Policy Research Working Paper 9000

Impacts of PROSPERA on Enrollment, School Trajectories, and Learning

Jere R. Behrman
Susan W. Parker
Petra Todd

WORLD BANK GROUP
Social Protection and Jobs Global Practice
September 2019

Policy Research Working Paper 9000

Abstract

Many studies have demonstrated that Mexico's conditional cash transfer program, PROSPERA, has substantial effects on educational attainment. Nevertheless, little evidence exists on whether increases in time spent in school have led to higher learning in the context of the poor areas where PROSPERA principally operates, which tend to have overall low school quality. This study combines data from nationwide achievement tests with administrative data on PROSPERA beneficiaries to estimate impacts on achievement tests. The analysis finds significant effects on learning, as measured by standardized achievement tests, on the order of magnitude of 0.05 to standard deviation, with larger effects for indigenous children. The analysis also confirms large effects on enrollment in secondary and high school, using administrative school enrollment data rather than self-reported household-level data, as generally used in previous studies. Finally, given the existence of several alternative tracks in secondary and high school, the study also examines where PROSPERA beneficiaries enroll. The findings show that most of the increase in enrollment occurs in tele-secondary schools and, at the high school level, in general high schools.

This paper is a product of the Social Protection and Jobs 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 atswparker@umd.edu and the coordinators of the studies may be contacted at cavilaparra@worldbank.org.

Impacts of PROSPERA on Enrollment, School Trajectories, and Learning[1]

Jere R. Behrman[*]
Susan W. Parker[†]
Petra Todd[‡]

Keywords: Achievement, CCTs, Evaluation, Mexico
JEL: I25, I28

[1] This paper is part of the research project "Studies of PROSPERA's long-term results" which was funded by the Government of Mexico through the National Coordination of PROSPERA (CNP) as part of a technical cooperation agreement with the World Bank. The World Bank team is grateful to the Government of Mexico, including the Ministry of Social Development (now Ministry of Welfare), Ministry of Education, and Ministry of Health for the close collaboration in this research project, including the provision of the main datasets. The World Bank team would also like to thank the National Council for the Evaluation of Social Development Policy (CONEVAL), the National Institute of Public Health, development partners such as the Inter-American Development Bank, and the national and international researchers that contributed in the Advisory Evaluation Group. This group was established to ensure technical quality and methodological rigor of the "Studies of PROSPERA's long-term results." These studies were coordinated by Maria Concepción Steta, WB Senior Social Protection Specialist; Clemente Ávila, WB Social Protection Economist; and Mónica Orozco, WB Consultant.

[*] University of Pennsylvania
[†] University of Maryland
[‡] University of Pennsylvania

I. Introduction

Conditional cash transfer (CCT) programs were first introduced in Brazil and Mexico more than two decades ago and have since spread around the world. These programs aim to alleviate current poverty and, in addition, reduce future poverty by augmenting human capital of children and youth from poor families to increase their lifetime earnings potential. The Mexican program PROSPERA (previously named PROGRESA and Oportunidades) began in 1997 and, in its first two years, was rigorously evaluated using a randomized design. The still ongoing program conditions transfers to poor families on children's school attendance and family visits to health clinics.

In a time when experimental evaluations of social policy in developing countries were rare, a large-scale randomized evaluation was carried out that demonstrated substantial program effects on human-capital accumulation and poverty alleviation. The program's novelty and the finding of substantial positive program impacts contributed to both a large scaling-up within Mexico and an impressive spreading of the program's key features to new programs around the world. Parker and Todd (2017) review the literature on the development of the program, the evaluation of its impacts, and the key findings.

Many previous evaluation studies examined impacts on school enrollment and, in some cases, on longer-term educational attainment. Impacts of PROSPERA on the education levels of its beneficiaries are studied in Schultz (2000, 2004), Behrman, Sengupta and Todd (2005), Behrman, Parker and Todd (2006), Todd and Wolpin (2007), and Attanasio, Meghir and Santiago (2012). These papers generally find positive program impacts on school enrollment and educational attainment.

An understudied evaluation topic, however, has been the effects of the program on learning and, in particular, whether increased school enrollment translates into greater academic achievement. School quality tends to be poor in the isolated, rural areas of Mexico we study (OECD, 2013; Mancera et al. 2007). If the schools that the PROSPERA beneficiaries are attending are of low quality, then it is conceivable that children could be enrolling in school in increasing numbers with little benefit in terms of learning and academic achievement. Previous studies of program effects on standardized achievement tests have suggested limited impacts on learning despite increased enrollment (Behrman, Parker and Todd, 2009; Baird et al, 2018), which is consistent with low returns to education. Additionally, much of the beneficiary population in rural areas is engaged in agriculture, where returns to schooling may be lower.[2]

Thus, despite numerous evaluation studies, little is known about the quantitative impacts of the PROSPERA program on learning and achievement. One challenge in assessing test score impacts is that standardized tests are typically only given to children enrolled in school and

[2] Additionally, even if scores on achievement tests do not show increases with the program, increases in education may still lead to better outcomes in the labor market through other educational effects of the program not captured by standardized achievement tests. Increased education may also impact other variables such as early marriage or childbearing.

PROSPERA brings children into school who might have otherwise stayed home or engaged in work. These children may be more likely to be in the lower part of the test score distribution. If selective school attendance is not taken into account, then estimated program impacts on test scores are likely to be downward biased.

A related caveat is that little is known about how impacts might vary by available school quality. Maintaining and/or improving the supply of schools and their quality are responsibilities of the Ministry of Public Education in Mexico. However, PROSPERA impacts in the area of education are intimately intertwined with school quality. First, schooling quality may, in fact, have deteriorated for schools where PROSPERA beneficiaries attend due to crowding induced by increased enrollment. Second, the extent of benefits accruing to children from attending school may depend on school quality. Third, changes in school quality may affect decisions about which schools to attend, which is more likely to be the case in urban areas and at higher school grade levels where households can choose among multiple schooling options.[3]

This study analyzes nationally applied achievement test data that have recentlybeen linked to administrative data on beneficiary receipt, socio-economic characteristics of children taking the tests and school quality characteristics to study how the PROSPERA program affects school enrollment, type of school attended and academic achievement. Nationwide standardized administrative test score data from schools (ENLACE National Evaluation of Academic Achievement in Schools "Evaluacion Nacional de Logro Academico en Centros Escolares") were merged with information on which students come from PROSPERA households and with contextual information on family background obtained from a school survey. These data allow study of how students from beneficiary and non-beneficiary households differ in terms of their enrollment decisions, their grade progression and in terms of learning and academic achievement, as measured on standardized tests in math and Spanish.

As previously noted, a major challenge in assessing program effects on learning and academic achievement is the problem of selection. In primary school (up through grade 6), attendance at school is nearly universal, but in subsequent grades students or their families decide both whether and where to attend school and what type of school to attend. There is also selection into test-taking, as the achievement tests available are only administered to students who are enrolled in school. Simple comparisons of mean test scores of students from beneficiary and non-beneficiary households might show that students from beneficiary households have lower scores, simply because weaker students are induced by the program to enroll in school and thus take the standardized achievement tests. In making comparisons between children from beneficiary and non-beneficiary households, one needs to account for how the program affects the composition of students present in school. The selection problem is dynamic as it occurs at each grade level

[3] We do not attempt a comprehensive definition of school quality but rather are limited to available information in our data sources. By school quality in this study we refer to school level characteristics which reflect available school resources including infrastructure, supplies and teacher characteristics and, given high inequalities in these variables nationwide, are presumably indicative of differences in resource allocation.

and the students at risk for dropping out in a particular grade are those who stayed in school from the previous grade.

In this study, we analyze nationwide data on standardized tests in math and Spanish combined with administrative data on beneficiary receipt, school characteristics and on student and parent characteristics to study program effects on 1) enrollment, 2) type of school attended and 3) academic achievement. We construct a panel of students in sixth grade enrolled in primary schools in high and very high poverty areas in 2008, whom we observe in 2009, 2011 and 2014. Our estimates derive both from fixed-effects regression models and propensity score matching models. We study whether impacts are heterogeneous by gender and by indigenous status. We include a rich set of control variables at the individual, household, school and locality level. We study impacts of PROSPERA in medium, poor and very poor areas, as defined by margination indexes developed by Conapo (National Mexican Population Council "Consejo Nacional de Poblacion"), where a majority of beneficiaries reside.

We begin by describing the program and our research hypotheses, then turn to previous literature, a description of our data and methodology, followed by results and conclusions.

II. Background and Hypotheses

PROSPERA began operating in small rural communities in 1997, following a macroeconomic crisis in Mexico in 1995, and was part of a transition towards implementing targeted anti-poverty programs and eliminating general food subsidies. It quickly grew over time and currently covers six million families, or about one-quarter of all families in Mexico. Although the program has expanded into urban areas, it remains largely rural, with about two-thirds of its household beneficiaries deriving from communities with fewer than 2,500 inhabitants.

The program conditions cash payments to families on children regularly attending schools and on family members visiting health clinics for checkups. Program take-up was exceedingly high when the program began, with 97 percent of families who were offered the program participating (Skoufias and Parker, 2001). Program rules allow students to fail each grade once, but if a student repeats a grade twice, the schooling benefits are discontinued permanently. The program also provides some additional subsidies for school supplies and a transfer linked to regular visits to health clinics. Children and youth age 21 and younger are eligible to receive the school subsidies.

Table 1 shows the monthly grant levels for children between the third grade and the twelfth grade in the second semesters of 1997 and 2003 (when the exchange rate was about 8 and 11 pesos per U.S. dollar, respectively). Originally, the program provided grants only for children between the third and ninth grades, but in 2001, the grants were extended to grades 10-12. At grades seven and above, the grants are slightly higher (by about 13 percent) for girls than boys, a response to historically lower enrollment rates among girls than among boys after primary school.

Specific monthly grant amounts range in 2003 from $US9.50 ($105 pesos) in the third grade of primary to about $US53 ($580 pesos) for boys and $US60 ($660 pesos) for girls in the third year of senior high school (grades 10-12). By the senior year of high school, the grant amount represents about two-thirds of Mexico's minimum wage. All monetary grants are given to the mother of the family, with the exception of scholarships for senior high school, which the youth can receive themselves subject to their mothers' authorization. The program is means tested, with both geographic and household-level targeting. The geographic targeting uses aggregate census indicators to select poor rural communities based on the marginality index.

Table 1: Monthly Amount of Schooling Grants, 1997 and 2003

	2nd semester 1997		2nd semester 2003	
	Boys	Girls	Boys	Girls
Primary school				
3rd year	60	60	105	105
4th year	70	70	120	120
5th year	90	90	155	155
6th year	120	120	210	210
Secondary school				
1st year	175	185	305	320
2nd year	185	205	320	355
3rd year	195	225	335	390
High school				
1st year			510	585
2nd year			545	625
3rd year			580	660
Max HH amount without high-schooler	550		950	
Max HH amount with high-schooler			1635	

Note: Amounts in nominal pesos. The peso-to-dollar exchange rate was exchange rate was roughly 8 in 1997 and 11 in 2003.
Source: www.prospera.gob.mx.

How might PROSPERA affect enrollment, school attended and achievement? In particular, conditional transfers should increase enrollment through both a price and an income effect. The price effect occurs because the subsidy reduces the shadow wage (or relative value) of children's time in activities other than school. There is also an income effect, because the subsidy increases a family's total potential income. Thus, both the price and income effect of the transfers work to increase a child's enrollment in school.

In addition to increasing the probability of enrolling in school, the program might also change the school in which a child enrolls. Part of the transfers might be spent on costs necessary to attend a better school, for instance in paying additional transportation costs for a school that is located further away from residence. The extent to which this is possible, of course, depends on school options within a reasonable distance. For example, if there is only one high school within 50 kilometers, it is unlikely that the transfers would affect the school attended. In this paper, we

study the observed school at which students enrolled and study whether the program affects the type of sub systems in which a child enrolls.[4] At the secondary level (seventh through ninth) there are several different types of secondary schools. The different types, broadly speaking, are 1) tele-secondary schools that use electronic media or television to provide education through distance learning and are generally located in more rural areas, 2) technical secondary schools that are vocational, 3) general secondary schools with traditional academic curriculums and 4) private schools. Similarly, at the high school level or medio superior the main options are 1) general high schools with traditional academic programs that prepare students to continue studying at university, 2) technical high schools that have academic focus but with technical and vocational training, 3) technical professional schools that aim to combine studies with work and 4) private schools.

Enrollment processes for public secondary and high schools are decentralized and carried out at the state level, with Federal regulation. Each state carries out its own enrollment process where each student generally takes an exam and fills out an application specifying the schools of his/her preference and a school is assigned to each student based on the student's performance, locality of residence and existence of siblings already enrolled at the school.

With regard to achievement, the central hypothesis is that the program, by increasing time spent in school, should lead to greater learning and thus better performance on standardized achievement tests. Such impacts might be limited however if increases in enrollment due to the program also lead to overcrowding in schools and thus reducing schooling quality. Further, schools with lower quality might have lower effects on learning than schools with higher quality. Finally, if PROSPERA affects the type of school chosen, for example by increasing the probability of enrolling in a higher quality school, this is also a possible mechanism through which the program might affect learning. We are unaware of impact studies that analyze the effect of the program on schooling supply and quality, but we consider this an important area for future research. A final consideration is the impact of PROSPERA on both parents and students with respect to learning. Parents might pressure students and/or students might feel pressured to attend school more regularly and to improve performance due to their beneficiary status, which might translate to increased effort on standardized testing.

III. Related Literature

Most quantitative impact studies of the effects of PROSPERA derive from the initial 1997 experimental evaluation in which 506 rural communities from seven Mexican states were randomized into a treatment and a control group. Of these communities, 320 were randomly assigned to receive benefits immediately and the other 186 to receive benefits later. Eligible households in

[4] An important question is the extent to which students are located close enough to schools such that there is in fact more than one plausible option for attending. Ideally such an analysis would construct a measure of available schools based on geographic distance and access to roads from the student's residence. We leave this possibility for future research.

the treatment localities began receiving program benefits in the spring of 1998, whereas eligible households in the control group began receiving benefits at the end of 1999, after which both groups continued to receive program benefits. Evaluation questionnaires (called ENCELs) were applied every 6 months over the time period 1997 to 2000, with a final follow up in 2003. The majority of existing studies focus on the initial experimental period, when the treatment group received benefits and the control group did not. Some follow-up studies post 2000 of this evaluation sample have been carried out, notably in 2003 and 2007. However, note that follow-up studies after 2000 of the original experimental design carry out differential exposure impact estimates comparing effectively two treatment groups, one which has about 2 additional years of benefits than the other.

Early evaluation studies using the experimental design and data (see Parker and Todd, 2017 for a description) demonstrated positive effects of PROSPERA in improving school enrollment, reducing grade repetition and increasing completed grades of schooling (Schultz, 2004; Behrman, Sengupta and Todd, 2005). Enrollment impacts are higher at the transition from primary to secondary school; however, younger children experience large reductions in grade repetition and better grade progression (Behrman, Sengupta and Todd, 2005). At the junior high school level (grades 7-9), the program reduces the dropout rate and also encourages re-entry among those who have dropped out. Both Schultz (2004) and Behrman, Sengupta and Todd (2005) use their short-term estimates to predict the effects on overall schooling attainment. Both studies predict an overall increase of 0.6 grade for a child receiving grants from primary school through junior high school (grades 3-9).

More recent studies have explored alternative methodologies and data to study longer-run program impacts. Behrman, Parker, Todd, (2009, 2011) find that extended time participating in the program leads to significant improvements in grades completed, about 1 full grade for children who participate in the program for 6 years beginning at ages 9 to 12, compared to non-participating children. Studies based on structural estimation also find program effects of a similar magnitude (Todd and Wolpin, 2006; Attanasio, Meghir and Santiago, 2012). Parker and Vogl (2018), using nationally representative Census data from 2010, find that long-term impacts on total schooling attainment for children who grow up in a household with PROSPERA benefits are 1.4 grades both for men and women, compared with those in households not offered the program during childhood.

Overall, these diverse studies provide consistent evidence that the program significantly increases schooling levels. The evidence from the medium- and longer-term evaluations currently suggests program impacts of increasing final levels of schooling attainment by between 15% and 20%. However, these previous studies are all based on self-reported information on school enrollment and attendance obtained from a household survey, raising concerns about whether there could be some reporting biases that are systematically related to program participation. For instance, beneficiary households, knowing that benefits are conditional on their children attending school might over-report their children's enrollment and attendance. Our study avoids

these potential biases in enrollment by constructing an alternative indicator of enrollment using administrative data based on attendance the day of the application of the ENLACE tests.

There are very few quantitative studies of Prospera on achievement test impacts, mainly because achievement tests were not part of the main data collection in the initial experimental evaluation. In 2003, a version of the Woodcock Johnson tests in mathematics and Spanish was applied. Comparing the test scores of the original treatment and control groups, Behrman, Parker and Todd (2009) find no impacts on achievement tests, despite significant impacts on increasing schooling attainment between the two groups.

Several reviews examine the evidence on how other CCT programs impact school achievement. An early study by Schady and Fizbein (2009) summarizes the effects of CCTs on achievement tests and argues that results have been disappointingly small. Since then, several additional studies show mixed effects. Baird et al. (2014) review education effects of both conditional and unconditional transfers and find that effect on achievement are "small, at best." Snilstveit et al. 2015 carry out a meta-analysis review of educational interventions, studying both conditional and unconditional cash transfer programs. They conclude that, while cash transfers have been effective at increasing school enrollment, overall, they have low effects on achievement.

One caveat to these conclusions is there are far fewer studies on the effects on achievement as there are on the effects on enrollment. Snilstveit et al. note the number of studies ofachievement is much lower than that for enrollment, in particular they review studies from 38 programs of the effects of transfers on enrollment, but only 11 programs have evaluation studies of their achievement impacts. Within these studies, however, there are examples of transfer programs that significantly affect attendance. For instance, Baird, McIntosh, and Ozler (2010) show significant impacts of 0.14 standard deviations on English achievement tests and 0.12 on math achievement tests after two years of operation of a CCT program in Malawi.

Our study builds on two recent studies which use the ENLACE data to study impacts of PROSPERA. Acevedo, Ortega and Székely (2018) use several sources of variation to estimate impact of PROSPERA on achievement tests, including comparing beneficiaries with non-beneficiaries and comparing beneficiaries with less time in the program to beneficiaries with more time in the program. Depending on the method used, they find varying sizes of impacts, ranging from effects on the order of 0.3 SD to insignificant effects. Our study uses different methodologies to estimate impact as well as combining socio-economic characteristics and school characteristics as conditioning variables and as potential sources of program impact heterogeneity.

De Hoyos, Estrada, and Vargas (2018) analyze a longitudinal panel of ENLACE data for a group of students that took the test in sixth grade in 2007 and then again in grade 9 and in grade 12. They also merge the grade 12 test scores with a special module of the Mexican labor force survey (ENOE) applied to individuals aged 18 to 20 years old in 2010. They find that the ENLACE test scores have strong predictive power for future education and labor market outcomes. They show that grade 6 learning outcomes are an important predictor of lower- and

upper-secondary on-time graduation and test scores. A one standard-deviation (SD) increase in grade 6 test scores is associated with an increase in the probability of on-time lower secondary graduation by 10 percentage points and (conditional on this) with an increase in learning outcomes by 0.6 SD. They also find a strong association between grade 6 test scores and grade 12 outcomes. Lastly, they find that grade 6 test scores and future outcomes have a strong relationship even when comparing individuals with identical family backgrounds, which is done using a subsample of twins.

IV. Data and Methodology

We are fortunate to have access to a number of administrative data sets that we combine to create a data set with information on achievement tests, characteristics of students, PROSPERA receipt and school quality. We now describe the data sets used and our methodology.

From the school years 2006-2007 to 2013-2014, the Mexican Secretariat of Public Education (SEP) applied the Evaluación Nacional de Logro Académico en Centros Escolares, ENLACE. The test gathered information at the end of each academic year on student performance in math, Spanish and a rotating subject for all third to ninth graders in private and public schools. Beginning in 2007-2008, ENLACE was also given to students in their final year of high school (grade 12). The test was intended to be a low-stakes assessment that would be informative about learning outcomes to SEP and to parents. The proportion of students who take the test is close to 90% (Hoyos, Estrada and Vargas 2018). The last school year of application of the test was 2013-2014. ENLACE was eventually replaced by PLANEA (National Plan for Evaluating Learning "Plan Nacional para la Evaluación de los Aprendizajes") during the Pena Nieto administration, this test however is applied only to a sample of students in each school.

An important issue related to using the ENLACE for this study is the possibility of cheating and, in particular, student copying on the ENLACE exams. When the government stopped applying the ENLACE tests, one of the stated concerns was over an increased incidence of suspected student copying. Unless the level of student copying was extremely high with the majority of students engaging in copying, the main way in which cheating might affect the results presented here is if there were differential rates of cheating between PROSPERA and non-PROSPERA students, e.g. that results on achievement tests would reflect both performance on the test and cheating. We are unaware of publicly available data either on the overall level of cheating detected by the Ministry of Education (SEP)/National Institute of Education Evaluation (INEE). Individual answer sheets would be necessary to directly analyze whether there are differential rates of cheating between PROSPERA and non-PROSPERA students based on standard statistical cheating detection methods developed in the education literature. (e.g. Wollak, 2003).

However, since the ENLACE exams are not used to calculate student grades, it is not obvious that PROSPERA and non-PROSPERA students would have different incentives for cheating on

the ENLACE. Furthermore, in conversations with INEE, the suggested proportion of students copying was reported to be in nearly all cases less than 10%. Thus, we consider unlikely that the presence of some cheating/copying biases our results on the ENLACE scores.

For our main analysis, we begin with the database of all sixth graders taking the ENLACE test in the school year 2007-2008, the second year in which ENLACE was applied. We focus on this group because it is the group for which we can merge the largest number of follow-up tests. In particular, we can merge to this database of sixth graders in 2007-2008, test results in 2008-2009 (seventh grade for most), 2010-2011 (ninth grade for most) and 2013-2014 (twelfth grade for most). In this way we construct a panel of achievement tests over a period of six years. Because program effects on enrollment and attendance in the primary level are lower than in subsequent levels (Behrman, Sengupta and Todd, 2005), we treat the 2007-2008 sixth grade scores as a baseline, that is unlikely to have been significantly impacted by PROSPERA although households in which the children in our database reside, may have been receiving grants before this period. To the extent achievement of PROSPERA children had already been affected in our data before the school year 2007-2008, the results we report here represent an underestimate of total program effects on achievement.[5]

This approach also allows us to treat taking the ENLACE test as a proxy for being enrolled in school. Although defining enrollment in this way underestimates the overall level of children enrolled in school, estimates of the impact of PROSPERA on enrollment will be unbiased assuming the implementation of the ENLACE does not differentially affect who attends school the day of the application with regard to PROSPERA receipt. This seems a reasonable assumption.

The construction of the panel allows us to study educational trajectories of youth between sixth grade and twelfth grade beginning with sixth graders in 2007-2008. We choose this panel as it allows following students for the longest time period possible with the ENLACE. We merge ENLACE exams from seventh grade (first grade of secondary school) in 2009-2010, ninth graders in 2011-2012 and twelfth graders in 2014-2015. Ideally, we would have had ENLACE scores for tenth graders in 2012-2013 so as to better analyze the transition from secondary to high school, but ENLACE tests were only administered in high schools to twelfth graders.

The ENLACE data also contain information on the type of school attended. Thus, we are also able to study program effects on enrollment in different types of schools in addition to achievement test results. There are a number of pathways youth may take, for instance, sixth graders may dropout or repeat sixth grade, or they may enter a tele-secondary, technical, general or private school. At the secondary level we study enrollment in technical, tele-secondary or general schools (less than 1 percent enroll in private secondary schools). At the high school level, we distinguish between general high schools, technical high schools, CONALEP high schools and private high schools.

[5] Data available for this version of the paper did not include students who continued to be enrolled but repeated a grade. Thus, these students are excluded in this analysis. Failure rates are, however, relatively low in sixth grade and beyond (Behrman, Sengupta and Todd, 2005).

With our panel of constructed ENLACE data, we merge information on program receipt at the individual level derived from administrative data. We define an individual's beneficiary status based on the receipt of a PROSPERA scholarship in the sixth grade. Because enrollment in primary school is near universal, we assume that if administrative data shows that a sixth grader was not receiving a PROSPERA grant, that sixth grader lives in a household that does not receive PROSPERA. Our definition thus of beneficiary versus non-beneficiary reflects status in sixth grade and the identification strategy compares sixth grade PROSPERA recipients with sixth grade non-recipients in all the years we study.

One might expect PROSPERA recipients to show higher poverty rates than non-recipients and this would imply lower rates of school enrollment between the groups in the absence of the program. We are fortunate to be able to also merge information on children and their households using Context Questionnaires that were applied to children at the time of the ENLACE application to sixth graders in 2007-2008 and this information provides valuable variables for controlling for differences in characteristics between PROSPERA recipients and non-recipients. The context questionnaires were applied only to a random subsample of schools, so we restrict our analysis to this subsample.

We also have data on characteristics of the schools that are collected using "911" forms, which are administrative forms that are filled out at the level of the school at the beginning and the end of each school year, which we also merge to our data. These data allow us to control for measures of quality and to potentially study if PROSPERA impacts are significantly different when children have access to different quality schools. We also use these variables to control for impacts of available school quality on achievement. The particular variables available varies by school year. We use those variables available at the beginning of our panel, for the school year 2007-2008 to construct the following school level indicators: student/teacher ratio, whether school is "multi-grado" meaning a teacher has more than 1 grade in his/her classroom, teacher education levels (measured byproportion of teachers who obtained their bachelor's degree of all teachers) and the availability of a computer.

Finally, we merge data at the community level from components of the margination index produced by Conapo based on the 2000 Mexican Population Census. The index is the normalized first principal component of nine locality-level population shares: the share illiterate, the share with less than primary school education, the share without a toilet, the share without electricity, the share without running water, the share without drainage, the share with crowding (few rooms per capita), the share with a dirt floor, and the share without a refrigerator. These indicators are also useful as control variables.

We now turn to a discussion of our methodology. Our methodology is non-experimental in that we derive a comparison group not though randomization but rather by choosing a comparison group that potentially represents a counterfactual for the beneficiary group in the absence of the program. In this study, we compare PROSPERA recipients and non-recipients to estimate program impacts. As we will see below, even when PROSPERA and non-PROSPERA students study in the same school, PROSPERA students typically have a lower socioeconomic status than

non-PROSPERA students, suggesting that PROSPERA students, in the absence of the program, would obtain lower education outcomes. We therefore include a rich set of control variables that aims to control for differences between beneficiaries and non-beneficiaries in socioeconomic status.

Studying effects on enrollment and type of school attended is straightforward empirically because we observe for all students whether they are enrolled in each year (as proxied by taking the ENLACE exam) and in which school. Studying program effects on ENLACE achievement tests is less straightforward, because students only take the tests if they are enrolled in school. That is, the test score outcome is not measured for students not in school. As previous studies have shown, the program likely affects post-primary school enrollment and potentially influences who takes the ENLACE tests. If the marginal students who enroll in school with the program but who would have dropped out without the program tend to be lower achieving (as is plausible), then estimates of test score impacts based on simple comparisons of beneficiary and non-beneficiaries are likely to be downward biased. It is important to control for enrollment changes induced by the program to obtain unbiased estimates of test score impacts.

Let X denote the subset of conditioning variables available, which include gender and region of residence. Also, let $D = 1$ if an individual is a program beneficiary, $D = 0$ if not. Let $E = 1$ if an individual is enrolled in school and $E = 0$ if not. Let ΔY_0 denote the test score gain (e.g. the gain in test scores from sixth grade to twelfth grade) if an individual did not receive the program and ΔY_1 the test score gain if an individual did receive the program. Note that only one of these states is observed for any given individual.

We are interested in the average increase in test score gains for individuals who participated in the program and who are enrolled in school, conditional on observed X.

$$E(\Delta Y_1 - \Delta Y_0 \mid X, D = 1, E = 1). \tag{1}$$

We observe $E(\Delta Y_1 \mid X, D = 1, E = 1)$ directly from the data. However, we need to infer $E(\Delta Y_0 \mid X, D = 1, E = 1)$, which is not observed. To do so, invoke an independence assumption

$$F(\Delta Y_0 \mid X, D = 1, E = 1) = F(\Delta Y_0 \mid X, D = 0, E = 1). \tag{2}$$

This assumption implies that the distribution of test score gains in the absence of treatment is independent of treatment assignment after conditioning on the X variables and on school enrollment ($E = 1$). In this paper, we use two different estimation approaches: OLS FE models and matching models. When the outcome measure corresponds to test scores, the outcome is measured as the test score gain. Therefore, the estimated test score results are essentially based on a difference-in-difference approach. When the outcome measure corresponds to school enrollment or type of school, the outcome is measured in levels but grade 6 characteristics (e.g. test score) are used as a conditioning variable.

The matching estimators are carried out in two steps. The first step is to estimate $Pr(D = 1|\ X,\ E = 1)$ (using probit or logistic regression). Then, in the second step, each individual in the treatment group is matched to an individual in the comparison group ($D = 1$) with a similar propensity score. We carry out two variations of matching that combine exact matching with matching on the propensity score. In addition to the propensity score, we exact match on age and state of residence, and, in another variant that we report, we exact match on age and school.

We also estimate two different OLS FE estimators. One includes school-level fixed effects and the other includes state-level fixed effects. The within-school estimates are attractive empirically, effectively estimating program impact by comparing PROSPERA recipients with non-recipients who attend the same primary school. However, in practice this identification strategy limits the sample studied, as only a minority of schools have both PROSPERA and non-PROSPERA students and, even in schools with both types of students, frequently there are small numbers in one category or another. For these reasons, we also provide similar estimations using within state rather than within school variation in beneficiary status and controlling for school and locality level characteristics.

We also analyze effects of receiving the program ($D = 1$) on school enrollment in ninth or twelfth grade, by estimating a model for school enrollment $Pr(E = 1 \mid X,\ D,\ Y_6)$, where Y_6 is the sixth grade test score and D is an indicator for program participation. For this estimation, we use probit models.[6] Similarly, we analyze the effects of participating in the program on the probability of attending certain types of schools.

We estimate effects on enrollment and on type of school enrolled in seventh grade, ninth grade and twelfth grade. For the achievement test outcomes, we take the difference between achievement tests scores in each year and the test score in the baseline year (2007-2008) as the dependent variable. The ENLACE tests are standardized in the same way (mean 500, standard deviation 100) for all ENLACE tests in primary and secondary school. However, high school ENLACE results are not standardized in the same way so that we cannot take the difference with the sixth grade ENLACE score. For the case of the twelfth grade scores, we construct a dummy variable measuring whether a student increased their category of classification in the ENLACE between sixth grade and twelfth grade.[7] (In addition to a continuous variable, the ENLACE tests are assigned to four categories defined as Insufficient, Basic, Good and Excellent for each year and for each test.) This of course reduces the potential variation that we can observe in the case of ENLACE scores in twelfth grade. That is, we will only be able to measure changes in categories but not improvements or reductions within categories.

Our set of control variables from the context questionnaires consists of the following: age, number of siblings, health problems, household size, whether lives with mother, whether lives

[6] Results are nearly identical using logit or linear probability models.

[7] Less than 3 percent of sixth grade students in 2008 attained the highest category, Excellent, for Math and Spanish. These students are excluded from the analysis as it is not possible for this group to improve their category rating.

with father, whether currently working, household ownership of DVD, television, a MP3 player, cellular phone, computer, calculator, whether English is spoken and whether the student aspires to study in a university. Preprogram variables measuring school quality from the Ministry of Education 911 forms (formats filled out by schools) include: student-teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade and average teacher schooling levels. These variables correspond to characteristics of the primary school attended in 2007-2008. In addition, we merge the locality-level margination index and its components to our data file. Community-level control variables include the margination index, the proportion of the population that is illiterate, the proportion of households without electricity, the proportion of households without sewage drainage, and the proportion of households without access to piped water. This set of control variables is used to estimate the propensity score model, additionally exact matching on gender and state is carried out. The type of matching is nearest neighbor matching with 1 match.[8]

V. Results

We present results based on two main samples: 1) sixth grade students in the school year 2007-2008 who attend primary schools in localities of medium, high or very high levels of margination[9] and 2) sixth grade students in the school year 2007-2008 who are enrolled in indigenous primary schools. For each sample we present results on enrollment, type of school and achievement tests in mathematics and Spanish.

We present three sets of results for each indicator:

1) State fixed-effects models with individual, school and locality level control variables
2) School fixed-effects models with individual control variables and
3) Nearest-neighbor matching estimates using individual, school and locality variables for the propensity score and exact matching on gender and state.[10]

We begin by presenting descriptive statistics for our samples.

[8] We also carried out matching estimations using 2 and 3 neighbors respectively. Results were very similar, and for brevity we provide only those based on 1 neighbor.

[9] The majority of rural Progresa recipients live in communities with medium to very high levels of marginality. We thus focus our analysis on primary schools located in these areas. Nationwide, about 70% of all Progresa sixth grade students attend schools in these areas, consistent with about 70% of the PROSPERA population living in rural areas. If the program increased the likelihood that sixth grade students traveled to schools outside to non-poor areas, this might impart a selection bias to restricting the sample in this way. We have also replicated the results here using the set of nationwide sixth grade students which necessarily includes urban students but avoids selecting the sample based on potential school selection and we obtain similar although slightly smaller results.

[10] For the matching estimators for all students, we use a random sample of 20% of all schools because of computation issues for the matching.

A. Characteristics of PROSPERA and non-PROSPERA children

Tables 2a and 2b present descriptive statistics comparing baseline characteristics (in 2007-2008 for sixth graders) for PROSPERA and non-PROSPERA students for the sample of sixth graders in schools located in communities with middle to very high margination indices and for the sample of sixth graders in indigenous primary schools, which we treat as a proxy for impacts on indigenous children. As expected, Table 2a shows there are clear socioeconomic differences favoring non-PROSPERA children. With respect to school characteristics, the schools that PROSPERA children attend tend to have fewer students per teacher, but teachers have lower education levels and schools have fewer resources, as measured by having a computer.

Table 2a: Descriptive Statistics of Beneficiaries and Non-beneficiaries
All 6th Graders in Medium to Very High Poverty Areas: School Year 2007-2008

Variables	Non-PROSPERA		PROSPERA	
	Mean	Std. Dev.	Mean	Std. Dev.
Student and Household Characteristics				
Age	11.76	0.78	11.93	0.79
# siblings	3.14	2.26	4.39	2.53
Have health problems	0.08	0.27	0.11	0.31
Spanish is first language	0.89	0.31	0.77	0.42
Currently works	0.66	0.47	0.53	0.5
HH has DVD	0.73	0.44	0.55	0.5
HH has TV	0.89	0.31	0.79	0.41
HH has MP3 player	0.47	0.5	0.33	0.47
HH has cell phone	0.6	0.49	0.39	0.49
HH has computer	0.27	0.45	0.09	0.29
HH has internet	0.15	0.36	0.05	0.23
HH has calculator	0.89	0.32	0.87	0.34
Aspires to attend college	0.56	0.5	0.36	0.48
Studies English	0.32	0.47	0.13	0.34
School Characteristics				
Student teacher ratio	28.63	13.44	24.99	13.05
Has computer	0.74	0.44	0.63	0.48
More than 1 grade per classroom	0.17	0.38	0.28	0.45
Proportion of teachers with university	0.75	0.28	0.68	0.35

Source: Authors calculations: ENLACE 2007-2008 merged with context questionnaires and school characteristics.

Table 2b presents the same information for the sample of sixth graders enrolled in indigenous schools. Although beneficiary children have lower socioeconomic status in this group as well, it is worth noting that differences are much smaller between beneficiary and non-beneficiary children than was true in the broader sample. The estimations below control for these variables

in an effort to control for differences between PROSPERA and non-PROSPERA children in the absence of the program. Given the higher poverty rate of PROSPERA children, we expect any remaining biases to have the effect of reducing the estimated impacts of PROSPERA on enrollment and on learning.

Table 2b: Descriptive Statistics of Beneficiaries and Non-beneficiaries
6th Graders 2008 Enrolled in Indigenous Schools.

	Non-PROSPERA		PROSPERA	
Variable	Mean	Std. dev.	Mean	Std. dev.
Student and Household characteristics				
Age	12.097	0.856	12.004	0.837
# siblings	4.334	2.515	4.781	2.534
Have health problems	0.157	0.364	0.159	0.366
Spanish is first language	0.485	0.5	0.419	0.493
Currently works	0.399	0.49	0.392	0.488
HH has DVD	0.51	0.5	0.45	0.497
HH has TV	0.664	0.473	0.629	0.483
HH has MP3 player	0.32	0.466	0.286	0.452
HH has cell phone	0.344	0.475	0.275	0.446
HH has computer	0.145	0.352	0.109	0.312
HH has internet	0.077	0.266	0.076	0.266
HH has calculator	0.826	0.379	0.811	0.392
Aspires to attend college	0.385	0.487	0.315	0.465
Studies English	0.117	0.321	0.098	0.298
School Characteristics				
Student teacher ratio	25.203	7.363	24.927	7.251
Has computer	0.485	0.5	0.497	0.5
More than 1 grade per classroom	0.26	0.439	0.27	0.444
Proportion of teachers with university	0.313	0.274	0.304	0.278

Source: Authors calculations: ENLACE 2007-2008 merged with context questionnaires and school characteristics.

B. Descriptive variables on outcomes

Table 3 shows the proportion of students enrolled after sixth grade, which we take as a baseline for our study. We present the proportion enrolled in seventh grade, ninth grade and twelfth grade, by gender and for the samples of schools in medium to very high marginalized areas and for students in indigenous primary schools during sixth grade. The Table shows enrollment, proxied by the proportion of students taking the ENLACE in that year. According to the Table, about 60% of males and 64% of female sixth graders are captured in the seventh grade ENLACE applied in the next academic year and about 30% of males and 32% of females are captured as taking the twelfth grade ENLACE. As mentioned before, this is an underestimate of enrollment

continuation as some enrolled youth will be absent the day of the ENLACE application and others may have repeated a grade/grades and thus not be in the ENLACE application. In our estimations, our objective will be the comparison of levels to derive impact. Still it is noteworthy in Table 3 that enrollment levels for indigenous continuation rates are even lower both for females and males than for the overall sample of children in medium to very poor areas. [11]

Table 3: Enrollment Rates in 7th Grade, 9th Grade and 12th Grade

6th Graders in 2008				
	Primary schools in Medium to Very High Marginalized Areas		Indigenous Primary Schools	
	Male	Female	Male	Female
7th grade	59.13%	63.79%	45.56%	44.92%
9th grade	59.22%	62.97%	50.16%	49.27%
12th grade	29.31%	32.43%	24.95%	24.96%

Source: Author's calculations using ENLACE databases.

Table 4 shows the distribution of sixth grade students who enroll in seventh grade and in which types of schools. The majority of indigenous children enroll in tele-secondary schools with about 20% enrolling in technical secondary schools. Only about 10% enroll in general secondary schools. The proportions in tele-secondary schools are somewhat lower for the overall sample of youth in marginalized areas at about 59% for boys and 46% for girls. Girls have a correspondingly higher probability of enrolling in a general school within our sample of schools located within marginalized areas.

Table 4: Secondary School Attended in 7th Grade

6th Graders in 2008				
	Primary Schools in Medium to Very High Marginalized Areas		Indigenous Primary Schools	
	Male	Female	Male	Female
General	20.78	32.75	10.3	11.11
Private	0.88	2.75	0.72	0.92
Tele-secondary	58.79	45.5	67.64	66.12
Technical	19.55	19	21.34	21.85

Source: Author's calculations using ENLACE databases.

Finally, Tables 5a and 5b show results in the mathematics ENLACE score between sixth grade and twelfth grade for the population that continue in school until twelfth grade, with separate tables for students in schools in marginalized and very marginalized areas in sixth grade and for students in indigenous primary schools in sixth grade. The tables make clear that a majority of children in both samples have poor results on the ENLACE test, with the vast majority attaining

[11] Table 3 shows that for children in indigenous primary schools, enrollment in ninth grade is higher than enrollment in seventh grade. This likely reflects higher absentee rates for the ENLACE in seventh versus ninth grade for this population.

levels of insufficient or basic both in sixth grade and in twelfth grade. There is however a fair amount of movement over time between categories including both improvements and falling to lower categories over time.

Table 5a: Mathematics ENLACE Scores, 6th Grade versus 12th Grade, Schools in Areas of Medium to Very High Margination

		12th grade score, 2013-2014				
		Insufficient	Basic	Good	Excellent	Total
6th Grade 2008	Insufficient	2,287	1,256	400	233	4,176
		54.77	30.08	9.58	5.58	100
	Basic	4,123	5,231	2,585	1,546	13,485
		30.57	38.79	19.17	11.46	100
	Good	563	1,287	1,381	1,653	4,884
		11.53	26.35	28.28	33.85	100
	Excellent	87	104	180	656	1,027
		8.47	10.13	17.53	63.88	100
	Total	7,060	7,878	4,546	4,088	23,572
		29.95	33.42	19.29	17.34	100

Source: Author's calculations, ENLACE standardized achievement tests.

Table 5b: Mathematics ENLACE Scores, 6th Grade versus 12th Grade
Indigenous Students
12th Grade Score, 2013-2014

		Insufficient	Basic	Good	Excellent	Total
6th Grade 2008	Insufficient	533	246	88	54	921
		57.87	26.71	9.55	5.86	100
	Basic	674	671	311	205	1,861
		36.22	36.06	16.71	11.02	100
	Good	56	108	94	136	394
		14.21	27.41	23.86	34.52	100
	Excellent	4	9	15	33	61
		6.56	14.75	24.59	54.1	100
	Total	1,267	1,034	508	428	3,237
		39.14	31.94	15.69	13.22	100

Source: Author's calculations, ENLACE standardized achievement tests.

C. Impact results

We now turn to the principal impact estimation results. We begin with the estimations for all sixth graders in primary schools located in areas with middle to very high levels of margination,

according to the margination index produced by CONAPO (Consejo Nacional de Poblacion) and then turn to the sample of sixth graders enrolled in indigenous schools in 2007-2008. Note that we follow sixth graders to whatever school they attend after sixth grade; the secondary and high schools that students attend thus are not restricted to be in marginalized areas.

Tables 6a to 6c presents the effects of PROSPERA on enrollment in seventh grade, ninth grade and twelfth grade. As described earlier, we measure enrollment using attendance the day of the ENLACE test application. One advantage of this measure is that it may have less biases than self-reported data on children's enrollment reported by parents in household surveys, which hasbeen the usual source of data for previous evaluation studies. Tables 6a and 6b show state fixed effects models and school fixed effects models. The results indicate that the impact of PROSPERA on enrollment in secondary school is about 8 percentage points for seventh grade, 9 to10 percentage points for ninth grade and decreases to about 5 percentage points for twelfth grade. The results based on propensity score matching (Table 6c) are similar but about 1 percentage point smaller for enrollment in seventh and ninth grade.

Table 6a: Impacts of PROSPERA on School Enrollment
All 6th Graders in Medium, High and Very Highly Marginalized Communities 2008
State Fixed Effects

	7th grade	9th grade	12th grade
Variables	(1)	(2)	(3)
PROSPERA beneficiary 2008	0.081***	0.101***	0.045***
	(0.006)	(0.006)	(0.006)
Female beneficiary	0.005	0	0.002
	(0.007)	(0.007)	(0.008)
Female	-0.014**	-0.022***	-0.013**
	(0.006)	(0.006)	(0.006)
Constant	1.039***	1.207***	0.806***
	(0.036)	(0.037)	(0.037)
Observations	62,870	62,870	62,870
R-squared	0.223	0.207	0.124

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis. Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Table 6b: Impacts of PROSPERA on School Enrollment
All 6th Graders in Medium, High and Very Highly Marginalized Communities
School Fixed Effects

	7th grade	9th grade	12th grade
Variables	(1)	(2)	(3)
PROSPERA beneficiary 2008	0.076***	0.091***	0.046***
	(0.005)	(0.006)	(0.006)
Female beneficiary	0.008	0.004	-0.002
	(0.007)	(0.007)	(0.007)
Female	-0.016***	-0.025***	-0.014**
	(0.006)	(0.006)	(0.006)
Constant	0.878***	1.050***	0.551***
	(0.031)	(0.033)	(0.034)
Observations	68,871	68,871	68,871
R-squared	0.428	0.332	0.245

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, and calculator.

Table 6c: Impacts of PROSPERA on School Enrollment:
All 6th Graders in Medium, High and Very Highly Marginalized Communities 6th graders in 2008
Nearest Neighbor Matching Models

	Enrollment		
Variables	7th grade 2009-2010	9th grade 2011-2012	12th grade 2014-2015
Propensity score matching with exact matching on age and state (number of matches)			
Girls	0.0688***	0.080***	0.0480**
	(0.018)	(0.0187)	(0.0231)
Observations	6,654	6,654	6,654
Boys	0.0670***	0.0888*	0.0447**
	(0.0212)	(0.022)	(0.0222)
Observations	4,853	4,853	4,853

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Variables which enter the propensity score include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Overall, these results are highly consistent with Schultz, 2004; Behrman, Sengupta and Todd, 2005, where enrollment was measured by self-reported household data. With respect to gender, the results do not show significant differences by gender in the enrollment impacts, that is, program impacts on enrollment are similar for both genders. Note that both Schultz, 2004 and Behrman, Sengupta and Todd, 2005 demonstrate higher effects on enrollment at the secondary level for girls than boys in the randomized evaluation sample carried out in 506 communities in seven states in the initial years of this evaluation. Nevertheless, follow-up results of this sample

(Behrman, Parker and Todd 2011) as well as analysis from other samples (Parker and Vogl, 2018) do not show gender differences in overall attainment. The sample we analyze in this paper is a national sample so that it is not necessarily surprising to have some differences between our results with those derived from a less representative evaluation sample.

Tables 7a through 7c present the impacts on the probability of attending secondary school by type of school. Note that we present unconditional results on type of school, so that changes in these variables include both the impact of more students enrolling in school and changes in the type of school conditional on enrollment. We present the impact of PROSPERA on enrolling in a tele-secondary school, a general secondary school and a technical school. Overall, the results suggest that the increase in secondary school enrollment we observed in Tables 6a to 6c leads to a larger proportional increase in enrollment in tele secondary schools. There is no effect on the probability of enrolling in general schools; this implies that the pool of students (now larger due to the enrollment effect) is just as likely to enroll in a general school with or without the program. There is however a reduction in the probability of attending a technical school as a result of the increase in the probability of attending a tele-secondary school. Thus, participating in PROSPERA is associated with an increase in enrollment and an increase in the probability of attending tele-secondary schools. The implications of this shift are important, as tele-secondary schools are generally considered to be an inferior option in the sense that these types of schools tend to be attended only when general and technical schools are not available.

Table 7a: Impacts of PROSPERA on Type of Secondary School in 7th grade
All 6th Graders in Medium, High and Very Highly Marginalized Communities 2008
State Fixed Effects

	Tele-sec	General	Technical
Variables	(1)	(2)	(3)
PROSPERA beneficiary 2008	0.099***	-0.008*	-0.091***
	(0.006)	(0.005)	(0.006)
Female beneficiary	0.011	0.003	-0.005
	(0.007)	(0.006)	(0.008)
Female	-0.004	-0.008*	0.006
	(0.006)	(0.005)	(0.007)
Constant	0.605***	0.255***	0.163***
	(0.035)	(0.029)	(0.039)
Observations	62,870	62,870	62,870
R-squared	0.215	0.2	0.144

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis. Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Table 7b: Impacts of PROSPERA on Type of School
All 6th Graders in Medium, High and Very Highly Marginalized Communities 2008
School Fixed Effects

Variables	Telesec (1)	General (2)	Technical (3)
PROSPERA beneficiary 2008	0.068***	0.006*	-0.072***
	(0.004)	(0.004)	(0.005)
Female beneficiary	0.014**	-0.007	-0.009
	(0.006)	(0.005)	(0.007)
Female	-0.008	0	0.009
	(0.005)	(0.004)	(0.006)
Constant	0.442***	0.258***	0.295***
	(0.026)	(0.022)	(0.031)
Observations	68,871	68,871	68,871
R-squared	0.561	0.531	0.454

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.

Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, and calculator.

Table 7c: Impacts of PROSPERA on Type of Secondary School:
Sample: 6th graders in 2008 in Medium, High and Very Highly Marginalized Communities
Nearest Neighbor Matching Models

Variables	Enrollment 7th grade Telesecondary	7th grade General school	7th grade Technical school
Propensity score matching with exact matching on age and state			
Girls	0.1263***	-0.0209	-0.0912***
	(0.0164)	(0.0149)	(0.0188)
Observations	6,654	6,654	6,654
Boys	0.104***	-0.014	-0.0920***
	(0.0212)	(0.0133)	(0.022)
Observations	4,853	4,853	4,853

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Variables which enter the propensity score include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Why is the increase in secondary school enrollment concentrated in tele-secondary schools? One obvious explanation is that these are the schools which are available in the poor rural areas that are the focus in this paper, and it is not feasible to reasonably attend other schools (due to distance).

However, it might also be the case that parents/students attempted to enroll their students in alternative types of schools but were not successful, perhaps due to crowding or limited spaces in other types of schools. Distinguishing the mechanisms that affect enrollment in the different types of schools is an important question for future research.

Tables 8a to 8c show similar tables for the effect on high school enrolled in the twelfth grade. These results show that there is an increase in the probability of attending a general high school and a corresponding reduction in the probability of attending a technical high school. There is no significant effect on the probability of attending a technical professional school or on the probability of attending a private school. Compared with the results for secondary schools, these results do not suggest that the enrollment effect leads students to shift towards lower quality schools. However, further examination of the characteristics of the schools that PROSPERA children attend would be useful. The results do not differ by gender.

Table 8a: Impacts of PROSPERA on Type of High School
All 6th Graders in Medium, High and Very Highly Marginalized Communities 2008
State Fixed Effects

Variables	General HS (1)	Technical HS (2)	Technical/Prof (3)	Private (4)
PROSPERA beneficiary 2008	0.049***	-0.049***	0	0
	(0.005)	(0.005)	(0.002)	(0.002)
Female beneficiary	-0.003	0.01	0.001	-0.008***
	(0.007)	(0.007)	(0.002)	(0.002)
Female	-0.003	0	-0.004***	0.007***
	(0.006)	(0.006)	(0.002)	(0.002)
Constant	0.649***	0.343***	0.013	-0.005
	(0.033)	(0.035)	(0.01)	(0.01)
Observations	62,870	62,870	62,870	62,870
R-squared	0.079	0.085	0.011	0.037

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Table 8b: Impacts of PROSPERA on Type of High School
All 6th Graders in Medium, High and Very Highly Marginalized Communities 2008
School Fixed Effects

Variables	General HS (1)	Technical HS (2)	Technical/Pr of (3)	Private (4)
PROSPERA beneficiary 2008	0.040***	-0.039***	0.001	-0.002
	(0.005)	(0.005)	(0.002)	(0.002)
Female beneficiary	-0.006	0.007	0.001	-0.003
	(0.007)	(0.007)	(0.002)	(0.002)
Female	-0.002	0.002	-0.004**	0.003**
	(0.005)	(0.006)	(0.002)	(0.002)
Constant	0.392***	0.559***	0.021**	0.028***
	(0.03)	(0.031)	(0.009)	(0.009)
Observations	68,871	68,871	68,871	68,871
R-squared	0.24	0.238	0.117	0.202

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, and calculator.

Table 8c: Impacts of PROSPERA on Type of High School:
6th Graders in 2008 in Medium, High and Very Highly Marginalized Communities
Nearest Neighbor Matching Models

Variables	General	Technical	Technical professional	Private
Propensity score matching with exact matching on age and state				
Girls	0.0519***	-0.0317*	-0.0039	-0.0109*
	(0.0163)	(0.0169)	(0.0045)	(0.005)
Observations	6,654	6,654	6,654	6,654
Boys	0.0895***	-0.01980*	0.0361	0.0017
	(0.0214)	(0.0123)	(0.01862)	(0.0049)
Observations	4,853	4,853	4,853	4,853

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Variables which enter the propensity score include: Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Finally, Tables 9a to 9c shows the difference-in-difference estimates of the impact of PROSPERA on ENLACE achievement tests for seventh grade, ninth grade and twelfth grade. For seventh and ninth grade, the coefficients report the increase in points achieved on the ENLACE exams due to the program. Because of the test score standardization, the estimates represent the impacts in standard deviations if one divides these coefficients by 100.

Table 9a: Impacts of PROSPERA on ENLACE tests
All 6th Graders in Medium, High and Very Highly Marginalized Communities 2008
State Fixed Effects

Variables	7th-6th score	9th-6th score	12th-6th level
Panel A: Spanish			
PROSPERA beneficiary 2008	4.806***	8.181***	-0.012
	(1.46)	(1.72)	(0.011)
Female beneficiary	2.641	-0.578	0.018
	(1.881)	(2.214)	(0.014)
Female	8.090***	0.313	-0.036***
	(1.583)	(1.867)	(0.011)
Constant	-16.145*	43.677***	0.761***
	(9.133)	(10.797)	(0.071)
Observations	40,382	39,540	19,914
R-squared	0.035	0.081	0.022
Panel B: Mathematics			
PROSPERA beneficiary 2008	4.944***	11.139***	-0.030**
	(1.614)	(2.195)	(0.012)
Female beneficiary	-0.918	-0.411	0.012
	(2.08)	(2.825)	(0.015)
Female	-2.187	2.071	-0.090***
	(1.75)	(2.383)	(0.012)
Constant	6.103	115.010***	0.936***
	(10.099)	(13.778)	(0.076)
Observations	40,382	39,540	19,440
R-squared	0.048	0.084	0.027

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

The results in these tables vary to some extent with the estimator used, but overall suggest some small positive effects on achievement with impacts on achievement in seventh and ninth grade averaging between 0.03 and 0.11 standard deviation, both on Spanish and on mathematics and both for girls and boys. With respect to the impacts in twelfth grade, recall that we use an alternative indicator that measures whether individuals increase at least one category in the ENLACE grading between sixth and twelfth grade. This indicator overall does not show significant effects, perhaps reflecting that there is less variation in the measure, although there is a slight unexpected and significant reduction in the probability of advancing categories in mathematics for 2 of the three models.

Table 9b: Impacts of PROSPERA on ENLACE tests
All 6th Graders in Medium, High and Very Highly Marginalized Communities 2008
School Fixed Effects

Variables	7th-6th score	9th-6th score	12th-6th score
	Panel A: Spanish		
PROSPERA beneficiary 2008	3.388**	4.619***	-0.014
	(1.354)	(1.524)	(0.011)
Female beneficiary	0.848	1.69	0.016
	(1.735)	(1.951)	(0.014)
Female	8.572***	-3.083*	-0.041***
	(1.471)	(1.658)	(0.012)
Constant	-18.606**	44.358***	0.785***
	(7.665)	(8.635)	(0.068)
Observations	43,245	42,914	21,372
R-squared	0.328	0.43	0.241
	Panel B: Mathematics		
PROSPERA beneficiary 2008	1.609	4.917***	-0.024**
	(1.361)	(1.841)	(0.012)
Female beneficiary	1.491	4.964**	0.013
	(1.744)	(2.358)	(0.015)
Female	-5.993***	-6.819***	-0.098***
	(1.478)	(2.003)	(0.012)
Constant	10.187	104.716***	0.949***
	(7.705)	(10.435)	(0.071)
Observations	43,245	42,914	20,875
R-squared	0.453	0.493	0.294

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, and calculator.

Table 9c: Difference in Difference Estimates of the Impact of PROSPERA on ENLACE Achievement Tests:
All 6th Graders in 2008 in Medium, High or Very Highly Marginated Communities 2008.

Variables	7th grade- 6th grade score	9th grade -6th grade score	Increase in category of score from 6th grade to 12th grade
Propensity score matching with exact matching on age and state			
Panel A: Spanish			
Girls	2.41	6.389	-0.0608
	(4.022)	(5.29)	(0.0335)
Boys	7.29	14.67**	0.0604
	(5.358)	(6.501)	(1.79)
Observations	4,565	4,415	2,253
Panel B: Mathematics			
Girls	7.478*	7.145	-0.0157
	(4.23)	(6.623)	(0.0499)
Boys	9.99*	14.097*	0.01
	(5.19)	(7.558)	(0.004)
Observations	3,072	2,908	1,494

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Variables which enter the propensity score include: Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

We now turn to results concentrating on children enrolled in indigenous primary schools. Tables 10a to 10c presents impacts on enrollment in seventh grade, ninth grade and twelfth grade for the set of children in indigenous schools in sixth grade. The results show important impacts on enrollment in all grades, and significantly larger than the impacts reported for all sixth graders in medium to very high marginalized areas. We observe increases in enrollment of 9 to 11 percentage points in seventh grade, 11 to 12 percentages points in ninth grade and about 8 to 9 percentage points in twelfth grade. There are no differences in these enrollment effects by gender.

Table 10a: Impacts of PROSPERA on School Enrollment
6th Graders Enrolled in Indigenous Primary Schools in 2008
State Fixed Effects

	7th grade	9th grade	12th grade
Variables	(1)	(2)	(3)
PROSPERA beneficiary 2008	0.116***	0.121***	0.078***
	(0.016)	(0.016)	(0.015)
Female beneficiary	-0.009	-0.011	-0.025
	(0.022)	(0.022)	(0.021)
Female	0.001	-0.004	0.011
	(0.02)	(0.02)	(0.019)
Constant	0.891***	0.921***	0.588***
	(0.087)	(0.088)	(0.084)
Observations	9,253	9,253	9,253
R-squared	0.277	0.259	0.129

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Table 10b: Impacts of PROSPERA on Enrollment
6th Graders Enrolled in Indigenous Primary Schools in 2008
School Fixed Effects

	7th grade	9th grade	12th grade
Variables	(1)	(2)	(3)
PROSPERA beneficiary 2008	0.089***	0.108***	0.086***
	(0.013)	(0.015)	(0.014)
Female beneficiary	0	0.001	-0.043**
	(0.018)	(0.02)	(0.019)
Female	0.007	-0.02	0.017
	(0.016)	(0.018)	(0.017)
Constant	0.640***	0.791***	0.383***
	(0.069)	(0.076)	(0.073)
Observations	11,200	11,200	11,200
R-squared	0.495	0.391	0.271

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, and calculator.

Table 10c: Impacts of PROSPERA on School Enrollment:
6th Graders Enrolled in Indigenous Primary Schools in 2008
Nearest-Neighbor Matching models

	Enrollment		
Variables	7th grade 2009-2010	9th grade 2011-2012	12th grade 2014-2015
Propensity score matching with exact matching on age and state			
Girls	0.097***	0.089***	0.039**
	(0.0199)	(0.019)	(0.018)
Observations	4,93	4,931	4,931
Boys	0.085***	0.127***	0.081***
	(0.018)	(0.018)	(0.017)
Observations	4,751	4,751	4,751

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Variables which enter the propensity score include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Turning to type of school, Table 11a through 11c shows effects on type of secondary school enrolled.

Table 11a: Impacts of PROSPERA on Type of Secondary School
6th Graders Enrolled in Indigenous Primary Schools in 2008
State Fixed Effects

	Tele-sec	General	Technical
Variables	(1)	(2)	(3)
PROSPERA beneficiary 2008	0.108***	-0.004	-0.105***
	(0.015)	(0.008)	(0.016)
Female beneficiary	0.002	-0.014	0.014
	(0.021)	(0.011)	(0.022)
Female	-0.009	0.014	-0.007
	(0.019)	(0.01)	(0.019)
Constant	0.595***	0.103**	0.265***
	(0.082)	(0.043)	(0.086)
Observations	9,253	9,253	9,253
R-squared	0.245	0.122	0.244

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Table 11b: Impacts of PROSPERA on Type of Secondary School
6th Graders Enrolled in Indigenous Primary Schools in 2008
School Fixed Effects

	Telesec	General	Technical
Variables	(1)	(2)	(3)
PROSPERA beneficiary 2008	0.075***	0	-0.072***
	(0.011)	(0.006)	(0.012)
Female beneficiary	0.009	-0.016*	0.005
	(0.015)	(0.008)	(0.017)
Female	-0.007	0.012	-0.004
	(0.014)	(0.007)	(0.015)
Constant	0.380***	0.072**	0.529***
	(0.06)	(0.032)	(0.064)
Observations	11,200	11,200	11,200
R-squared	0.556	0.454	0.529

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, and calculator.

Table 11c: Impacts of PROSPERA on Type of Secondary School:
6th Graders Enrolled in Indigenous Primary Schools in 2008
Nearest Neighbor Matching Models

	Enrollment		
Variables	7th grade Telesecondary	7th grade General school	7th grade Technical school
	Propensity score matching with exact matching on age and state		
Girls	0.097***	-0.011	-0.083***
	(0.017)	(0.0098)	(0.019)
Observations	4,931	3,931	3,931
Boys	0.105***	-0.018*	-0.087***
	(0.017)	(0.009)	(0.018)
Observations	4,751	4,751	4,751

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Variables which enter the propensity score include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

As was the case for the sample of children in medium to very high areas of poverty, the impact of the program is to increase the probability of enrollment in tele-secondary versus other schools.

The probability of attending a general school does not change and the probability of attending a technical school decreases. Program impacts do not differ by gender, with the exception of the school fixed effects specification where the interaction between female and being a beneficiary reduces the probability of attending a general secondary school.

With respect to enrollment in high school (Tables 12a through 12c), the results are similar as was the case for the entire population in poor and very poor areas; that is, the program increases the probability of attending a general high school and reduces the probability of attending a technical school. No differences in the size of the effects by gender are observed.

Table 12a: Impacts of PROSPERA on Type of High School
6th Graders Enrolled in Indigenous Primary Schools in 2008
State Fixed Effects

Variables	General HS (1)	Technical (2)	Tec Prof (3)	Private (4)
PROSPERA beneficiary 2008	0.083***	-0.082***	-0.001	0
	(0.014)	(0.015)	(0.003)	(0.003)
Female beneficiary	-0.026	0.025	0.003	-0.002
	(0.02)	(0.02)	(0.004)	(0.004)
Female	0.008	-0.008	-0.002	0.002
	(0.018)	(0.018)	(0.004)	(0.003)
Constant	0.362***	0.616***	0.005	0.017
	(0.078)	(0.08)	(0.017)	(0.015)
Observations	9,253	9,253	9,253	9,253
R-squared	0.101	0.104	0.029	0.007

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Table 12b: Impacts of PROSPERA on Type of High School
6th Graders Enrolled in Indigenous Primary Schools in 2008
School Fixed Effects

Variables	General HS (1)	Technical (2)	Tec Prof (3)	Private (4)
PROSPERA beneficiary 2008	0.076***	-0.074***	-0.001	-0.001
	(0.013)	(0.013)	(0.003)	(0.002)
Female beneficiary	-0.031*	0.032*	0.002	-0.002
	(0.017)	(0.018)	(0.004)	(0.003)
Female	0.01	-0.012	0	0.002
	(0.016)	(0.016)	(0.003)	(0.003)
Constant	0.254***	0.716***	0.007	0.023*
	(0.068)	(0.069)	(0.015)	(0.012)
Observations	11,200	11,200	11,200	11,200
R-squared	0.263	0.261	0.135	0.157

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, and calculator.

Table 12c: Impacts of PROSPERA on Type of High School:
6th Graders Enrolled in Indigenous Primary Schools in 2008
Nearest-Neighbor Matching Models

Variables	12th grade General	12th grade Technical	12th grade Technical professional	12th grade Private
Propensity score matching with exact matching on age and state				
Girls	0.046***	-0.046***	0.004	-0.0037
	(0.017)	(0.018)	(0.004)	(0.0037)
Observations	4,931	4,931	4,931	4,931
Boys	0.074***	-0.066***	- 0.0062	-0.0015
	(0.017)	(0.017)	(0.004)	(0.0037)
Observations	4,751	4,751	4,751	4,751

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Variables which enter the propensity score include: Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

Finally, Table 13a through 13c shows impact on achievement tests for indigenous students and shows some mixed evidence. Impacts on ENLACE scores in Spanish are positive and significant ranging up to 0.13 standard deviations on achievement tests up through ninth grade. In fact, these effects are larger than those observed for the overall population of beneficiary children in marginalized areas, consistent with the higher enrollment impacts that were also observed. However, effects on mathematics achievement are smaller and less precise/significant.

Table 13a: Impacts of PROSPERA on ENLACE Tests
6th Graders Enrolled in Indigenous Primary Schools in 2008
State Fixed Effects

Variables	7th-6th score	9th-6th score	12th-6th level
	Panel A: Spanish		
PROSPERA beneficiary 2008	9.321*	12.644**	-0.023
	(5.211)	(6.142)	(0.037)
Female beneficiary	-3.54	-0.779	0.056
	(7.085)	(8.346)	(0.049)
Female	10.664*	-0.023	-0.052
	(6.477)	(7.631)	(0.045)
Constant	-101.052***	-9.61	0.661***
	(26.207)	(30.968)	(0.193)
Observations	4,439	4,616	2,402
R-squared	0.078	0.15	0.053
	Panel B: Mathematics		
PROSPERA beneficiary 2008	9.314	10.673	0.009
	(5.756)	(7.819)	(0.038)
Female beneficiary	-5.422	-2.401	-0.004
	(7.826)	(10.626)	(0.05)
Female	7.077	7.663	-0.108**
	(7.154)	(9.715)	(0.046)
Constant	-83.487***	-15.03	0.914***
	(28.947)	(39.426)	(0.197)
Observations	4,439	4,616	2,370
R-squared	0.094	0.147	0.07

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

With respect to twelfth grade achievement tests there are no significant effects on having a higher classification on achievement in twelfth grade with respect to the classification in sixth grade. Potential selection issues in enrollment of course become worse at the high school level than at the secondary level and so this may explain in part a lack of significant results.

Table 13b: Impacts of PROSPERA on ENLACE tests
6th Graders Enrolled in Indigenous Primary Schools in 2008
School Fixed Effects

Variables	7th-6th score	9th-6th score	12th-6th level
	Panel A: Spanish		
PROSPERA beneficiary 2008	9.893**	8.680*	-0.021
	(4.364)	(4.748)	(0.036)
Female beneficiary	-5.553	-2.514	0.028
	(5.926)	(6.416)	(0.047)
Female	10.006*	1.907	-0.011
	(5.439)	(5.887)	(0.043)
Constant	-5.195	50.291**	0.670***
	(20.881)	(22.769)	(0.176)
Observations	5,234	5,757	2,897
R-squared	0.411	0.546	0.315
	Panel B: Mathematics		
PROSPERA beneficiary 2008	7.185	2.262	0.018
	(4.439)	(5.807)	(0.035)
Female beneficiary	-3.579	1.045	-0.007
	(6.028)	(7.848)	(0.046)
Female	0.025	-0.578	-0.111***
	(5.533)	(7.201)	(0.042)
Constant	18.868	99.985***	0.690***
	(21.239)	(27.847)	(0.172)
Observations	5,234	5,757	2,863
R-squared	0.499	0.583	0.383

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Control variables include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, and calculator.

Table 13c: Difference-in-difference Estimates of the Impact of PROSPERA on ENLACE Achievement Tests:
6th Graders Enrolled in Indigenous Primary Schools in 2008

Variables	Propensity score matching with exact matching on age and state		
	7th grade- 6th grade score	9th grade -6th grade score	Increase in category of score from 6th grade to 12th grade
Panel A: Spanish			
Girls	3.77	10.98*	0.011
	(5.87)	(6.69)	(0.042)
Boys	7.21	8.16	0.07
	(5.82)	(7.08)	(0.05)
Observations	2,387	2,474	1,326
Panel B: Mathematics			
Girls	1.64	0.522	-0.002
	(6.18)	(8.58)	(0.043)
Boys	7.26	13.25	0.015
	(6.1)	(8.53)	(0.053)
Observations	2,297	2,365	1,183

Note: Estimates significance level *** p<0.01, ** p<0.05, * p<0.1. We report standard errors in parenthesis.
Variables which enter the propensity score include: age, number of siblings, household size, lives with mother, lives with father, health problems, whether currently working, household ownership of DVD, television, MP3 player, cellular, phone, computer, calculator, student teacher ratio, whether the school has a computer, whether the school has rooms with more than one grade, average teacher education levels and marginality index components. Community level variables include the margination index, proportion illiterate, proportion of households without piped water, proportion without electricity and proportion without drainage.

In summary, the results for sixth grade children enrolled in schools in medium to very high marginalized areas show significant increases in enrollment, effects which persist through high school. With respect to secondary schools attended the results indicate that the program overall increases the probability of attending a tele-secondary school and reduces the probability of attending a technical school. At the high school level, the program increases the probability of attending a general high school and reduces the probability of attending a technical high school. There is no effect of the program on enrollment in private schools. With respect to effects on achievement, the results are generally small and variable according to the estimator used. On balance, the results are statistically significant but not large in magnitude, although, due to selection in enrollment, the estimates are susceptible to downward bias.

Results for indigenous children show similar trends, but somewhat larger. Enrollment effects are on the order of 10 to 12 percentage points increases in the probability of attending secondary school and on the order of 8 percentage points for high school. Effects on achievement were larger in Spanish at about 0.1 standard deviation but generally not significant in the area of mathematics. Effects of 0.1 standard deviation however compare favorably to a recent survey of impacts of educational interventions around the world (see https://www.povertyactionlab.org/research-resources/cost-effectiveness).

VI. Conclusions

This paper uses nationwide information on ENLACE achievement tests to test whether the PROSPERA program, which has been shown in numerous evaluation studies to increase school enrollment, leads to significant increases in learning. We construct a panel of sixth graders and follow their achievement up through twelfth grade, comparing achievement of beneficiaries to non-beneficiaries. The data set we construct also allows us to estimate impacts of PROSPERA on enrollment and on the type of school in which PROSPERA beneficiaries enroll. We study impacts by gender for the entire sample of sixth graders in schools in medium to very high poverty communities and for the subsample of sixth graders in indigenous schools, which we use as a proxy to measure impacts for the indigenous.

Estimated enrollment effects confirm large effects of participating in the PROSPERA program. Although positive effects have been documented before, our results are innovative in that they are based on administrative rather than parental self-reported information on children's school attendance. The results are however quite similar to previous studies based on parental reports, showing important and significant effects on enrollment both in secondary and high school, for girls and boys and for the indigenous and non-indigenous. The analysis of type of school enrolled showed that, for secondary schools, the program increases the probability of enrolling in a tele-secondary school but does not affect the probability of enrolling in a general school. At the high school level, the probability of attending a general school increases relative to the probability of attending a technical school.

To estimate effects on achievement, we use specifications where the outcome measure is the difference in test scores. These differenced models essentially allow for individual fixed effects, which are assumed to be sufficient to control for selectivity in test-taking. Such control is necessary because some children are induced to take the achievement tests by the PROSPERA program. Our results indicate increases in achievement on the order of 0.03 to 0.1 standard deviations both for boys and girls in the sample of children in marginalized areas and on the order of 0.1 on Spanish tests, with statistically insignificant results in mathematics for the indigenous in secondary schools. We did not observe changes in achievement on the twelfth grade ENLACE, but we were constrained by being unable to construct a continuous measure of differences in achievement across years unlike the cases for seventh and ninth grade. In sum, these results provide evidence of small positive effects on achievement due to PROSPERA and provide evidence that the increases in school enrollment led to greater academic achievement. The achievement impacts we estimate are conditional on an attained level of schooling. Thus, they capture effects of PROSPERA participation on achievement that are distinct from the general effect of schooling on achievement.

Several questions remain for further research, particularly with regard to school quality. What is the relationship between type of school attended and achievement and how are effects on achievement interrelated with type of school attended and school quality? What determines the school at which beneficiaries and non-beneficiaries enroll and what are the effects of available

supply, distance and the enrollment processes in determining observed secondary and high school choices? And finally, why do indigenous children show significant effects on learning in Spanish but not in mathematics?

REFERENCES

Acevedo Ivonne, Araceli Ortega and Miguel Székely. 2018. Rendimiento Escolar y Transiciones Laborales con Transferencias Condicionadas en México. Mimeo.

Araujo, M. Caridad, M. Bosch, and N. Schady. 2017. "Can Cash Transfers Help Households Escape an Inter-Generational Poverty Trap?" In C.B. Barrett, M.R. Carter, and J.P. Chavas, eds., *The Economics of Asset Accumulation and Poverty Traps*. Chicago: University of Chicago Press.

Attanasio, Orazio, Costas Meghir and Ana Santiago 2012. Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa. *The Review of Economic Studies*, 79 (1): 37-66.

Backhoff , E. y Contreras Roldán, S. 2014. "Corrupción de la medida e inflación de resultados de enlace", Revista Mexicana de Investigación Educativa, número 63, volumen xix, octubre-diciembre. http://www.comie.org.mx/v1/revista/portal.php?idm=es&sec=SC03&&sub=SBB&criterio=ART63011.

Baird, S., Ferreira, F., ¨ Ozler, B., and Woolcock, M. 2014. "Conditional, Unconditional and Everything in Between: A Systematic Review of the Effects of Cash Transfer Programs on Schooling Outcomes," *Journal of Development Effectiveness*, Vol. 6(1), pp. 1-43.

Barham, T., K. Macours, and J.A. Maluccio. 2017. "Are Conditional Cash Transfers Fulfilling Their Promise? Schooling, Learning, and Earnings After 10 Years." Mimeo, Paris School of Economics.

Behrman, J.R., P. Sengupta, and P.E. Todd. 2005. "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment." *Economic Development and Cultural Change* 54: 237-276.

Behrman, J.R., S.W. Parker and P.E. Todd. 2009. "Medium-Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico." In S. Klasen and F. Nowak-Lehmann, eds., *Poverty, Inequality, and Policy in Latin America*. Cambridge, Massachussets: MIT Press, pp. 219–270.

Behrman, J.R., S.W. Parker and P.E. Todd. 2011. Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A Five-Year Follow-up of PROGRESA/Oportunidades." *Journal of Human Resources* 46(1): 93-122.

Behrman, Jere R. and Barbara L. Wolfe. 1984. "The Socioeconomic Impact of Schooling in a Developing Country" *Review of Economics and Statistics*, 66:2 (May), 296-303.

Behrman, Jere R. and Barbara L. Wolfe. 1987b. "Investments in Schooling in Two Generations in Pre-Revolutionary Nicaragua: The Roles of Family Background and School Supply", *Journal of Development Economics*, 27:1-2 (October), 395-420 (reprinted in International Trade, Investment,

Macro Policies and History: Essays in Memory of Carlos F. Diaz-Alejandro (eds., Pranab Bardhan, Jere R. Behrman, and Albert Fishlow), Amsterdam: North-Holland, 1987), 395-420.

Bobonis, G.J. 2011. "The Impact of Conditional Cash Transfers on Marriage and Divorce." *Economic Development and Cultural Change* 59(2): 281-312.

Calderon, David. 2015. Debate EVALUACIÓN, VALIDEZ Y RELEVANCIA Un comentario crítico a Corrupción de la medida e inflación de los resultados de ENLACE", de Eduardo Backhoff y Sofía Contreras Roldán. RMIE, 2015, VOL. 20, NÚM. 64, PP. 295-308 (ISSN: 14056666).

Card, D. and A. Krueger. 1992a. "School Quality and Black-White Relative Earnings: A Direct Assessment", *Quarterly Journal of Economics*, 107, No. 1 (Feb., 1992), pp. 151-200

Card, D. and A. Krueger. 1992b. "Does School Quality Matter: Returns to Education and the Characteristics of Public Schools in the United States", *Journal of Political Economy*, 100, 1, 1-40.

De Hoyos, Rafael, Ricardo Estrada, Maria JoseVargas. 2018. "Predicting Individual Wellbeing Through Test Scores: Evidence from a National Assessment in Mexico". manuscript.

Elacqua, Gregory. 2006. "Enrollment Pracitices in Response to Vouchers: Evidence from Chile", Occasional Paper 125, Columbia University, Teachers College, National Center for Study of Privatization in Education.

Evans, William N. and Robert Schwab. 1995. "Finishing High School and Starting College: Do Catholic Schools Make a Difference", *Quarterly Journal of Economics*, Vol. 110, No. 4, 941-974.

Fiszbein, A., and N.R. Schady. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington, DC: World Bank.

Gallego, Francisco. 2002. "Competencia y Resultados Educativos: Teoría y Evidencia Para Chile", *Cuadernos de Economía*, 39, 118, 309-352.

Heckman, James, Anne Layne-Farrar and Petra Todd. 1996. Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings, *Review of Economics and Statistics*, November.

Heckman, James J., Stixrud, Jora and Urzua, Sergio. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior", *Journal of Labor Economics*, Vol. 24 Issue 3, p411-482, 72p.

Mancera Corcuera C, Serna Hernández L, Priede Schubert A. Modalidad educativa y organización multigrado como factores asociados con las brechas de aprendizaje de los becarios del Programa Oportunidades (primaria y secundaria en 2007). En: Evaluación externa del Programa Oportunidades

2008. A diez años de intervención en zonas rurales (1997-2007). Tomo iii, El reto de la calidad de los servicios: resultados en educación. México, DF: Secretaría de Desarrollo Social, 2008.

OECD. 2013. *PISA In Focus: What Makes Urban Schools Different*. Paris: Organization for Economic Co-operation and Development.

Parker, S., and P. Todd. 2017. "Conditional Cash Transfers: The Case of Progresa/Oportunidades." *Journal of Economic Literature* 55(3): 866-915.

Parker, S., and T. Vogl. 2018. "Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico." NBER. Feb. 2018.

Schultz, T.P. 2004. "School Subsidies for the Poor: Evaluating a Mexican Strategy for Reducing Poverty", *Journal of Development Economics,* 74(1): 199-250.

Skoufias, Emmanuel, and Susan W. Parker. 2001. Conditional Cash Transfers and Their Impact on Child Work and Schooling: Evidence from the PROGRESA Program in Mexico, *Economia*, 2 (1): 45-86.

Snilstveit, B, Stevenson, J, Phillips, D, Vojtkova, M, Gallagher, E, Schmidt, T, Jobse, H, Geelen, M, Pastorello, M, and Eyers, J. 2015. Interventions for improving learning outcomes and access to education in low- and middle- income countries: a systematic review, 3ie Systematic Review 24. London: International Initiative for Impact Evaluation (3ie).

Wollack, James. 2003. Comparison of answer copying indices with real data. Journal of Educational Measurement, 40:189–205.

World Bank. 2018, ASPIRE: The Atlas of Social Protection Indicators of Resilience and Equity.