Yes
FE Department	Yes	No	No	No	No	No	No	No
Constant	14.403***	14.343***	14.231***	1.416***	1.477***	1.517***	1.534***	1.666***
Observations	2,074,267	2,073,506	522,076	2,097,731	2,097,731	2,097,731	531,835	530,944
R^2	0.087	0.108	0.132	0.113	0.116	0.117	0.168	0.249

Notes: The table shows the coefficients of the regressions corresponding to equation 1 and 2. Graduate takes the value of 1 if the student graduated from higher education; Candidate takes the value of 1 if the student attended more than 90 percent of the program but did not receive the degree; incomplete is the student who is absent for two or more consecutive semesters without registering a degree and has less than 90 percent of the program; Apprentice is the student I found as a SENA intern. Experience is the time elapsed since the time of the secondary school degree and the apparition in the Social Security. The score on the ICFES test is the percentile per student. Urban school, Public school, Income of household, self-employed, and public servant are dummies for the individual; the last two are dynamic in time. Income is in thousands of dollars of 2013 (1 US dollar = 1869.1 Colombian pesos in 2013). In columns (1), (2), and (3), the dependent variable is expressed in logarithm, the others are measured in levels. Columns (7) and (8) analyze the results for the 2002 cohort. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Results with formal workers with post-grades

	Log Income (1)	Log Income (2)	Log Income (3)	Income (4)	Income (5)	Income (6)	Income (7)	Income (8)
Bachelor	0.179***	0.151***	0.066	-0.123***	-1.787***	-2.578***	0.385	0.633
	(0.003)	(0.003)	(0.137)	(0.026)	(0.032)	(0.038)	(0.566)	(0.562)
Diploma	0.395***	0.336***	0.228	1.145***	-0.671***	-1.550***	1.425**	1.680***
	(0.019)	(0.019)	(0.139)	(0.077)	(0.080)	(0.083)	(0.571)	(0.567)
Master	0.648***	0.532***	0.483***	2.020***	0.143	-0.758**	2.665***	2.806***
	(0.083)	(0.084)	(0.160)	(0.301)	(0.301)	(0.302)	(0.646)	(0.642)
Candidate	0.019***	0.030***	-0.102	0.226***	0.227***	0.255***	2.600***	2.623***
	(0.006)	(0.006)	(0.138)	(0.024)	(0.024)	(0.024)	(0.568)	(0.565)
Incomplete	0.004*	0.016***	-0.111	0.136***	0.125***	0.147***	2.480***	2.511***
	(0.002)	(0.002)	(0.137)	(0.008)	(0.008)	(0.008)	(0.568)	(0.564)
Active	-0.028***	-0.009	-0.134	0.121***	0.105***	0.134***	2.447***	2.467***
	(0.010)	(0.010)	(0.138)	(0.039)	(0.039)	(0.039)	(0.569)	(0.566)
Apprenticeship	-0.334***	-0.326***	-0.260***	-0.788***	-0.780***	-0.782***	-0.767***	-0.747***
	(0.002)	(0.002)	(0.004)	(0.010)	(0.010)	(0.010)	(0.018)	(0.019)
Female	-0.053***	-0.053***	-0.014***	-0.426***	-0.429***	-0.429***	-0.257***	-0.316***
	(0.002)	(0.002)	(0.003)	(0.006)	(0.006)	(0.006)	(0.013)	(0.014)
Years after secondary	0.206***	0.205***	0.209***	0.342***	0.342***	0.342***	0.298***	0.290***
	(0.001)	(0.001)	(0.003)	(0.005)	(0.005)	(0.005)	(0.011)	(0.011)
Years after secondary^2	-0.006***	-0.006***	-0.005***	0.004***	0.001***	0.001***	0.013***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Secondary test score	0.002***	0.002***	0.003***	0.006***	0.007***	0.006***	0.008***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban school	0.018***	0.038**	0.015***	0.100	0.092	0.096	0.041***	-0.681***
	(0.002)	(0.019)	(0.004)	(0.072)	(0.072)	(0.072)	(0.014)	(0.153)
Public school	-0.059***	0.042*	-0.044***	0.237***	0.234***	0.228***	-0.248***	0.869***
	(0.002)	(0.022)	(0.003)	(0.084)	(0.084)	(0.084)	(0.014)	(0.185)
More than 3mmw	0.054***	0.021***	0.038***	0.208***	0.221***	0.218***	0.464***	0.219***
	(0.002)	(0.003)	(0.004)	(0.012)	(0.012)	(0.012)	(0.018)	(0.022)
Self-employeed	-0.083***	-0.087***	-0.181***	-0.962***	-0.977***	-0.976***	-1.731***	-1.704***
	(0.002)	(0.002)	(0.004)	(0.009)	(0.009)	(0.009)	(0.017)	(0.017)
Public servant	0.208***	0.183***	0.154***	-0.447***	-0.490***	-0.488***	-1.414***	-1.270***
	(0.009)	(0.009)	(0.016)	(0.028)	(0.028)	(0.028)	(0.052)	(0.053)
Graduate X Score				0.019***		0.016***	0.018***	0.013***
				(0.000)		(0.000)	(0.000)	(0.001)
Graduate X Years after secondary					0.394***	0.382***	0.295***	0.296***
					(0.004)	(0.004)	(0.005)	(0.005)
FE School	No	Yes	No	Yes	Yes	Yes	No	Yes
FE HEI	No	No	Yes	No	No	No	Yes	Yes
FE Department	Yes	No	No	No	No	No	No	No
Observations	2,074,267	2,073,506	522,076	2,097,731	2,097,731	2,097,731	531,835	530,944
R^2	0.087	0.108	0.132	0.114	0.116	0.117	0.168	0.250

Notes: The table shows the coefficients of the regressions corresponding to equation 1 and 2. Bachelor's degree if the student did not continue his or her studies after obtaining the higher education degree; Diploma if the student enrolled in a diploma; Master's degree if the student enrolled in a master's degree program; and Phd if the student enrolled in a doctorate program. Graduate takes the value of 1 if the student attended more than 90 percent of the program but did not receive the degree; incomplete is the student who is absent for two or more consecutive periods without registering a degree and has less than 90 percent of the program; Apprentice is the student I found as a SENA intern. Experience is the time elapsed since the time of the secondary school degree and the apparition in the Social Security. The score on the ICFES test is the percentile per student. Urban school, Public school, Income of household, self-employed, and public servant are dummies for the individual; the last two are dynamic in time. Income is in thousands of dollars of 2013 (1 US dollar = 1869.1 Colombian pesos in 2013). In columns (1), (2), and (3), the dependent variable is expressed in logarithm, the others are measured in levels. Columns (7) and (8) analyze the results for the 2002 cohort. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Tertiary education premium by type of graduate and candidate at year 5 and year 10

	At year 5				At year 10			
	Graduated late	Graduated on time	Candidate	Sheepskin effect	Graduated late late	Graduated on time on time	Candidate	Sheepskin effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-2.10	9.64	0.30	9.34	27.05	33.16	6.75	26.41
	(2.75)	(1.59)	(2.85)		(2.34)	(1.65)	(2.97)	
Male	-8.92	0.96	-3.27	4.23	26.26	31.22	4.36	26.86
	(2.83)	(1.72)	(2.60)		(2.35)	(1.87)	(2.67)	
High score	-8.97	8.43	-4.01	12.44	33.22	42.05	16.97	25.09
	(3.20)	(2.13)	(3.58)		(2.58)	(2.17)	(3.61)	
Low score	-4.90	3.30	1.41	1.90	18.68	26.10	1.36	24.75
	(2.61)	(1.45)	(2.32)		(2.33)	(1.58)	(2.42)	
More than 3 minimun wages	-14.05	3.78	-7.64	11.42	26.30	35.88	1.88	34.00
	(3.89)	(2.74)	(4.73)		(3.20)	(2.92)	(4.50)	
Less <= 3 minimun wages	-4.51	4.54	0.05	4.50	21.88	29.67	5.25	24.43
	(2.34)	(1.32)	(2.11)		(2.05)	(1.40)	(2.21)	
Public	-3.28	2.30	-1.57	3.87	21.51	30.07	9.20	20.88
	(2.83)	(1.57)	(2.55)		(2.54)	(1.74)	(2.79)	
Private	-9.23	8.85	-0.61	9.46	27.15	33.28	0.53	32.75
	(2.70)	(1.72)	(2.82)		(2.13)	(1.72)	(2.69)	
Urban	-7.24	5.73	-1.98	7.70	26.75	34.14	6.94	27.20
	(2.21)	(1.34)	(2.16)		(1.83)	(1.40)	(2.20)	
Rural	-3.66	4.55	2.05	2.50	20.83	25.64	-1.73	27.36
	(4.16)	(2.34)	(3.95)		(3.50)	(2.47)	(4.05)	
Self-employment	-3.29	0.68	-2.02	2.71	7.80	11.44	-1.46	12.90
	(3.56)	(2.35)	(3.82)		(2.53)	(2.13)	(3.25)	
No self-employment	-5.93	5.68	-0.11	5.80	28.42	35.68	6.07	29.61
	(2.24)	(1.30)	(2.14)		(1.97)	(1.43)	(2.28)	
Certified	-8.62	9.53	0.00	9.53	32.84	42.76	15.37	27.39
	(6.07)	(5.35)	(0.00)		(8.62)	(8.53)	(9.20)	
No Certified	0.00	10.30	1.49	8.81	17.90	24.10	0.00	24.10
	(0.00)	(3.02)	(3.49)		(3.31)	(2.88)	(0.00)	
University	0.69	16.64	5.80	10.84	18.20	28.13	0.00	28.13
	(5.06)	(4.82)	(5.52)		(3.51)	(3.43)	(0.00)	
Community college	2.44	-11.08	-10.00	-1.08	0.00	3.54	-19.64	23.18
	(10.40)	(7.18)	(7.47)		(0.00)	(7.41)	(7.70)	
STEM	0.00	26.23	18.41	7.83	12.54	22.21	-4.62	26.84
	(0.00)	(5.57)	(6.22)		(8.74)	(8.41)	(8.94)	
No STEM	-5.20	3.52	-3.39	6.91	23.72	29.47	2.92	26.55
	(2.23)	(1.30)	(2.32)		(1.87)	(1.38)	(2.33)	

Notes: Standard errors in parentheses.