Estimating the Medium-Term Impact of a Scholarship Program Using Standardized Test Data

Marlene Saint Martin[1] & Miguel Szekely Pardo[1]

[1] Center for Education and Social Studies (CEES), Mexico City, Mexico

Correspondence: Miguel Szekely Pardo, Center for Education and Social Studies (CEES), Mexico City, Mexico. E-mail: mszekely@prodigy.net.mx

Abstract

In 2006, the BBVA Foundation in Mexico designed a scholarship and mentoring program that targeted vulnerable lower secondary education students in municipalities with high-intensity migration rates. We follow applicants who started lower secondary education in 2009 to estimate the impact of the Program on the probability of graduating from the next level of education. We use the records of a standardized test undertaken by all students in Mexico to identify applicants who were still enrolled three years after the Program. Our results show the Program had a positive and significant effect on upper secondary completion.

Keywords: dropout rates, secondary education, scholarship program, mentoring, education policy, standardized test, enrollment, attendance, Mexico

1. Introduction

1.1 Dropouts in Secondary Education

According to data from the National Institute for the Evaluation of Education (INEE for its acronym in Spanish), coverage at the primary level of education for the 2016–2017 academic year in Mexico was close to 100 percent. In contrast, coverage for the lower and upper secondary levels of education was 84.3 and 63.8 percent, respectively (Note 1). Although in these last two levels, attendance rates increased by 10 and 9 percentage points, respectively, since the 2010–2011 academic year, clearly universalization, which is mandated by the Constitution, is far from achieved (Note 2).

There are several reasons why coverage still shows a significant gap in lower and upper secondary education. The main reason, however, appears to be the education system's limited capacity to retain students throughout each academic level.

The data indicates that 97.1 percent of students who successfully graduated from primary school, enrolled in the next level of education (Note 3). Similarly, 99.3 percent of students who graduated from lower secondary education, registered at the upper secondary level. However, when examining the information within each level, 4.2 and 12.8 percent of students drop out during lower and upper secondary education, respectively, every year (Note 4).

Additionally, the recent standardized test PLANEA (National Plan for the Evaluation of Learning) reveals that students in Mexico have low levels of academic achievement. The 2015 edition of the test shows that over 75 and 89 percent of students enrolled in the last grade of lower secondary education have either an insufficient level of learning or only an essential domain of Spanish and Mathematics, respectively (Note 5).

Different data show that low education quality is likely one of the main factors, besides financial need, closely related to high dropout rates. According to the National Survey of School Dropout in Upper Secondary Education (ENDEMS, 2011), 36.4 percent of young people reported they had dropped out of school due to financial difficulties, while the second most cited reason was that they didn't enjoy studying (Note 6). Similarly, the Ministry of Social Development (SEDESOL) and the National Institute of Public Health carried out a survey among observations between the ages of 14 to 27 whose families were affiliated to the national conditional cash transfer program PROGRESA (Note 7). In this survey, 43.9 percent had dropped out because of financial problems, while 19.3 percent (the second-highest response) mentioned their dislike for school.

1.2 The Scholarship and Mentoring Program: Impact on Upper Secondary Education Completion

In 2006, the BBVA Foundation in Mexico designed a scholarship and mentoring program for lower secondary education students. The Program extended its coverage to the upper secondary education level in 2013, and into higher education in 2016. The scholarship is granted to low-income students and is conditional on academic permanence and performance, as well as attendance to mentoring sessions.

Students get paired with an employee who serves voluntarily as their mentor throughout the three years of lower secondary education (Note 12). Students check in bi-monthly with their mentors during the academic year and discuss their performance in school, among other issues. The Foundation encourages mentors to explicitly urge students to improve their academic performance and to continue studying until higher education.

Initially, the Program prioritized the allocation of scholarships to students in municipalities with high immigration intensity, but later, as the Program quickly expanded, it focused instead on preventing school dropout. In 2006, the "pilot" version of the Program awarded 600 scholarships for students enrolled in lower secondary education in 6 municipalities across the country (Note 13). In 2015, the Program reached out to students across 22 states, providing 50,000 and 18,000 scholarships for lower and upper secondary education, respectively. In the 2019–2020 school year, students from all states and municipalities in the country were able to apply online to the Program.

The objective of this study is to evaluate the impact of the Program on graduation rates for upper secondary education. For this purpose, we follow a quasi-experimental evaluation strategy, which consists in taking advantage of the census design of the standardized National Evaluation of Academic Achievement in School Centers (ENLACE) during the 2006–2014 period and of PLANEA (2015). The databases for these standardized tests allow us to identify students that received the treatment, and to compare their trajectory with that of applicants with similar characteristics who were eligible but could not access the Program due to limited funding.

1.3 Literature Review: CCTs and Scholarship Programs

Mexico was one of the pioneer countries in introducing conditional cash transfer programs (CCTs) starting with PROGRESA in 1997. CCTs are characterized by providing financial support to families living in poverty, conditioning the aid's delivery to school attendance, among other factors. Similarly, the Foundation's Scholarship Program grants financial assistance through a scholarship, conditioning its delivery to school attendance, academic performance, and participation in the mentoring sessions.

Bastagli et al. (2016) analyze the evolution of RCT programs in Latin America, showing its rapid expansion during the first decade of the 21st century after the success of the Mexican PROGRESA program (Note 8). The authors compile the results of evaluations for more than 56 programs in 30 different countries and conclude that these programs are overall successful in increasing school attendance and schooling significantly. However, the results on the development of skills and knowledge are inconclusive. Along the same lines, Molina-Millan et al. (2016) conduct a review of the long-term effects of CCTs in Latin America (Note 9) and conclude that there is a positive and significant effect of the programs on school attendance—in programs in Mexico, Nicaragua, and Colombia—but not on academic achievement.

In Mexico, PROGRESA—called Oportunidades in 2003 and Prospera in 2014—was implemented in rural communities in 1997 with an experimental design (Note 10). The results of the experiment suggest positive and statistically significant short-term effects on the enrollment rates of the beneficiaries (Note 11). The impact of PROGRESA on the probability of enrolling in high school for students who were in sixth grade when it started, was 14.8 percentage points for girls and 6.5 percentage points for boys.

Behrman, Parker, and Todd (2009; 2011) use the Household Socioeconomic Characteristics Survey (ENCASEH) and the ENCEL panel survey of PROGRESA to determine the impact of different degrees of exposure. The evaluation results suggest that participants with greater exposure have 0.2 more academic degrees completed than those with less exposure.

Similarly, Parker, Rubalcava, and Teruel (2012) use the National Household Living Level Survey (ENNViH 2002 and 2005) to assess the medium-term effect of the program. The authors estimate a model of differences in differences with which they compare young people in communities that were selected to participate in the program in the early years (1997–1998) with young people in selected communities that received the program after 2004. The results suggest that the program has a positive effect by increasing the schooling levels of the former by 0.5 years.

Parker and Vogl (2018) use Difference in Differences (DID) to compare students with different lengths of

exposure to the program. The program, which begins before students are 12 years of age, is associated significantly with better educational and professional outcomes. The results suggest that students who had been exposed to the program the longest achieved 1.4 more years of schooling. For women, longer exposure increases participation and wages, as well as the number of hours invested in work, and it enhanced the transition from informal to formal employment—understood as access to social security. For men, the effects are also positive and significant, but to a lesser extent.

Acevedo, Ortega, and Székely (2018), use the ENLACE standardized test to follow a generation of students who received support from PROSPERA. They find a positive association between the time of exposure to the Program and finishing upper secondary education—according to the authors, each year of exposure is associated with a 12.5 and 12.2 percentage points increase in the probability of completing the academic trajectory to upper secondary education.

Behrman, Parker, and Todd (2018) build a panel of PROGRESA beneficiary students with the databases of the standardized test ENLACE, following academic trajectories from the 6th grade of elementary school until the 3rd grade of upper secondary education. The authors find a significant effect on enrollment for the first grade of lower secondary school—8 percentage points, the third grade of lower secondary school—10 percentage points, and the last grade of upper secondary education—5 percentage points.

As for other scholarship programs implemented in Mexico, De Hoyos et al. (2019) take advantage of the randomized design of the Scholarship Program of Upper Secondary Education (PROBEMS) by the Ministry of Education. The authors assess the impact on graduation rates, using ENLACE as a proxy to measure whether students graduated from this education level—similar to the strategy of the present study. The evaluation found a significant impact of 0.04 percentage points on the probability of graduating from high school for beneficiaries who were above the 70th percentile in the ENLACE Mathematics test at the upper secondary level. The authors consider that students who reach high school with learning deficiencies are at a disadvantage, increasing the risk of dropout.

Finally, Dustan (2018) uses DID to assess the impact of the "Prepa Sí" scholarship program in Mexico City. The study finds the program didn't have an effect on students' graduation from upper secondary education or on their academic achievement. One possible factor behind these results identified by the author is that the monetary value of the scholarship was lower than the opportunity cost of continuing to study in an urban context. The absence of mechanisms to verify students' attendance and permanence might have also been part of the problem.

2. Method

2.1 Connecting Databases

The objective of the present evaluation is to estimate the impact of the BBVA Foundation's scholarship program on students' probability of graduating from upper secondary education. To perform this analysis, we use students' participation in a standardized test (Note 14) —which was undertaken by all students in the country approximately two months before graduation—as a proxy for identifying students who were still enrolled.

We assume students spend three years in lower secondary school and another three years in upper secondary school. Then, the first generation of the program—the 2006 generation-, should have participated in the 2012 edition of ENLACE at the end of their last year of upper secondary education, if still enrolled. As of 2016, however, the standardized test—now called PLANEA—was conducted only in a sub-sample of schools. Thus, we focus on the beneficiaries that were admitted between 2006 and 2009 as these generations should have graduated from upper secondary education between 2012 and 2015 (the years for which the standardized data is available at the national level).

The students' ID (Unique Population Registry Code, CURP) is the link to connect the administrative records of the Foundation with the standardized tests databases. In the administrative records for 2009, 95 percent of applicants registered a valid 18-digit ID (Note 15). However, in the data from previous years, less than 60% of students registered a valid ID, complicating the search for their records in the standardized tests databases. Therefore, we focus on the generation that started lower secondary education in 2009. This generation, if enrolled, would have graduated from lower secondary education in 2012, and continued into upper secondary school until graduation in 2015.

Students admitted in 2009 were only eligible to receive the scholarship through their three-year lower secondary education trajectory, as the program had not expanded yet to the next academic level (Note 16). As the treatment group concluded their participation three years before the outcome of interest, the evaluation is purposely measuring the medium-term effects of the scholarship program.

Finding medium-term effects seems feasible, given the program's emphasis on academic achievement. While the financial support of the scholarship is an important motivator, the element that distinguishes the program from other interventions is the mentoring component: providing an aspirational figure outside of the family circle that supports and encourages the student. Mentors also urge students to think about their professional future, asking about the profession they would like to pursue, and highlighting the role of education beyond lower and upper secondary school. This distinctive element could be the main force driving positive effects over student performance in the long run (Note 17).

2.2 Applicant Pool (2009)

To apply, prospective students and their families had to fill out a socioeconomic survey and presented official documentation (Note 18). Based on this information, the Foundation constructed a ranking of the applicants, giving priority to those with higher GPAs, greater economic need, and relatives who had emigrated to the United States of America or Europe (Note 19). Starting at the top of the ranking, the BBVA Foundation offered admission until all the budget was exhausted. The applicants that were not selected, but complied with the program's criteria, appear in the administrative records under the "Reserve" status. On the other hand, applicants who failed to comply with the eligibility criteria appear under the "Not Accepted" status.

Students under the "Accepted" status received a monthly scholarship—around 50 dollars—on a debit card and were paired up with a mentor. These students had to remain enrolled in a public lower secondary school, keep an academic average above 8.5 (Note 20), and attend the bi-monthly appointments with their mentors. Admitted students who did not comply with these conditions were discharged from the program and appear in the administrative records under the "Disenrolled" status. After a student lost the scholarship, the Foundation would contact the next student in the ranking. The new beneficiary would, therefore, shift from the "Reserve" status into the "Substitute" status (Note 21).

2.3 Sample Size

In 2009, the BBVA Foundation admitted 5,149 students into the Program (approximately 38% of the eligible applicant pool). During the students' trajectory through lower secondary education, however, 1,736 were discharged from the Program and subsequently replaced with applicants under the "Reserve" status. Table 1 shows the distribution of applicants by status.

Table 1. Distribution of applicants by status (initial sample)

Status	2009 generation of applicants	
	Number of observations	Percentage
Accepted	3,413	25.1%
Disenrolled	1,736	12.8%
Substitutes	1,582	11.7%
Reserves	6,854	50.5%
Total	13,585	100%

Note. Authors' calculations based on administrative data.

In order to estimate the effect of the Program, the administrative data and the standardized test data are merged using students' ID. Out of the 13,585 observations that represent the eligible pool of applicants, 95.5 percent registered a valid ID (Table 2). Table 3 shows the results of balance tests in which the total sample—13,585 observations—is compared to the sample with valid IDs -12,926 observations. The only significant difference—at the 5 percent level—between the samples is the percentage of observations with a migrant relative. The difference between samples, however, is minimal—0.01 percentage points, and the rest of the socioeconomic variables are balanced.

Table 2. Percentage of applicants with valid ID by status

Status	2009 generation of applicants		
	Observations	Observations with a valid ID	Percentage of observations with a valid ID
Accepted	3,413	3,331	97.6%
Disenrolled	1,736	1,621	93.4%
Substitutes	1,582	1,540	97.34%
Reserves	6,854	6,434	93.9%
Total	13,585	12,926	95.5%

Note. Authors' calculations based on administrative data.

Table 3. Balance tests between the total sample (St) and the sample with valid ID records (Sid)

2009 generation of applicants					
Variables	N St	N Sid	Media St	Media Sid	Difference
Monthly income	13585	12926	$4102.50	$4088.92	$13.58
Number of people in the family	13582	12926	4.79	4.78	0.00
Per capita monthly income	13539	12884	$922.21	$920.27	1.93
GPA (primary level)	13416	12760	9.20	9.20	0.00
A family member has migrated to the US or Europe	13585	12926	0.30	0.31	-0.01 **
Remittances	4069	4069	$880.81	$880.81	0.00
House is owned by the family	13582	12924	0.57	0.58	0.00
The floor is ground or cement	13582	12924	0.63	0.64	0.00
Water is connected to public service	13582	12924	0.70	0.70	0.00
Lighting is connected to public service	13582	12924	0.99	0.99	0.00
Home has at least one bathroom	13582	12924	0.85	0.86	0.00
Home has at least three rooms	13582	12924	0.71	0.71	0.00
Telephone line is connected	13582	12924	0.47	0.47	0.00
Has a car	13582	12924	0.34	0.34	0.00
Has a computer	13582	12924	0.16	0.16	0.00

Note. Authors' calculations based on administrative data. *** Significant difference at the 1 percent level; ** Significant difference at the 5 percent level; * Significant difference at the 10 percent level.

Finally, Table 4 shows the distribution by the status of the applicant pool with a valid ID. This sample represents the observations from which the treatment and the control group will be drawn. The applicants under the "Reserve" status represent around 50% of the applicant pool. Considering that the quasi-experimental methods that will be used to evaluate the program depend on suitable matches between the treatment and control group, the considerable pool of "Reserves" is likely to guarantee quality matches.

Table 4. Distribution of the sample with a valid ID by status

Status	2009 generation of applicants	
	Observations	Percentage
Accepted	3,331	25.77%
Disenrolled	1,621	12.54%
Substitutes	1,540	11.91%
Reserves	6,434	49.78%
Total	12,926	100%

Note. Authors' calculations based on administrative data.

2.4 Research Design

The administrative records for 2009 show 15,682 students applied to the program. We created different treatment groups based on program exposure (Table 5). The first treatment group (T1) is made up of all applicants under the "Accepted" status. Through this group, we expect to capture the effect of receiving the "complete" dose of treatment (3 years of scholarship and mentoring). The second treatment group (T2) is formed by applicants who were exposed at least two years to the program. Similarly, the third treatment group (T3) includes applicants who

were exposed at least for one year.

The objective of forming groups T2 and T3 is to examine whether the results vary with a different "dose" of the treatment. In terms of the ranking of applicants, observation under the "Disenrolled" and "Admitted" status are similar. However, there are unobservable differences between these applicants as some kept the scholarship, and others lost it. Admitted students lose the scholarship for lowering their GPA or for dropping out of school. Therefore, any effect for T2 and T3 is expected to be lower than that for T1. Finally, we define the T4 as the treatment group that includes only applicants under the "Substitute" status—students initially under the "Reserve" status that ended up receiving the scholarship because an admitted applicant lost it.

Table 5. Treatment description

Treatments			2009 generation of applicants		
Name	Exposure to the program	Status	Number of observations with a valid ID	% with a valid ID found in PLANEA 2015	% with a valid ID found in ENLACE 2014
T1	Three years	Accepted	3,331	69.2%	1.7%
T2	Two years or more	Accepted, Substitutes, and Disenrolled	3,608	67.05%	1.7%
T3	One year or more	Accepted, Substitutes, and Disenrolled	4,879	62.4%	1.5%
T4	Between two and 34 months	Substitutes	1,540	71.7%	1.6%

Note. Calculations performed by the authors based on administrative and standardized test data.

Table 6 describes the control group, which is formed by the applicants with "Reserve" status. Despite the differences between groups, there are advantages of using the "Reserves" as controls. Indeed, the applicants lived in the same municipalities, were enrolled in the same public lower education institutions, had a GPA of at least 8.5 by the end of primary school, demonstrated financial need, and submitted an application to participate in the program (which made them eligible to receive benefits).

Table 6. Introducing the control group

Control group			2009 generation of applicants		
Name	Exposure to the program	Status	Number of observations with a valid ID	% with a valid ID found in PLANEA 2015	% with a valid ID found in ENLACE 2014
C	0	Reserve	6,434	61.3%	1.4%

Note. Calculations performed by the authors based on administrative and standardized test data.

2.5 Propensity Score Matching and OLS Regression

As explained before, the assignment to the Program was not random. Thus, we consider quasi-experimental methods, which are feasible since the entry to the Program among the eligible population was determined through an exogenous factor—the budget allowed for a pre-determined number of scholarships. Given that admission into the Program is based on a ranking built around observable characteristics, we will present results using Propensity Score Matching. We also conduct OLS regressions as a sensibility analysis, to observe if and how results differ between methods.

Propensity Score Matching first estimates the probability of receiving the treatment—the propensity score- based on observable characteristics. One must use variables that are theoretically related to the treatment and the variable of interest—avoiding variables that perfectly predict treatment allocation, variables that are influenced by the treatment, and variables that are related to treatment allocation but not the variable of interest.

One algorithm commonly used to pair observations is the Nearest Neighbor Matching, in which the treatment observations are paired with the closest control observations in terms of their propensity score. A caliber can be set to avoid "low quality" pairs. If the diameter of the caliber is minimal, some observations in the treatment group might not find a "partner", decreasing the size of the sample. On the other hand, if the diameter is too wide, the quality of the pairs is likely to fall. After the pairing process, the average treatment effect in the treated observations is estimated—the average difference concerning the outcome between the treatment and control groups (Note 22).

We considered using Regression discontinuity (RDD) to estimate the effects of the Program, as the BBVA Foundation constructed a ranking to determine the applicants' priority. However, the administrative data hold no records of the ranking or the exact formula that was used to calculate it.

The following equation represents the OLS model used to estimate the effect of the program on the probability of students' graduating from upper secondary education. The treatment groups—T1, T2, T3, and T4—and the control group remain the same. Given the non-random assignment of the treatment and the differences between applicants across status, the estimates should be biased. In this regard, the control group has, on average, a higher family per capita income than the observations that received the treatment. Thus, we expect the estimates to be biased against the defined treatment groups. However, since the model includes the variables that are used to calculate the propensity score as covariates, the OLS estimates, and the PSM estimates should not differ much.

$$Y_i = \alpha_0 + \beta_1 T_i + \beta_2 Female_i + \beta_3 Income_i + \beta_4 GPA_i + \beta_5 Migrant_i + u_i \qquad (1)$$

Y_i is the probability of student $_i$ graduating from upper secondary education; T_i is a dummy variable that takes the value of one if student $_i$ belongs to the treatment group; $Female_i$ indicates that student $_i$ is female; $Income_i$ stands for the natural logarithm of the student $_i$ family per capita income; GPA_i is a dichotomous variable that takes the value of one if student $_i$'s average grade is equal to or greater than 9; $Migrant_i$ takes the value of one if a relative of student $_i$ has emigrated to the United States or Europe; and u_i represents the error term.

3. Results

3.1 Baseline Data

To better understand the differences between the status in the applicant pool, Table 7 shows an analysis of means based on administrative data (Note 23). The differences in means across status provide a clear picture of the admissions criteria. Indeed, reflecting the priority built upon the ranking, students under the "Accepted" and "Disenrolled" status come from families with a lower average income, which suggests that students in the treatment group would be at a disadvantage in terms of the learning environment to which they are exposed at home, as compared to those remaining in the "Reserve" status. Also, students under the "Accepted" status have a higher GPA at baseline; this is of relevance as students that lost the scholarship have a lower GPA than students who did not (in 0.21 points). Given that our control group stems from the observations under the "Reserve" status, it is essential to note the differences between this group and the rest. These observations have, on average, the highest monthly and per capita income, and the lowest probability of having a migrant relative, which would be expected to be positively correlated with the probability of completing school trajectories. Their GPA is, on average, 9.19, 0.07 points below the "Accepted" status' GPA and 0.14 above the "Disenrolled" status'.

The analysis of means demonstrates the non-experimental nature of the treatment assignment. As they are, comparing observations under the "Accepted" and "Reserve" status, for example, would yield biased estimates. The differences in means also highlight the admission criteria used for ranking students for their inclusion into the program, which is based on income, family history of migration, and GPA.

Table 7. Socioeconomic variables by status

2009 generation of applicants				
Variable	Accepted	Substitutes	Disenrolled	Reserves
Monthly income	$3,087.08	$3,887.42	$3,036.23	$4,921.04
Number of people in the family	5	5	5	5
Per capita monthly income	$713.92	$791.54	$711.13	$1,110.58
GPA (primary level)	9.26	9.31	9.05	9.19
A family member has migrated to the US or Europe	0.65	0.21	0.56	0.10
Remittances	$993	$547.6	$,1069	$397
House is owned by the family	0.56	0.59	0.52	0.59
The floor is ground or cement	0.67	0.66	0.69	0.60
Water is connected to public service	0.68	0.69	0.67	0.72
Lighting is connected to public service	0.98	0.99	0.98	0.99
Home has at least one bathroom	0.83	0.86	0.82	0.88
Home has at least three rooms	0.67	0.72	0.68	0.74
Telephone line is connected	0.46	0.44	0.43	0.49
Has a car	0.29	0.34	0.24	0.39
Has a computer	0.13	0.14	0.11	0.19

Note. Authors' calculations based on administrative data.

3.2 Estimation of the Propensity Score

To measure the probability of program applicants graduating from upper secondary school, we exploit the census nature of the standardized tests ENLACE and PLANEA. These tests were undertaken by students—if enrolled—in the last two months before graduating from upper secondary education. Thus, their participation in the test is interpreted as a proxy for upper secondary education completion. In this section, we present the results of the impact evaluation for the BBVA Foundation scholarship program on this variable, using Propensity Score Matching (PSM), and Ordinary Least Squares Regression (OLS). We expect the results to differ between methods as OLS estimates will be biased, given the non-random assignment of the treatment.

The first step in PSM is to calculate the propensity score—the probability of receiving the treatment. Based on this score, the treatment and control groups are paired up and compared to estimate the Average Treatment Effect on the Treated (ATET). For the calculation of the propensity score, we used as covariates the gender of the applicants, the natural logarithm of their family's per capita income, and a dichotomous variable that takes the value of one when the applicants' primary grade average is equal to or higher than 9.

The estimated propensity score meets the balance property for all defined treatments. Figures 1 to 4 show the density distribution of the propensity score for the treatment groups and the control group. The distributions between groups are similar, guaranteeing the common support of the score and, thus, quality pairs for the observations in the treatment group.
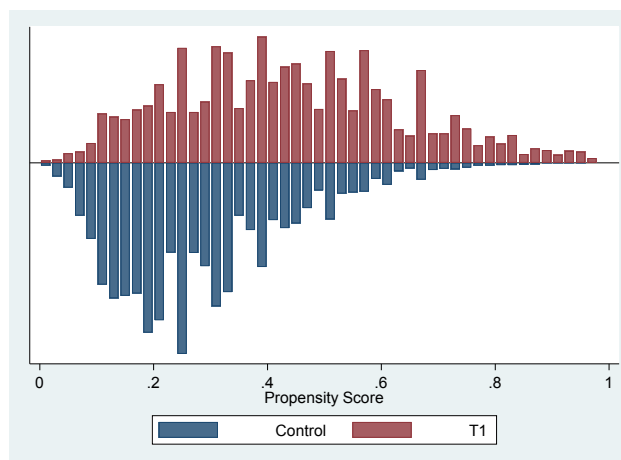

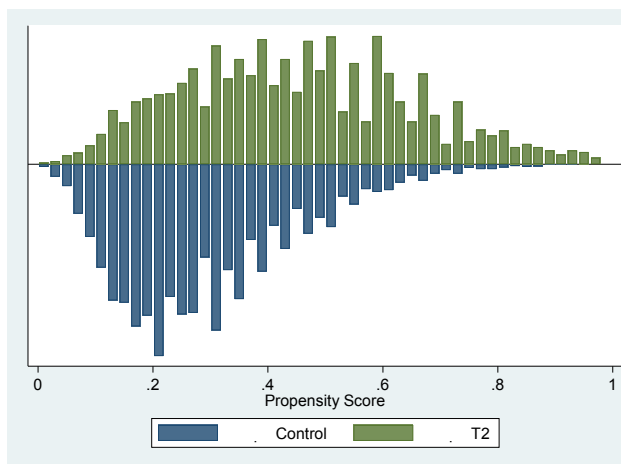Figure 1. Density distribution for the propensity score (T1 and Controls)

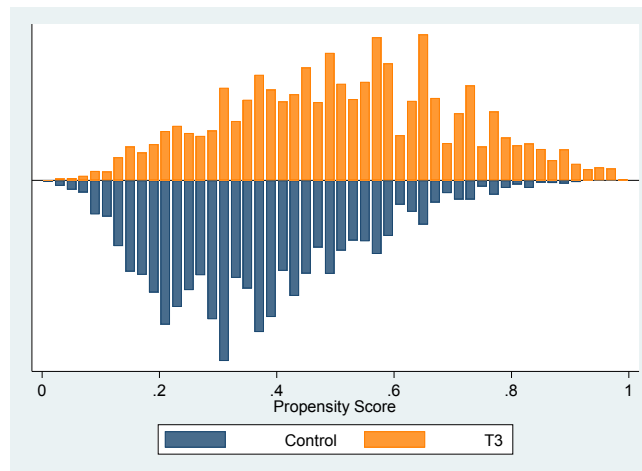Figure 2. Density distribution for the propensity score (T2 and Controls)

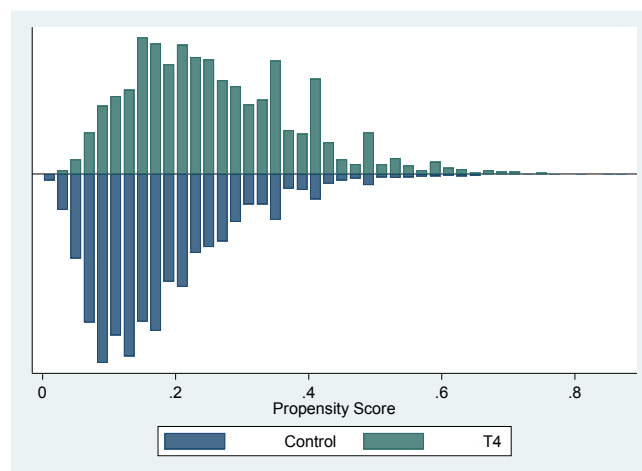Figure 3. Density distribution for the propensity score (T3 and Controls)

Figure 4. Density distribution for the propensity score (T4 and Controls)

After calculating the propensity score, the pairs between groups are formed. For robustness, we use three different pairing algorithms to calculate the ATET (Table 8): Nearest Neighbor matching (NN) with at least one pair—NN (1), two pairs—NN (2), and three pairs—NN (3) for each observation in the treatment groups. Likewise, the pairs were required to be at a maximum distance—a caliper of 0.05 in terms of the propensity score to ensure quality pairs (Note 24).

3.3 PSM and OLS estimates

Table 8 shows the OLS and PSM estimates for the probability of completing upper secondary education by each treatment group. According to PSM, on average, students in T1 are 10 percentage points more likely than the observations in the control group—applicants with "Reserve" status—to complete upper secondary education significantly at the 1 percent level (Note 25). The results between the other matching algorithms are equivalent, which is indicative of their robustness. As for the OLS regressions, everything else constant, the treatment significantly increases the probability of completing upper secondary education in 9.7 percentage points (at the 1 percent level). The results between methods are very similar, even across treatments.

As for T2, we expect results to be somewhat weaker than those of T1, given that there are observations in this group that received an incomplete dose of treatment. However, given the small number of observations under this circumstance, it is expected that the difference between results will not be much (Note 26). Table 8 shows that, according to the PSM estimates, the probability of completing high school is, on average, 8 percentage points higher for observations in the T2 group than for observations in the control group at a 1 percent level of significance. As for the OLS estimates, the treatment increases significantly the probability of students completing

upper secondary education in 7.4 percentage points at the 1 percent level, holding everything else constant.

Table 8 also shows the PSM and OLS results—the probability of completing upper secondary education—for applicants in T3. This group is formed by applicants who received the program for at least a year -26 percent of the observations received the treatment between one and two years, and 69.3 percent received the full dose (Note 27). According to the PSM estimates, applicants in T3 have a 3-percentage point higher probability of completing upper secondary education than the observations in the control group significantly at the 5 percent level (Note 28). The OLS method estimates that this treatment increases the probability of completing upper secondary education on 3.6 percentage points significantly at the 1 percent level, everything else constant.

Finally, Table 8 presents the PSM and OLS results for T4, which is formed by 1,540 observations under the "Substitute" status. Since these observations were not selected in the first round, they are the most similar to the control group. However, by construction, T4 is a very heterogeneous group in terms of the time they were exposed to the treatment, as over 50 percent of the group had the scholarship between one and three months. Therefore, these results should be interpreted as the effect of being eligible after the first round of admissions. Considering that the observations in T4 are from a significantly higher socioeconomic background than the observations in T1, the estimated difference is expected to be higher for T4. The PSM estimates show that, on average, the observations in T4 have a 12-percentage point higher probability of graduating from upper secondary education than the control group at the 1 percent level of significance. As for the OLS estimates, the treatment appears to increase the probability of completion in 11 percentage points at the 1 percent level of significance, everything else constant.

Table 8. Average treatment effect on the treated (ATET) and OLS coefficients—probability of graduating from upper secondary education

		Propensity Score Matching (PSM)				Ordinary Least Square
		Caliper = 0.05				(OLS) estimates
		NN(1)	NN(2)	NN(3)		
T1:	ATET	0.10***	0.10***	0.10***	β_1	0.0969***
Three years of treatment		(0.01)	(0.01)	(0.01)		(0.0136)
("Accepted")	N_C	6244	6244	6244	N_C	6,244
	N_{T1}	3157	3157	3156	N_{T1}	3,157
T2:	ATET	0.08***	0.08***	0.08***	β_1	0.0736***
Two years or more of treatment		(0.01)	(0.01)	(0.01)		(0.0126)
	N_C	6,244	6244	6244	N_C	6,244
	N_{T2}	3374	3374	3374	N_{T1}	3,778
T3:	ATET	0.03**	0.04***	0.03***	β_1	0.0363***
One year or more of treatment		(0.01)	(0.01)	(0.01)		(0.0116)
	N_C	6244	6244	6244	N_C	6,244
	N_{T3}	4571	4571	4571	N_{T1}	4,975
T4:	ATET	0.12***	0.12***	0.12***	β_1	0.112***
"Substitutes"		0.02	0.01	0.01		(0.0143)
	N_C	6,244	6,244	6,244	N_C	6,244
	N_{T4}	1,489	1,489	1,489	N_{T1}	1,489

Note. For PSM, the Abadie-Imbdens standard errors are shown in parentheses. For OLS, the standard errors are shown in parentheses. The OLS regressions are presented fully in Appendix 1. *** Significant at the 1 percent level; ** Significant at the 5 percent level; *Significant at the 10 percent level.

4. Discussion

The main takeaway of the impact evaluation is the evidence of positive and significant medium-term effects of the BBVA Foundation Scholarship Program on school completion at the upper secondary level in Mexico, among students with low socioeconomic background, high academic performance, and presence of migrants in the family.

Between 2006 and 2015, all students in the last year of upper secondary education undertook the national standardized test ENLACE or PLANEA, and we exploited the census nature of these evaluations to determine if

applicants were still enrolled three years after the intervention officially concluded. The control group was formed using applicants under the "Reserve" status. These observations did not receive the Program despite their eligible status because the budget had been exhausted. As for admitted students, exposure to the Program in terms of the length of time during which they received the treatment was not homogenous. After calculating the months of exposure for each student, we created different treatment groups by using administrative data.

Since there could be inherent differences between the control group and the treatment groups—which work in favor of the control group-, we used Propensity Score Matching and Ordinary Least Squares Regression to minimize and compare selection bias between estimates. We expect the estimates, if biased, to be underestimating the effect of the Program as the control group has, on average, a higher economic income. According to our results, which do not vary significantly among the two methods, exposure during lower secondary education increases significantly the probability of students completing the next level of education. The first treatment group, formed by students who received the full dose of the program, had, on average, a 10-percentage point higher probability of graduating from upper secondary education than those that were also eligible, but were never incorporated. As exposure decreases, so do the estimated probability of completion, although the difference between treatment and control remains positive and statistically significant.

The results stand out considering that other government scholarship programs implemented in Mexico—and referred to in the text—which targeted upper secondary education students, have had no impact on the probability of their beneficiaries completing this level of education. The BBVA Foundation Program heavily revolves around motivating students to maintain a strong academic performance both through financial support and mentoring, which distinguishes the initiative to alternatives that only provide scholarships.

Acknowledgements

References

Acevedo, I., Ortega. A., & Székely, M. (2018). *Rendimiento Escolar y Transiciones Laborales con Transferencias Condicionadas en México*. Unpublished manuscript.

Ana Luiza Neves de, H. B., & Carlos Henrique, L. C. (2013). *Bolsa Família, Occupational Choice and Informality in Brazil*. Working Papers 118, International Policy Centre for Inclusive Growth. Retrieved from https://ideas.repec.org/p/ipc/wpaper/118.html

Angrist, J., Bettinger, E., & Kremer, M. (2006). Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia. *American Economic Review*, *96*(3), 847–862. https://doi.org/10.1257/aer.96.3.847

Attanasio, O., De Hoyos, R., & Meghir, C. (2019). *Targeting High School Scholarships to the Poor: The Impact of a Program in Mexico*. NBER Working Papers 26023, National Bureau of Economic Research, Inc. https://doi.org/10.3386/w26023

Barham, T., Macours, K., & Maluccio, J. (2017). *Are Conditional Cash Transfers Fulfilling Their Promise? Schooling, Learning, and Earnings After 10 Years*. CEPR Discussion Papers 11937, C.E.P.R. Discussion Papers. Retrieved from https://ideas.repec.org/p/cpr/ceprdp/11937.html

Barrera-Osorio, F., Linden, L., & Saavedra, J. (2017). *Long term educational consequences of alternative conditional cash transfer designs: Experimental evidence from Colombia* (forthcoming). American Economic Journal: Applied Economics. https://doi.org/10.3386/w23275

Behrman, J. R., & Hoddinott, J. (2005). Programme evaluation with unobserved heterogeneity and selective implementation: The Mexican PROGRESA impact on child nutrition. *Oxford Bulletin of Economics and Statistics*, *67*(4), 547–569. https://doi.org/10.1111/j.1468-0084.2005.00131.x

Behrman, J. R., Parker, S. W., & Todd, P. E. (2009). Medium-Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico. *Poverty, Inequality, and Policy in Latin America*, 219–270. https://doi.org/10.7551/mitpress/9780262113243.003.0008

Behrman, J., Parker, S., & Todd, P. (2011). Do School Subsidy Programs Generate Lasting Benefits? A Five-Year Follow-Up of Oportunidades Participants. *Journal of Human Resources*, *46*(1), 93–122. https://doi.org/10.3368/jhr.46.1.93

Behrman, J., Parker, S., & Todd, P. (2018). *Impacts of Prospera on Enrollment, School Trajectories and*

Learning. Working paper, forthcoming. https://doi.org/10.1596/1813-9450-9000

Brooks, J. M., & Ohsfeldt, R. L. (2013). Squeezing the balloon: propensity scores and unmeasured covariate balance. *Health Services Research*, *48*(4), 1487–1507. https://doi.org/10.1111/1475-6773.12020

Cecchini, S., & Atuesta, B. (2017). *Programas de transferencias condicionadas en América Latina y el Caribe: tendencias de cobertura e inversión*. CEPAL. Serie: Serie Políticas Sociales No. 224. LC/TS.2017/40.

De Hoyos, R., Estrada, R., & Vargas, M. J. (2018). *Predicting Individual Wellbeing Through Test Scores: Evidence from a National Assessment in Mexico*. Research Department working papers 1187, CAF Development Bank of Latinamerica. https://doi.org/10.1596/1813-9450-8459

Dustan, A. (2018). *Can Large, Untargeted Scholarships Decrease High School Dropout? Evidence from Mexico City's Prepa Sí*. Working Paper. Department of Economics, Vanderbilt University. Retrieved from https://andrewdustan.com/pdf/dustan_prepasi.pdf

Fiszbein, A., & Schady, N. (2009). *Conditional cash transfers: reducing present and future poverty*. World Bank Publications. https://doi.org/10.1596/978-0-8213-7352-1

Garrido, M. M., Kelley, A. S., Paris, J. R., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, *49*(5), 1701–1720. https://doi.org/10.1111/1475-6773.12182

González de la Rocha, M. (2008). *A diez años de intervención en zonas rurales* (1997–2007). Tomo I, Efectos de Oportunidades en áreas rurales a diez años de intervención. México, DF: Secretaría de Desarrollo Social, 2008.

Hanushek, E., & Woessmann, L. (2007). *The Role of School Improvement in Economic Development*. CESifo Working Paper No. 1911. https://doi.org/10.3386/w12832

Heinrich, C., Maffioli, A., & Vazquez, G. (2010). *A primer for applying propensity-score matching*. Washington, D.C.: Inter-American Development Bank. Retrieved from https://publications.iadb.org/en/primer-applying-propensity-score-matching

Kausel, E. (2015). Joshua D. Angrist and Jörn-Steffen Pischke. Mastering 'Metrics: The Path from Cause to Effect. (Book Review). *Personnel Psychology*, *68*, 931–933. https://doi.org/10.1111/peps.12116_2

Kugler, A. D., & Rojas, I. (2018). *Do PTMCs Improve Employment and Earnings in the Very Long-Term? Evidence from Mexico*. National Bureau of Economic Research 24248. https://doi.org/10.3386/w24248

Molina-Millan, T., Barham, T., Macours, K., Maluccio, J. A., & Stampini, M. (2016). *Long-Term Impacts of Conditional Cash Transfers in Latin America: Review of the Evidence*. Washington, D.C.: Inter-American Development Bank. Retrieved from https://publications.iadb.org/en/publication/12565/long-term-impacts-conditional-cash-transfers-latin-america-review-evidence

Parker, S., Rubalcava, L., & Teruel, G. (2012). *Do conditional cash transfer programs improve work and earnings among its youth beneficiaries? Evidence after a decade of a Mexican cash transfer program*. Unpublished manuscript.

Parker, S., & Vogl, T. (2018). *Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico*. National Bureau of Economic Research. https://doi.org/10.3386/w24303

Reynolds, T., Anderson, C., Biscaye, P., Coomes, D., Madsen, T., & Ebeling, E. (2017). *Review of Long-Term Impacts of Cash Transfer Programs*. EPAR Technical Report #359.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Saavedra, J. E. (2016, May). The effects of conditional cash transfer programs on poverty reduction, human capital accumulation and wellbeing. In *United Nations Expert Group Meeting:"Strategies for eradicating poverty to achieve sustainable development for all" convened in New York on June* (pp. 1–3). https://doi.org/10.1007/978-3-319-53094-9_1

Schultz, P. (2000). *Impact of PROGRESA on school attendance rates in the sampled population* (No. 600-2016-40129).

Skoufias, E. (2006). *Progresa y su efecto sobre el bienestar de las familias rurales de México* (Vol. 139). Intl

Food Policy Res Inst.

StataCorp. (2013). *Stata: Release 13*. Statistical Software. College Station, TX: StataCorp LP.

Todd, J., & Winters, P. (2011). The Effect of Early Interventions in Health and Nutrition on On-Time School Enrollment: Evidence from the Oportunidades Program in Rural Mexico. *Economic Development and Cultural Change*, *59*(3), 549–581. https://doi.org/10.1086/658347

Villa, J. M. (2018). The continuous treatment effect of an antipoverty program on children's educational attainment: Colombia's Familias en Accion. *Review of Development Economics*. https://doi.org/10.1111/rode.12380

Notes

Note 1. Source: INEE, "La educación obligatoria en México, Informe 2019".

Note 2. There is a challenge as well in preschool coverage (72.3 percent).

Note 3. Absorption rate (2016-2017). Source: Estadísticas Educativas del Sistema Nacional de Información Estadística Educativa (SNIE).

Note 4. Source: Estadísticas Educativas del Sistema Nacional de Información Estadística Educativa (SNIE; 2016–2017).

Note 5. In the last grade of upper secondary education, the rates are 62 and 89.5 percent for both subjects, respectively.

Note 6. Source: Secretaría de Educación Pública (2012); Reporte de la Encuesta Nacional de Deserción en la Educación Media Superior. The sample consisted of 2,549 observations that had dropped out of upper secondary education, 5,686 observations that were enrolled in upper secondary education, and 4,779 observations that had never been enrolled in upper secondary education.

Note 7. Source: Secretaría de Desarrollo Social (2013). ESJóvenes 2013. Análisis descriptivo del Cuestionario de Seguimiento a jóvenes de PROSPERA Programa de inclusión social.

Note 8. Mentors in upper secondary education and higher education hold a certification in Mentoring.

Note 9. Guanajuato, Michoacán, and Zacatecas.

Note 10. See Cecchini and Atuesta (2017); Conditional Cash Transfers in Latin America and the Caribbean: Coverage and Investment Trends.

Note 11. Molina-Millán et al. (2016) define long-term effects as those that imply human capital accumulation and/or effects that are observed in another phase of the cycle of life.

Note 12. The program included money transfers, the provision of products with high nutritional content for minors, access to health and nutrition services, and education scholarships. These elements were offered subject to school attendance and health check-ups, as well as the participation of the head of household -mothers- to training sessions. By 2014, the program had expanded the scholarships to higher education.

Note 13. See Shultz (2000); Impact of Progresa on school attendance rates in the sampled population.

Note 14. Considering the period of interest for the evaluation, we are interested in the editions of ENLACE or PLANEA that were conducted as a census during the years of ENLACE 2012, ENLACE 2013, ENLACE 2014 and PLANEA 2015.

Note 15. It was only after 2009 that the BBVA Foundation incorporated an online system to register the applicants' information. When the IDs were entered electronically into the system, the quality checks reduced errors.

Note 16. The program expanded to upper secondary education in 2013. Thus, these beneficiaries would have graduated from the upper secondary level in 2016. The edition of PLANEA for 2016 is no longer a "census", so the evaluation strategy is no longer feasible.

Note 17. While some mentors are more involved than others, the Foundation makes particular emphasis on following up the school trajectory of every beneficiary through a community-based support service, led by Bank personnel. Those performing these services engage in school visits, coordinate student-mentor meetings, and acts as a "coach" available to students and their families during their participation in the Program more flexibly and informally, as compared to the student-mentor relation.

Note 18. These include documents such as the student's report card and proof of enrollment in a public lower secondary school, among others. Also, only students who were finishing the 6th grade of elementary school at the time of the application were considered in the admissions process.

Note 19. It should be noted that throughout the history of the program, the questions in the application have evolved to improve not only the quality of the information but also to strengthen the selection process. The same can be said of the weight given to these criteria.

Note 20. In Mexico, grades can go from 5 to 10. Grades between 6 and 10 are considered approbatory, while a 5 is regarded as a fail.

Note 21. Those not fulfilling the requirements for participation into the program, where evidently *not* considered as possible "substitutes".

Note 22. Although the Propensity Score Matching method is usually used to design a control group that minimizes selection bias, the matching paradox occurs when, after matching, the bias between the groups increases, invalidating the results of the estimates, Therefore, it is essential to verify that the differences between the groups are smaller with respect to the variables used to estimate the propensity score, after pairing. In the same way, the estimation of the impact by the propensity score matching method requires that the assumption of conditional independence be met, which refers to the fact that, given a set of observed characteristics, all the results are independent of the allocation to the treatment. In other words, it is assumed that there are no other characteristics, other than those used to estimate the propensity score, that influence treatment allocation. The second assumption is that there is a common support region in terms of the estimated probability of receiving the intervention—propensity score—under which comparisons are made between the treatment and control groups. The common support ensures the existence of close observations between the treatment and control groups based on the estimated score. Annex B develops the model and its assumptions with mathematical notation.

Note 23. Given that treatment assignment was based on a rule or priority ranking, we do not conduct balance tests.

Note 24. Pairs were also formed imposing a caliper of 0.1—equivalent to 20 percent of the standard deviation of the natural logarithm of the propensity score—and the results are the same.

Note 25. 3,331 observations with three years of scholarship and mentoring. By imposing the use of observations inside the common support, as well as the caliber and the minimum number of pairs per observation in the treatment, the sample is reduced to 3,157 observations for NN (1) and NN (2), and 3,156 for NN (3), which represents a very slight decrease.

Note 26. By imposing the use of the observations in the common support, the difference in the number of observations between groups T1 and T2 is reduced to 217.

Note 27. 26 and 5 percent of the observations in T3 are applicants under the "Disenrolled" and "Substitute" status.

Note 28. T3 is reduced to 4,571 observations when imposing the use of observations in the common support.

Appendix A
OLS regressions

Table A1. OLS regressions for T1

	(1) Graduated from upper secondary education	(2) Graduated from upper secondary education	(3) Graduated from upper secondary education	(4) Graduated from upper secondary education	(5) Graduated from upper secondary education
Treatment 1	0.0818***	0.0785***	0.106***	0.0976***	0.0969***
	(0.0101)	(0.0101)	(0.0111)	(0.0111)	(0.0136)
Female		0.0116	0.0127	0.00891	0.00891
		(0.00989)	(0.00996)	(0.00998)	(0.00998)
Per capita income (ln)			0.0572***	0.0511***	0.0509***
			(0.00929)	(0.00934)	(0.00945)
GPA (9 or higher)				0.100***	0.101***
				(0.0105)	(0.0105)
Migrant relative (1 or more)					0.00113
					(0.0134)
Constant	0.627***	0.624***	0.230***	0.208***	0.208***
	(0.00591)	(0.00835)	(0.0647)	(0.0648)	(0.0653)
Observations	9,765	9,718	9,524	9,401	9,401
R-squared	0.007	0.006	0.011	0.020	0.020

Note. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A2. OLS regressions for T2

	(1) Graduated from upper secondary education	(2) Graduated from upper secondary education	(3) Graduated from upper secondary education	(4) Graduated from upper secondary education	(5) Graduated from upper secondary education
Treatment 2	0.0551***	0.0521***	0.0791***	0.0717***	0.0736***
	(0.00962)	(0.00963)	(0.0105)	(0.0105)	(0.0126)
Female		0.0135	0.0149	0.0104	0.0103
		(0.00963)	(0.00971)	(0.00972)	(0.00972)
Per capita income (ln)			0.0581***	0.0516***	0.0519***
			(0.00907)	(0.00912)	(0.00920)
GPA (9 or higher)				0.105***	0.105***
				(0.0102)	(0.0102)
Migrant relative (1 or more)					-0.00338
					(0.0126)
Constant	0.627***	0.623***	0.222***	0.200***	0.198***
	(0.00595)	(0.00827)	(0.0632)	(0.0633)	(0.0637)
Observations	10,416	10,366	10,153	10,022	10,022
R-squared	0.003	0.003	0.007	0.018	0.018

Note. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A3. OLS regressions for T3

	(1) Graduated from upper secondary education	(2) Graduated from upper secondary education	(3) Graduated from upper secondary education	(4) Graduated from upper secondary education	(5) Graduated from upper secondary education
Treatment 3	0.0113	0.00890	0.0359***	0.0307***	0.0363***
	(0.00898)	(0.00899)	(0.00981)	(0.00982)	(0.0116)
Female		0.0132	0.0146	0.0101	0.0101
		(0.00917)	(0.00925)	(0.00926)	(0.00926)
Per capita income (ln)			0.0573***	0.0506***	0.0518***
			(0.00860)	(0.00864)	(0.00874)
GPA (9 or higher)				0.102***	0.102***
				(0.00965)	(0.00965)
Migrant relative (1 or more)					-0.0106
					(0.0116)
Constant	0.627***	0.623***	0.227***	0.209***	0.202***
	(0.00601)	(0.00814)	(0.0599)	(0.0600)	(0.0605)
Observations	11,668	11,612	11,350	11,219	11,219
R-squared	0.000	0.000	0.004	0.014	0.014

Note. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4. OLS regressions for T4

	(1) Graduated from upper secondary education	(2) Graduated from upper secondary education	(3) Graduated from upper secondary education	(4) Graduated from upper secondary education	(5) Graduated from upper secondary education
Treatment 4	0.105***	0.104***	0.123***	0.111***	0.112***
	(0.0135)	(0.0135)	(0.0140)	(0.0142)	(0.0143)
Female		0.00244	0.00337	0.000335	0.000322
		(0.0110)	(0.0110)	(0.0110)	(0.0110)
Per capita income (ln)			0.0620***	0.0546***	0.0549***
			(0.0112)	(0.0113)	(0.0113)
GPA (9 or higher)				0.0978***	0.0977***
				(0.0116)	(0.0116)
Migrant relative (1 or more)					-0.00728
					(0.0172)
Constant	0.627***	0.629***	0.202***	0.190**	0.189**
	(0.00593)	(0.00884)	(0.0776)	(0.0783)	(0.0783)
Observations	7,974	7,927	7,894	7,733	7,733
R-squared	0.008	0.008	0.011	0.020	0.020

Note. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B

Propensity Score Matching

Based on the identification of the observations with treatment (T = 1), without treatment (T = 0), and the set of observed characteristics X, a propensity score P(X) is estimated, such that:

$$P(X) = Pr(T=1|X) \tag{1}$$

where P (X) is the probability of having received the treatment, conditioned to characteristics X. To calculate this

probability, it is recommended to use variables that are related to the treatment and the variable of interest. One must avoid variables that perfectly predict the treatment allocation, variables that are influenced by the treatment, and variables that are related to the treatment allocation, but not with the variable of interest. Performing P-based pairing (X) has the advantage of reducing the dimensionality that would result from making the matching using the X features separately.

After matching, the differences regarding the outcome between the observations in the treatment group and the control group allow to estimate the average treatment effect on the treated (Average Treatment Effect on the Treated):

$$ATET = E_{P(X)\,|\,T=1} \{E\,[Y^T\,|\,T=1, P\,(X)] - E\,[Y^C\,|\,T=0, P\,(X)]\} \quad (2)$$

PSM requires conditional independence which means that given a set of observed characteristics X, the potential results Y are independent of the treatment allocation T. In other words, it is assumed that there are no other characteristics, other than those used to estimate the propensity score, that influence treatment allocation.

$$(Y_i^T, Y_i^C) \perp T_i\,|\,X_i \quad (3)$$

where Y_i^T, and Y_i^C are the outcomes for the treatment and control group, respectively, and where the X variables are not affected by the treatment. If the ATET is calculated instead of the average treatment effect (ATE), then condition (3) relaxes to:

$$Y_i^C \perp T_i\,|\,Xi \quad (4)$$

The second assumption is that there is a common support region in terms of the estimated probability of receiving the intervention -propensity score- in which comparisons are made between the treatment and control groups. The common support ensures the existence of similar observations between the treatment and control groups, based on the estimated score:

$$0 < Pr(T=1|X) < 1 \quad (5)$$

This condition can be relaxed for the estimation of the ATET to:

$$P\,(Ti = 1|Xi) < 1 \quad (6)$$

Appendix C

Covariate balance pre and post matching

Table C1. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T1 and C; Matching algorithm: NN(1)

	Raw	Matched		
Num of obs	9401	6314		
Treated obs	3157	3157		
Control obs	6244	3157		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.12	-0.02	0.94	1.02
Grade average > =9 (promedio_9omas)	0.12	-0.01	0.90	1.01
Per capita income (ln_ingpercapita)	-0.82	0.00	1.44	1.02

Table C2. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T1 and C; Matching algorithm: NN(2)

	Raw	Matched		
Num of obs	9401	6314		
Treated obs	3157	3157		
Control obs	6244	3157		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.12	-0.02	0.94	1.01
Grade average > =9 (promedio_9omas)	0.12	-0.01	0.90	1.01
Per capita income (ln_ingpercapita)	-0.82	-0.01	1.44	1.02

Table C3. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T1 and C; Matching algorithm: NN(3)

	Raw	Matched		
Num of obs	9400	6312		
Treated obs	3156	3156		
Control obs	6244	3156		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.12	-0.02	0.94	1.01
Grade average > =9 (promedio_9omas)	0.13	0.01	0.90	0.99
Per capita income (ln_ingpercapita)	-0.82	0.00	1.43	1.02

Table C4. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T2 and C; Matching algorithm: NN(1)

	Raw	Matched		
Num of obs	9618	6748		
Treated obs	3374	3374		
Control obs	6244	3374		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.12	-0.02	0.94	1.01
Grade average > =9 (promedio_9omas)	0.11	-0.02	0.92	1.02
Per capita income (ln_ingpercapita)	-0.83	-0.01	1.43	1.01

Table C5. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T2 and C; Matching algorithm: NN(2)

	Raw	Matched		
Num of obs	9618	6748		
Treated obs	3374	3374		
Control obs	6244	3374		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.12	-0.01	0.94	1.01
Grade average > =9 (promedio_9omas)	0.11	-0.04	0.92	1.04
Per capita income (ln_ingpercapita)	-0.83	-0.01	1.43	1.03

Table C6. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T2 and C; Matching algorithm: NN(3)

	Raw	Matched		
Num of obs	9618	6748		
Treated obs	3374	3374		
Control obs	6244	3374		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.12	-0.02	0.94	1.01
Grade average > =9 (promedio_9omas)	0.11	-0.02	0.92	1.02
Per capita income (ln_ingpercapita)	-0.83	-0.01	1.43	1.02

Table C7. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T3 and C; Matching algorithm: NN(1)

	Raw	Matched		
Num of obs	10815	9142		
Treated obs	4571	4571		
Control obs	6244	4571		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.06	-0.01	0.97	1.01
Grade average > =9 (promedio_9omas)	0.05	-0.03	0.97	1.02
Per capita income (ln_ingpercapita)	-0.82	-0.01	1.41	1.03

Table C8. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T3 and C; Matching algorithm: NN(2)

	Raw	Matched		
Num of obs	10815	9142		
Treated obs	4571	4571		
Control obs	6244	4571		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.06	-0.01	0.97	1.01
Grade average > =9 (promedio_9omas)	0.05	-0.02	0.97	1.02
Per capita income (ln_ingpercapita)	-0.82	0.00	1.41	1.03

Table C9. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T3 and C; Matching algorithm: NN(3)

	Raw	Matched		
Num of obs	10815	9142		
Treated obs	4571	4571		
Control obs	6244	4571		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.06	-0.01	0.97	1.01
Grade average > =9 (promedio_9omas)	0.05	-0.03	0.97	1.02
Per capita income (ln_ingpercapita)	-0.82	-0.01	1.41	1.03

Table C10. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T4 and C; Matching algorithm: NN(1)

	Raw	ing Alg		
Num of obs	7733	2978		
Treated obs	1489	1489		
Control obs	6244	1489		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.17	0.00	0.91	1.00
Grade average > =9 (promedio_9omas)	0.22	0.01	0.81	0.99
Per capita income (ln_ingpercapita)	-0.69	0.00	0.81	0.99

Table C11. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T4 and C; Matching algorithm: NN(2)

	Raw	Matched		
Num of obs	7733	2978		
Treated obs	1489	1489		
Control obs	6244	1489		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.17	0.00	0.91	1.00
Grade average > =9 (promedio_9omas)	0.22	-0.01	0.81	1.02
Per capita income (ln_ingpercapita)	-0.69	0.00	0.81	1.01

Table C12. Standardized differences and variance ratio for pre- and post-matching covariates. Groups: T1 and C; Matching algorithm: NN(3)

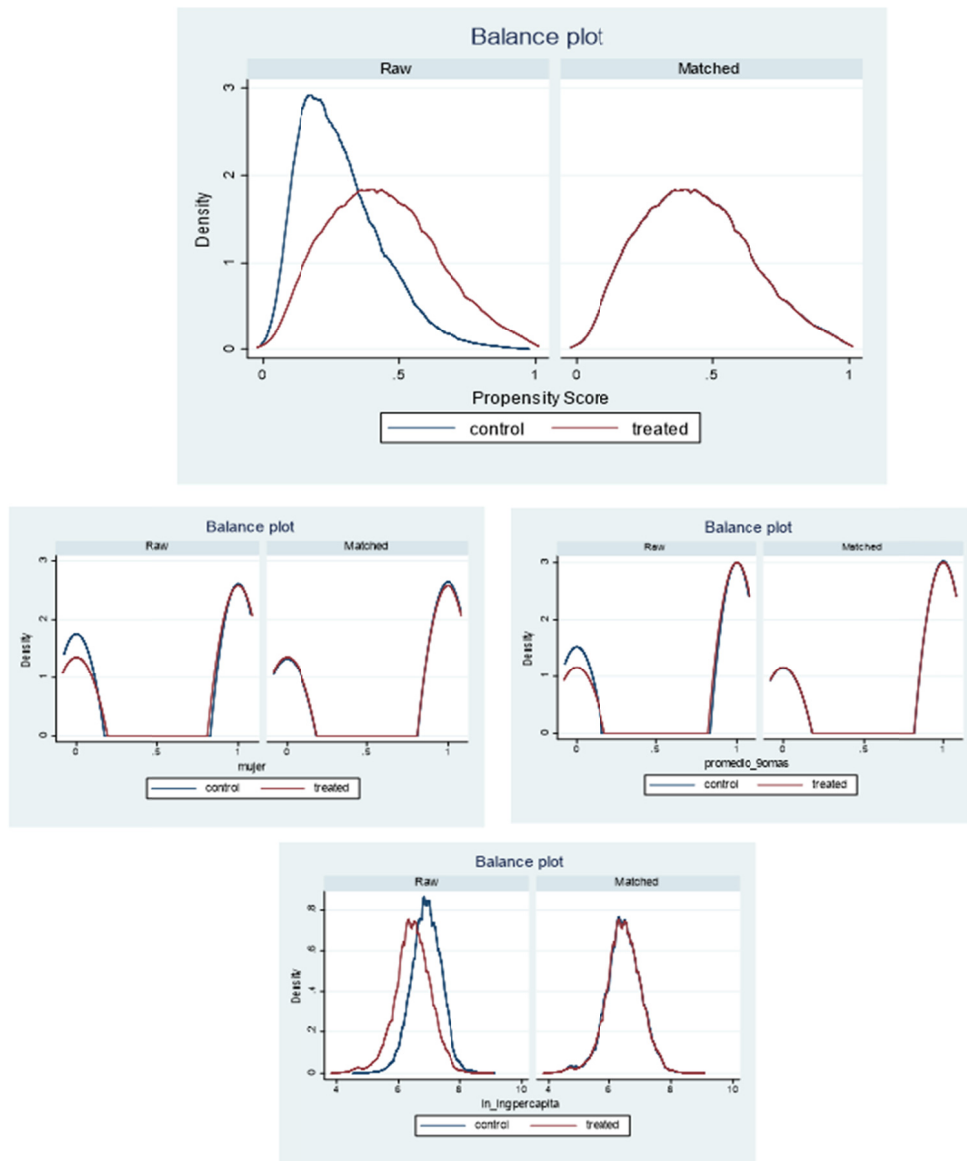|	Raw	Matched		
Num of obs	7733	2978		
Treated obs	1489	1489		
Control obs	6244	1489		
	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Sex (mujer)	0.17	0.01	0.91	1.00
Grade average > =9 (promedio_9omas)	0.22	-0.02	0.81	1.03
Per capita income (ln_ingpercapita)	-0.69	-0.01	0.81	0.99

Figure C1. Balance of covariates pre- and post-matching. Groups: T1 and C. Matching Algorithm: NN(1)
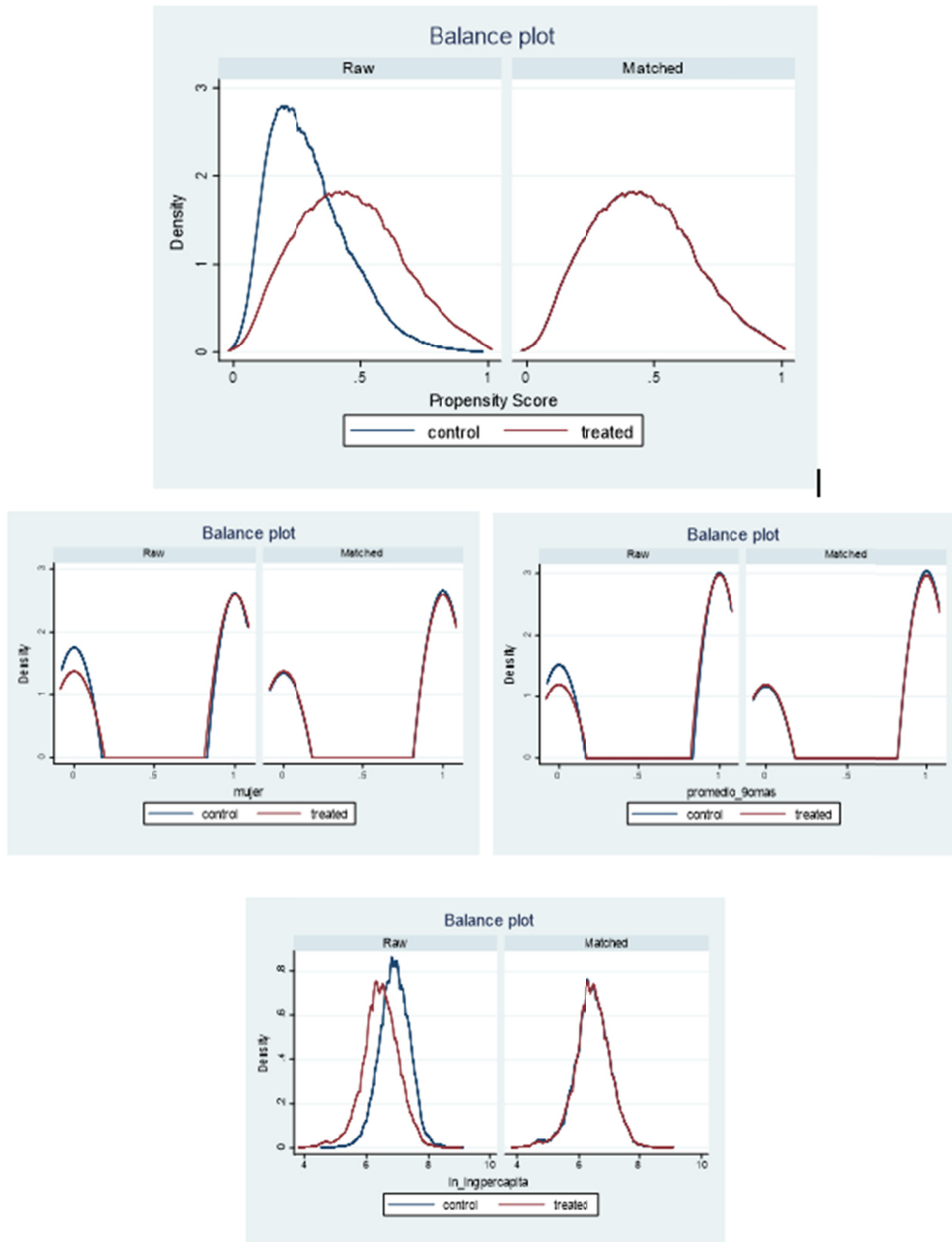
Figure C2. Balance of covariates pre and post matching. Groups: T2 and C. Matching Algorithm: NN(1)
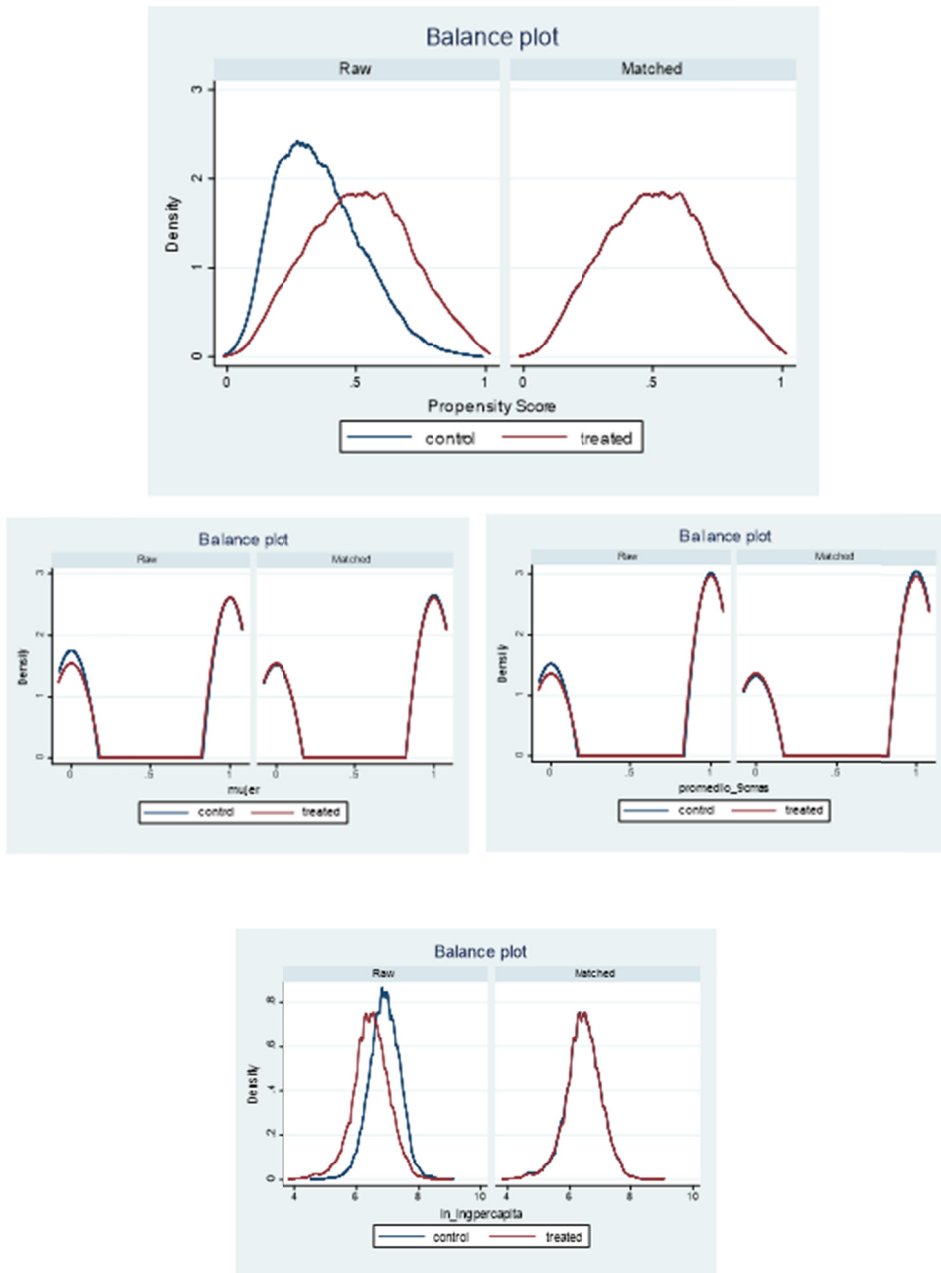
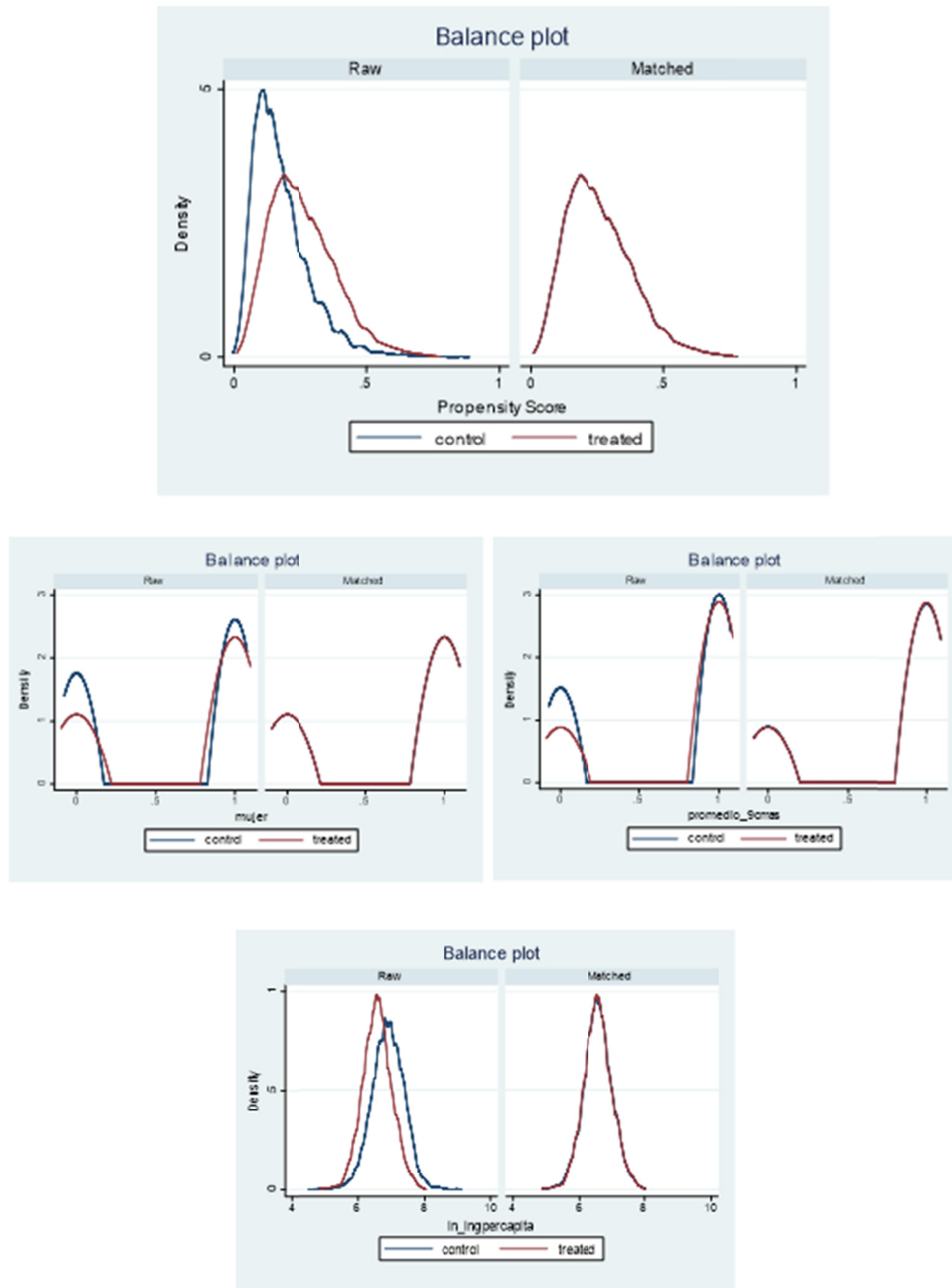Figure C3. Balance of covariates pre and post matching. Groups: T3 and C. Matching Algorithm: NN(1)

Figure C4. Balance of covariates pre and post matching. Groups: T4 and C. Matching Algorithm: NN(1)

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.