Policy Research Working Paper 9717

Labor Market Effects of Short-Cycle Higher Education Programs

Lessons from Colombia

Maria Marta Ferreyra
Camila Galindo
Sergio Urzúa

WORLD BANK GROUP
Latin America and the Caribbean Region
Office of the Chief Economist
June 2021

Policy Research Working Paper 9717

Abstract

This paper estimates the heterogeneous labor market effects of enrolling in higher education short-cycle (SC) programs. Expanding access to these programs might affect the behavior of some students (compliers) in two margins: the expansion margin (students who would not have enrolled in higher education otherwise) and the diversion margin (students who would have enrolled in bachelor's programs otherwise). These responses are quantified by exploiting local exogenous variation in the supply of higher education institutions (HEIs) facing Colombian high school graduates in an empirical multinomial choice model with several instruments. Estimates indicate that the presence of at least one HEI specialized in SC programs in the vicinity of the student's high school municipality increases SC enrollment by 3.7–4.5 percentage points (40–50% of the SC enrollment rate). The diversion margin largely drives this effect. For female compliers, enrollment in SC programs increases formal employment relative to the next-best alternative. For male compliers, in contrast, it lowers formal employment and wages. These results should alert policymakers of the unexpected consequences of higher education expansionary policies.

This paper is a product of the Office of the Chief Economist, Latin America and the Caribbean Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at mferreyra@worldbank.org.

Labor Market Effects of Short-Cycle Higher Education Programs: Lessons from Colombia

Maria Marta Ferreyra* Camila Galindo† Sergio Urzúa‡§

Originally published in the Policy Research Working Paper Series on *June 2021*.
This version is updated on *February 2023*.
To obtain the originally published version, please email prwp@worldbank.org.

*The World Bank, mferreyra@worldbank.org.

†Department of Economics, University of Chicago, camilagalindo@uchicago.edu.

‡Department of Economics, University of Maryland, surzua@umd.edu.

§We thank Omar Arias, Marina Bassi, Matias Busso, Gabriel Demombynes, Lelys Dinarte, Guido Kuersteiner, Macarena Kutscher, Jack Mountjoy, Martin Rama, Matteo Ruzzante, Kevin Stange, and Ulrich Zachau for comments on earlier drafts of this paper, as well as participants at the University of Maryland's microeconomics seminar (2019) and The World Bank's author's workshop for the book "The Fast Track to New Skills : Short-Cycle Higher Education Programs in Latin America and the Caribbean" (2020). Nicolas Torres Franco provided excellent research assistance. We thank the Ministry of Education of Colombia for granting us access to the data.

1 Introduction

In today's fast-changing labor market, expanding higher education is viewed around the world as a pathway to provide or update skills . While popular, bachelor's degrees are not always effective or affordable (Rodríguez et al., 2016; Ferreyra et al., 2021), which might explain the increasing interest in short-term programs (typically lasting two or three years) in many countries. For the U.S., recent evidence shows that these programs can benefit students in terms of educational attainment and earnings (Acton, 2020; Bettinger and Soliz, 2016; Denning, 2017; Jepsen et al., 2014; Marcotte, 2019; Minaya and Scott-Clayton, 2020), with effects mainly driven by students who would not have enrolled in higher education otherwise (Mountjoy, 2022). Evidence on these programs is scant, however, for the developing word.

This paper estimates the heterogeneous labor market effects of enrolling in short-term programs (henceforth, short-cycle programs) in Colombia.[1] To this end, we exploit the recent growth of these programs in Colombia. Consistent with evidence from developed countries, we document that these programs attract students who would have not enrolled in higher education otherwise (the *expansion* or *democratization* margin). However, and in contrast to previous findings, we also show that they induce an even larger proportion of students to divert away from bachelor's and into short-cycle programs (the *diversion* margin) (Leigh and Gill, 2003; Rouse, 1995). We characterize the students who respond to the expansion of short-cycle (SC) programs along either margin (compliers) and estimate the programs' effects on employment and wages, highlighting their heterogeneity across student groups.

Our interest in Colombia is twofold. First, Colombia has experienced a fast higher education expansion over the last decades (Carranza and Ferreyra, 2019), which might have affected the labor market returns to higher education (Camacho et al., 2017; González-Velosa et al., 2015). Second, while SC programs are relatively less popular in Latin America than the rest of the world, Colombia is an exception. By 2018, approximately 30% of its total higher education enrollment corresponded to SC programs (Ferreyra et al., 2021). Yet, to the best of our knowledge, there is neither evidence of their impact on labor market outcomes nor a characterization of the type of student that benefits from them. We fill these two gaps.

We address the multiple challenges associated with identifying SC programs' heterogeneous impacts. First, high-school graduates self-select into their preferred option

[1] UNESCO uses the term "short-cycle tertiary education programs" for level-5 programs in the International Standard Classification of Education (ISCED 2011). They encompass community college programs lasting two years in the U.S., and the two- and three-year programs in Colombia discussed in this paper.

(enrolling in a bachelor's program, enrolling in a SC program, or not enrolling in higher education), which might bias OLS estimates of labor market impacts. Second, while a standard Instrumental Variables (IV) approach could address self-selection issues in homogeneous treatment effect settings, under general conditions it would not identify the average effect of SC programs (Heckman and Urzúa, 2010). More precisely, in our essential heterogeneity setting, a standard IV strategy would only identify the effect of SC programs on compliers relative to the next-best option that the student would have otherwise chosen–enrolling in a bachelor's program or not enrolling at all in higher education–without identifying the separate margins of choice. In other words, a standard IV strategy would identify the effect of SC programs relative to a mixture of the other two options. Only under additional assumptions could one decompose this effect into its sub-components and provide the underlying pairwise comparisons (as in Kline and Walters, 2016; Hull, 2018). Rather than impose additional assumptions, we exploit rich student longitudinal data and local exogenous variation in higher education supply to estimate the share of students who respond along each margin of choice–expansion and diversion–while also estimating SC programs' labor market effects relative to the next-best option.

Evidence on these margins is novel for developing countries. While other studies have explored higher education returns in developing countries, they have either compared returns across countries (Ferreyra et al., 2017, 2021) or, within a country, have estimated returns for new *versus* existing programs (Camacho et al., 2017), by program and field (Ferreyra et al., 2020), or for selective institutions (Barrera-Osorio and Bayona-Rodríguez, 2019). Unlike previous studies, we quantify the margins of choice and argue that these are critical to who benefits from SC programs, and why. Consider, for instance, a policymaker who builds a new higher education institution (HEI) specialized in SC programs. This might lead some students (the compliers) to enroll in SC programs even though they would not have done it otherwise, perhaps leading the policymaker to believe that all compliers would benefit from the new SC programs and improve their labor market outcomes. Whether this happens, however, depends on students' next-best option. For a student who would not have attended higher education, the SC program might indeed raise human capital and improve labor market outcomes. For a student who would have attended a relatively good bachelor's program and for whom the SC program might be an "undermatch", labor market outcomes might become worse. Yet, for a student who would have otherwise attended a low-quality bachelor's program and for whom the SC program might provide a better match, labor market outcomes might actually improve. Because individuals differ in their next-best option based on their personal background and local opportunities, the policymaker should not necessarily expect average positive effects from the SC expansion but rather heterogeneous gains and losses–as we indeed find.

We exploit the local supply of higher education programs in Colombia as a source of exogenous variation for instruments. We use estimated propensity scores of enrollment choices to estimate the share of expansion and diversion compliers under a behavioral assumption (partial monotonicity, as in Mountjoy (2022) and Mogstad et al. (2021)) that allows us to isolate the effect of changes in the local supply of HEIs specialized in offering SC programs while holding the supply of other HEIs constant. To improve over standard two-stage least squares (TSLS) methods and identify the parameters of interest, we posit a multinomial model of enrollment choices, multiple IVs, and essential heterogeneity.[2]

We use administrative information for the universe of Colombia's high school graduates in 2005. We observe their high school exit exam (*Saber 11*) scores, higher education enrollment choices, labor market outcomes between 2008 and 2013 (employment status and wages), and socioeconomic information. We link this data to municipality-level information on higher education supply, and focus on whether there is at least one HEI specialized in SC programs within the 10-kilometer radius of the student's high school municipality. Hence, we compare the enrollment choices of high school students in municipalities with and without access to at least one specialized institution offering SC programs. Moreover, we examine whether the existence of such HEIs has different effects depending on the local supply of HEIs offering bachelor's programs.

Our findings show that the presence of SC-specialized HEIs raises SC enrollment by $3.7-4.5$ percentage points (pp). This is a sizable increase, representing 40-50% of the sample enrollment rate in SC programs.[3] At least 70% of this increase is due to students who would have otherwise attended a bachelor's program (the diversion margin), while the remaining students react along the expansion margin. The diversion share is higher among men than women ($75-83\%$ versus $70-74\%$, respectively). In addition, most of

[2] Models with essential heterogeneity are characterized by heterogeneous responses to treatments and self-selection into treatments with partial knowledge of idiosyncratic responses. In these models, Heckman and Urzúa (2010) show that if an instrument affects only one margin of choice (in our case, enrolling in a SC program vs. not doing so), IVs estimate the effect of one option vs. the next-best (in our case, the effect of enrolling in a SC program relative to a mixture of not enrolling in higher education and enrolling in a bachelor's program). Also, Mountjoy (2022) shows that multivariate TSLS (multiple margins of choice and multiple IVs) estimate effects that are a combination of many margins of choice and do not necessarily identify well-defined effects. In our case, where there are three margins of choice (not enrolling in higher education vs. enrolling in SC programs, not enrolling in higher education vs. enrolling in bachelor's programs, and enrolling in bachelor's programs vs. enrolling in SC programs), multivariate TSLS would estimate a combination of effects for the three margins. Our empirical strategy takes into account these complexities.

[3] Our estimated effects are in the range of others in the literature, obtained for the U.S. with other sources of variation. For Texas, Mountjoy (2022) finds that a 10-mile increase in the distance between the student's high school and the nearest two-year college decreases enrollment in two-year colleges by 4.3 percentage points. Acton (2020) finds that a \$1000-tuition reduction in local community colleges increases their enrollment by 3.5 percentage points.

the reaction to local availability is concentrated among students residing in municipalities that have HEIs available as well. This is consistent with the fact that, during our sample period, SC programs in Colombia were mostly offered in large and medium cities–which already offered bachelor's programs– and highlights the importance of interacting the instruments. Relative to the average SC student, compliers come from more advantaged backgrounds and are better prepared for higher education as measured by their *Saber 11* scores. Male compliers, in turn, are more advantaged than their female counterparts. For females, we uncover a positive impact (31 pp) from choosing a SC program on formal employment, yet find negative effects on formal employment and wages (24 percentage points and 23 percent, respectively) for males. Our results are robust to different specifications. These findings, together with the high estimated diversion share, suggest that SC programs might provide a better match than bachelor's programs for females but an "undermatch" for males.

The fact that diversion prevails over expansion in Colombia contrasts with the U.S., where the reverse takes place.[4] This may be because students who enroll in community college in the U.S. typically intend to transfer to a four-year institution (Acton, 2020; Denning, 2017; Kane and Rouse, 1999) and view community college as the first step towards a four-year degree. In contrast, SC degrees in Colombia are usually viewed as terminal and rarely offer a bachelor's program pathway.[5] In addition, community colleges in the US are often located outside large cities and therefore attract students from medium and small localities without other higher education options. In contrast, during our sample period SC programs in Colombia were mostly offered in large and medium cities, where bachelor's programs were already being offered.

The remainder of the paper is organized as follows. Section 2 discusses the institutional background of Colombia's higher education system. Section 3 describes data sources, summary statistics, and identifying variation. Section 4 introduces our empirical strategy and discusses how to identify student shares along the expansion and diversion margins. Section 5 presents results and robustness checks, and section 6 concludes.

[4]For instance, Mountjoy (2022) finds that about two-thirds of students would react to changes in proximity to two-year college by switching along the expansion margin; the remaining 30% would divert from four- into two-year colleges. Denning (2017) finds that lower two-year college tuition induces students who would not have enrolled otherwise to attend a two-year college, with no evidence of students substituting two- for four-year college degrees.

[5]In our sample, only 6% of SC students graduate from a bachelor's program. In the U.S., about 30% of community college students graduate from a four-year institution (Levesque, 2018).

2 Institutional Background

Colombia's higher education system offers bachelor's programs (four- to six-years long) and SC programs. The latter (akin to asssociate's degrees in the U.S.) encompass technical and technological programs, lasting two and three years respectively. The system includes several HEI types: universities, university institutes, technological institutes, and technical-professional institutes. Universities and university institutes are allowed to offer either bachelor's or SC programs, while technological institutes and technical professional institutes are only allowed to offer SC programs. In what follows we use the terms "SC HEI" to refer to institutions specialized in SC programs (technological institutes and technical professional institutes), and "Other HEIs" to refer to institutions that offer bachelor's (and perhaps SC) programs (university and university institutes).

During the early 2000s, Colombia experienced a sizable increase in higher education enrollment (Camacho et al., 2017; Carranza and Ferreyra, 2019) due to an expansion in the number of programs offered by existing institutions as well as the opening of new institutions and campuses. As a result, the number of municipalities with at least one HEI increased by about 20% between 2000 and 2004.[6] We focus on local availability of higher education as defined by the presence of HEIs in the vicinity of each municipality. To this end, we identify HEIs within a 10-kilometer radius from each municipality's centroid.

Figure 1 displays the number of municipalities with SC HEIs and other HEIs in the 2000-2006 period. While the local supply of both HEI types grew, it grew faster for SC HEIs. Between 2000 and 2004 alone, the number of municipalities with SC HEIs grew by 40% relative to 2000, yet the number of municipalities with other HEIs grew by only 10%. These differential trends define the relative availability of higher education programs within a specific community. Our analysis seeks to uncover the effect of opening SC HEIs.

We focus on higher education supply as of 2004, one year before students in our sample graduated from high school. Of the 300 HEIs operating at the time, only one-third were SC HEIs. SC programs in Colombia are also provided by SENA (*Servicio Nacional de Aprendizaje*). This public institution has been providing workforce training for many decades; it only started offering SC programs in 2003 and had a negligible SC market share by 2004. It is not an HEI and operates outside the purview of the Ministry of Education. As a result, we do not include it in our analysis.

Of the 918 municipalities in our sample, 14% had an SC HEI within a 10 km radius (Table 3, municipality-level panel). This figure was 28% for other HEIs. Moreover, during our period of analysis, the vast majority of SC HEIs were located in cities that already

[6] Own calculations using data from the National System of Information on Higher Education (SNIES, in Spanish).

included other HEIs. Since the local HEI supply is likely to affect students' enrollment choices, we use as instruments the availability of SC HEIs and other HEIs.

3 Data

We use individual-level administrative data on student background characteristics, high school exit exam test scores, higher education enrollment choices and trajectories, and formal labor market outcomes for the universe of high school graduates in 2005. Individual records from the high school exit exam (*Saber 11*) contain test scores for math, reading, science, physics, history, chemistry, geography, and philosophy, as well as student age, gender, and socioeconomic characteristics (e.g., mother's education, household income level, and number of siblings) at the time of the test. Higher education enrollment, trajectories, and completion come from the System for Dropout Prevention of Higher Education (SPADIES, in Spanish).

The initial sample includes the universe of Saber 11 test-takers in 2005, restricted to those who were 14-24 years old (93.5% of the full set of test takers). We further restrict the sample to students with data on all socioeconomic variables and test scores. We match this sample with higher education enrollment data from SPADIES, and set an enrollment window of five years after high school graduation. We restrict graduation to take at least two and four years for those in SC programs and bachelor's programs, respectively. In addition, we restrict graduation age to the 19-30 year-old range. We remove students who enrolled in more than four programs between 2005 and 2015 (less than one percent of students).

Wages and formal employment for higher education graduates are available in the Labor Market Observatory for Education (OLE, in Spanish) for 2008-2013.[7] For higher education graduates, we restrict entry age in the labor market to be at least 20 years. Our datasets do not include labor market information for students who did not enroll in higher education, enrolled but dropped out, or enrolled but were still in college by 2013. For these students, we use Colombia's Integrated Household Survey (GEIH, in Spanish) from the Socioeconomic Database for Latin America and the Caribbean (SEDLAC) to impute labor market participation and wages between 2008-2013. We estimate regressions of formal labor market participation and wages as a function of gender, age, household characteristics, an indicator of urban area, and region fixed effects for the cohort of individuals who were 14-24 years old in 2005, and use predicted labor market outcomes for the imputations (see Appendix A.3 for further details).

[7] About 3% of students who enroll in higher education graduate after 2013 and we therefore lack OLE data for them. We classify students who graduate after 2013 as having incomplete higher education. Dropping them from the estimation sample does not affect our results (not shown).

We combine administrative records at the student level with information on higher education supply from the National System of Information on Higher Education (SNIES, in Spanish), which contains HEI-level information such as HEI type, geographic location, and number and types of programs offered. Using SNIES, we classify HEIs into those that can only offer SC programs (SC HEIs) and those than can offer all programs (other HEIs). We compute binary indicators for the existence of these HEI types in our students' high school municipalities by using HEIs' geographic locations reported in SNIES. We link our dataset to the Municipal Panel of the Center of Studies on Economic Development (CEDE, in Spanish), which contains municipality-level information such as population (urban, rural, and total), GDP, and area. We match the municipal- and student-level data using the municipality identifier of students' high school. Last, we restrict our estimating sample to students in high schools with more than 20 students. The final dataset consists of $328,358$ students, or 81% of all *Saber 11* test-takers in 2005.

3.1 Descriptive statistics

As Table A1 shows, 53.8% of high school graduates in the sample are women, 27.2% come from low-income households (with income lower than the minimum wage), and 42.5% have mothers with at most primary education. More than half of high school graduates (61.2%) did not enroll in higher education ("NE students").[8] Meanwhile, 9.4% enrolled in SC programs ("SC students") and 29.4% in a bachelor's program ("BP students"). Further, 3.6% and 14.8% of high school graduates obtained a SC degree and bachelor's degree, respectively, and half (49.3%) participated in the formal labor market in 2013.

Table 1 compares SC, BP, and NE students. Among the three student groups, NE students come from the most disadvantaged backgrounds and have the lowest *Saber 11* average performance. Relative to BP students, on average SC students are less likely to be female. They are also slightly older; belong to larger, lower-income households; have less educated mothers; and perform worse in *Saber 11*. The share of students who complete any degree is higher among BP than SC students by almost 10 percentage points, while the share of BP students transferring out of their initial enrollment choice is lower than that of SC students (1.2% and 16.2%, respectively). Moreover, less than half of SC students who transfer to a bachelor's program earn a bachelor's degree. This is consistent with the fact that SC programs in Colombia are usually viewed as terminal degrees and rarely offer a pathway towards a bachelor's degree. In terms of labor market outcomes, SC and BP students participate in the formal market at similar rates (51.7%

[8]These students do not match to SPADIES (the dataset of higher education enrollment) or OLE (the dataset of labor market earnings). SPADIES does not contain information on SENA students; therefore, we classify them as NE. Since they account for a negligible share of higher education enrollment in 2005, this does not pose a problem for our analysis.

and 54.7%, respectively) but NE students participate less (47%). Average monthly wages for BP students are about 23% higher than for SC students and almost twice as high as for NE students.

Table 2 presents average characteristics of students' high school municipalities. Relative to SC and BP students, on average NE students live in municipalities that are smaller and less urban, with lower GDP and higher homicide rates. SC students, in turn, come from larger, denser, slightly more urban, and higher-GDP municipalities than BP students. As explained below, these patterns are related to the geographic location of SC HEIs.

3.2 Identifying variation

To capture variation in local supply of higher education, we use SNIES data from 2004 (again, one year prior to high school graduation in our sample). We first identify a student's high school location as given by its administrative division or municipality and find the coordinates for its geographic centroid. We then construct the binary variables Z_j and H_j, where j denotes the municipality of interest. Z_j is equal to one if there is at least one SC HEI within a 10-km radius around municipality j's centroid, and zero otherwise. H_j is equal to one if there is at least one HEI offering bachelor's (and perhaps SC) degrees–namely, at least one HEI belonging to the "Other HEIs" category.[9]

As expected, the local supply of HEIs is correlated with students' enrollment choices (bottom panel of Table 2). We observe that 73.7% of SC students went to high school in a municipality where there was at least one SC HEI, but this share drops to 64.6% and 56.2% for BP and NE students, respectively. In contrast, approximately the same share (84%) of SC and BP students attended high school in municipalities where other HEIs were available. This is because, as of 2005, institutions providing bachelor's programs (captured by $H = 1$) were already widespread across cities of all sizes but SC HEIs (captured by $Z = 1$) were mostly present in large and medium-sized cities. Indeed, 77.5% of all students attended high school in municipalities with institutions providing bachelor's programs but only 60% did so in municipalities with SC HEIs.

Each column of Table 3 denotes a combination of the binary variables of higher education local supply, Z and H. While more than half of students in our sample lived in municipalities with both HEI types, 20% lived in municipalities without any higher education supply, and three percent in municipalities with only SC HEIs. The availability of SC HEIs is related to SC enrollment, as the share of SC students is higher in municipalities including an SC HEI than in others, regardless of the supply of other HEIs (8.8%

[9]Neither H nor Z include SENA, which is not an HEI. We focus exclusively on in-person programs and exclude distance or virtual programs.

vs. 6.2% when $H = 0$, and 11.6% vs. 6.3% when $H = 1$). Nonetheless, the difference in SC enrollment related to the availability of SC HEIs (equal to 2.6 and 5.3 percentage points, respectively, for $H = 0$ and $H = 1$) is larger when other HEIs are available than when they are not–an issue to which we return below.

Students in our sample are spread across 918 municipalities. Students without any kind of HEI were spread across 624 small municipalities (population below $400,000$). In contrast, students with both HEI types were spread across 94 municipalities, and half of them lived in large cities (population above 2.2 million).

4 Empirical Strategy

We focus on two labor market outcomes: probability of working in the formal labor market and (log) average monthly wages, both measured in 2013. Let Y_{ij} denote labor market outcomes for student i in municipality j. Let D_{SCij} be a binary variable that equals one if student i in municipality j enrolls in a SC program, and zero otherwise.

We use IVs to control for self-selection into programs. In particular, we use the binary variables that capture variation in local supply of higher education, Z_j and H_j. Both instruments follow the logic of cost-shifters (as in Card, 1995 and Mountjoy, 2022).[10]

To formalize this setting, we introduce the following notation. Let $\tilde{Z} = (Z_j, H_j)$ with $Z_j \in \{0,1\}$ and $H_j \in \{0,1\}$, with support $\mathcal{Z} = \{0,1\} \times \{0,1\}$. Let D_{dij} be a binary variable denoting student i's enrollment choice. Thus, D_{dij} equals one if student i chooses option $d \in \{SC, BP, NE\}$, and zero otherwise, where SC, BP, and NE denote SC program enrollment, bachelor's program enrollment, and no higher education enrollment, respectively. Let $Y_{ij}(0), Y_{ij}(1)$ denote potential outcomes for high school student i under two different regimes: "No enrollment in SC program" ($D_{SCij} = 0$) and "Enrollment in SC programs" ($D_{SCij} = 1$), respectively. Similarly, define $D_{SCij}(\tilde{z})$ as a potential enrollment choice that equals one if student i chooses a SC program when the instruments in $\tilde{Z}$ are equal to $\tilde{z}$, where $\tilde{z} \in \{0,1\} \times \{0,1\}$. For instance, $D_{SCij}(Z_j = 1, H_j = 1) = 1$ denotes students who would enroll in a SC program when all HEI types are available.

Following the literature, identification relies on two assumptions. First, we assume conditional independence: $(Y_{ij}(0), Y_{ij}(1), D_{SCij}(\tilde{z})) \perp\!\!\!\perp \tilde{Z} | X$ where X contains student characteristics (age, gender, number of siblings, household income level, mother's education, and *Saber 11* test scores) as well as characteristics of the student's high school municipality. This assumption states that, conditional on the student's background and

[10] It is reasonable to assume that students with a SC HEI in the vicinity of their municipality face lower enrollment costs than students without it. Enrollment costs can be related to access (*e.g.*, having a nearby HEI reduces transportation costs) or information (*e.g.*, having a nearby HEI makes it easier for the student to learn about its offerings).

municipality characteristics, whether or not there exists a SC HEI in a 10-km radius is uncorrelated with future outcomes. Plus, it also states that the existence of a SC HEI in the student's high school municipality affects her future labor market outcomes only by affecting her higher education enrollment choice.[11] Second, we assume partial monotonicity (Mountjoy, 2022; Mogstad et al., 2021). In our setting, this states that every student should be (weakly) induced towards SC programs as Z_j rises, conditional on H_j. Formally, we assume that $D_{SCij}(0, H_j) \leq D_{SCij}(1, H_j)$. Thus, $Z_j = 1$ weakly induces students to choose SC programs regardless of the existence of other HEIs in municipality j's vicinity. Under these assumptions, we interpret our point estimates as a weighted average of local average treatment effects (LATEs) for the two subpopulations of interest (compliers along the diversion and expansion margins).[12]

We estimate the impact of SC programs on the labor market outcomes of individuals who would choose SC programs in response to the local supply of SC HEIs (Z_j), controlling for the local supply of other HEIs (H_j). We use Z_j as an instrument and control for the other instrument, H_j, in the first and second stage (Mogstad et al., 2021).[13] Our first-stage equation is:

$$D_{SCij} = \gamma_0^{SC} + \gamma_1^{SC} Z_j + \gamma_2^{SC} H_j + \gamma_X^{SC} X_{ij} + \epsilon_{SCij}, \tag{1}$$

where X_{ij} includes age, gender, number of siblings, household income level, mother's education, *Saber 11* test scores, and characteristics of the student's high school municipality (listed in Table 2). Recall that all variables in X_{ij} are measured prior to the enrollment decision. The outcome (second-stage) equation is:

$$Y_{ij} = \beta_0 + \beta_{SC} D_{SCij} + \beta_H H_j + \beta_X X_{ij} + u_{ij}, \tag{2}$$

[11] HEIs' endogenous location represents a threat to our identification strategy. Suppose, for example, that HEIs tend to locate in large municipalities. Then, as we consider smaller radii around the centroid of the student's high school municipality, the identifying variation could correlate more with factors associated with living near a city. If those factors drove our findings, our point estimates would be sensitive to different radii values. Section 5.3 shows that our estimates are stable to radii changes. Even if we cannot rule out the possible link between HEI supply and proximity to cities, our estimates could be viewed as lower bounds of SC program effects if individuals farther from the municipality's centroid gained more than others from these programs.

[12] Recent work underlines the role of covariates in the interpretation of TSLS estimates as LATEs. Słoczyński (2021) and Blandhol et al. (2022) show that only under flexible specifications and strong monotonicity assumptions can one interpret TSLS estimates as convex combinations of conditional (on covariates) LATEs. While we recognize this caveat, the approach in this paper entails the appropriate TSLS specification under multiple instruments. For robustness, we also present results using a saturated model for covariates (which includes interactions between H_j and the covariates in X_{ij}); see Section 5.3.

[13] By using one instrument separately while controlling for the other, we guarantee that the weights on different effects for complier groups are positive (Heckman et al., 2006; Mogstad et al., 2021). This approach isolates the variation in Z_j, which denotes the policy-relevant effect of an expansion in the local supply of SC HEIs.

where ϵ_{SCij} and u_{ij} are unobserved components, such as student preferences and unobserved ability.

The previous setting does not account for the potential interaction between the two instruments. However, the effect of Z_j on D_{SCij} might depend on the supply of other HEIs, H_j. For instance, the opening of an SC HEI might induce more students to choose a SC program in municipalities without other HEIs–where they introduce the very first higher education option–than in municipalities with other HEIs. To examine this possibility, we estimate a saturated version of our model in which we also control for the interaction between the IVs in equations (1) and (2). These equations estimate β^{SC} for $H_j = 0$. To estimate β^{SC} when $H_j = 1$, we transform H_j into $\tilde{H}_j = H_j - 1$ and estimate TSLS in a model where Y_{ij} is written as a function of D_{SCij}, $\tilde{H}_j$, $\tilde{H}_j \times Z_j$ and X_{ij}; and D_{SCij} is determined by Z_j, $\tilde{H}_j$, $\tilde{H}_j \times Z_j$ and X_{ij}. In this way, Z_j becomes the excluded instrument (from the second-stage). This approach produces the same coefficients as those obtained by implementing a Wald estimator of the total effect of Z_j and $H_j \times Z_j$ on Y_{ij} divided by the total effect of Z_j and $H_j \times Z_j$ on D_{SCij}. Section 5 discusses estimates from the different specifications and confirms the importance of interacting the instruments.

4.1 Expansion and diversion margins

We are interested in β_{SC}, the labor market effect of enrolling in a SC program. For a given student, the effect depends on her counterfactual choice– enrolling in a bachelor's program (BP) or not enrolling in higher education at all (NE)–which can in turn vary across individuals.

We define two groups of SC students according to their counterfactual choices: (i) students who would not have enrolled in higher education (choosing *NE* instead) and (ii) students who would have enrolled in a bachelor's program (choosing *BP* instead). Rouse (1995) defines the former as the *democratization* or *expansion margin* (students who enter higher education via SC programs), while the latter is referred to as the *diversion margin* (students who divert from BP into SC programs). Changes in the local supply of SC HEIs could shift students along those margins into SC. Students who would change their choices due to Z_j changes are defined as compliers. Our setting includes two complier groups: along the expansion and diversion margin.

Depending on their margin, compliers might experience different effects from enrolling in a SC program. Let $LATE_{SC-BP}$ and $LATE_{SC-NE}$ represent the local effects of SC programs for compliers along the diversion and expansion margin, respectively. Previous literature (Heckman et al., 2006; Heckman and Urzúa, 2010) has shown that, when compliers switch away multiple initial states, a standard (univariate) instrumental variable approach identifies the effect of one option versus the next best–in our case, SC versus a mixture of NE and BP. Formally,

$$\beta_{SC} = \omega \times \text{LATE}_{SC-BP} + (1-\omega) \times \text{LATE}_{SC-NE}, \tag{3}$$

where ω is the share of compliers along the diversion margin, or diversion share. We estimate the LATE of SC programs vs. the next best, β_{SC}, by estimating equations (1) and (2) using TSLS. While we are able to identify the diversion share, the margin specific LATEs cannot be separately identified without additional assumptions.[14] Nonetheless, if the estimated share of compliers along one of the margins is low, one can view the LATE of SC programs vs. the next best as largely reflecting the local effect for students along the other margin.

4.2 Complier shares

Next, we describe how to estimate ω in equation (3) for the case of one binary instrument under partial monotonicity.[15] We assume that the enrollment choice, D_{dij} with $d \in \{SC, BP, NE\}$, depends on H_j, Z_j and X_{ij}. We also follow the logic of partial monotonicity and impose two additional assumptions. First, $D_{BPij}(0, H_j) \geq D_{BPij}(1, H_j)$, so that having access to SC HEIs makes BP weakly less preferable. Second, $D_{NEij}(0, H_j) \geq D_{NEij}(1, H_j)$, so that access to SC HEIs makes NE weakly less preferable. We can then identify the share of compliers switching at each margin as:

$$\omega = \frac{E[D_{BPij}|Z_j = 0, H_j] - E[D_{BPij}|Z_j = 1, H_j]}{E[D_{SCij}|Z_j = 1, H_j] - E[D_{SCij}|Z_j = 0, H_j]},$$

$$1 - \omega = \frac{E[D_{NEij}|Z_j = 0, H_j] - E[D_{NEij}|Z_j = 1, H_j]}{E[D_{SCij}|Z_j = 1, H_j] - E[D_{SCij}|Z_j = 0, H_j]},$$

and the total share of compliers as $E[D_{SCij}|Z_j = 1] - E[D_{SCij}|Z_j = 0]$.

To estimate complier shares, we first estimate propensity scores for every enrollment choice in $d \in \{SC, BP, NE\}$. We posit the following linear probability model:

$$D_{dij} = \gamma_0^d + \gamma_1^d Z_j + \gamma_2^d H_j + \gamma_X^d X_{ij} + \epsilon_{dij}, \tag{4}$$

where D_{dij} is a binary variable equal to one if the student chooses option d, and ϵ_{dij} is the error term. Note that, for the probability of enrolling in SC programs, equation (1)

[14] See, for instance, Hull (2018); Kirkeboen et al. (2016); Kline and Walters (2016).

[15] Mountjoy (2022) presents a formal derivation of the shares and estimates them using continuous instruments.

is equivalent to equation (4) when $d = SC$.[16]. Taken together, the estimated diversion ($\hat{\omega}$) and expansion shares ($1 - \hat{\omega}$) are given by:

$$\hat{\omega} = \frac{-\hat{\gamma}_1^{BP}}{\hat{\gamma}_1^{SC}} \text{ and } 1 - \hat{\omega} = \frac{-\hat{\gamma}_1^{NE}}{\hat{\gamma}_1^{SC}}. \tag{5}$$

In the saturated model, where we include the interaction of Z_j and H_j, the estimated ω might vary depending on H_j. For students without other HEIs nearby ($H_j = 0$), we would expect a larger expansion share than for those with other HEIs. As described above, when a SC HEI opens in a market without other HEI types, students who would not have entered higher education (due, perhaps, to the absence of local higher education options) might be now induced to enter. In contrast, in markets where BPs were already available ($H_j = 1$) and where students could have entered the higher education system to pursue them, the opening of SC HEIs might not attract as many new students into higher education. Instead, it might attract some who would have otherwise chosen BPs but might prefer the shorter, more practical SC programs.

To characterize the type of student who responds to a SC expansion, we compute for compliers the average of some variables in X_{ij}. Let $\tilde{X}_{ij}$ denote a variable in X_{ij}. To estimate compliers' average for $\tilde{X}_{ij}$, we follow Abadie (2002) and compute:

$$Pr(\tilde{X}_{ij}|D_{SCij}(1, H_j) - D_{SCij}(0, H_j) = 1) = \frac{E(D_{SCij}\tilde{X}_{ij}|Z_j = 1, H_j) - E(D_{SCij}\tilde{X}_{ij}|Z_j = 0, H_j)}{E(D_{SCij}|Z_j = 1, H_j) - E(D_{SCij}|Z_j = 0, H_j)}.$$

In practical terms, we estimate a modified TSLS where the first-stage is equation (1) and the second stage is the regression of $D_{SCij}\tilde{X}_{ij}$ on D_{SCij}, Z_j, and H_j (and the interaction of the instruments, for the saturated model). We cluster standard errors at the high school municipality level.

5 Main Results

We first estimate the probability of enrolling in SC programs, bachelor's programs, or not enrolling in a HEI as a function of the instruments associated with the local supply of HEIs and the covariates (equation (4)). The corresponding estimates are shown in Tables 4 and Table A2, respectively.

For each enrollment choice, Table 4 presents estimates for two different specifications: with Z_j and H_j as controls ("No-interaction" column), and with Z_j, H_j, and $Z_j \times H_j$ as controls ("Interaction" column). From the no-interaction results, we find that having

[16] We also estimate propensity scores using a multinomial logit model. See Section 5

a SC HEI in a 10-km radius increases the probability of choosing SC programs by 3.7 pp and decreases the probability of choosing bachelor's programs by 2.9 pp, but has a negligible (-0.8 pp) and marginally significant effect on the probability of not enrolling in higher education.[17,18]

The effect of having a SC HEI in the high school municipality is concentrated among students whose municipalities also include other HEI types. The "Interaction" column in Table 4 shows that the direct effect of Z_j on the probability of choosing a SC program is not statistically significant and is negligible. In contrast, the interaction effect is substantive (3.9 pp) and statistically significant. In the bottom panel of Table 4, we compute the total effect of having an SC HEI when there are other HEIs in the municipality ($H_j = 1$), and observe a large increase (4.5 pp) in the probability of choosing a SC program. This is accompanied by a 3.2-pp decrease in the probability of choosing a bachelor's program and a 1.2-pp decrease in the probability of not enrolling in higher education.[19]

5.1 Margins and complier shares

Based on the previous estimates, Table 5 presents the share of students along the expansion and diversion margin, the total share of compliers, and the diversion share. In the "No Interaction" columns, the expansion and diversion margins correspond to the coefficient on Z_j in the "Not Enrolled" and "Bachelor's Program" columns of Table 4, respectively. Similarly, in the "Interaction" columns of Table 5, they correspond to the "Total Effects of Z when H=1" from Table 4. We present results for all students as well as separately by gender.

The estimated diversion share, $\hat{\omega}$, is in the 0.7-0.8 range, overall and by gender. In other words, most compliers respond along the diversion margin. The overall share of compliers is higher among men than women (5.3% v. 3.8% respectively in the interaction model), which is consistent with the lower share of female students in SC than bachelor's programs. Since the diversion share is at least 70% across all samples and models, the expansion margin captures less than 30% of compliers. We argue, therefore, that the estimated labor market effects presented below are largely driven by diversion compliers.

[17] In terms of the covariates, results in Table A2 suggest that females have a lower probability of enrolling in SC programs than males, and students with more siblings are less likely to select a SC program. An increase in *Saber 11* scores renders students more likely to choose BP and less likely to choose SC programs or NE.

[18] Table A3 reports marginal effects from a multinomial logit model instead of a linear probability model. Compared to the results displayed in Table 4, we observe minor differences in the instruments' marginal effects.

[19] Although we do not focus on the direct effect of H_j, it is worth noting that other types of HEIs seem to shift students along the expansion margin (going from NE to BP) as well as the diversion margin (going from SC to BP). However, the total variation in enrollment choices induced by changes in H_j is about 33% lower than the corresponding variation from changes in SC supply.

5.2 Labor market effects

To recover the labor market effects of SC programs for compliers (relative to the next-best alternative), we use TSLS to estimate the model described by equations (1) and (2). Table 6 reports estimates for the two outcomes of interest (employment and monthly wages in the formal labor market) for all our specifications and samples. Panel A contains point estimates of β_{SC} for the full sample. It shows that the effects of SC programs vs. the next-best alternative are generally larger, in absolute value, when interactions are included than when they are not. For the full sample, SC effects on formal employment are not significant, but the SC effects on (log) wages are negative and significant. For instance, the interaction model indicates that a SC program lowers average monthly wages by about 23 percent relative to the next-best alternative.[20] For all outcomes and specifications we include diagnostic tests for weak instruments (Kleibergen-Paap and Cragg-Donald F-statistics), which are all well above the corresponding critical values.

We find heterogeneous effects by gender. Female compliers benefit from SC programs, at least in terms of employment opportunities (panel B of Table 6). Their probability of formal employment rises by 31 pp (in the interaction model) while effects on their wages are not significant. For males, in contrast, effects are negative for both outcomes (panel C). Choosing an SC lowers their formal employment probability by 24 percentage points and their average monthly salaries by about 23 percent. Based on these estimated effects as well as the large diversion share for males and females, we conclude that female compliers might experience employment gains from choosing an SC rather than a bachelor's program, but male compliers might experience employment and wage losses.

These stark differences might be driven by pre-higher education differences between male and female compliers. Thus, we estimate average pre-higher education enrollment characteristics by gender. Figure 2 displays the estimated share of compliers with an above-median *Saber 11* score. We compare five different groups: all bachelor's program students, all SC program students, all compliers, male compliers, and female compliers. In general, compliers outperform the average SCP student but not the average bachelor's program student. Male compliers outperform female compliers in all subjects except Reading and, notably, the share of male compliers scoring above the median in mathematics is almost as high as among bachelor's students. While compliers are more advantaged than the average SC program student in terms of household income and mother's level

[20] Table A4 presents point estimates obtained via OLS (regressing Y on D_{SC} and X) and via a "standard" TSLS, where both Z_j and H_j are excluded from the outcome equation). First, as expected, OLS estimates are substantially different from TSLS estimates. Second, TSLS estimates are similar to those in Table 6, which might be expected if the IV weights are positive when both IVs were excluded. We confirm that this is indeed the case using the test for positive IV weights in Mogstad et al. (2021) (results available upon request). The largest differences between the TSLS estimates and those in Table 6 are in the interacted model, but the results are largely consistent in terms of magnitude and significance.

of education (see Figure 3), male compliers are more advantaged than female compliers in these dimensions as well. Relative to the average SC program student, compliers are more likely to live in medium-sized cities and less likely to live in big cities, without noticeable differences between males and females in this regard.

Overall, then, female compliers are more disadvantaged and less academically proficient than male compliers. Their next-best option may have been a low-quality bachelor's program, relative to which SC programs may have provided a better match and improved employment outcomes. In contrast, male compliers may have diverted away from good bachelor's programs and into SC programs that constituted an "undermatch" to their academic preparation and background, thereby worsening their labor market outcomes.

5.3 Robustness

Different radii. An identification concern is that all variation might come from students right at, or near, the 10 km radius. If that is the case, our findings might be sensitive to the radius selection. The composition of compliers might change if a larger radius induces a higher share of students to enroll and creates more compliers along the expansion margin. Since a 10-km radius yields a large diversion share, our results on SC enrollment might be interpreted as lower bounds if expansion compliers are more prevalent at other radii.

To assess the robustness of our findings, we estimate the model using different radii values. Figure 4 (top panel) shows that the share of students with available HEIs increases as the radius increases from 5km to 20km (first plot to the left), the share of compliers decreases but remains stable after 10km (middle plot), and the diversion share also increases (top-right plot). These results are expected. For instance, a larger radius of HEI availability raises enrollment costs (e.g., transportation costs to the HEI) and therefore lowers the share of compliers. By the same logic, expansion compliers are more prevalent at smaller radii because enrollment costs are lower. The top-right plot compares the diversion share with and without interactions. The diversion share is less sensitive to the radius choice in the interaction than the no-interaction model.

The bottom panel of Figure 4 displays TSLS estimates of labor market effects for different radii. Despite changes in the share of compliers and the diversion share, estimates are stable. In the case of employment effects (left panel), although the effects change sign for larger radii, they are not statistically significant in any specification. Similarly, estimates for log wages (right) fluctuate in magnitude and significance but remain negative across all radii.[21]

[21]We present the share of compliers and TSLS estimates for different radii for females and males in Figures A1 and A2 in the Appendix. We observe relatively stable estimates across radii values,

SENA: A Threat to Identification? We have not included SENA students in our analysis because SENA captured a negligible share of higher education enrollment in the early 2000s, and SENA students are not included in the student-level higher education dataset, SPADIES. Students in SENA, therefore, have been classified as "not enrolled" for our estimation. In principle, this could affect the interpretation of our findings.

Figure A3 presents the evolution of higher education enrollment in Colombia by program type. Panel (a) excludes SENA enrollment and leads to the conclusion that higher education enrollment has grown mainly through bachelor's programs. This is consistent with SC programs having a small expansion margin. However, the story changes when we include SENA enrollment, which expands dramatically between 2005 and 2015 (panel b). Further, the distribution of SC students across cities changes over time and when SENA enrollment is included. The left side of Figure A4 shows that, before 2005, most SC enrollment outside SENA was concentrated in large cities; only later–and very slowly–did it rise in medium and small cities. SENA, however, was particularly aggressive in those markets (right side). A similar picture emerges when looking at number of SC programs rather than enrollment (Figure A5).

The fact that most SC programs were offered in large cities in the early 2000s helps explain our small expansion effects. Since those cities already had HEIs offering bachelor's programs, those students might not have been persuaded to enter higher education just because of the opening of an SC HEI. In contrast, in smaller cities without other HEIs, such openings might have induced higher education entry. Data from later years (including more SC HEIs in smaller cities), especially if including SENA students, might have uncovered a larger expansion margin. Our results, therefore, illustrate what might happen when SC programs expand in markets in which other, more widely preferred higher education options are already available.

Fully saturated model. We investigate the sensitivity of our main findings to model specification. In particular, we estimate a *fully* saturated model where H_j interacts with the entire vector of covariates in X_{ij}. Table A8 displays estimates for this model, which can be compared to those in Table 6. We do not observe large discrepancies across specifications.

particularly for females. We also present additional results for radii greater than 20 km in Tables A5, A6, and A7 in the Appendix. Our preferred findings (with a 10 km radius) remain largely stable across radii up to 20km. After that, the share of students with SC HEIs is above 70%, and above 90% for other HEIs. Since the majority of students have HEIs available at radii greater than 20 km, such large radii wipe out most of the identifying variation and render the estimates less informative.

6 Conclusion

In this paper, we estimate heterogeneous economic returns to SC programs in Colombia. We account for the self-selection of students across multiple higher education alternatives (bachelor's programs, SC programs, and not enrolling in higher education). In line with the literature, we show that the effects of SC programs largely differ depending on the fallback alternative (next-best option) of students. We use an Instrumental Variables (IV) approach, where identification comes from the supply of SC programs at the municipal level. More specifically, we define a binary instrument to indicate whether a municipality has an institution specialized in offering SC programs in a 10-kilometer radius.

We present three main findings. First, we document 3.7-4.5 pp increase in SC program enrollment as the local supply of these programs rises. Students who respond to the supply increase (compliers) are concentrated in municipalities where bachelor's programs were already available. Second, at least 70% of compliers would react along the diversion margin, switching from bachelor's programs into SC programs. The remaining 30% would react along the expansion margin, entering the higher education system. These patterns are similar for males and females. Third, females experience gains from SC programs in terms of formal employment while males experience losses in formal employment and wages. Based on the high diversion share and the fact that males are more advantaged and better prepared academically than females, we argue that SC programs might constitute an "undermatch" for males, who might have otherwise attended good bachelor's programs, but a better match for females, who might have otherwise attended low-quality bachelor's programs.

While different from findings for the U.S., our findings can be reconciled with those. Using data from Texas, Mountjoy (2022) finds that the expansion margin share is about two-thirds, much higher than our 30%. One possible explanation for this discrepancy is that SC programs (typically taught at community colleges in the U.S.) are often viewed as an entry point into the higher education system, with students intending to transfer to bachelor's programs. In Colombia, in contrast, SC degrees are viewed as terminal and rarely offer a pathway towards a bachelor's program, as they mostly focus on facilitating the school-to-work transition. Another possible explanation is that many community colleges in the U.S. are located outside large cities, where they constitute the only higher education option for local students. In contrast, during our sample period SC programs in Colombia were mostly offered in large and medium cities that already other higher education options, leaving little room for newly entering students into the system. Further, Mountjoy (2022) finds that compliers along the expansion margin are more likely than those on the diversion margin to benefit in terms of average earnings and years of education. Nonetheless, in policy simulations that raise the diversion share above 70%, he finds a negative net effect of SC programs, which is in line with our results for male compliers. All in all, our findings underscore the fact that a seemingly "good" policy such as expanding SC programs may have unintended consequences depending on students' other options and background characteristics, and may therefore require thoughtful design.

References

Abadie, A. (2002). Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models. *Journal of the American Statistical Association*, 97(457):284–292. Publisher: [American Statistical Association, Taylor & Francis, Ltd.].

Acevedo, K. and Bornacelly, I. (2014). Panel Municipal del CEDE. *Documento Centro de Estudios sobre Desarrollo Económico*, 2014-26.

Acton, R. (2020). Effects of Reduced Community College Tuition on College Choices and Degree Completion. *Education Finance and Policy*, pages 1–71.

Barrera-Osorio, F. and Bayona-Rodríguez, H. (2019). Signaling or better human capital: Evidence from Colombia. *Economics of Education Review*, 70:20–34.

Bettinger, E. and Soliz, A. (2016). *Returns to Vocational Credentials: Evidence from Ohio's Community and Technical Colleges. A CAPSEE Working Paper*. Center for Analysis of Postsecondary Education and Employment.

Blandhol, C., Bonney, J., Mogstad, M., and Torgovitsky, A. (2022). When is TSLS Actually LATE? Working Paper 29709, National Bureau of Economic Research. Series: Working Paper Series.

Camacho, A., Messina, J., and Uribe Barrera, J. (2017). The Expansion of Higher Education in Colombia: Bad Students or Bad Programs? SSRN Scholarly Paper ID 2921965, Social Science Research Network, Rochester, NY.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling, Aspects of labour market behaviour: essays in honour of John Vanderkamp. ed. *LN Christofides, EK Grant, and R. Swidinsky.*

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

Denning, J. T. (2017). College on the Cheap: Consequences of Community College Tuition Reductions. *American Economic Journal: Economic Policy*, 9(2):155–188.

Ferreyra, M. M., Avitabile, C., Botero Álvarez, J., Haimovich Paz, F., and Urzúa, S. (2017). *At a Crossroads : Higher Education in Latin America and the Caribbean*. Directions in Development—Human Development. World Bank, Washington, DC.

Ferreyra, M. M., Dinarte, L., Melguizo, T., and Sanchez Diaz, A. M. (2020). Estimating the Contribution of Short-Cycle Programs to Student Outcomes in Colombia.

Ferreyra, M. M., Dinarte, L., Urzúa, S., and Bassi, M. (2021). *The Fast Track to New Skills: Short-Cycle Higher Education Programs in Latin America and the Caribbean*. World Bank, Washington, DC. Accepted: 2021-05-19T15:16:32Z.

González-Velosa, C., Graciana, R., Sarzosa, M., and Urzúa, S. (2015). Returns to Higher Education in Chile and Colombia \textbar Publications. *IDB WORKING PAPER SERIES*, IDB-WP-587.

Heckman, J. J., Urzua, S., and Vytlacil, E. (2006). Understanding Instrumental Variables in Models with Essential Heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432.

Heckman, J. J. and Urzúa, S. (2010). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics*, 156(1):27–37.

Hull, P. (2018). IsoLATEing: Identifying Counterfactual-Specific Treatment Effects with Cross-Stratum Comparisons. SSRN Scholarly Paper ID 2705108, Social Science Research Network, Rochester, NY.

Jepsen, C., Troske, K., and Coomes, P. (2014). The Labor-Market Returns to Community College Degrees, Diplomas, and Certificates. *Journal of Labor Economics*, 32(1):95–121.

Kane, T. J. and Rouse, C. E. (1999). The Community College: Educating Students at the Margin between College and Work. *Journal of Economic Perspectives*, 13(1):63–84.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of Study, Earnings, and Self-Selection. *The Quarterly Journal of Economics*, 131(3):1057–1111.

Kline, P. and Walters, C. R. (2016). Evaluating Public Programs with Close Substitutes: The Case of Head Start. *The Quarterly Journal of Economics*, 131(4):1795–1848.

Leigh, D. E. and Gill, A. M. (2003). Do community colleges really divert students from earning bachelor's degrees? *Economics of Education Review*, 22(1):23–30.

Levesque, E. M. (2018). Improving community college completion rates by addressing structural and motivational barriers. *The Brookings Institution*.

Marcotte, D. E. (2019). The Returns to Education at Community Colleges: New Evidence from the Education Longitudinal Survey. *Education Finance and Policy*, 14(4):523–547.

Minaya, V. and Scott-Clayton, J. (2020). Labor Market Trajectories for Community College Graduates: How Returns to Certificates and Associate's degrees Evolve Over Time. *Education Finance and Policy*, pages 1–62.

Mogstad, M., Torgovitsky, A., and Walters, C. R. (2021). The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables. *American Economic Review*, 111(11):3663–3698.

Mountjoy, J. (2022). Community Colleges and Upward Mobility. *American Economic Review*, Forthcoming.

Rodríguez, J., Urzúa, S., and Reyes, L. (2016). Heterogeneous Economic Returns to Post-Secondary Degrees: Evidence from Chile. *Journal of Human Resources*, 51(2):416–460. Publisher: University of Wisconsin Press.

Rouse, C. E. (1995). Democratization or Diversion? The Effect of Community Colleges on Educational Attainment. *Journal of Business & Economic Statistics*, 13(2):217–224.

Słoczyński, T. (2021). When Should We (Not) Interpret Linear IV Estimands as LATE? *arXiv:2011.06695 [econ, stat]*. arXiv: 2011.06695.

Tables and Figures

Table 1: Summary Statistics, by Higher Education Choice

Variable	**Short-Cycle Program**	**Bachelor's Program**	**Not Enrolled**
Female	0.481	0.548	0.542
Age at Saber 11	16.782	16.550	17.292
Siblings	2.347	2.132	3.040
	Household Income Level		
<1 Minimum Wages (MW)	0.187	0.149	0.343
1-2 MW	0.506	0.353	0.463
2-3 MW	0.205	0.201	0.131
2-3 MW	0.102	0.296	0.063
	Mother's Education Level		
Primary	0.368	0.221	0.532
Secondary	0.431	0.325	0.357
Short-Cycle program	0.136	0.178	0.072
At least Bachelor's program	0.066	0.276	0.039
	Standardized Test Scores from the High School Exit Exam		
Math	0.003	0.361	-0.149
Reading	0.132	0.555	-0.240
Biology	0.044	0.506	-0.218
Physics	0.005	0.340	-0.144
History	0.070	0.464	-0.202
Chemistry	0.030	0.504	-0.214
Geography	0.056	0.424	-0.182
Philosophy	0.064	0.407	-0.173
	Educational Attaintment		
Short-Cycle Incomplete	0.495		
Short-Cycle Complete	0.342	0.012	
Bachelor's program Incomplete	0.101	0.502	
Bachelor's program Complete	0.062	0.485	
	Formal Labor Market Outcomes		
Works in 2013	0.517	0.547	0.473
Avg. Monthly Wage (2013)	909,590.9	1,120,830.8	548,288.1
N	30,803	96,504	201,051

Note: The sample corresponds to the universe of students who took the high school exit exam (Saber 11) in 2005. The information on educational attainment comes from SPADIES (System for Dropout Prevention of Higher Education), and labor market outcomes for higher education graduates are from OLE (Labor Market Observatory for Education). For high school graduates and those with incomplete higher education, we impute labor market participation and wages using household survey data from SEDLAC (Socioeconomic Database for Latin America and the Caribbean); see Appendix A.3 for details. Avg. Monthly wages are in Colombian pesos (COP).

Table 2: Municipality-level Summary Statistics, by Higher Education Choice

Variable	Short-Cycle Program	Bachelors' Program	Not Enrolled	All
Total GDP (billions COP)	27.379	25.114	18.853	21.493
Ratio of Urban/Rural Population	1.403	1.364	1.030	1.163
Homicide rate (per 1,000 inhabitants)	0.354	0.356	0.393	0.379
Total Population (in millions)	2.279	2.085	1.597	1.805
Area (per 10,000 squared km)	0.090	0.105	0.098	0.099
Availability of Higher Education Institutions (HEI)				
Z: HEIs only offer SC programs	0.737	0.646	0.562	0.603
H: Other type of HEIs	0.844	0.842	0.732	0.775
N	30,803	96,504	201,051	328,358

Note: The sample corresponds to the universe of students who took the high school exit exam (Saber 11) in 2005. The information on local characteristics comes from the Municipal Panel from CEDE (Center of Studies on Economic Development). The information on higher education programs and institutions is from SNIES (National System of Information on Higher Education).

Table 3: Local Supply of Higher Education Institutions (HEIs), at the Student- and Municipality-level

	Other HEI=0 $H=0$		Other HEI=1 $H=1$	
	HEI only offers SC=0 $Z=0$	HEI only offers SC=1 $Z=1$	HEI only offers SC=0 $Z=0$	HEI only offers SC=1 $Z=1$
Variable	**Student level**			
Share of students	19.9%	3.0%	20.1%	57.1%
N	64,053	9,917	66,302	188,086
Higher education enrollment				
Not Enrolled (NE)	73.6 %	68.6%	61.8%	56.4%
Short-Cycle Program (SC)	6.2%	8.8%	6.3%	11.6%
Bachelor's Program (BP)'	20.3%	22.6%	31.9%	32.0%
Total	100%	100%	100%	100%
Municipality size (population)				
Small municipality (<0.4 million)	100%	69.7%	87.9%	24.1%
Medium city (0.4-2.2 million)		30.3%	12.1%	28.9%
Big city (>2.2 million)				46.9%
Total	100%	100%	100%	100%
Municipality level				
Share of municipalities	68.0%	3.9%	17.9%	10.2%
N	624	36	164	94

Note: The sample corresponds to the universe of students who took the high school exit exam (Saber 11) in 2005. The information on population size comes from the Municipal Panel from CEDE (Center of Studies on Economic Development). The information on higher education programs and institutions is from SNIES (National System of Information on Higher Education).

Table 4: Probability of Enrollment in Higher Education as a Function of the Local Supply of Higher Education Institutions (HEIs)

	Short-Cycle Program		Bachelor's Program		Not Enrolled	
Variable	No-interaction	Interaction	No-interaction	Interaction	No-interaction	Interaction
Z: HEI only offers SC programs (1)	0.037***	0.006	-0.029***	-0.017**	-0.008*	0.011
	(0.003)	(0.005)	(0.005)	(0.007)	(0.005)	(0.008)
H: Other type of HEI	-0.005*	-0.013***	0.023***	0.026***	-0.018***	-0.013**
	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.005)
HEI only offers SC × Other HEI (2)		0.039***		-0.015*		-0.024**
		(0.006)		(0.009)		(0.010)
Total effect of **Z** when *Other HEI=1*: HEI only offers SC + HEI only offers SC × Other HEI						
(1)+(2)		0.045***		-0.032***		-0.012**
		(0.004)		(0.005)		(0.005)
N	328,358	328,358	328,358	328,358	328,358	328,358
R^2	.021	.021	.242	.242	.226	.226

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, clustered at the high school level. Each choice probability is estimated with a separate Linear Probability Model. The "No-interaction" columns present results without interacting the higher education supply variables ("HEI only offers SC" and "Other type of HEI"); the "Interaction" columns present results interacting "HEI only offers SC" and "Other type of HEI". All regressions include region fixed effects as well as the set of covariates, X_{it}. Table A2 in the Appendix reports the estimated coefficients on X_{it}.

Table 5: Compliers, Margins, and Diversion Share

Expansion margin NE→ SC (1)		Diversion margin BP→ SC (2)		Complier share SC (1)+(2)		Diversion share $\hat{\omega}$ (2)/(1)+(2)	
No-interaction	Interaction [Other HEI=1]	No-interaction	Interaction [Other HEI=1]	No-interaction	Interaction [Other HEI=1]	No-interaction	Interaction [Other HEI=1]
Full sample [N = 328,358]							
0.008*	0.012**	0.029***	0.032***	0.037***	0.045***	0.790***	0.724***
(0.005)	(0.005)	(0.005)	(0.005)	(0.003)	(0.004)	(0.116)	(0.110)
Female [N = 176,651]							
0.008	0.012*	0.022***	0.026***	0.030***	0.038***	0.741***	0.692***
(0.005)	(0.006)	(0.005)	(0.006)	(0.004)	(0.004)	(0.173)	(0.157)
Male [N = 151,707]							
0.008	0.013**	0.038***	0.040***	0.046***	0.053***	0.831***	0.756***
(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.114)	(0.110)

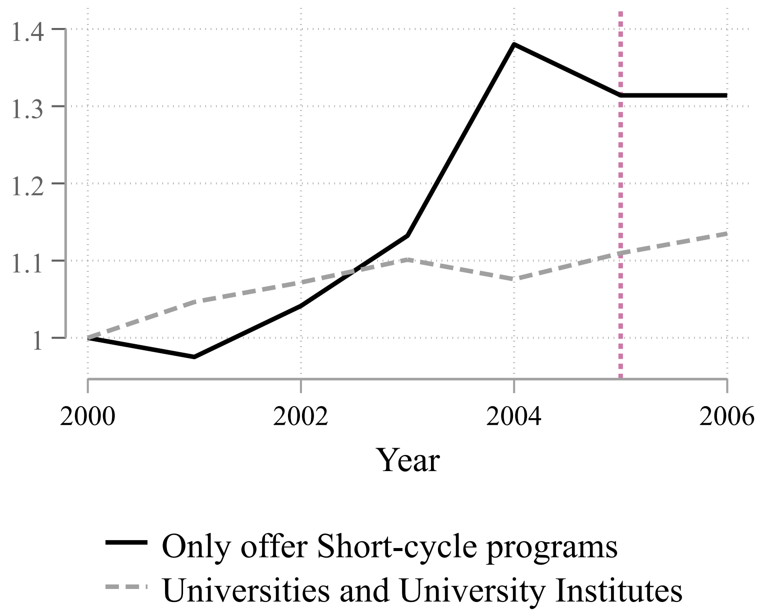
Note: This table shows the estimated share of students who react to the variation in the local supply of short-cycle HEIs by switching from not enrolling towards short-cycle programs in block (1) and by diverting from bachelor's and into short-cycle programs in block (2), as well as the total complier share in block (1)+(2). The last block shows the diversion share, $\hat{\omega}$, which is the ratio of (2)/(1)+(2). The "No-interaction" columns present results without interacting the variables of supply of higher education ("HEI only offers SC" and "Other type of HEI"); the "Interaction [Other HEI=1]" columns present results interacting "HEI only offers SC" and "Other type of HEI" and evaluating the effects at "Other type of HEI=1". * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses, clustered at the high school level.

Table 6: LATEs of Short-cycle (SC) Programs vs. the Next-best Option

	Prob(Working)		Log Monthly Wage	
	No-interaction	Interaction [Other HEI=1]	No-interaction	Interaction [Other HEI=1]
A. Full sample				
SC vs. next-best	0.068	0.014	-0.534***	-0.234*
	(0.121)	(0.112)	(0.154)	(0.125)
N	328358	328358	156823	156823
Kleibergen-Paap F-stat	115.912	124.939	97.900	115.894
Cragg-Donald F-stat	601.700	726.542	309.562	399.201
B. Female				
SC vs. next-best	0.501***	0.311**	-0.460*	-0.154
	(0.182)	(0.156)	(0.265)	(0.214)
N	176651	176651	31872	31872
Kleibergen-Paap F-stat	67.811	80.459	22.143	35.639
Cragg-Donald F-stat	235.768	305.882	46.495	68.561
C. Male				
SC vs. next-best	-0.237*	-0.243**	-0.434***	-0.228**
	(0.122)	(0.123)	(0.128)	(0.114)
N	151707	151707	124951	124951
Kleibergen-Paap F-stat	100.798	101.022	108.271	113.189
Cragg-Donald F-stat	373.543	426.592	300.575	358.374

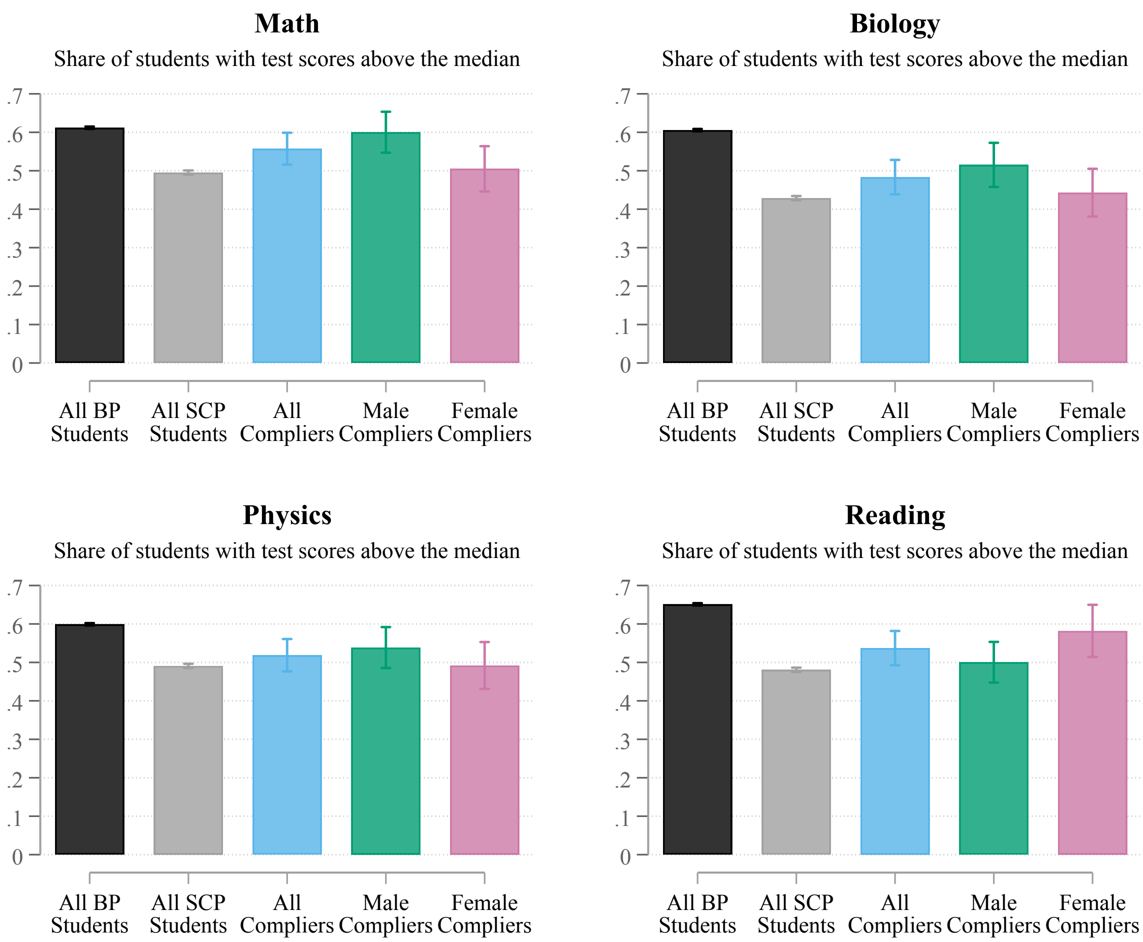
Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses, clustered at the high school level. The next-best options refers to a combination of enrollment options different from short-cycle programs (that is, choosing a bachelor's program or choosing not to enroll in higher education). The "No-interaction" columns present TSLS estimates without interacting the variables of supply of higher education ("HEI only offers SC" and "Other type of HEI"); the "Interaction [Other HEI=1]" columns present TSLS estimates interacting "HEI only offers SC" and "Other type of HEI" and evaluating the effects at "Other type of HEI=1".

Figure 1: Trends in Local Availability of Higher Education Institutions, by Institution Type (base year=2000)



Note: The figure displays the trend in the number of municipalities that have at least one higher education institution available in a 10 km radius around their centroid. The trends use 2000 as the base year (i.e., each series starts at 1 in 2000). Separate trends are shown for institutions specialized in short-cycle programs and for other institutions (universities and university institutions). The municipalities included are those where students in the sample attended high school. The dotted vertical line (2005) denotes the year of high school graduation for students in the sample. Source: SNIES.

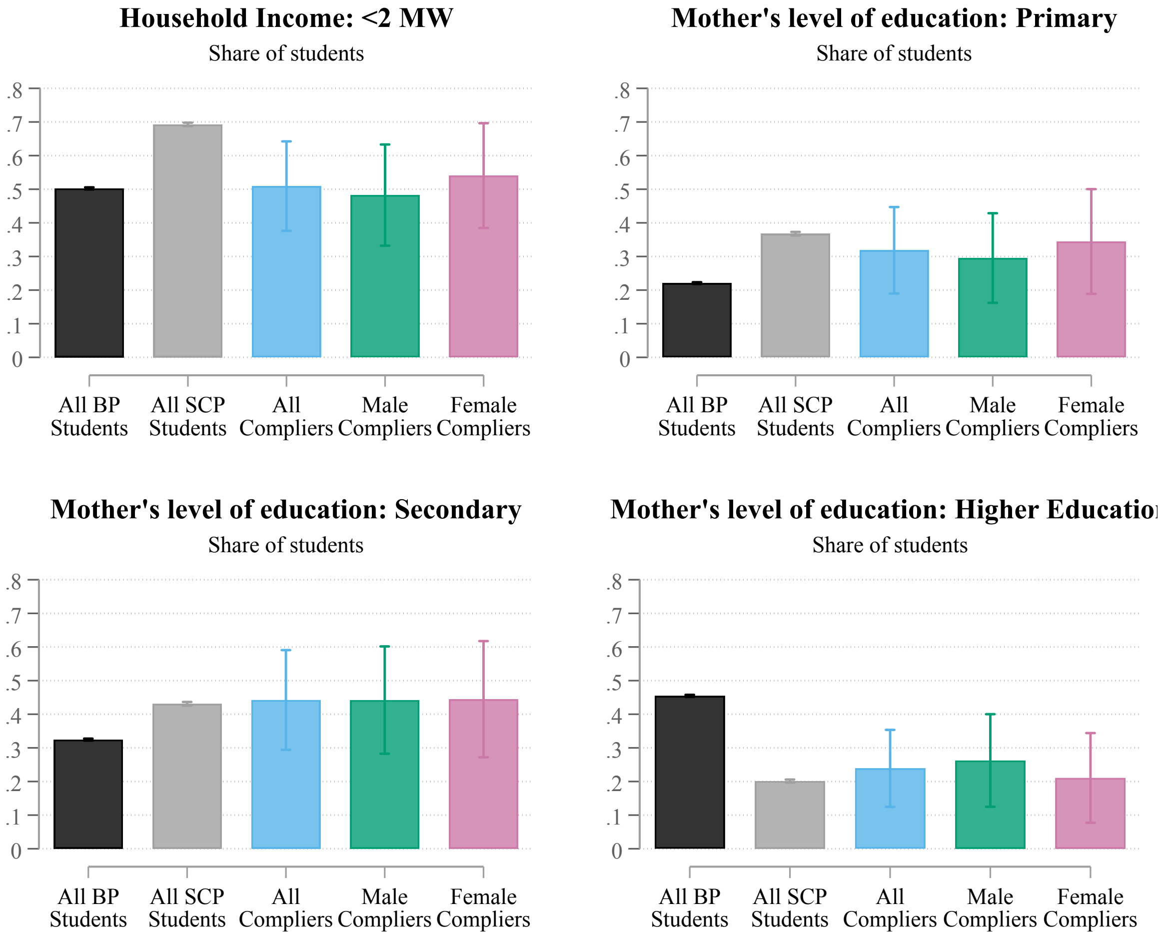
Figure 2: Saber 11 Performance for Compliers and Other Students



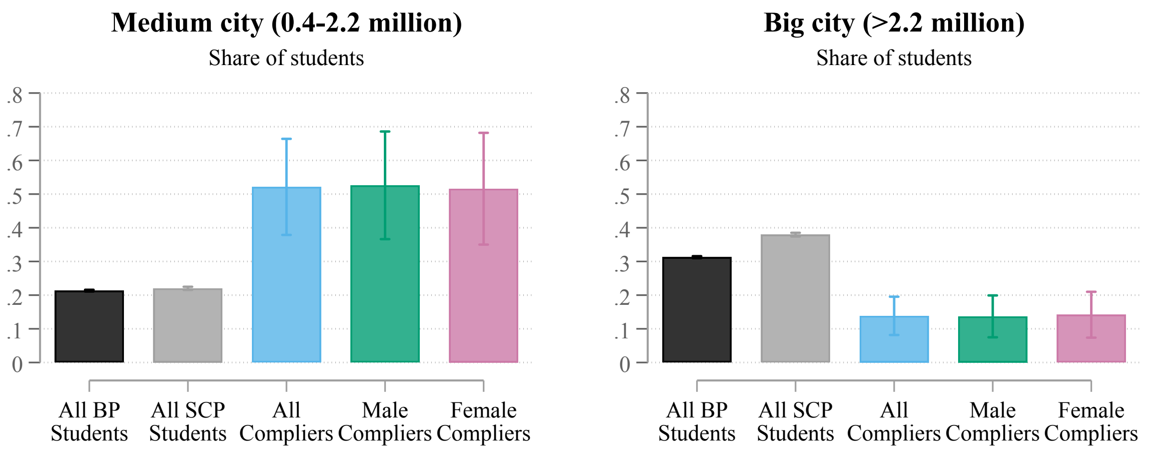
Note: The figure shows the share of students with Saber 11 scores for bachelor's program (BP) students, short-cyle program (SCP) students, compliers, and compliers by gender.

Figure 3: Average Household Characteristics and Municipality Size for Compliers and Other Students

(a) Household level

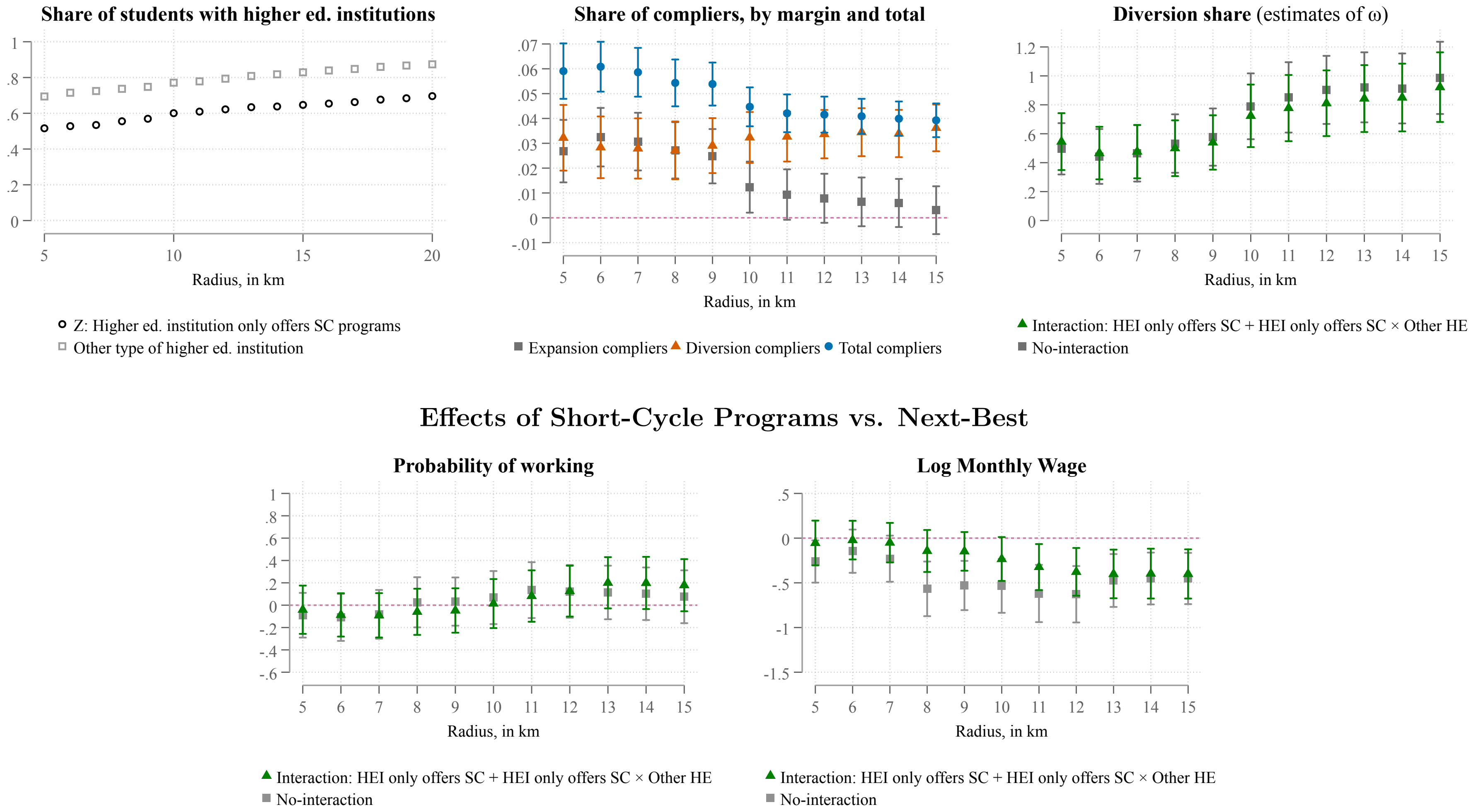


(b) Municipality size (population)



Note: The figure shows average characteristics and municipality size of bachelor's program (BP) students, short-cyle program (SCP) students, compliers, and compliers by gender.

Figure 4: Share of Compliers and Coefficient Stability for TSLS Results at Different Radii of HEIs Availability



Note: The figure shows estimates of complier shares and labor market outcomes at different radii for HEI availability.

Appendix

A.1 Additional Tables and Figures

Table A1: Summary Statistics

Variable	Mean	SD
Female	0.538	0.499
Age at Saber 11	17.026	1.589
Siblings	2.708	1.655
Household Income Level		
<1 Minimum Wages (MW)	0.272	0.445
1-2 MW	0.435	0.496
2-3 MW	0.158	0.365
2-3 MW	0.135	0.342
Mother's Education Level		
Primary	0.425	0.494
Secondary	0.354	0.478
Short-Cycle program	0.109	0.312
At least Bachelor's program	0.111	0.315
Higher Education Enrollment		
Not enrolled (NE)	0.612	0.487
Short-Cycle program (SC)	0.094	0.292
Bachelor's program (BP)	0.294	0.456
Educational Attainment		
High School Graduate	0.612	0.487
Short-Cycle Incomplete	0.046	0.210
Short-Cycle Complete	0.036	0.185
Bachelor's Incomplete	0.157	0.364
Bachelor's Complete	0.148	0.356
Formal Labor Market Outcomes		
Works in 2013	0.493	0.499
Avg. Monthly Wage (2013) [N=163,670]	767,614.9	626,834.1
N	328,358	

Note: The sample corresponds to the universe of students who took the high school Exit Exam in 2005. The information on educational attainment comes from SPADIES (System for Dropout Prevention of Higher Education), and labor market outcomes for higher education graduates are from OLE. For high school graduates and those with incomplete higher education, we impute formal labor market participation and experience using household survey data from SEDLAC (Socioeconomic Database for Latin America and the Caribbean). Monthly wages from OLE (Labor Market Observatory for Education) are in Colombian pesos (COP).

Table A2: First-stage Results: Estimated coefficients associated with covariates, X_{it}

Variable	Short-Cycle Program	Bachelor's Program	Not Enrolled
Female	-0.0250***	0.0319***	-0.0069***
	(0.0016)	(0.0021)	(0.0023)
Age at Saber 11	-0.0077***	-0.0214***	0.0291***
	(0.0004)	(0.0006)	(0.0007)
Siblings	-0.0025***	0.0005	0.0020
	(0.0007)	(0.0015)	(0.0016)
Household Income			
1-2 MW	0.0163***	0.0194***	-0.0357***
	(0.0026)	(0.0040)	(0.0042)
>2 MW	0.0106***	0.0579***	-0.0684***
	(0.0040)	(0.0061)	(0.0072)
Mother's level of education			
Secondary	0.0190***	0.0624***	-0.0813***
	(0.0028)	(0.0039)	(0.0044)
Higher Education	-0.0161***	0.2620***	-0.2459***
	(0.0042)	(0.0077)	(0.0078)
Standardized Test Scores from the High School Exit Exam			
Math	-0.0040***	0.0205***	-0.0165***
	(0.0006)	(0.0008)	(0.0009)
Reading	0.0066***	0.0450***	-0.0516***
	(0.0007)	(0.0010)	(0.0011)
Biology	-0.0020***	0.0294***	-0.0274***
	(0.0007)	(0.0009)	(0.0010)
Physics	-0.0030***	0.0154***	-0.0124***
	(0.0006)	(0.0008)	(0.0009)
History	0.0004	0.0239***	-0.0243***
	(0.0007)	(0.0009)	(0.0010)
Chemistry	-0.0036***	0.0280***	-0.0244***
	(0.0007)	(0.0010)	(0.0010)
Geography	-0.0015**	0.0221***	-0.0206***
	(0.0006)	(0.0008)	(0.0009)
Philosophy	0.0013**	0.0159***	-0.0172***
	(0.0006)	(0.0009)	(0.0009)
HS municipality characteristics			
Total GDP (billions COP)	-0.0020***	0.0022***	-0.0002
	(0.0005)	(0.0006)	(0.0007)
Ratio of Urban/Rural Population	-0.0109***	-0.0051*	0.0159***
	(0.0018)	(0.0026)	(0.0027)
Homicide rate (per 1,000 inhabitants)	-0.0123***	-0.0059	0.0182***
	(0.0033)	(0.0057)	(0.0058)
Total Population (in millions)	0.0324***	-0.0315***	-0.0009
	(0.0068)	(0.0082)	(0.0098)
Area (per 10,000 squared km)	-0.0033	0.0525***	-0.0492***
	(0.0057)	(0.0190)	(0.0187)
N	328,358	328,358	328,358
R^2	0.0207	0.2425	0.2258

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, clustered at the high school level. All regressions include region fixed effects. Each choice probability is estimated with a separate Linear Probability Model. The model also includes Z_j and H_j (instruments) as controls. The estimated coefficients associated with these variables are reported in Table 4.

Table A3: Probability of Higher Education Enrollment and Local Supply of Higher Education Institutions (HEIs)–Multinomial Logit Model

	Short-cycle Program		Bachelors Program		Not Enrolled	
Variable	No-interaction	Interaction	No-interaction	Interaction	No-interaction	Interaction
Z: HEI only offers SC programs (1)	0.036***	0.007	-0.030***	-0.021***	-0.006	0.013
	(0.003)	(0.005)	(0.004)	(0.008)	(0.004)	(0.008)
H: Other type of HEI	-0.002	-0.015***	0.019***	0.023***	-0.017***	-0.009
	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)
HEI only offers SC $\times$ Other HEI (2)		0.037***		-0.011		-0.026***
		(0.006)		(0.009)		(0.009)
Total effect of **Z** when *Other HEI=1*: HEI only offers SC + HEI only offers SC $\times$ Other HEI						
(1)+(2)		0.043***		-0.032***		-0.011**
		(0.0034)		(0.005)		(0.005)
N	328,358	328,358	328,358	328,358	328,358	328,358

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, clustered at the high school level. All regressions include region fixed effects and controls at the municipal level. The probability of enrollment choice is estimated jointly for the three options (short-cycle program, bachelor's program, not enrolled) using a multinomial logit model. The table presents marginal effects.

Table A4: OLS and TSLS Estimates of Labor Market Effects of SC Programs (excluding all instruments from the outcome equation)

	Prob(Working)			Log Monthly Wage		
	OLS	TSLS		OLS	TSLS	
		No-interaction	Interaction		No-interaction	Interaction
Full sample						
SC vs. next best	-0.041***	0.028	-0.020	0.123***	-0.559***	-0.201
	(0.005)	(0.118)	(0.111)	(0.004)	(0.155)	(0.126)
N	328,358	328,358	328,358	156,823	156,823	156,823
Kleibergen-Paap F-stat		59.928	43.988		52.007	44.995
Cragg-Donald F-stat		304.647	244.760		154.866	133.990
Female						
SC vs. next best	0.177***	0.467***	0.248	-0.060***	-0.662***	-0.123
	(0.005)	(0.179)	(0.155)	(0.007)	(0.256)	(0.195)
N	176,651	176,651	176651	31,872	31,872	31,872
Kleibergen-Paap F-stat		34.660	28.012		12.674	12.881
Cragg-Donald F-stat		118.764	103.424		27.749	28.367
Male						
SC vs. next best	-0.252***	-0.268**	-0.270**	0.151***	-0.439***	-0.239**
	(0.006)	(0.116)	(0.116)	(0.005)	(0.129)	(0.116)
N	151,707	151,707	151,707	124,951	124,951	124,951
Kleibergen-Paap F-stat		52.158	36.365		62.884	47.508
Cragg-Donald F-stat		189.819	144.573		150.291	119.489

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, clustered at the high school level. The next-best refers to a combination of enrollment options different from short-cycle programs (that is, choosing a bachelor's program or choosing not to enroll in higher education). The "No-interaction" columns present TSLS estimates without interacting the variables of supply of higher education ("HEI only offers SC" and "Other type of HEI"); the "Interaction" columns present TSLS estimates interacting "HEI only offers SC" and "Other type of HEI".

A.2 Robustness Tables and Figures

Table A5: TSLS Estimates at Different Radii for HEIs Availability

	Prob(Working)		Log Monthly Wage	
	No-interaction	Interaction [Other HEI=1]	No-interaction	Interaction [Other HEI=1]
5km radius				
SC vs. next best	-0.090	-0.041	-0.262**	-0.054
	(0.102)	(0.110)	(0.120)	(0.128)
N	328,358	328,358	156,823	156,823
Kleibergen-Paap F-stat	132.475	107.477	97.567	79.023
Cragg-Donald F-stat	1047.501	951.160	517.462	435.045
10km radius				
SC vs. next best	0.068	0.014	-0.534***	-0.234*
	(0.121)	(0.112)	(0.154)	(0.125)
N	328,358	328,358	156,823	156,823
Kleibergen-Paap F-stat	115.912	124.939	97.900	115.894
Cragg-Donald F-stat	601.700	726.542	309.562	399.201
15km radius				
SC vs. next best	0.075	0.179	-0.451***	-0.400***
	(0.121)	(0.119)	(0.146)	(0.140)
N	328,358	328,358	156,823	156,823
Kleibergen-Paap F-stat	119.956	128.393	109.672	116.733
Cragg-Donald F-stat	545.676	614.267	289.828	325.909
20 km radius				
SC vs. next best	0.173	0.266**	-0.323**	-0.310**
	(0.130)	(0.132)	(0.153)	(0.152)
N	328,358	328,358	156,823	156,823
Kleibergen-Paap F-stat	117.630	113.783	112.304	112.782
Cragg-Donald F-stat	481.007	489.387	257.668	269.969
25km radius				
SC vs. next best	0.147	0.245*	-0.149	-0.174
	(0.140)	(0.143)	(0.158)	(0.161)
N	328,358	328,358	156,823	156,823
Kleibergen-Paap F-stat	108.043	102.964	105.072	102.202
Cragg-Donald F-stat	420.490	411.108	235.321	234.633
30 km radius				
SC vs. next best	0.040	0.072	-0.799***	-0.811***
	(0.180)	(0.179)	(0.245)	(0.249)
N	328,358	328,358	156,823	156,823
Kleibergen-Paap F-stat	64.887	66.075	61.187	60.319
Cragg-Donald F-stat	234.982	239.897	124.541	122.768

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, clustered at the high school level. The next-best refers to a combination of enrollment options different from short-cycle programs. The "No-interaction" columns present TSLS estimates without interacting the variables of supply of higher education ("HEI only offers SC" and "Other type of HEI"); the "Interaction" columns present TSLS estimates interacting "HEI only offers SC" and "Other type of HEI".

Table A6: TSLS Estimates at Different Radii of HEIs availability - Female Students

	Prob(Working)		Log Monthly Wage	
	No-interaction	Interaction [Other HEI=1]	No-interaction	Interaction [Other HEI=1]
5km radius				
SC vs. next best	0.288**	0.312**	-0.494**	-0.585*
	(0.136)	(0.146)	(0.226)	(0.329)
N	176,651	176,651	31,872	31,872
Kleibergen-Paap F-stat	94.611	80.247	26.048	15.757
Cragg-Donald F-stat	468.507	434.920	77.590	39.157
10km radius				
SC vs. next best	0.501***	0.311**	-0.460*	-0.154
	(0.182)	(0.156)	(0.265)	(0.214)
N	176,651	176,651	31,872	31,872
Kleibergen-Paap F-stat	67.811	80.459	22.143	35.639
Cragg-Donald F-stat	235.768	305.882	46.495	68.561
15km radius				
SC vs. next best	0.351**	0.481***	-0.371*	-0.356
	(0.163)	(0.166)	(0.208)	(0.237)
N	176,651	176,651	31,872	31,872
Kleibergen-Paap F-stat	78.191	83.319	34.840	27.703
Cragg-Donald F-stat	239.577	264.888	63.643	50.006
20 km radius				
SC vs. next best	0.432***	0.521***	-0.323	-0.303
	(0.167)	(0.178)	(0.208)	(0.226)
N	176,651	176,651	31,872	31,872
Kleibergen-Paap F-stat	79.515	74.982	35.517	31.300
Cragg-Donald F-stat	225.668	218.442	62.018	53.272
25km radius				
SC vs. next best	0.351**	0.408**	-0.301	-0.275
	(0.176)	(0.184)	(0.204)	(0.220)
N	176,651	176,651	31,872	31,872
Kleibergen-Paap F-stat	72.695	69.724	37.673	33.270
Cragg-Donald F-stat	198.914	191.235	64.386	56.141
30 km radius				
SC vs. next best	0.046	0.080	-0.554**	-0.515**
	(0.224)	(0.225)	(0.220)	(0.220)
N	176,651	176,651	31,872	31,872
Kleibergen-Paap F-stat	49.180	49.563	36.731	36.270
Cragg-Donald F-stat	122.988	124.525	57.422	56.229

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, clustered at the high school level. The next-best refers to a combination of enrollment options different from short-cycle programs. The "No-interaction" columns present TSLS estimates without interacting the variables of supply of higher education ("HEI only offers SC" and "Other type of HEI"); the "Interaction" columns present TSLS estimates interacting "HEI only offers SC" and "Other type of HEI".

Table A7: TSLS Estimates at Different Radii of HEIs availability - Male Students

	Prob(Working)		Log Monthly Wage	
	No-interaction	Interaction [Other HEI=1]	No-interaction	Interaction [Other HEI=1]
5km radius				
SC vs. next best	-0.387***	-0.368***	-0.205**	0.068
	(0.101)	(0.114)	(0.104)	(0.105)
N	151,707	151,707	124,951	124,951
Kleibergen-Paap F-stat	103.441	80.452	98.008	84.861
Cragg-Donald F-stat	581.003	518.954	460.146	442.506
10km radius				
SC vs. next best	-0.237*	-0.243**	-0.434***	-0.228**
	(0.122)	(0.123)	(0.128)	(0.114)
N	151,707	151,707	124,951	124,951
Kleibergen-Paap F-stat	100.798	101.022	108.271	113.189
Cragg-Donald F-stat	373.543	426.592	300.575	358.374
15km radius				
SC vs. next best	-0.164	-0.106	-0.366***	-0.300**
	(0.139)	(0.134)	(0.135)	(0.123)
N	151,707	151,707	124,951	124,951
Kleibergen-Paap F-stat	95.635	102.423	112.382	128.873
Cragg-Donald F-stat	309.527	354.802	263.436	319.641
20 km radius				
SC vs. next best	-0.032	0.045	-0.268*	-0.207
	(0.165)	(0.160)	(0.146)	(0.139)
N	151,707	151,707	124,951	124,951
Kleibergen-Paap F-stat	90.383	90.493	118.563	126.543
Cragg-Donald F-stat	257.401	274.893	230.473	258.657
25km radius				
SC vs. next best	-0.022	0.103	-0.081	-0.086
	(0.182)	(0.182)	(0.155)	(0.153)
N	151,707	151,707	124,951	124,951
Kleibergen-Paap F-stat	82.968	79.259	108.251	110.094
Cragg-Donald F-stat	224.263	223.159	203.731	212.379
30 km radius				
SC vs. next best	-0.021	0.055	-0.813***	-0.816***
	(0.267)	(0.262)	(0.257)	(0.259)
N	151,707	151,707	124,951	124,951
Kleibergen-Paap F-stat	45.649	47.038	55.423	55.345
Cragg-Donald F-stat	114.493	118.034	92.434	92.594

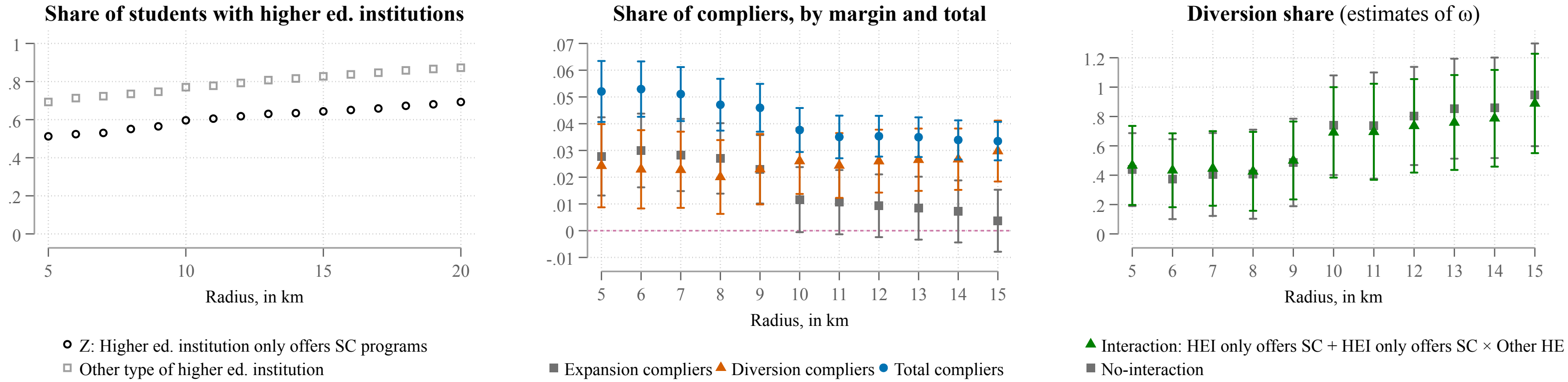
Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, clustered at the high school level. The next-best refers to a combination of enrollment options different from short-cycle programs. The "No-interaction" columns present TSLS estimates without interacting the variables of supply of higher education ("HEI only offers SC" and "Other type of HEI"); the "Interaction" columns present TSLS estimates interacting "HEI only offers SC" and "Other type of HEI".

Table A8: TSLS Estimates Without Interacting H and Z, Interacting H and Z, and Interacting H with All the Covariates

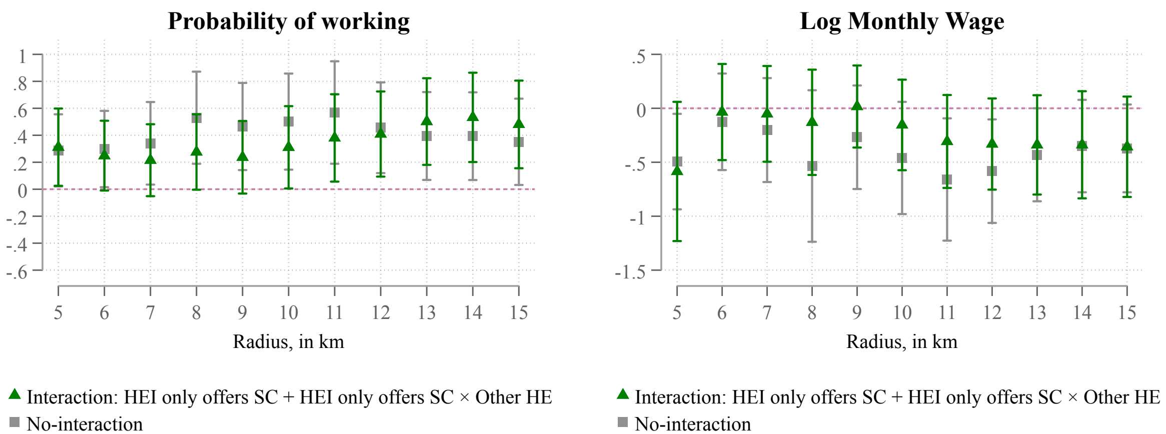
	Prob(Working)			Log Monthly Wage		
	No-interaction	Interaction [Other HEI=1]		No-interaction	Interaction [Other HEI=1]	
Full sample						
SC vs. next best	0.068	0.014	0.053	-0.534***	-0.234*	-0.406***
	(0.121)	(0.112)	(0.117)	(0.154)	(0.125)	(0.139)
N	328,358	328,358	328,358	156,823	156,823	156,823
Kleibergen-Paap F-stat	115.912	124.939	106.205	97.900	115.894	94.879
Cragg-Donald F-stat	601.700	726.542	582.540	309.562	399.201	317.810
Female						
SC vs. next best	0.501***	0.311**	0.247	-0.460*	-0.154	-0.333
	(0.182)	(0.156)	(0.171)	(0.265)	(0.214)	(0.238)
N	176,651	176,651	176,651	31,872	31,872	31,872
Kleibergen-Paap F-stat	67.811	80.459	65.608	22.143	35.639	29.640
Cragg-Donald F-stat	235.768	305.882	235.507	46.495	68.561	57.288
Male						
SC vs. next best	-0.237*	-0.243**	-0.115	-0.434***	-0.228**	-0.357***
	(0.122)	(0.123)	(0.123)	(0.128)	(0.114)	(0.126)
N	151,707	151,707	151,707	124,951	124,951	124,951
Kleibergen-Paap F-stat	100.798	101.022	89.857	108.271	113.189	95.231
Cragg-Donald F-stat	373.543	426.592	351.818	300.575	358.374	291.807
Interaction with H	No	Only Z	Z and X	No	Only Z	Z and X

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, clustered at the high school level. The next-best refers to a combination of enrollment options different from short-cycle programs. The "No-interaction" columns present TSLS estimates without interacting the higher education supply variables ("HEI only offers SC" and "Other type of HEI"); the "Interaction" columns present TSLS estimates interacting "HEI only offers SC" and "Other type of HEI", and interacting these two variables in addition to also interacting "Other type of HEI" with all the covariates ("Z and X").

Figure A1: Exposure to HEIs, Share of Compliers, and Labor Market Effects for TSLS Estimates at Different Radii of HEIs Availability, among Female Students

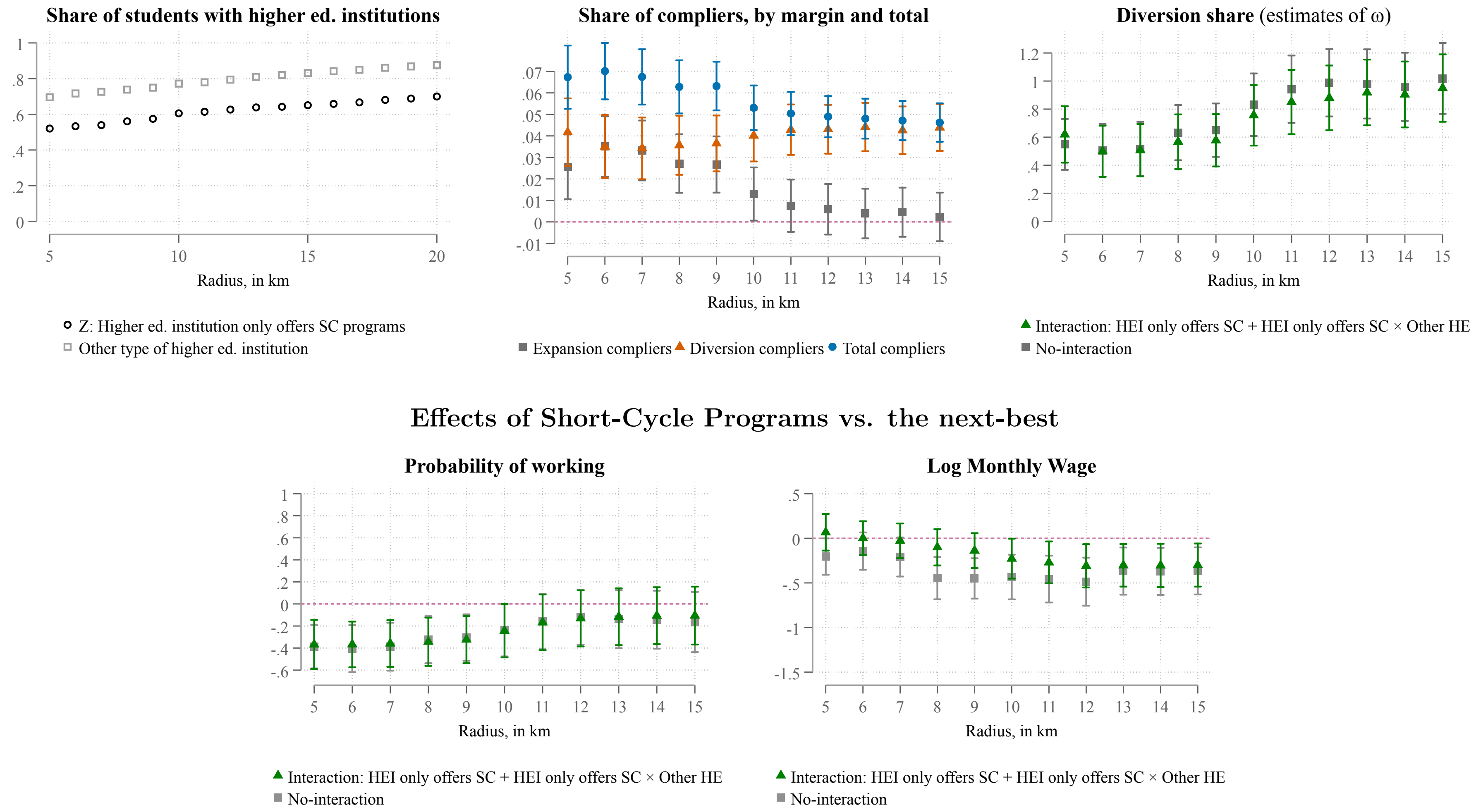


Effects of Short-Cycle Programs vs. the next-best



Note: The figure shows the share of students with exposure to HEI (top left plot), estimates of complier shares (top center plot), estimates of diversion shares (top right plot), and TSLS estimates for the probability of working (bottom left plot) and the log of monthly wage (bottom right plot) at different radii of HEIs availability.

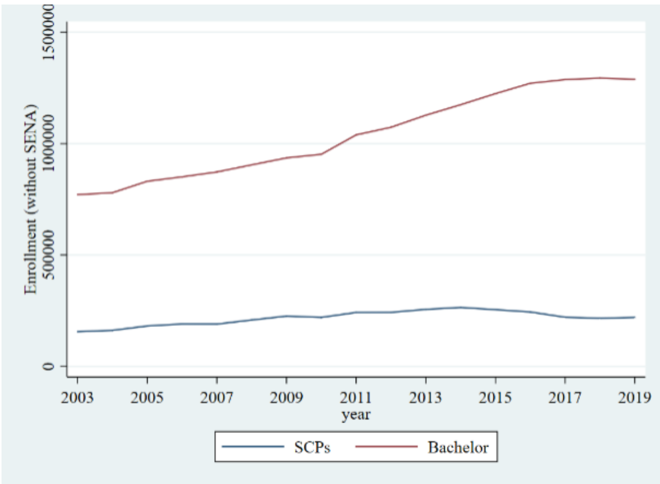
Figure A2: Exposure to HEIs, Share of Compliers, and Labor Market Effects for TSLS Estimates at Different Radii of HEIs Availability, among Male Students



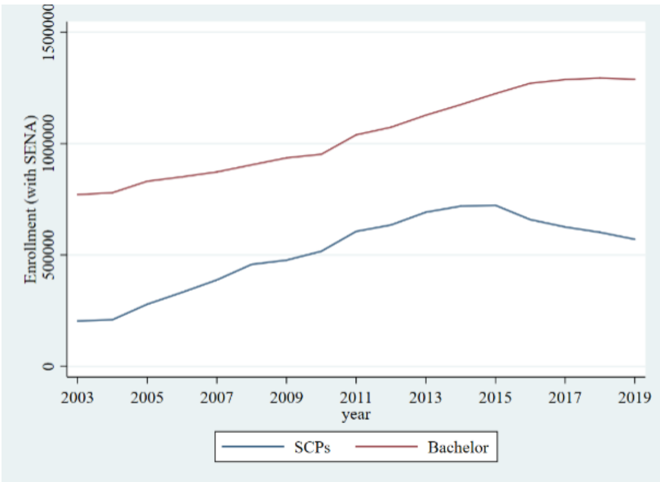
Note: The figure shows the share of students with exposure to HEI (top left plot), estimates of complier shares (top center plot), estimates of diversion shares (top right plot), and TSLS estimates for the probability of working (bottom left plot) and the log of monthly wage (bottom right plot) at different radii of HEIs availability.

Figure A3: Total Enrollment in Short-cycle and Bachelor's Programs With and Without SENA

(a) Enrollment without SENA



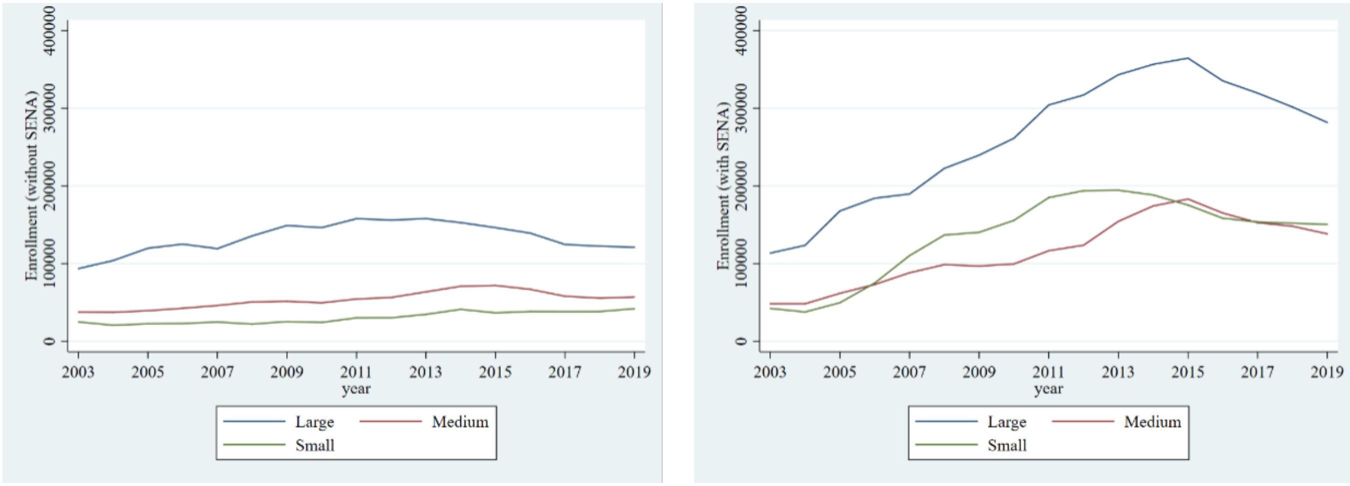
(b) Enrollment with SENA



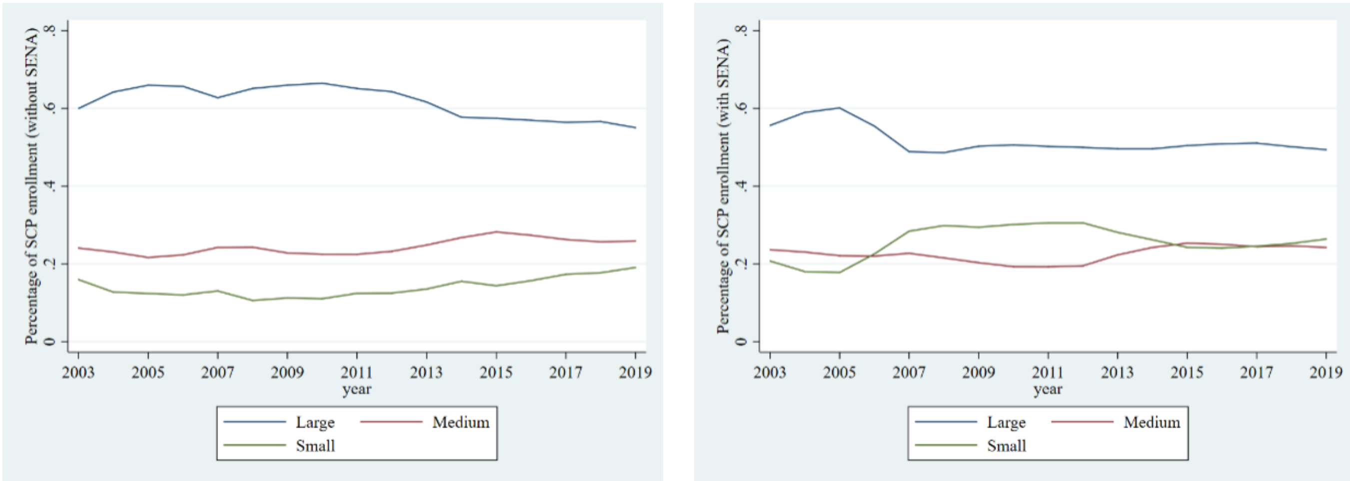
Note: Own calculations using data from SNIES (*Sistema Nacional de Información de la Educación Superior*, in Spanish). We exclude programs taught at distance or virtually.

Figure A4: Enrollment in Short-cycle Programs With and Without SENA, by City Size

(a) Total Enrollment (left: without SENA; right: with SENA)



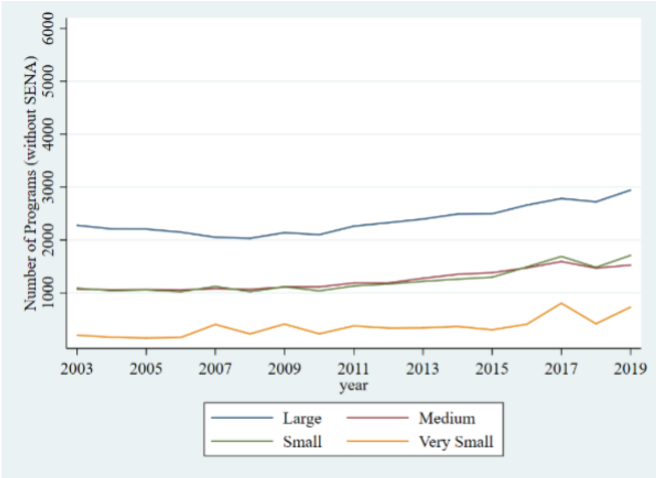
(b) Distribution of Enrollment (left: without SENA; right: with SENA)



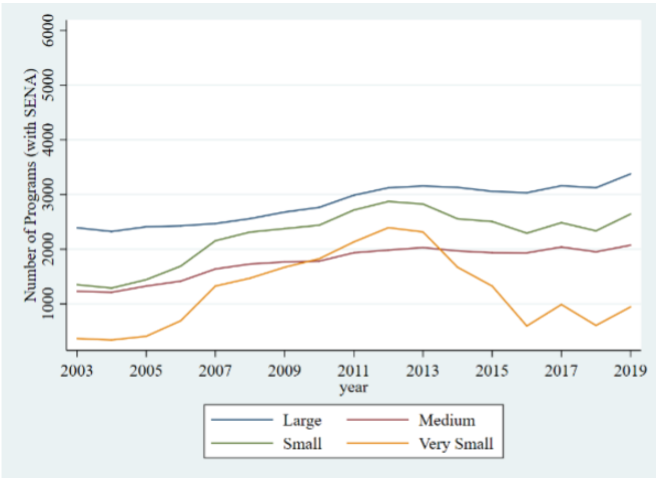
Note: Own calculations using data from SNIES (*Sistema Nacional de Información de la Educación Superior*, in Spanish). Large cities: Population above 2.5 million; Medium cities: Population between 400,000 and 2.5 million; Small cities: Population below 400,000.

Figure A5: Number of Short-cycle Programs with and without SENA, by City Sizes

(a) Without SENA



(b) With SENA



Note: Own calculations using data from SNIES (*Sistema Nacional de Información de la Educación Superior*, in Spanish). Large cities: Population above 2.5 million; Medium cities: Population between 400,000 and 2.5 million; Small cities: Population between 50,000 and 400,000; Very small cities: Population below 50,000.

A.3 Imputation of wages and formal employment

To impute formal employment and average monthly wages for high school graduates and students with incomplete higher education, we use household survey data between 2008 and 2013. In particular, we use the set of homogenized household surveys in SEDLAC, known as the Integrated Household Survey project (GEIH, in Spanish). We restrict each sample to individuals who were 14-24 old by 2005. Let Y denote the labor market outcome of interest. We posit a linear regression model of Y on a set of controls. For formal employment, we estimate a linear probability model with gender, age, number of household members, household income level, a dummy that takes the value of one if the individual lives in an urban area, and regional fixed effects as controls. Estimates are reported in Table A9. For average wages we estimate quantile regressions and predict coefficients for the 25th, 50th, and 75th percentile while controlling for gender, age, age squared, number of household members, household income level, a dummy that takes the value of one if the individual lives in an urban area, and regional fixed effects. Estimates are displayed in Tables A10 and A11.

Table A9: Regression Results: Probability of Formal Employment in 2013

Variables	HS Graduates	HE Incomplete
Male	0.350***	0.135***
	(0.014)	(0.017)
Age	0.013***	0.027***
	(0.003)	(0.003)
Number of members in main household	-0.019***	-0.028***
	(0.005)	(0.005)
$1-2$ MW	0.080***	0.140***
	(0.027)	(0.043)
$2-3$ MW	0.154***	0.254***
	(0.031)	(0.043)
> 3 MW	0.190***	0.340***
	(0.031)	(0.040)
Urban area	0.020	-0.030
	(0.022)	(0.035)
Region: Oriental	0.068**	0.040*
	(0.028)	(0.023)
Region: Central	-0.023	0.038*
	(0.025)	(0.021)
Region: Pacifica	0.000	-0.015
	(0.028)	(0.025)
Region: Bogotá	0.067*	0.067**
	(0.034)	(0.027)
Observations	4,722	8,141

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Estimates produced using data from the Integrated Household Survey (GEIH, in Spanish) in SEDLAC (Socioeconomic Database for Latin America and the Caribbean).

Table A10: Regression Results for High School Graduates: Hourly wages in main occupation (2013)

Variables	q25	q50	q75
Male	461.627***	434.299***	172.838
	(100.972)	(105.942)	(143.657)
Age	437.637	131.550	-6.856
	(284.343)	(273.086)	(328.158)
Age^2	-8.053	-1.763	0.952
	(5.356)	(5.101)	(6.104)
Number of members in main household	-117.627***	-121.458***	-173.951***
	(16.687)	(17.445)	(24.616)
$1-2$ MW	1,203.005***	1,318.647***	1,507.471***
	(112.909)	(128.563)	(158.184)
$2-3$ MW	1,419.765***	1,598.469***	2,040.413***
	(150.009)	(138.681)	(180.284)
>3 MW	2,023.496***	2,102.490***	2,847.308***
	(142.385)	(153.329)	(277.423)
Urban area	-236.529**	-315.042**	-756.664***
	(102.791)	(151.966)	(237.149)
Region: Oriental	315.999**	367.526***	197.873
	(135.960)	(123.872)	(142.295)
Region: Central	155.621	236.217	354.958**
	(128.751)	(152.566)	(164.369)
Region: Pacifica	-110.810	-113.379	-47.057
	(135.958)	(116.014)	(231.480)
Region: Bogotá	655.060***	510.973***	326.354*
	(189.608)	(150.499)	(196.395)
Constant	-4,739.014	-359.963	2,686.551
	(3,726.144)	(3,609.618)	(4,344.162)
Observations	3,090	3,090	3,090

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Estimates produced using data from the Integrated Household Survey (GEIH, in Spanish) in SEDLAC (Socioeconomic Database for Latin America and the Caribbean).

Table A11: Regression Results for Incomplete Higher Education: Hourly Wages in Main Occupation (2013)

Variables	q25	q50	q75
Male	240.217***	342.924***	423.750**
	(89.161)	(86.817)	(184.014)
Age	680.753**	210.822	-61.922
	(265.982)	(295.035)	(509.594)
Age^2	-11.772**	-2.368	3.341
	(4.981)	(5.542)	(9.509)
Number of members in main household	-208.221***	-277.203***	-438.176***
	(26.766)	(21.133)	(41.265)
$1-2$ MW	1,503.986***	1,853.146***	1,342.269***
	(182.953)	(159.451)	(265.630)
$2-3$ MW	2,170.080***	2,771.923***	2,532.960***
	(192.120)	(173.614)	(278.575)
>3 MW	3,077.149***	3,808.032***	4,681.549***
	(151.087)	(160.508)	(285.479)
Urban area	-408.050***	-308.839***	-229.820
	(117.800)	(92.751)	(269.067)
Region: Oriental	341.604**	368.984***	687.261**
	(141.283)	(134.636)	(293.303)
Region: Central	31.376	-175.898*	-285.977
	(119.496)	(90.773)	(193.988)
Region: Pacifica	152.830	183.826	103.817
	(139.976)	(154.098)	(225.862)
Region: Bogotá	378.608***	243.773**	155.136
	(127.699)	(124.243)	(227.950)
Constant	-8,028.344**	-1,587.394	3,238.408
	(3,506.050)	(3,893.089)	(6,742.760)
Observations	5,378	5,378	5,378

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Estimates produced using data from the Integrated Household Survey (GEIH, in Spanish) in SEDLAC (Socioeconomic Database for Latin America and the Caribbean).