Policy Research Working Paper 8826

Can Scholarships Increase High School Graduation Rates?

Evidence from A Randomized Control Trial in Mexico

Rafael de Hoyos
Orazio Attanasio
Costas Meghir

WORLD BANK GROUP
Education Global Practice
April 2019

Policy Research Working Paper 8826

Abstract

This paper studies the impact of PROBEMS, a scholarship program in Mexico aimed at improving graduation rates and test scores among upper secondary school students from poor backgrounds. The identification strategy is the random allocation into the program, which took place in 2009. The strategy allows measurement of the effects of PROBEMS on test scores and graduation rates three years later in 2012. The paper finds that, on average, the program has no discernible impact on graduation rates or math or Spanish test scores. The size of the sample allows investigation of the reasons for this disappointing result. The paper finds that the program is substantially mis-targeted, with the majority of the recipients not coming from the most disadvantaged families. However, the most plausible explanation for the absence of positive impacts is that many eligible students do not seem to have the minimum learning level to face successfully the academic requirements of upper secondary school. An important policy implication is that a well-targeted scholarship program should be complemented with a remedial education intervention.

This paper is a product of the Education Global Practice. 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 rdehoyos@worldbank.org.

Can Scholarships Increase High School Graduation Rates? Evidence from a Randomized Controlled trial in Mexico*

Rafael de Hoyos† Orazio Attanasio‡ Costas Meghir§

Keywords: scholarships, dropout, high school, RCT.

JEL codes: I25, D80, O12

*We are especially indebted to Elizabeth Monroy and Juan Manuel Martínez de la Calle who were responsible for project and data management at the Mexican Secretariat of Public Education (SEP) at the time of implementing the project. Juan Bedoya provided excellent research assistance. The views expressed here are those of the authors alone, and do not necessarily reflect the opinions of the World Bank.

†Corresponding author, Education Unit for Latin America, The World Bank. E-mail: rdehoyos@worldbank.org.

‡Department of Economics, University College London. Email: o.attanasio@ucl.ac.uk.

§Department of Economics, Yale University. Email: c.meghir@yale.edu

1 Introduction

Improving high school or upper secondary graduation rates for children from lower income backgrounds is often seen as important for offering access to opportunity and improving intergenerational mobility. Indeed, in many circumstances low attendance is an important issue of economic efficiency and has implications for aggregate economic growth ((Hanushek and Woessmann 2012), (WorldBank 2018)). As such it has motivated policy in numerous countries, where financial support is offered for upper secondary attendance that goes beyond providing tuition free. The underlying reason that is often cited is that certain high ability but low SES students are liquidity constrained, which generates a wealth gradient in attendance even conditional on ability ((Abbott et al. 2019), (Belley and Lochner 2007)). Such liquidity constraints can arise both because parents may be reluctant to fund education when they are unlikely to be repaid and/or because willing parents are themselves poor and constrained in their ability to fund their children's education. This can imply that children with potentially high returns do not obtain enough schooling. Both from efficiency and equity considerations, targeting such programs appropriately is likely to be of first-order importance.

In light of such arguments, the Mexican government introduced a new scholarship program in 2007 (PROBEMS) targeted to poor upper secondary students, with the objective of increasing graduation rates as well as learning outcomes. As there was excess demand from eligible students for such a program, the scholarships were allocated through a lottery system in 2009. Based on this randomization, we estimate the impact of the intervention on beneficiary students' probability of graduation and on their test scores at the end of upper secondary (2012).

The main result we obtain is that the scholarship program had no effects either on upper secondary graduation rates or on performance in the standardized test. However, to better understand the mechanisms behind the results of the exercise, one has to take into account a variety of factors, ranging from different motivations for school enrollment, to the possibility of heterogeneous effects in different groups of the population.

This surprising result could be due to miss-targeting of the program. As a result of such a conditional transfer we would expect two different groups to increase their graduation rates: those who have high net returns but are liquidity constrained and those who have low net returns and would not have enrolled in upper secondary (optimally) but because of the conditionality are incentivized to do so. If the program is badly targeted and liquidity constraints are not an issue the question remains whether we can expect an impact (albeit possibly inefficient) from those whose net returns become positive as a result of the transfer. While this may be the case, many of these children may have been unprepared for the relative rigors of upper secondary either because of earlier low quality of education or because of low effort in anticipation of dropping out.

To try to better understand the absence of an impact, we obtain estimates of the effects of the intervention for different groups of youths, divided on the basis of their socioeconomic background, their motivation at baseline and their academic *readiness*. Our results show that most of the impacts for different groups are not significantly different from zero. Even among the poorest students and among highly motivated ones–proxied by the highest education level they expect to achieve the scholarship was ineffective. The only subgroup of students where the scholarship had a positive and significant impact on the probability of graduation were those with relatively high initial test scores. For students at the top tercile of the test score distribution at baseline, being awarded a scholarship increases the probability of graduation by 4 percentage points.

One possible interpretation of the results we find is that individuals from extremely poor backgrounds are at a disadvantage which is not exclusively economic. It might be that their set of basic skills presents important deficiencies that cannot be alleviated by a scholarship. It may also be that they lack the confidence in their own academic skills or that they perceive the returns from upper secondary education to be low. An alternative and possibly more productive use of public resources would be to fill in the skills and knowledge gap among poor and less poor students before upper secondary, possibly improving the education these youths receive in primary and junior high school. A remediation program at the start of upper secondary (or during the summer between lower and upper secondary), providing basic skills to kids from marginalized backgrounds could be a more

effective use of the resources devoted to PROBEMS.

The rest of the paper is organized as follows. In Section 2 we discuss the Mexican context and describe the intervention. In particular, in Section 2.1, we discuss the upper secondary or high school system in Mexico, Section 2.2 describes the PROBEMS scholarship system, including the trends in number of beneficiaries. Section 3 describes the evaluation design, the baseline and follow-up data. Section 4 presents the main results. Section 5 presents a discussion of some of the reasons that explain the lack of effects of the intervention. Finally, Section 6 concludes with the policy implications of the results.

2 Context and Intervention

Mexico, like other middle-income countries, has reached almost universal enrollment rates in primary school (grades 1 to 6) and lower secondary school (grades 7 to 9) school. However, its education system still faces important challenges, especially in upper secondary school. For instance, around 35 of every 100 students who enroll in upper secondary will never graduate. Among those who graduated from high school in 2015, more than 60% attained insufficient levels in math according to the national standardized test, Planea ((INEE 2017b)).[1] Many of the students dropping out or finishing high school but with insufficient skills come from poor or marginalized households. Therefore, upper secondary education dropouts and low achievement levels have important implications for Mexico's long-term economic growth and income disparities ((de Hoyos et al. 2016)).

2.1 Upper Secondary Education in Mexico

The upper secondary education system in Mexico (EMS for its acronym in Spanish) consists of 4.9 million students, typically between 15 to 18 years old, in grades 10th, 11th and 12th. The EMS system is large and complex with several service providers and types of degree programs. EMS is offered by four different providers: 1) the federal government (accounting for 21.8% of total enrollment), 2) the state governments (47.4%), 3) publicly financed autonomous universities (12.3%), and 4) private entities. EMS offers three types

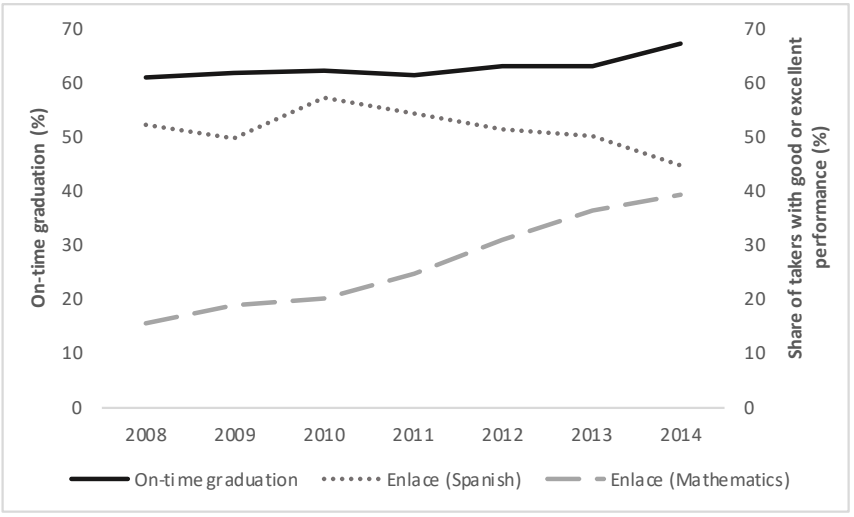
[1] http://publicaciones.inee.edu.mx/buscadorPub/P2/A/328/P2A328.pdf

of degree programs: *general* – preparing students for higher education, *technological* – preparing students both for the labor market and for higher education, and *technical* – emphasizing technical and vocational education (INEE 2017a)).

Although graduation rates and learning outcomes in EMS have experienced an improvement during the last 10 years, their levels are consistently low (see Figure 1). According to the official statistics from Mexico's National Institute for the Evaluation of Education (INEE for its acronym in Spanish),[2] in 2014 only 67% of students graduated three years after enrolling in EMS, with this share being significantly higher among females (70%) than males (62%). Graduation rates vary across types of degree programs with *general* schools showing the highest (69%), followed by *technological* schools with rates very close to the national average and *technical* schools showing the lowest (54%). According to INEE, more than 60% of the cumulative dropouts throughout the three years of EMS take place during the first year. Household survey data show that enrollment in EMS among youths aged 15 to 18 varies substantially across household income deciles. For instance, in 2012, only 13.5 percent of youth aged 15 to 18 in the poorest households were enrolled in EMS versus an enrollment rate of 95 percent among the richest households. The disparity in enrollment rates across the distribution raises the question of whether liquidity constraints have an important role to play. Moreover, information from the 2011 *EMS School Dropout Survey* shows that more than a third of the 2,549 EMS dropouts surveyed declared that economic constraints were the main reason for leaving school ((SEP 2012)).

[2]See http://publicaciones.inee.edu.mx/buscadorPub/P1/B/115/P1B115.pdf

Figure 1: Graduation Rates and Learning Outcomes



The EMS system is characterized by strict promotion criteria. Students must pass five of eight disciplinary subject areas and practical modules. Otherwise they have to repeat the semester. Students who fail three or fewer subject areas can enroll in the next semester, but they have to attend and pass intensive courses (the so-called *regularizacion*) during a fixed time window. In addition, students must satisfactorily complete all their subject areas and modules within 10 semesters after enrolling in EMS, otherwise they lose the right to re-enroll. Partly as a result of the strict promotion rules, there are very high grade and subject repetition rates, 15.3 percent and 31.3 percent, respectively in 2013.[3]

2.2 The PROBEMS Scholarship Program

In the context of a major EMS curricular reform[4], with the aim of reducing school dropouts and improving learning outcomes, in 2007 the Secretariat of Public Education (SEP for its acronym in Spanish) introduced a new scholarship program, PROBEMS. The program targeted poor EMS students mostly in urban areas since, at the time, a national conditional cash transfer program, *Progresa/Oportunidades* later renamed as *Prospera*, already benefited poor EMS students in rural areas.[5] The number of beneficiaries under

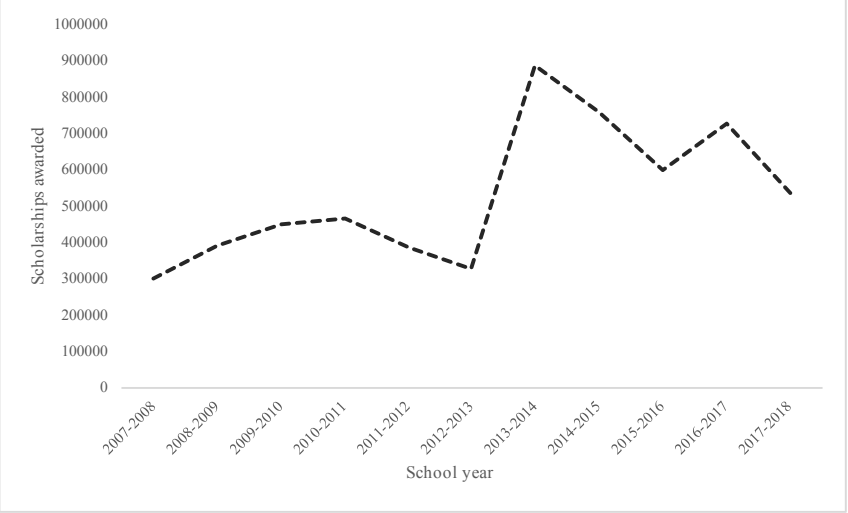
[3]Students who fail three or more subjects for two consecutive semesters have to repeat the entire grade.

[4]For more information on the EMS reform of 2008, see SEP (2008): http://cosdac.sems.gob.mx/portal/index.php/riems.

[5]The conditional cash transfer program has since then expanded to many urban areas.

PROBEMS increased substantially between 2007 and 2014, from less than 300,000 to almost 900,000 (see Graph 2), representing 23% of the total enrollment in EMS in 2014. The budget assigned to the program in 2014, the year when PROBEMS had the largest number of beneficiaries, was close to US $410 million, representing 23% of the total EMS federal budget, excluding salaries.

Figure 2: Evolution in the Number of PROBEMS Beneficiaries



PROBEMS had three different types of scholarships: "support", "retention" and "excellence". The difference between the three types of scholarships was determined by the grade point average (GPA) during the year prior to applying to the program (the last year of lower secondary in the case of candidates about to enter EMS). The "excellence" scholarship required a minimum GPA of 9 out of 10, "retention" a GPA between 8 and 8.9 and a minimum GPA of 6 (the passing mark) was required under the "support" modality. To incentivize students' efforts, monthly transfers were marginally larger under the "excellence" modality than under the "retention" modality and this was marginally larger than the transfer under the "support" modality. Transfers also varied between gender, grade and type of degree program. Table 1 shows the monthly transfer in 2009 under the different modalities, by grades and gender. The overall average monthly transfer of a PROBEMS scholarship was $716 MX or $56 US dollars in 2009, slightly less than the

extreme poverty line of that same year.[6]

Table 1: Monthly Transfers by Type of Scholarship

	Grade	**Man**	**Woman**
Support (GPA 6.0 to 7.9)	1o	500	525
	2o	525	575
	3o and 4o	575	625
Retention (GPA 8.0 to 8.9)	1o	650	700
	2o	700	750
	3o and 4o	750	790
Excellence	ME1 (GPA 9.0 to 9.4)	850	900
	ME2 (GPA 9.5 to 9.7)	900	950
	ME3 (GPA 9.8 to 10)	950	1000

All amounts expressed in Mexican pesos of 2009 (13 pesos to the US$ at the time)

PROBEMS' operating rules in 2009, the year when the lottery took place, defined the following eligibility criteria to be considered as a potential beneficiary:

1. Be enrolled in a public EMS school.
2. In the case of applicants who were about to enter EMS, present their lower secondary graduation certificate. For students continuing EMS studies, they must have a passing mark in all subject areas of the previous semester.
3. Not be enrolled in any other government program providing scholarships.
4. Have a household per capita income below the "assets" poverty line defined by CONEVAL as $1,413 MX and $2,102 per person, per month for households located in rural and urban areas, respectively.
5. Apply for a PROBEMS scholarship through the program's portal.[7]

[6] In 2009 the National Council of Social Policy Evaluation (CONEVAL) defined three poverty lines in Mexico: "food intake" (pobreza alimentaria) or extreme poverty line, "capabilities" (pobreza de capacidades) and "assets" (pobreza patrimonial), with values of $949 MX, $1,164 MX, and $1,905 MX, per person, per month, respectively.

[7] https://www.becasmediasuperior.sep.gob.mx

To assign the PROBEMS scholarships, SEP issues one and sometimes two calls for applications per year depending on budget availability. The call for applications is usually issued in March of each year, targeting students starting or continuing EMS in the following academic year (August). Interested candidates start their application process by filling in a form capturing household income, the availability of household assets, parents' education and basic academic information such as the school and grade of enrollment. All applicants must have a personal identifier known as the CURP, issued by the Government of Mexico.[8] SEP uses the CURP to identify applicants who are beneficiaries of the conditional cash transfer program (PROGRESA / Oportunidades) at the time of applying, which automatically excludes them from PROBEMS. The self-declared household income and assets information included in the on-line application format is used to rank students in terms of their likelihood of being below the "assets" poverty line. Finally, as stated by the call for applications, SEP gives priority to first-year EMS students. The likelihood of being poor, budget availability and EMS grade of enrollment determined a list of pre-selected candidates.

When a candidate is pre-selected–usually one month after the call for applications is closed, he or she is notified by email with instructions on how to complete the application process. The next step is to enroll in EMS or continue to the second or third grade, and present, at the school, all the relevant documentation to validate the information uploaded in the program's portal. School directors receive the list of pre-selected candidates via email and are responsible for validating the information provided by the candidates through the program's portal. In particular, the school director has to validate that the student is indeed enrolled in his or her school, that the student has no subject areas of previous semesters below the passing mark, and validate the GPA as declared by the student. Once the school director validates the information, the candidates receive a second email notifying them that they have been awarded a scholarship. SEP issues debit cards under the name of the beneficiary. As a final step the student, together with their father, mother or legal guardian, collects the debit card at the closest branch of the commercial bank

[8]Mexican citizens have a unique personal identifier, known as *Clave Única de Registro Poblacional, CURP*, formed by an algorithm combining name, surname, date of birth, sex, state of birth, plus 2 randomly generated digits.

participating in the program.

3 Evaluation Design

Our evaluation strategy relies on the random assignment into the program. The call for applications in March 2009 received close to 70,000 valid applications of which more than 50,000 of them fulfilled the eligibility criteria. However, SEP's budget was enough for awarding between 40,000 to 45,000 scholarships, depending on the distribution by type of scholarship. Thus, for the purposes of evaluation SEP authorized a randomized assignment of the scholarship in a pre-selected eligible population of 12,000 students: 6,000 students were randomly selected to be part of the treatment group and the same number formed the control group. All 12,000 students received an email from SEP at the end of April 2009 notifying them that they were granted or rejected a scholarship according to the random assignment. PROBEMS is assigned on a per-year basis, therefore the compliance with the random assignment was only valid for one school year, so all students, treatments and controls, could apply to receive a scholarship in the call for applications in March 2010.

Since PROBEMS prioritized first-year EMS students, 60% of the eligible candidates were completing lower secondary (9th grade) while submitting their application in March 2009. Therefore 3,648 students in the treatment and 3,673 in the control group, respectively, were about to start EMS in August of 2009. For the purposes of this paper, we will concentrate on this subsample. Focusing on first-year students and following them through the three school years of EMS, estimates the impact of one year of exposure to the program at the beginning of EMS, versus not having this monetary support, on graduation rates and learning outcomes measured by the standardized test ENLACE in 12th grade.

Figure 3 presents a timeline of the evaluation design including the important dates:

1. Call for applications to PROBEMS on March 2009. Candidates fill the program's online application form.

2. Randomization among eligible students happened between end of March and April

2009.

3. Candidates were informed of the outcome of their application in April 2009.

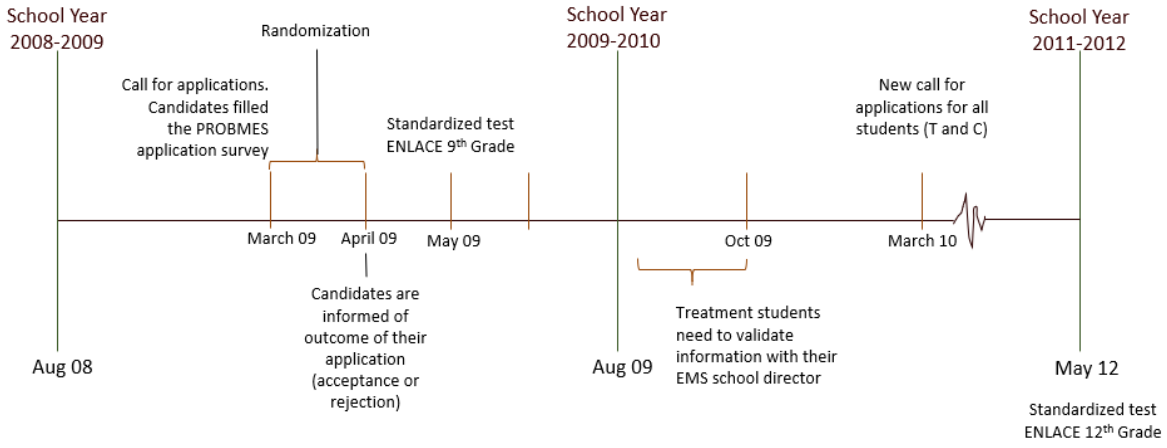
4. The national standardized test ENLACE 9th grade was applied to all students finishing lower secondary in May 2009.

5. To receive the scholarship, school directors of first-year treatment students had to validate all the information provided during the application between August and end of September 2009.

6. A new call for applications was issued in March 2010 and all students, regardless of their treatment status could apply.

7. The national standardized test ENLACE 12th grade was applied to all students finishing EMS in May 2012 and May 2013. These are the administrative data used to measure the outcome variables: graduation rates and math and Spanish test scores.

Figure 3: Timeline of the PROBMES Impact Evaluation Strategy



3.1 Evaluation Data

In order to measure students' characteristics at baseline, we use two sources. First, we rely on the form completed by all applicants in March 2009, through the PROBEMS portal. This survey included self-reported information on household income, education

of parents, number of family members living in the same household, the availability of household assets as well as gender, age, geographic location and GPA of the applicant, among others.

Second, to measure students' ability at baseline, we use the information from the census-based ENLACE 9th grade, measuring math and language achievement levels. From 2007 to 2013, ENLACE was administered to all students in 3rd to 9th grades and those finishing 12th grade. The test had no consequences either on graduation or on student's GPA. The score of ENLACE is normalized to have a mean of 500 and a standard deviation of 100.

Using the CURP we were able to merge the baseline application format with the micro data from ENLACE 9th grade which was taken by applicants in May 2009. For ease of exposition and interpretation, when presenting the results, we renormalize the ENLACE results to a mean of zero and a standard deviation of 1. Table 2 shows the baseline characteristics, distinguishing between students in the treatment and the control groups. In the top panel we report the socioeconomic characteristics measured through the baseline format completed by all applicants, in the bottom panel the administrative information on 9th grade test scores. Overall, the characteristics of the treatment and control groups are well balanced in line with the randomized design of the evaluation.

Outcome variables are measured through ENLACE 12th grade administered in May 2012 to all students finishing EMS. Additionally, our evaluation sample was also merged with ENLACE 12th grade of 2013 to identify students that graduated from high school but that had a lag of one year. The participation of students in our sample, who entered EMS in August 2009, in ENLACE 12th grade of 2012 or 2013 is used as a proxy for upper secondary graduation.[9] The same end-of-secondary-school test is also used to measure the impact of PROBEMS scholarship on math and Spanish test scores.

[9]In what follows, when we refer to ENLACE 12th grade, which includes the information from the tests administered in 2012 and 2013. Of the total 7,321 students in our evaluation sample, 4,978 where identified in ENLACE 12th grade of 2012, a proxy for on-time graduation. An additional 262 students were identified in ENLACE 12th grade of 2013, these students also graduated but with a lag of one year. So most of the effects on outcome "graduation" are, indeed, on-time graduation. For a discussion on the reliability of this measure as a proxy for EMS graduation, see (Dustan et al. 2017), (Avitabile and de Hoyos 2018).

Table 2: Baseline Characteristics by Treatment Status

Variable	Treatment		Control		T=C	N
	Mean	SD	Mean	SD	p-value	
Panel A: Household and location						
Mud floors	0.04	0.19	0.04	0.19	0.86	7,320
Refrigerator	0.91	0.29	0.91	0.29	0.79	7,321
Piped water	0.92	0.27	0.93	0.26	0.54	7,321
Piped water inside	0.82	0.38	0.82	0.38	0.99	7,321
Toilet exclusive for the household	0.82	0.39	0.83	0.37	0.10	7,321
Piped water toilet	0.77	0.42	0.77	0.42	0.79	7,321
Electricity	0.98	0.15	0.97	0.17	0.20	7,321
Blender	0.87	0.34	0.87	0.33	0.60	7,321
Gas stove	0.93	0.25	0.93	0.25	0.79	7,321
Radio	0.57	0.49	0.56	0.50	0.49	7,321
Time to school (hours)	0.17	0.81	0.17	0.81	0.97	7,321
Expenditure transportation to school	54.69	63.63	53.19	61.97	0.31	7,321
Urban	0.46	0.50	0.46	0.50	0.99	7,321
Panel B: Other characteristics						
Income	634.79	496.06	631.09	490.81	0.75	7,321
Age	15.26	0.44	15.27	0.44	0.39	7,220
Male	0.55	0.50	0.56	0.50	0.40	7,220
Spanish score ENLACE 2009	0.01	1.01	-0.01	0.99	0.30	7,321
Math score ENLACE 2009	0.01	1.01	-0.01	0.99	0.50	7,321

Note: p-value derive from a OLS regression on the treatment variable

3.2 Empirical Strategy

To estimate the causal impact of providing a PROBEMS scholarship to first-year EMS students on education outcomes, we estimate the following equation:

$$Y_i = \beta_0 + \beta_1 D_i + \gamma' X_i + u_i \tag{1}$$

where Y_i is either the presence of student i in ENLACE 12th grade or his or her test score recorded in ENLACE. D_i is an indicator dummy that takes the value 1 if student i is assigned to the treatment group, 0 otherwise. β_1 measures the intention-to-treat effect of the scholarship on education outcomes. X_i is a vector of baseline covariates measured at the individual level and includes age and gender of the student, math and Spanish test scores in 9th grade ENLACE, a dummy controlling for rural areas of the school where the student attended 9th grade, regional dummies and self-declared household income and the availability of household assets.

We standardize all test scores using the mean and the standard deviation observed in the control group. In order to address the inference issues related to the presence of multiple learning outcomes, we consider the effect on a composite score, defined by the simple average of the standardized scores in math and Spanish. When we consider multiple separate hypotheses, we compute stepdown p-values that correct for multiple hypothesis testing based on the method by (Romano and Wolf 2005).

4 Results

4.1 Education Outcomes

The main results of the paper are summarized in Table 3. We present the effects of PROBEMS on four education outcomes: EMS graduation–proxied by students present in the standardized test ENLACE 12th grade–math test scores, Spanish test scores, and a simple average of math and Spanish scores. We ran two specifications for each of these four outcomes, one without controls and a second one including the controls as defined above. Regardless of the specification, we do not find statistically significant effects of the

PROBEMS scholarships on any of the four education outcomes. All the effects are very small and not statistically different from zero, even though our sample is large enough to detect small impacts.

Table 3: Impact of PROBEMS on EMS Education Outcomes

Outcome variable	ENLACE (Y/N)		Math		Spanish		Average	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.012	0.008	0.006	-0.016	0.018	-0.004	0.012	-0.010
	(0.011)	(0.011)	(0.029)	(0.023)	(0.028)	(0.023)	(0.026)	(0.026)
(RW p-value)	0.68	0.90	0.91	0.90	0.91	0.90	0.91	0.90
Controls		Yes		Yes		Yes		Yes
N	7,321	7,220	5,050	4,988	5,050	4,988	5,050	4,988
Mean Dep. Control group	0.68	0.69	0.07	0.07	-0.03	-0.03	0.02	0.02
SD Dep. Control group	0.47	0.46	1.00	1.00	0.98	0.99	0.90	0.90

Romano-Wolf p-values (RW). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

RW p-values for two groups of coefficients, columns (1), (3), (5) and (7) for estimations with no controls and columns (2), (4), (6) and (8) for estimations including controls The full set of controls include age, dummies for sex and area (urban-rural) and test scores on Enlace 9, one for spanish and other one for mathematics. ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise. Math and Spanish refer to the ENLACE score in 2012. Average refers the average of spanish and math scores in 2012.

The large sample of our evaluation allows for the analysis of heterogeneous effects of the scholarships. It could be that the overall effect is zero but some particular groups did experience an improvement in education outcomes as a result of the cash transfer.[10]Tables A1 and A2 in Appendix A show the effects of the PROBEMS scholarship on graduation rates by age, rural versus urban, gender and geographical region. The scholarships did not increase either girls' or boys' probability of graduating. It also did not have any impact among students enrolled in rural or urban lower secondary schools at the time of applying to the scholarship. The only effect that is significant (p-value=0.06) is among slightly older candidates, those who were 16 years old at the time of applying for the scholarship. Among this population subgroup–who have a significantly lower probability of graduating compared to younger students–PROBEMS is increasing the probability of graduation by 4 percentage points, on a mean graduation rate of 62%. Table A2 also in Appendix A,

[10]By using stepdown p-values adjusted for multiple testing, we avoid the pitfalls of data mining that could lead us to false positives from such heterogeneity analysis.

presents the effects by geographical region showing zero impact of the scholarship in the 5 regions defined. Although not presented here, the lack of statistically significant effects on graduation rates by age, area, sex and region hold for math and Spanish test scores.

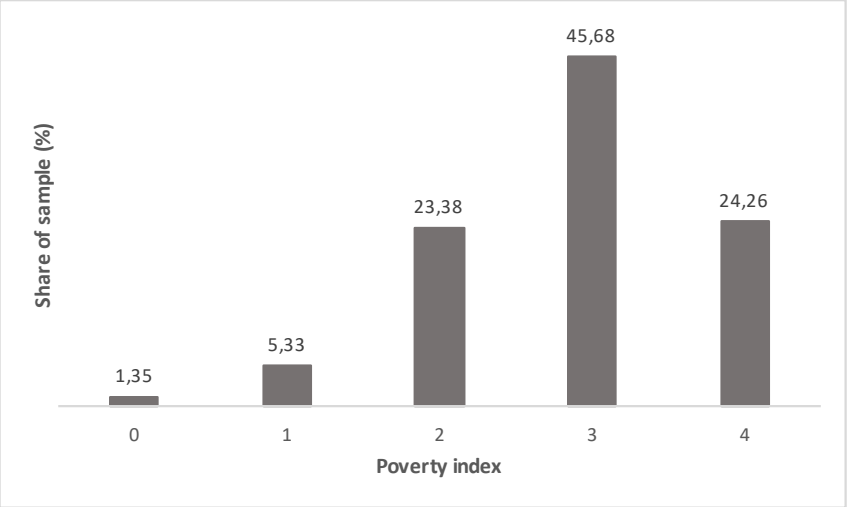
5 Why are PROBEMS scholarships ineffective?

It is possible that the lack of effects of the scholarships on education outcomes is the result of poor targeting, poor students' motivation, weak academic conditions among scholarship recipients or a combination of these. In this section, we perform further analysis to identify the effects of these three potential explanations behind the lack of impact of the scholarship program.

5.1 Targeting

To test if poor targeting explains the lack of effects of the scholarships, we rely on information from the PROBEMS application portal to compute an index of self-declared household assets. The index of household assets (IHA) is the sum of four variables, indicating, respectively, the following basic characteristics of the dwelling: (1) concrete floors, (2) running water inside the house, (3) toilets connected to sewerage, and (4) concrete roof. A zero is used to indicate the absence of the amenity and a one its presence. The IHA index, therefore, has a maximum value of 4 and minimum of 0. The distribution of the IHA is shown in Figure 4. Few students self-declared that they live in a household with zero or only one of the basic dwelling characteristics; close to 70% of students have only two or three of the basic needs met, and 23% of students in our sample declared living in a household that meets all four basic needs.

Figure 4: Share of Students in the Different Categories of the IHA



To test for heterogenous impacts by socio-economic status the evaluation sample was divided in two groups according the value of the IHA: those with a value of the IHA of 2 or less and the rest. Table 4 shows the effects of PROBEMS on graduation rates, by socioeconomic status in a specification with and without controls. The effects of PROBEMS are zero, even within the group of more marginalized individuals (IHA of 2 or below). As a complementary exercise we estimate the effects of PROBEMS among individuals with a value of the IHA of 0 or 1 (the most marginalized) and the results are shown in Table A6 in the Appendix. For students with a value of the IHA of zero or one, in the specification without controls, being awarded a scholarship increases the probability of graduation by 8 percentage points (p-value of 0.06), on a mean value of 60% among the control group. However, when the controls are included in the specification, the positive effect remains but is no longer statistically significant (p-value of 0.104). The lack of statistically significant effects among the group of poorest students could be explained by the relatively small number of students with a value of the IHA of 0 or 1 (483). We also test for heterogenous effects on test scores by different levels of IHA, splitting the sample as indicated in Table 4 and found very small and statistically insignificant effects.[11]

To identify how well targeted PROBEMS was, we use the Mexican national household

[11]The results are available from the authors upon request.

Table 4: Impact of PROBEMS on EMS Graduation by Levels of Household Assets

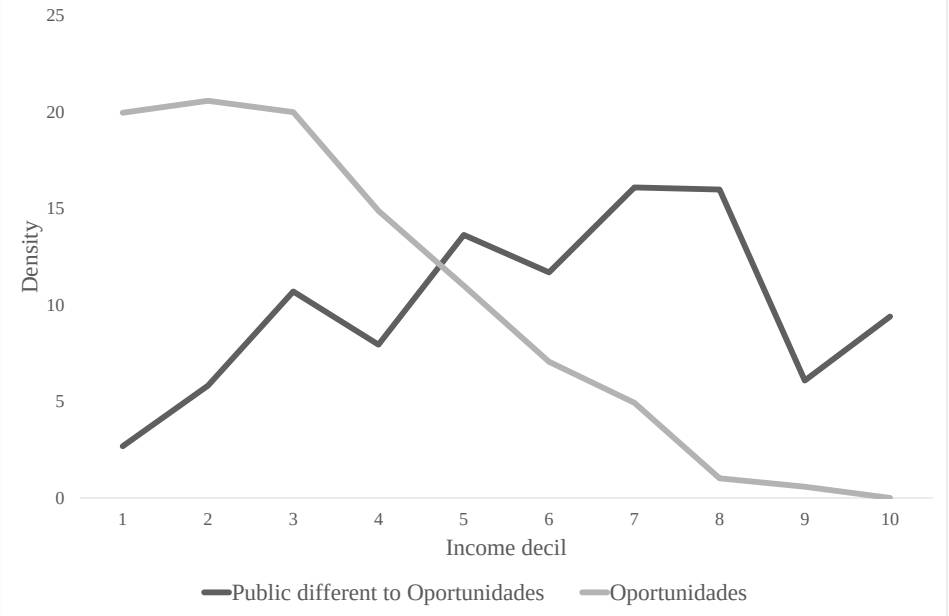
Outcome variable	ENLACE (Y/N)			
	Household index (0, 1 & 2)		Household index (3 & 4)	
	(1)	(2)	(3)	(4)
Treatment	0.006	0.005	0.015	0.010
	(0.020)	(0.020)	(0.013)	(0.013)
Controls		Yes		Yes
N	2,201	2,176	5,120	5,044
Mean Dep. Control group	0.67	0.68	0.69	0.69
SD Dep. Control group	0.47	0.47	0.46	0.46
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$				

The full set of controls include age, dummies for sex and area (urban-rural) and test scores on Enlace 9, one for spanish and other one for mathematics. ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise. The houseld index takes a value from 0 to 4 depending if the student's household has cement roof, cement floor, piped water and piped water toilet.

survey ENIGH, which measures incomes and expenditures, to identify, self-declared, scholarship holders and divide them between recipients of "Oportunidades" and "other government program". Using ENIGH 2010[12] we identify households with students enrolled in EMS and, among this subsample, those that declared having an "Oportunidades" scholarship or a scholarship from another government program. Figure 5 shows the share of scholarship beneficiaries by decile of the distribution of per capita household income. Overall, "Oportunidades" scholarships seem to be well-targeted, benefiting the poorest EMS students. However, beneficiaries of "other government programs", a category basically capturing PROBEMS, tend to be located in deciles 5 to 9 of the income distribution. The lack of targeting combined with the marginally positive effects among the most marginalized suggest that PROBEMS could increase the likelihood of finishing high school among the poorest students, but the program's targeting mechanism was not very effective in reaching the poorest, or they were already covered by the conditional cash transfer program (Oportunidades).

[12]ENIGH is collected every two years and there was no survey in 2009, the year when PROBEMS applicants filled the online format with the socio-economic information.

Figure 5: Income distribution (Oportunidades vs. Public different to Oportunidades)



5.2 Motivation

In this section we test the role played by motivation (or lack thereof) in explaining the ineffectiveness of the PROBEMS scholarships. A random sample of around 10% of the ENLACE takers were administered a survey, the so called *ENLACE de contexto* that elicits a broad set of information about student sociodemographic characteristics, such as student effort, family composition and aspirations. We merge the information of the *ENLACE de contexto* applied to students in our evaluation sample when they sat for the ENLACE 9th grade in 2009 to test a student motivation hypothesis.

The merged sample includes only 564 individuals, a sample substantially smaller than our original one but still well-balanced between treatment and control and showing the same lack of treatment effects as with the entire sample. Students' motivation is proxied by the, self-declared, highest degree they expected to achieve, the options being: lower secondary, EMS, technical higher education, university or postgraduate. Of the students in our sample, 80% expected to finish, at least, a university degree and 47% expected

Table 5: Heterogenous effect by motivation

	Enlace (Y/N)	Math	Spanish	Enlace (Y/N)	Math	Spanish
	(1)	(2)	(3)	(4)	(5)	(6)
Motivated (M)	0.127**	0.345**	0.331**			
	(0.054)	(0.134)	(0.141)			
Highly Motivated (HM)				0.210***	0.530***	0.346*
				(0.065)	(0.180)	(0.192)
T x Motivation	-0.019	0.045	-0.035	0.001	0.063	-0.058
	(0.075)	(0.185)	(0.196)	(0.093)	(0.257)	(0.274)
Treatment (T)	0.013	0.141	0.025	-0.006	0.084	0.037
	(0.052)	(0.133)	(0.141)	(0.083)	(0.237)	(0.253)
(RW p-value Treatment)	0.96	0.63	0.96	0.98	0.97	0.98
Observations	564	410	410	564	410	410
Romano-Wolf p-values (RW) *** p<0.01, ** p<0.05, * p<0.1						

RW p-values for two groups of coefficients, columns (1), (2), (3) and columns (4), (5), (6).

to get a postgraduate degree. We classify these two groups as "motivated students" and "highly motivated students", respectively.

Table 5 shows the results of a specification including, separately, our two definitions of motivation on EMS graduation and test scores. As expected, motivated and highly motivated students have a higher probability of graduating and better test scores at the end of EMS. However, the treatment effect of PROBEMS on graduation and test scores remains zero in these new specifications and the interaction between the treatment and motivation is not statistically significant. In other words, even among motivated and highly motivated students, the scholarship is not relevant for increasing their likelihood of finishing EMS or obtaining better grades. The results remain when we use the self-declared hours spent doing homework (from the *ENLACE de contexto*) as an alternative proxy for "motivated students".

5.3 Academic *Readiness*

We now consider the role of school readiness and whether this can explain the lack of impact. Academic *readiness* is measured by the proficiency level at the end of lower sec-

ondary (grade 9). We use, separately, math and Spanish results of ENLACE 9th grade of 2009–the year when the candidates applied for a scholarship–to divide our sample in terciles of the distribution of test scores. We ran two specifications within each of the terciles, with and without controls, to measure the effects of the scholarship on the probability of graduating and test scores at the end of EMS. The effects of PROBEMS on the probability of graduation, by tercile of the 9th grade math distribution, are summarized in Table 6.

The results show positive and statistically significant effects (p-value of 0.02) of the scholarship on the probability of graduation among those candidates in the top tercile of the 9th grade math test scores distribution (i.e. those with academic *readiness*). The inclusion of controls does not change the results. For candidates with academic *readiness*, receiving a scholarship increases their probability of graduation by 4 percentage points (of a mean level of 76% among the control group). We found no effects within the lowest and middle terciles of the 9th grade test scores distribution. There are also no effects of the scholarships on learning outcomes by tercile of 9th grade math or Spanish test scores.[13] Candidates with academic *readiness*, those at the top tercile of the 9th grade math test scores, have the largest probability of finishing EMS (76%), as oppose to those in the bottom tercile (57%). It seems that the scholarship is only effective as long as the candidates have academic *readiness*, otherwise the requirements of EMS are such that a scholarship is irrelevant to address this constraint. This is consistent with other school interventions that seem to build on prior success (Machin, McNally, and Meghir 2010).

In Appendix Table A5, we cross academic readiness with socio-economic status measured by the IHA. We found that the positive effects of PROBEMS on EMS graduation rates among students with academic *readiness* hold both for poor and non-poor households, but are statistically significant only for non-poor households. The lack of statistically significant effects among poor students that show academic *readiness* could be explained by the relatively low number of poor students in the upper tercile of the 9th grade test score distribution.

[13] Results available upon request.

Table 6: Impact on high school graduation by performance in ENLACE 2009 (math)

Outcome variable	ENLACE (Y/N)					
	Lowest tercile		Middle tercile		Highest tercile	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.004	-0.003	-0.004	-0.005	0.040**	0.037**
	(0.020)	(0.020)	(0.018)	(0.018)	(0.017)	(0.017)
Controls		Yes		Yes		Yes
N	2,486	2,452	2,449	2,418	2,386	2,350
Mean Dep. Control group	0.57	0.58	0.72	0.72	0.76	0.76
SD Dep. Control group	0.49	0.49	0.45	0.45	0.43	0.43
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$						

The full set of controls include age, dummies for sex and area (urban-rural) and test scores on Enlace 9, one for spanish and other one for mathematics. ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise. The three groups are formed using the terciles of performance in ENLACE math 2009.

5.4 Interpretation

The results presented here are consistent with other evidence on the effects of scholarships and conditional cash transfers in EMS education outcomes in Mexico. The most recent evidence identifies the importance of targeting to make scholarships effective. For instance, (Parker and Vogl 2018) rely on a difference-in-difference estimator to show that young children exposed to Progresa-Oportunidades, a well-targeted conditional cash transfer program, increase the likelihood of graduating from EMS. Using administrative data and a difference-in-difference identification strategy, (Araujo et al. 2018) also show a positive and significant effect on EMS graduation associated with the expansion of Oportunidades into urban areas. However, when the cash transfer is not well-targeted it has no effects on EMS graduation rates, as shown by (Dustan 2018) evaluation of "Prepa Si" a universal scholarship program for EMS students in Mexico City. (Dustan 2018) also finds that "Prepa Si" had no impact on test scores measured by the end of upper secondary ENLACE.

Increasing student motivation by providing monetary incentives linked to math performance or information on the returns to high school does not increase EMS graduation rates, as shown in (Behrman et al. 2015) and (Avitabile and de Hoyos 2018), respectively. These results are in line with those presented here, suggesting that the lack of effects of PROBEMS scholarships in Mexico is not explained by students' low motivation, the scholarships are ineffective even among students with higher motivation. However, our results show that the program's weak targeting mechanism can partly account for its ineffectiveness, which is consistent with (Dustan 2018), (Parker and Vogl 2018) and (Araujo et al. 2018). The most important explanation behind PROBEMS' lack of effects is the large share of first-year EMS students without academic *readiness*. For 2 of 3 eligible candidates for a scholarship, income is not the binding constraint and therefore the monetary transfer will not do much to improve their likelihood of graduating from upper secondary. For the majority of first-year EMS students, the real problem lies in the profound academic deficiencies that students carry over from basic education.

6 Conclusions and Policy Recommendations

In this paper, we present evidence of the impact of a scholarship program aimed at poor households with the intention of improving upper secondary school graduation rates and test scores. The results we obtain show that, by and large, the program was ineffective.

We explore this negative result in detail, to understand the reasons behind it. We show that the program was not effectively targeted, as students from poor households were a minority among its beneficiaries. We also show that the ineffectiveness of the program is not explained by the lack of students' motivation: the program has no impact both among students with high and low motivation. Finally, we show that the only positive impact we find, even after correcting for possible inference biases due to multiple hypothesis testing, is among students with academic *readiness*, that is, those students who finish lower secondary school with relatively good grades.

These results are important not only because they point out the ineffectiveness of PROBEMS,

an important and expensive program, but also because they indicate possible solutions to improve the educational outcomes of students from poor families in Mexico. It seems that programs such as the one we evaluated have to be complemented by interventions that improve the quality of education in lower secondary school and, probably, even earlier in pre-primary and primary school. Only then one can hope to improve the educational outcomes of poor students by improving graduation and learning at the upper secondary level.

These results are consistent with other results for Mexico and other low- and middle-income countries. It has been shown, for instance, that conditional cash transfers improve enrollment past primary school (and in particular in lower secondary) (see, for instance (Todd and Wolpin 2006) and (Attanasio, Meghir, and Santiago 2012)). But it is not clear that the quality of the education these students receive has improved.

Our results also show that education policy and education interventions should be seen as a whole, with interventions aimed at both increasing the demand but also the quality of the supply offered. The formation of skills and human capital, which is obviously important to reduce inequality and the intergenerational transmission of poverty, should be seen as a dynamic process that starts very early in the life cycle, possibly even before children start formal education. Interventions aimed at secondary education should be complemented and preceded by interventions aimed at improving the ability of poor students to exploit education opportunity during adolescence.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. Violante (2019). Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy Forthcoming*.

Araujo, M. C., S. Martinez, M. A. Martínez, M. Pérez, and M. Sánchez (2018). ¿se mejora la escolaridad con becas de mayor monto?: La evidencia de las localidades urbanas de méxico. *IDB Publications (Working Papers) 8777, Inter-American Development Bank*.

Attanasio, O. P., C. Meghir, and A. Santiago (2012). Education choices in mexico: Using a structural model and a randomized experiment to evaluate progresa. *Review of Economic Studies, Oxford University Press 79*(1), 37–66.

Avitabile, C. and R. de Hoyos (2018). The heterogeneous effect of information on student performance: Evidence from a randomized control trial in mexico. *Journal of Development Economics, Elsevier 135*(C), 318–348.

Behrman, J. R., S. W. Parker, P. E. Todd, and K. I. Wolpin (2015). Aligning learning incentives of students and teachers: Results from a social experiment in mexican high schools. *Journal of Political Economy 123*(2), 325–364.

Belley, P. and L. Lochner (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human Capital* (1), 37–89.

de Hoyos, R., H. Rogers, and M. Székely (2016). Out of school and out of work : Risk and opportunities for latin america's ninis. Technical report, World Bank, Washington, DC.

Dustan, A. (2018). Can large, untargeted scholarships increase urban high school graduation rates? evidence from mexico city's prepa sí. Technical report, Department of Economics, Vanderbilt University.

Dustan, A., A. de Janvry, and E. Sadoulet (2017). Flourish or fail?: The risky reward of elite high school admission in mexico city. *Journal of Human Resources, University of Wisconsin Press 52*(3), 756–799.

Hanushek, E. and L. Woessmann (2012, December). Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of Economic Growth, Springer 17*(4), 267–321.

INEE (2017a). Panorama educativo de méxico. indicadores del sistema educativo nacional 2017. educación básica y media sperior. Technical report, Instituto Nacional de Evaluación para la Educación, Ciudad de México, México.

INEE (2017b). Resultados planea en educación media superior. Technical report, Instituto Nacional de Evaluación para la Educación, Ciudad de México, México.

Machin, S., S. McNally, and C. Meghir (2010). Resources and standards in urban schools. *Journal of Human Capital 4*, 365–393.

Parker, S. W. and T. Vogl (2018). Do conditional cash transfers improve economic outcomes in the next generation? evidence from mexico,. *NBER Working Papers 24303, National Bureau of Economic Research, Inc.*.

Romano, J. P. and M. Wolf (2005). Stepwise multiple testing as formalized data snooping. *Econometrica 73*(4), 1237–1282.

SEP (2008). Reforma integral de la educación media superior. Technical report, Secretaría de Educación Pública, Ciudad de México, México.

SEP (2012). Encuesta nacional de deserción en la educación media superior. Technical report, Secretaría de Educación Pública, Ciudad de México, México.

Todd, P. E. and K. I. Wolpin (2006). Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a behavioral model of child schooling and fertility. *American Economic Review 96*(5), 1384–1417.

WorldBank (2018). World development report (wdr) 2018: Learning to realize education's promise. Technical report, World Bank Group, Washington, D.C.

A Additional Results

Table A1: Heterogenous Impact of PROBEMS on EMS Graduation

Outcome variable	ENLACE (Y/N)					
Group variable	Age		Area		Sex	
	15 years	16 years	Rural	Urban	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.001	0.041*	0.017	0.007	-0.000	0.019
	(0.012)	(0.022)	(0.015)	(0.015)	(0.017)	(0.014)
N	5,308	1,912	3,962	3,359	3,213	4,007
Mean Dep. Control group	0.71	0.62	0.66	0.72	0.68	0.69
SD Dep. Control group	0.45	0.48	0.48	0.45	0.47	0.46
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$						

ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise.

Table A2: Impact of PROBEMS on EMS Graduation, by Region

Outcome variable	ENLACE (Y/N)				
	Northwest	Northeast	West	Center	Southeast
	(1)	(2)	(3)	(4)	(5)
Treatment	0.002	0.024	0.030	-0.008	0.004
	(0.028)	(0.020)	(0.024)	(0.026)	(0.024)
N	1,096	2,077	1,586	1,291	1,271
Mean Dep. Control group	0.70	0.69	0.61	0.68	0.75
SD Dep. Control group	0.46	0.46	0.49	0.47	0.43
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$					

ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise. Northwest includes Baja California, Baja California Sur, Chihuahua, Sinaloa and Sonora; Northeast includes Coahuila, Durango, Nuevo León, San Luís Potosí and Tamaulipas; West includes Aguascalientes, Colima, Guanajuato, Jalisco, Michoacán, Nayarit, Queretaro and Zacatecas; Center includes Ciudad de México, Guerrero, Hidalgo, México, Morelos, Puebla and Tlaxcala; Southeast includes Campeche, Chiapas, Oaxaca, Quinatan Roo, Tabasco, Veracruz and Yucatán.

Table A3: Impact on high school graduation by performance in ENLACE 2009 (Spanish)

Outcome variable	ENLACE (Y/N)					
	Lowest tercile		Middle tercile		Highest tercile	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.002	-0.007	0.016	0.019	0.021	0.021
	(0.019)	(0.019)	(0.019)	(0.019)	(0.017)	(0.017)
Controls		Yes		Yes		Yes
N	2,605	2,565	2,335	2,307	2,381	2,348
Mean Dep. Control group	0.58	0.59	0.70	0.70	0.78	0.78
SD Dep. Control group	0.49	0.49	0.46	0.46	0.42	0.42
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$						

The full set of controls include age, dummies for sex and area (urban-rural) and test scores on Enlace 9, one for spanish and other one for mathematics. ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise. The three groups are armed using the terciles of performance in ENLACE spanish 2009.

Table A4: Impact on high school graduation by performance in ENLACE 2009 (average)

Outcome variable	ENLACE (Y/N)					
	Lowest tercile		Middle tercile		Highest tercile	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.015	0.007	-0.012	-0.011	0.037**	0.036**
	(0.020)	(0.020)	(0.019)	(0.019)	(0.017)	(0.017)
Controls		Yes		Yes		Yes
N	2,564	2,522	2,369	2,349	2,388	2,349
Mean Dep. Control group	0.56	0.57	0.72	0.72	0.77	0.77
SD Dep. Control group	0.50	0.50	0.45	0.45	0.42	0.42
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$						

The full set of controls include age, dummies for sex and area (urban-rural) and test scores on Enlace 9, one for spanish and other one for mathematics. ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise. The three groups are armed using the terciles of performance in ENLACE average 2009.

Table A5: Impact by academic readiness with socio-economic status

Outcome variable	ENLACE (Y/N)					
	Non-poor households			Poor households		
	(1) Highest tercile	(2) Middle tercile	(3) Lowest tercile	(4) Highest tercile	(5) Middle tercile	(6) Lowest tercile
Treatment	0.036**	-0.006	-0.004	0.135	0.027	0.088
	(0.017)	(0.019)	(0.021)	(0.086)	(0.069)	(0.069)
N	2,268	2,284	2,280	118	165	206
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$						

ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise. The three groups are formed using the terciles of performance in ENLACE math 2009. Poor households imply IHA of 0 or 1, Non-por households imple IHA of 2 or more.

Table A6: Impact of PROBEMS on EMS Graduation by Levels of Household Assets

Outcome variable	ENLACE (Y/N)			
	Household index (0 & 1)		Household index (2,3 & 4)	
	(1)	(2)	(3)	(4)
Treatment	0.081*	0.069	0.008	0.004
	(0.043)	(0.043)	(0.011)	(0.011)
Controls		Yes		Yes
N	489	483	6,832	6,737
Mean Dep. Control group	0.60	0.61	0.69	0.69
SD Dep. Control group	0.49	0.49	0.46	0.46
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$				

The full set of controls include age, dummies for sex and area (urban-rural) and test scores on Enlace 9, one for spanish and other one for mathematics. ENLACE (Y/N) takes the value of 1 if the student took the 12th grade exam in 2012 and 0 otherwise. The household index takes a value from 0 to 4 depending if the student's household has cement roof, cement floor, piped water and piped water toilet.