Prospering through *Prospera*: CCT Impacts on Educational Attainment and Achievement in Mexico*

Jere R. Behrman Susan W. Parker Petra E. Todd Weilong Zhang

October 1, 2021

Abstract

This paper develops and estimates a dynamic model of student enrollment, school choice, academic achievement and grade progression to evaluate the impacts of Mexico's conditional cash transfer program *Prospera* on educational outcomes over grades 4-9. Academic achievement is measured by nationwide standardized test scores in mathematics and Spanish. Enrollment decisions are the outcomes of sequential decisions at each age from individuals' feasible choice sets, determined by the types of schools locally available and local-labor-market opportunities. The achievement production function has a value-added structure. Model parameters are estimated by maximum likelihood using nationwide administrative test-score data (the ENCEL data) combined with survey data from students and parents, census labor-market data, and geo-coded school-location data. The estimation approach controls for selective school enrollment in different types of schools, grade retention and unobserved heterogeneity. The results show that the *Prospera* program increases school enrollment and academic achievement for program beneficiaries in lower-secondary school grades (grades 7-9). The average test-score impacts are 0.09-0.13 standard deviations in mathematics and 0.03-0.05 standard deviations in Spanish. Students from the most disadvantaged backgrounds experience the largest impacts. The availability of telesecondary distance-learning schools is shown to be an important determinant of the *Prospera* program's impacts on educational outcomes.

*Jere R. Behrman is the William R. Kenan, Jr. Professor of Economics and Sociology and a Research Associate of the University of Pennsylvania's Population Studies Center, Population Aging Research Center and Penn Institute for Development Research. Susan W. Parker is a Professor of Public Policy at the University of Maryland and Affiliate Professor at CIDE in Mexico City. Petra Todd is the Edmund J. and Louise W. Kahn Term Professor of Economics at the University of Pennsylvania and a member of the University of Pennsylvania Population Studies Group, NBER and IZA. Weilong Zhang is an Assistant Professor at the University of Cambridge. We are grateful for financial support from NSF award #1948943 and from the University of Pennsylvania School of Arts and Sciences internal grants "Making a Difference in Diverse Communities" and the "Dean's Global Inquiries Fund". We also thank Gabrielle Vasey, Rodrigo Deiana, Elizaveta Brover, Pinar Gupta, Mira Potter-Schwartz, and Erika Trevino-Alvarado for research assistance. We thank Hector Robles Vaquez for preparing databases and for assistance in working with these data. We thank Miguel Szekely for helpful conversations about the Mexican educational system. This paper was presented at the University of Cambridge, the University of Arizona, the University of Maryland, the University of Pennsylvania, Stanford Institute for Theoretical Economics (2021) and University of Glasgow. We thank Eric French, Magne Mogstad, and Christopher Taber for suggestions.

1 Introduction

Conditional cash transfer (CCT) programs were first introduced in Brazil and Mexico 25 years ago and have since spread around the world. The programs aim to alleviate current poverty and reduce future poverty by encouraging investment in the human capital of children and youth from poor families. In the early stages of the program, a large-scale randomized evaluation of the Mexican *PROGRESA* program demonstrated substantial impacts on educational attainment, child work and family income. These findings contributed to both a large scaling-up within Mexico and an impressive adoption of similar programs around the world.[1] [2]

Several studies examined the impacts of *PROGRESA/Oportunidades/Prospera* on school enrollment and, in some cases, on longer-term educational attainment (See, e.g., Schultz (2004), Behrman et al. (2005b), Behrman et al. (2005a), Behrman et al. (2009), Todd and Wolpin (2006), and Attanasio et al. (2012)). This literature demonstrated positive CCT program impacts on school enrollment and educational attainment. However, a longstanding concern has been whether and to what extent increased school enrollment translates into higher academic achievement, presumably important for the program to significantly affect earnings potential.

With newly available data, we are now able to examine the effects of the *Prospera* program (the program name during the time of our data collection) not only on enrollment and educational attainment but also on academic achievement in mathematics and Spanish. Nationwide standardized and longitudinal administrative test score data (called the *ENLACE* data) as well as complete enrollment rosters were merged with administrative information on which students come from *Prospera* households and on school locations. They were also merged with survey information obtained from students and their parents. These data allow the study of how students' Prospera beneficiary status affects their school enrollment, grade progression and academic achievements over time.

A major challenge in evaluating CCT program impacts on academic achievements is selective school enrollments.[3] In primary school (up through grade 6), school enrollment rates are high and dropout is rare. However, dropouts increase substantially after grade 6 when students choose whether to enroll in lower-secondary school and also in which type of school to enroll. The ENLACE standardized tests are administered in schools, so test scores are only observed for enrolled children. It has been shown that the *Prospera* CCT program induced students from high-poverty backgrounds who were at high risk for dropping-out to stay in school longer. If weaker students remain in school, then average test scores could fall as a result of more marginal students being included in the testing. This selection problem arises whenever tests are administered in school and poses a challenge in evaluating program effects on academic achievement, regardless of whether the data are experimental or nonexperimental.[4] The

[1] CCT programs with similar features have now been implemented in over 60 countries on five continents. See Fiszbein and Schady (2009).

[2] Parker and Todd (2017) review the literature on the development, evaluation, and findings of the *PROGRESA/Oportunidades/Prospera* program. In 2018, the program was rolled back by the administration of Lopez Obrador and substituted with scholarship programs from SEP.

[3] Cameron and Heckman (1998), Cameron and Heckman (2001) and Glewwe (2002) discuss this selection problem in the context of analyzing the determinants of educational outcomes.

[4] This selection problem also affects cross-country comparisons of standardized tests, such as PISA test scores. The PISA tests are given at age 15 and, in some countries, a significant fraction of children have dropped-out by that age.

selection problem is dynamic as it occurs at each grade and the students at risk for dropping-out in a particular grade depend on the sample that stayed in school from the previous grade.

The goal of this paper is to examine how the *Prospera* program affects school-going and academic achievements, accounting for dynamic selection. To this end, we develop and estimate a dynamic model of students' school progression that incorporates decision-making in each grade with regard to enrollment, school choice, and dropping-out as well as grade- and subject-varying production functions for academic achievement. The model begins when students finish fourth grade and continues until the ninth grade, which is the end of lower-secondary education in Mexico.[5] About one third of students dropout before finishing lower-secondary school (grade 9). Among those who finish grade 9, only about half eventually finish higher-secondary school (grade 12). Our model captures work-related motivations for dropping-out using local labor-market information on the potential wages that students could earn if they dropped-out. We use geo-coded school locational information to account for the set of schools locally available to students and the travel distances to different kinds of schools. Very few students in the *Prospera* program attend private schools, so our analysis focuses on students who are either not attending school or attending public schools.

As noted, our multi-equation modeling framework includes production functions for achievements in mathematics and Spanish that are specified as value-added models with coefficients that vary by grade, school type attended, and grade-retention status. The linked dynamic panel specification captures the notion that skill accumulation at one stage in life raises skill attainment at other stages, which has been shown to be essential to characterizing human-capital-skill formation processes.(See, e.g. Cunha et al. (2006)). In primary school, most children attend schools close to home. Their parents can choose from general primary schools or bilingual indigenous schools, depending on local availabilities. In lower-secondary school, students/parents can choose from up to three types of public schools – general schools, technical schools, or telesecondary schools – all of which are academically oriented. Technical schools differ from general schools in incorporating vocational/technical educational components in the curriculum. Telesecondary schools are distance-learning schools that largely serve students living in rural communities and that enroll almost 20% of lower-secondary school students. Children/youth from *Prospera* beneficiary families attend telesecondary schools in greater proportions than average. Section two provides more detail on how these schools differ from general schools.

The model we estimate controls for selection arising from school-enrollment, dropout, grade-retention, and school-type choices, all of which potentially affect students' grade progressions and academic achievements. The model incorporates rich observed heterogeneities by allowing student gender and family demographics to affect preferences for schooling as well as the production functions for achievements. It also incorporates unobserved heterogeneity in the form of unobserved discrete multinomial types, as in Heckman and Singer (1984) and Cameron and Heckman (1998).[6] These types enter multiple model equations and, in doing so, allow for correlated error structures. The data also contain information on a small percentage of students suspected to have copied answers on the

[5] There are no nationwide standardized tests in the 10th and 11th grades.

[6] As discussed in Todd and Wolpin (2003) and Rivkin et al. (2005), in estimating value-added models, it is important to control for unobserved inherent student abilities, corresponding to cognitive skills, personality traits, or motivation, that may affect a child's achievement growth.

multiple-choice standardized tests. Copying induces one-sided measurement error in the dependent, the lagged independent variable, or both, which we take into account in the estimation. Model parameters are estimated by maximum likelihood where the outcomes at different ages/grades are school enrollment, mathematics and Spanish test scores, dropping-out, and grade retention.

We use the estimated model to evaluate how *Prospera* beneficiary status affects schooling progressions and academic achievements in different grades. In particular, we simulate school-choice decisions, working decisions and academic-achievement with and without the *Prospera* program, for different groups of children. There are multiple channels through which *Prospera* participation can affect these outcomes. First, past participation may increase lagged achievements, which can facilitate present learning. For example, greater comprehension of sixth-grade mathematics can facilitate learning and comprehension of the seventh-grade curriculum.[7] Second, contemporaneous program participation can directly affect learning if the program encourages regular school attendance, student engagement and study efforts. There are two reasons why we might expect the program to influence students in this way. *Prospera* program rules stipulated that children must attend school at least 85% of days and can only fail a grade once to receive the cash transfers. Additionally, *Prospera* transfers could have reduced the pressure on children/youth to work in labor markets while in school and thereby allow for greater focus on schoolwork.

Our analysis yields a number of findings regarding *Prospera*-program effects and the effectiveness of different school types. In primary-school grades (grades 5 and 6), we do not find evidence of program impacts on test scores. In lower-secondary grades (grades 7, 8 and 9), however, there are positive and statistically significant impacts on test scores with larger overall average impacts in mathematics (0.09-0.13 standard deviations) than in Spanish (0.03-0.05 standard deviations). Also, program-participation effects accumulate over time and increase with longer exposure. Interestingly, we find significantly larger test-score impacts for children from more disadvantaged backgrounds. The estimated average impacts are similar by gender.

Our analysis also shows that telesecondary schools are important determinants of *Prospera*-program impacts. The production-function estimates indicate that telesecondary schools are effective in producing test-score gains for the students attending them, particularly in mathematics. When we use the estimated model to analyze the effect of removing the telesecondary option from the set of schooling options, we find that, in the absence of these distance-learning schools, the dropout rate would be substantially higher and average educational-attainment levels lower. Our results are consistent with the difference-in-difference analysis of Navarro-Sola (2019) that found that the expansion of telesecondary schools led to substantial increases in educational attainments for local students.

We also find that failing to control for the dynamic selection would lead to underestimation of the program's impact. In particular, we estimate a simpler value-added model by grades, without controlling for selection from multiple sources (dropout, school choice, grade retention). Comparing the results to those derived from our richer model, the cumulative program impacts are noticeably smaller. We identify two sources that cause these downward biases. First, the program causes students at the margin of dropping-out to stay in school longer and not controlling for dropping-out leads to down-

[7] Cunha et al. (2006) term this feature of cognitive achievement production functions "self-productivity."

ward bias in the impact estimates. Second, the simpler model does not allow heterogeneous impacts across different types of schools. Consequently, it does not capture that *Prospera* beneficiaries attend telesecondary schools in greater numbers and that these schools are particularly effective in teaching mathematics. Overall, we find the richer modeling framework is required to capture heterogeneous program impacts and to control for sources of selection bias.

The paper develops as follows. Section two provides a brief description of the Mexican school system and of the datasets used in this study. Section three discusses three different strands of related literatures. Section four describes the model and section five the estimation approach. Section six presents the empirical results. In section seven, we compare the different types of schools in terms of measured quality and in terms of student engagement to explore possible mechanisms underlying the estimated *Prospera* program impacts. Section eight compares the estimates obtained from our modeling approach to those derived from a simpler model to examine the importance of controlling for selection. Section nine concludes.

2 Background and related literature

2.1 Mexican educational system and child-labor laws

The Mexican educational system consists of three levels: primary, secondary, and tertiary education. Formal basic education includes preschool, primary school (grades 1-6), and lower-secondary school (grades 7-9), all of which are compulsory. However, compulsory schooling laws are not well-enforced. Many children dropout before completing grade 9, particularly children from lower-SES families, indigenous backgrounds and rural areas.

Primary school is considered to be a part of "Basic Education" and public primary schools are free of charge. The Secretariat of Public Education (SEP) standardizes curriculum content, which includes Spanish, mathematics, natural sciences, history, geography, art, and physical education. The National Institute for Assessment of Education (INEE) monitored standards during our period of study. Secondary school is divided into lower-secondary school (grades 7-9) and upper-secondary school (grades 10-12). Lower-secondary school is free and students may follow either a general academic track or a technical track, which has more of a vocational focus. Both tracks are designed to prepare students for further education. There are fewer lower-secondary schools than primary schools and attending lower-secondary schools often requires traveling some distance from home, particularly for children living in more remote areas. Public schools do not generally provide transportation for students. Upper-secondary education (grades 10-12) is not compulsory. Many upper-secondary schools are affiliated with large public universities, while others are SEP or state-controlled, and there are also private options. At the tertiary level, the Mexican educational system is similar to the United States' system, with many different programs and degree options.

The Mexican Constitution prohibits child labor for minors under 14 years of age. However, the child-labor laws are not well enforced. 8% of children age 12 report working for pay in the 2010

Mexican census data.[8]

2.2 *ENLACE* test score data

From 2006 to 2013, the SEP applied the *Evaluación Nacional de Logro Académico en Centros Escolares*, called the *ENLACE*. The test gathered information at the end of each academic year on student performance in mathematics, Spanish and a rotating subject for all third-to-ninth graders in private and public schools. Beginning in 2008, *ENLACE* was also given to students in their final year of upper-secondary 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 test completion rate is close to 90%. As described by De Hoyos et al. (2018), 15.1 million students in 136,000 schools took the examination in 2013, the last year the test was applied.

De Hoyos et al. (2018) analyze a longitudinal panel of *ENLACE* data for a group of students that took the test in grade 6 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 educational and labor-market outcomes. Additionally, -hey 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 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 standard deviations. They also show 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 sub-sample of twins.

3 Related literatures

Our analysis builds on three different strands of related literatures. One is the literature that estimates the impacts of the Mexican CCT program (*PROGRESA/Oportunidades/Prospera*) on educational outcomes. The second is the literature that develops and implements value-added models as a way of studying the dynamics of academic achievements. The third is the literature that develops and estimates school-choice models. Our modeling framework incorporates empirical strategies developed in each of these literatures.

3.1 Studies of the effects of *PROGRESA/Oportunidades/Prospera* on educational outcomes

As previously noted, Mexico's CCT program PROGRESA was initially evaluated using a randomized experiment, which ran for two years in rural areas. Early evaluation studies using the experimental data

[8]Based on the authors' tabulations.

demonstrated positive effects on improving school enrollment, reducing grade repetition and increasing completed grades of schooling (Schultz (2004), Behrman et al. (2005b) and Parker and Todd (2017)). Enrollment impacts were higher at the transition from primary to lower-secondary school; however, younger children experienced reductions in grade repetition and better grade progression (Behrman et al. (2005b)). At the lower-secondary school level (grades 7-9), the program reduced dropout rates. Both Schultz (2004)) and Behrman et al. (2005b) used their short-term estimates to predict overall schooling attainment effects. Both studies predicted an overall increase of 0.6 grades for a child receiving PROGRESA grants from primary school through lower-secondary school (grades 3-9).

More recent studies have explored alternative methodologies and data to study longer-run *Prospera*-program impacts. Behrman et al. (2009) and Behrman et al. (2011)) used matching estimators applied to data gathered in urban and semi-urban areas and found that extended time participating in the program led to significant improvements in grades completed, about one full grade for children who participate for six years beginning at ages nine to 12, compared to non-participating children. Studies based on estimation of structural dynamic models found program effects of a similar magnitude (Todd and Wolpin (2006) and Attanasio et al. (2012)). Using national census data from 2010, Parker and Vogl (2018) found that long-term impacts on total schooling attainment for children who grew-up in a household with *Prospera* benefits were 1.0-1.6 grades for women and 0.6-1 grades for men, compared with children from households not offered the program.

Overall, these diverse studies provide consistent evidence that the program significantly increased schooling levels. However, two important dimensions have not yet been studied. First, almost all existing studies are based on self-reported school-enrollment information from household-level surveys and are therefore potentially subject to reporting biases. Beneficiary households, knowing that benefits are conditional on their children's regular school attendance, might over-report their children's enrollment. Second, prior studies did not analyze academic achievement impacts, because the original data collection did not include achievement test scores.[9] This study uses the administrative education data that should not be subject to reporting biases. The data contain information on enrollment at the beginning of the school year, on attendance on the day of the *ENLACE* test administration and on the test score results. In addition, we make use of survey data on individual and household characteristics that were collected each year from students and parents for a large random sub-sample of schools.

There are several reviews of transfer programs that survey the evidence of enrollment and achievement impacts. An early book by Fiszbein and Schady (2009) summarized the effects of CCT programs implemented in various countries on achievement tests and concluded that results have been disappointingly small. Subsequent studies have found somewhat mixed evidence. Baird et al. (2014) reviewed educational effects of both conditional and unconditional transfers and concluded that achievement impacts are "small, at best." Snilstveit et al. (2017) carried-out a meta-analytic review of educational interventions and concluded that while cash transfers have been effective at increasing school enrollment, they have modest effects on achievement. The programs examined include both conditional and

[9] In 2003, a version of the Woodcock Johnson tests in mathematics and Spanish was applied to a single cross-section. Using these data, Behrman et al. (2009) found no impacts of PROGRESA participation on achievement, based on comparing test scores of the original treatment and control groups. However, their study was limited in its ability to control for selective school enrollment due to the absence of any lagged test-score data.

unconditional cash transfer programs. There are much fewer studies of achievement than there are of enrollment. For instance, Snilstveit et al. (2017) reviewed 38 studies of the effects of transfers on enrollment; only 11 of the programs analyzed effects on achievement.

Some transfer programs have been shown to significantly affect achievement. For instance, Baird et al. (2014) found significant impacts of 0.14 standard deviations on English achievement tests and 0.12 standard deviations on mathematics achievement tests after a Malawi CCT program operated for two years. Barham et al. (2013) used the randomized phase-in of the Red de Proteccion Social CCT program in Nicaragua to study long-term effects on educational attainment and learning for boys 10 years later. They found a half-grade increase in schooling and substantial gains (approximately 0.25 standard deviations) in mathematics and language achievement scores. Using data from an RCT, Hadna and Kartika (2017) found statistically significant effects of a CCT program called *Program Keluarga Harapanin* in Indonesia on three subjects (Bahasa Indonesia, mathematics and English) as well as national mathematics examinations for junior-high-school students.[10]

3.2 Value-added modeling literature

There is a large literature dating back at least to Summers and Wolfe (1977) that uses value-added models to estimate school, teacher and parental contributions to student learning. Many studies interpret the model as an educational production function, which guides the choice of included variables and interpretation of their effects. The theoretical justification is the cumulative-achievement model developed by Boardman and Murnane (1979), Hanushek (1979), Todd and Wolpin (2003), Cunha et al. (2006) and Cunha et al. (2010), where current achievement is a function of a student's entire history of educational and family inputs. As those papers discuss, skill formation is a life-cycle process and early skill accumulation facilitates later accumulation. Assumptions on the cumulative model can be used to justify a value-added specification; in some cases the assumptions are testable (see, e.g., Todd and Wolpin (2003)).[11]

There is some debate in the literature about whether value-added models should be used to measure the effectiveness of individual teachers. A few recent papers compare estimates of teacher effectiveness derived from value-added models to test-score gains derived under experimental conditions (where students were randomly assigned to teachers) and find the models to be reliable.(e.g., Kane and Staiger (2008), Kane et al. (2013), Chetty et al. (2014a)). Value-added test-score gains have also been shown to be highly predictive of adult outcomes (Chetty et al. (2014b)). Our data are actually richer than the administrative data that are often used to estimate value-added models, because we have detailed student, family, teacher and school-level information derived from surveys.

[10]Other studies of the effects of CCT programs on cognitive-achievement measures include Andersen et al. (2015), Paxson and Schady (2010) and Simoes and Sabates (2014).

[11]Usually, it is assumed that the a lagged achievement score is a sufficient statistic for lagged input histories. Lagged input histories, particularly back to the start of education, are very rarely available in datasets.

3.3 The school-choice literature

As previously described, children/youth in Mexico can choose from multiple primary/secondary school options depending on the schools that are locally available. The school district usually assigns students to a default school that is close to their home. If they do not wish to attend that school, they can submit an application with a request to attend a different school.[12] There are two papers studying the school choices/school quality using the same datasets as used in this paper. Vasey (2021) develops a static-choice model over discrete school-work alternatives (school only, school combined with work, and work only) for students who finished primary school. She models the optimal choice of study effort under the first two alternatives and uses the estimated model to study the effects of better child-labor-law enforcement. Borgheson and Vasey (2021) use the same data and a marginal-treatment-effects (MTE) approach to evaluate the effectiveness of telesecondary schools relative to other types of schools for 7th graders. The treatment-on-the-treated (TT) and average-treatment effects (ATE) estimates show that telesecondary schools are effective in increasing achievement.[13]

The school-choice model that we develop and estimate builds on an extensive literature using multinomial discrete-choice models to model school choice. An early paper by Neal (1997) estimated the effects of Catholic schools on student outcomes in the United States using an instrumental-variables approach to address endogeneity of the Catholic-school attendance decision. Another paper by Altonji et al. (2005) estimated the effect of Catholic schools using a method that bounds the potential selectivity bias due to unobserved variables by the degree of selection on observed variables.

There are a number of school-choice models estimated using Chilean administrative test-score data. McEwan (2001) and Sapelli and Vial (2002) analyzed test-score differences in public and private schools in Chile using control-function approaches to account for selection. Gallego and Hernando (2009) analyzed school-choice determinants in Chile and found that proximity to schools and school-test scores were the two most important attributes that families considered when choosing schools. Additional studies developing school-choice frameworks to estimate the effects of school vouchers include Rouse (1998), Figlio and Rouse (2006) for the United States and Hsieh and Urquiola (2006), Bravo et al. (2010), Angrist et al. (2002), and Angrist et al. (2006) for Latin America.

A separate strand of the school-choice literature develops models for the purposes of studying parents' preferences for school quality and analyzing the welfare effects of school policies. For example, Epple et al. (2018) estimated parents' preferences for each school in their choice set and used the model to study the welfare consequences of a superintendent closing a certain percentage of schools.[14].

[12]The procedure for choosing schools varies somewhat across states and schooling levels, but students (or their parents) typically submit a ranking of up to three alternative schools and their eventual assignment depends in part upon space availability and whether they have siblings attending that school.

[13]Vasey and Borgheson are Penn graduate students who work on our research team.

[14]See also research on estimating parental preferences by Hastings et al. (2009) and, more recently, Neilson et al. (2019)

4 Model

Our modeling framework combines a model of school attendance decisions at different ages/grades and at different types of schools with a model of academic achievement in mathematics and Spanish. We specify test-score gains from year-to-year using a value-added framework that relates current achievement to lagged achievement, to family and school inputs into the learning process, and to student-level unobserved heterogeneity (e.g. arising from ability or preferences). Our framework also allows for dropping-out of school, which is substantial in the Mexican context, and for grade retention. Our empirical framework is quasi-structural. The educational production function has a structural interpretation as a technology relating inputs to outputs. However, the school-choice model is reduced form. The school-enrollment choices and grade-retention outcomes likely reflect decisions of students, parents and school administrators, and it would be difficult to separately model the decisions of all of these entities and how they interact.

Individuals are indexed by i, $i = 1, .., n$ and each model period corresponds to one school year. In primary school, students/parents can choose to attend one of two types of schools: general ($j = 1$) or indigenous (bilingual) ($j = 4$), depending on the types locally available (within 5 km). After grade 6, students who continue with school make decisions about what type of lower-secondary school to attend from the following options: general ($j = 1$), telesecondary ($j = 2$), or technical ($j = 3$). Let $D_{ija} = 1$ if the individual is enrolled in school type j ($j \in \{1, 2, 3, 4\}$) at age a, else $D_{ija} = 0$. Let $D_{i0a} = 1$ if the individual does not enroll in school at age a, else $D_{i0a} = 0$. Let $I_{ia}^{Pass} = 1$ if the individual passes the grade at age a, else $I_{ia}^{Pass} = 0$. We assume enrollment and school-choice decisions are made at the end of the school year. The passing outcome is also realized at the end of the school year, prior to the other decisions being made. The enrollment decision D_{i0a} and school choices D_{ija} are made simultaneously. Let G_{ia} denote the grade that the individual is eligible to attend at age a, which increases by one if the student passes the current grade:

$$G_{i,a+1} = G_{ia} + I_{ia}^{Pass}.$$

We assume that the opportunity costs of being enrolled is the wage w_{ia} that individuals could earn, given their characteristics and geographic locations. Because the test-score databases do not contain information on wages for children/youth who do work, we use the 2010 Mexican census to impute wages to individuals, given their age, schooling attainment, and demographics as well as their region of residence. The imputation procedure includes a selection correction to account for selective labor-force participation. For details, see appendix B.

Lastly, let $P_i = 1$ denote if a child/youth comes from a *Prospera*-beneficiary family, else $P_i = 0$. Eligibility is determined on the basis of family poverty status. Families who learn that they are eligible for the program usually opt to participate, because the program-transfer amounts are substantial.[15] They receive separate subsidies for each child depending on the child's grade-level, gender and whether they attend school.

[15] Prior research showed that transfers received under the program led to an average increase of 20% in family income. See Parker and Todd (2017).

Our model also incorporates permanent unobserved heterogeneity, arising either from preferences or endowments, that enter into all the model components. Following Heckman and Singer (1984) and Cunha and Heckman (2008), we model the unobserved heterogeneity as random effects drawn from a discrete multinomial distribution. Let $\mu_{il} = 1$ if individual i is of type l, =0 else, where $l \in \{1, ..., L\}$ and the proportion of each type is estimated.[16] The unobserved types are assumed to be independent of the initial conditions, $\Omega_i(0)$.

The initial conditions include family background, achievement test scores at grade 4 (when test scores first become available for this sample), age, gender, a marginality index describing the local poverty level, urban/rural residence status, the state of residence, the set of primary and secondary schools available to each child, the minimum distances needed to travel to different types of schools and whether the family is participating in *Prospera*.[17] The time-varying state-space elements in any given time period, $\Omega_i(a)$, consist of whether attended school last year, type of school attended, grades completed thus far, whether retained in the last grade, and lagged achievement-test scores in mathematics and Spanish. We denote the total observed state space, which includes the initial conditions and the time-varying state-space elements, by $\tilde{\Omega}_i(a) = \{\Omega_i(0), \Omega_i(a)\}$.

We next describe how achievements in mathematics and Spanish evolve with school attendance at different types of schools. Let $m = 1$ denote mathematics, $m = 2$ Spanish and g denote the grade level. The value-added model is grade-specific and school-type (j) specific. It also depends on whether the student passed the previous grade $I^{Pass}_{i,a-1} = 1$ or is repeating the grade $I^{Pass}_{i,a-1} = 0$.[18] Let Z^A_{ia} denote the vector of observed characteristics of the youth and of the family that enter the achievement production function.

$$A^m_{ijal} = \delta^{mgI}_{0jl} + A_{i,a-1}\delta^{gI}_{1j} + \delta^{mgI}_{2j} P_i + Z^A_{ia}\delta^{mgI}_{3j} + \omega^{mgI}_{ija}. \tag{1}$$

In this equation, δ^{mgI}_{0jl} is a type-specific intercept that allows for unobserved heterogeneity (l denotes the type). δ^{mgI}_{2j} captures the impact of the *Prospera* program. $A_{i,a-1} = \left\{A^1_{i,a-1}, A^2_{i,a-1}\right\}$ is a 2×1 vector including both the mathematics score and the Spanish score from the previous period $a - 1$ and δ^{gI}_{1j} the vector of associated coefficients. We assume that the error terms in the achievement production functions, conditional on the unobserved types, are iid and normally distributed.

Next, we discuss the schooling choices in each grade. As previously described, students have different choice sets at different grades. At the initial age (a_f, corresponding to grade 4), students choose whether to attend general primary school ($j = 1$) or indigenous (bilingual) primary school ($j = 4$), depending on the local availabilities. Upon finishing primary school at the end of grade 6, students decide whether to attend one of the three types of lower-secondary schools (general ($j = 1$), telesecondary ($j = 2$), or technical ($j = 3$)) or to dropout ($j = 0$), again depending on local lower-

[16] Alternatively, we could impose a specific distribution for μ, for example, a mixture of normal distributions. However, evidence suggests that our current approach performs better. Using Monte Carlo simulations, Mroz (1999) shows the discrete-type assumption performs as well as the normal assumption when the true distribution is normal. When the true distribution is not normal, however, the discrete-type method performs better in terms of precision and bias.

[17] A full list of family background variables includes parental schooling attainments, parental employment statuses, whether both mother and father are present in the household, number of household members, first language spoken at home, internet accessibility, computer accessibility and years of preschool education.

[18] If a student repeats a grade, then the lagged test score pertains to the same grade as in the current time period and would therefore have a different associated coefficient from the case where the lag pertains to the previous grade.

secondary school availabilities (within 10 km). At other grade levels (7,8,9), students choose whether to continue in school or to dropout.[19] The choice set J_{ia}^g at different grades can be summarized as:

$$j_{ia} \in J_{ia}^g = \begin{cases} \{1,4\} \cap M_i^1 & G_a = 4 \ \& \ a = a_f \\ \{0,1,2,3\} \cap M_i^2 & G_{a-1} = 6 \ \& \ I_{i,a-1}^{Pass} = 1 \\ \{0, j_{i,a-1}\} & G_{a-1} \geq 7 \end{cases} \tag{2}$$

where a_f is the age when the student enters into the sample at grade 4. M_i^1 denotes the available local primary-school types and M_i^2 denotes the available local lower-secondary school types.[20] We assume the probability of choosing a schooling option j depends on the student's unobserved type μ_l, the grade level G_{ia}, lagged achievement $A_{i,a-1}$, *Prospera*-beneficiary status P_i, demographic and family-background characteristics, $Z_{ia}^D \in \tilde{\Omega}(a)$, As previously described, we also imputed a potential hourly wage that the child could earn given their characteristics and region of residence, w_{ia}, which is included in Z_{ia}^D. (See Appendix A.3 for details.)

Lastly, we also assume that the school choice depends on the local supplies of different types of schools, S_{ija}. For the lower-secondary school-choice problem, the supply variables are the distance to the closest lower-secondary school of type j and the total number of schools of type j within a 10km radius. For the primary choice problem the supply variable is the number of schools of each type within a 5km radius. Assuming a random-utility model at each age for the options available to that child (depending on his/her grade and geographic location) and assuming Type I extreme-value errors (taste heterogeneity), we obtain that the probability of choosing option j is multinomial logistic:

$$\Pr(D_{ija} = 1 | \tilde{\Omega}(a), \mu_l) = \begin{cases} \frac{\exp(\mu_{0jl}^g + A_{ia-1}\phi_{1j}^g + P_i\phi_{2j}^g + Z_{ia}^D\phi_{3j}^g + S_{ija}\phi_{4j})}{\sum_{j' \in J_a^g} \exp(\mu_{0j'l}^g + A_{ia-1}\phi_{1j'}^g + P_i\phi_{2j'}^g + Z_{ia}^D\phi_{3j'}^g + S_{ij'a}\phi_{4j'})} & \text{if } j_{ia} \in J_a^g \\ 0 & \text{if } j_{ia} \notin J_a^g \end{cases} \tag{3}$$

There are three variables that only enter the schooling-choice equation and do not enter other model equations, which provide natural exclusion restrictions: (i) the imputed hourly wage w_{ia}, (ii) the distance between the individual's primary school and lower-secondary school and (iii) the local supply of schools of each type. These variables affect the probabilities of attending different types of schools but do not directly enter the test-score equations. Below, we describe how the exclusion restrictions are useful for identifying model parameters.

Lastly, we specify a probabilistic model for whether a student passes a grade, which depends on the unobserved type μ_l, the current school type attended j_{ia}, academic knowledge as proxied by the achievement scores A_{ia}, the grade level G_{ia}, *Prospera*- beneficiary status P_i as well as some demographic and family background characteristics, $Z_{ia}^I \in \tilde{\Omega}(a)$:

$$Pr(I_{ia}^{Pass} = 1 | \tilde{\Omega}(a), \mu_l) = \Phi(\gamma_{0l}^g + A_{ia}\gamma_1^g + \gamma_2^g P_i + j_{ia}\gamma_3^g + Z_{ia}^I\gamma_4^g). \tag{4}$$

[19] Because school enrollment is very high during primary school, we assume students do not dropout during primary grades. Therefore, they make no choices about continuing in school until the end of grade 6.

[20] In particular, $M_i^1 \in \{\{1\},\{4\},\{1,4\}\}$ and $M_i^2 \in \{\{1,2,3\},\{1,2\},\{1,3\},\{2,3\},\{1\},\{2\},\{3\}\}$.

We allow the coefficients of the passing probability to be grade-specific. γ_{0l}^{g} is a unobserved type-specific intercept.

4.1 Model interpretation

As previously noted, the value-added achievement function can be interpreted as a production-function technology. Our school-choice model (3), can be interpreted as an approximation to a decision rule derived from a structural model. As in a dynamic discrete-choice schooling model, students' schooling decisions are stage-dependent and the educational choices in early grades affect their options in later grades. However, we are agnostic about the precise rules used by agents to make decisions.[21] In our case, the school-choice probability represents the difference between the value functions associated with alternative schooling-work options.

There are a few advantages of adopting this kind of "quasi-structural" approach. First, if multiple individuals are involved in the school-choice decision (students, parents, school administrators) then implementing a multiple-agent model structurally would be difficult. Second, estimating a dynamic structural schooling model that includes academic achievement is complicated because of the two lagged test scores in the state space. The high dimensional state space would require the use of emax approximation methods. Lastly, the model that we estimate allows for considerable flexibility in how conditioning variables affect decision-making at different ages.

A limitation of this type of quasi-structural approach, however, is that our model does not distinguish between current and future utility components and therefore cannot be used to study effects of policies that might be implemented in the future, such as the effects of transfer amounts given only upon completing certain grades. This is not a limitation in our context, because our interest centers on understanding the effects of the family participating in the *Prospera* program, which we observe in the data.

4.2 Identification

Manski (1988) and Heckman and Navarro (2007) consider semiparametric identification of binary-choice index models in static and dynamic settings.[22] In our model, prior to grade 6, there is no dropout decision and there is only the binary choice over two types of schools. Manski (1988) established identification of binary-choice index models under an assumption that unobserved variables are fully independent or quantile (e.g. median) independent of the regressors, which is satisfied in our model. At the end of grade 6, 7 and 8, students decide whether to dropout. The dropout decisions at the end of grades 7 and 8 are only observed for students who did not dropout in prior grades. In this case, the dropout probability can be semiparametrically identified using the identification-in-the-limit

[21] Todd and Wolpin (2006) and Attanasio et al. (2012) estimate structural schooling models to analyze the effects of the *Progresa* program on enrollment. Leite et al. (2011) develop a model to evaluate impacts of the Brazilian *Bolsa Escola* CCT program. Those models consider school enrollment/working decisions but do not consider the choice of school type or academic achievement. One hallmark of the full structural approach is the future discounted emax function, which is used to explicitly model agent expectations about costs and returns. For discussion of how to approximate the decision rules, see, e.g. Keane et al. (2011); Heckman and Navarro (2007); Heckman et al. (2018).

[22] Also, see Heckman et al. (2016), Heckman et al. (2018) for other applications and extensions.

argument of Heckman and Navarro (2007). Their theorem allows for general error structures that can be correlated over time for individuals but are assumed to be independent across individuals and of regressors in the initial time period. Our permanent-transitory (type) structure is a special case. Their proof requires transition-specific exclusion restrictions. In our context, the imputed wages, distances to different types of schools and local-school availability constitute these exclusion restrictions.

The value-added achievement model is a standard dynamic panel data model, of the kind discussed in Arellano and Bond (1991). In grades 7, 8 and 9, however, the achievement outcomes are only observed for the select group of students who enroll in school. We can again use an identification-in-the-limit strategy and base identification in the subset of students with characteristics such that they advance to the next grade with probability close to one. Within that subgroup, the school type selection can be controlled using either a parametric or semiparametric control function, with the latter requiring exclusion restriction(s) that vary the school type choice and that do not enter the outcome equations directly (which we have).

Value-added achievement models could in principal be identified allowing for individual fixed effects. However, we observe individuals at most 6 years, which is not a long enough panel to reliably estimate fixed effects. This is an important rationale for adopting the unobserved multinomial types specification.

5 Estimation

5.1 Some measurement issues

In this section, we discuss two issues related to test-score measurement. First, some enrolled students are missing test-score data, because they were absent the day of the test, perhaps due to illness. We use the notation $I_{ia}^{miss} = 1$ if the student's test score at age a is missing, else $I_{ia}^{miss} = 0$. Second, students' true test-score performances may be mis-measured if there is cheating. Usually, researchers do not have information on cheating, so the issue of potential cheating in estimating value-added models is usually ignored. However, SEP uses a statistical algorithm to detect potential cheating (in the form of copying) and includes in the database information on which students are likely to have cheated.[23]

To account for possible test-score distortion due to cheating, we specify a measurement equation for the relationship between the true test score A_{ia}^m and the observed test score $\tilde{A}_{ia}^m$:

$$\tilde{A}_{ia}^m = \begin{cases} \left(1 + c_{ija}^{mgI} \zeta_{ia}^m I_{ia}^{cheat}\right) A_{ia}^m & \text{if } I_{ia}^{miss} = 0 \\ \text{Not observed} & \text{if } I_{ia}^{miss} = 1 \end{cases} \tag{5}$$

where the effect of cheating is captured by the term $c_{ija}^{mgI} \zeta_{ia}^m$. c_{ija}^{mgI} captures the average cheating effect for a given subject m, grade g and retention status I, while $\zeta_{ia}^m \sim \log\text{normal}(-0.5\sigma_\zeta^2, \sigma_\zeta^2)$ captures the

[23] In particular, SEP estimates a student's probability of cheating using both the K-index and the Scrutiny method. If the software detects a possible cheating case from the response patterns, a note is added to the student's report (Gonzalez, 2015). Note that the exam is not deleted, and there are no consequences to the students; the cheating factor is purely informative (SEP, 2010).

randomness of the cheating effect.[24] When students do not cheat ($I_{ia}^{cheat} = 0$) and the test scores are not missing ($I_{ia}^{miss} = 0$), then, under this specification, the observed test score equals the true test score:

$$\tilde{A}_{ia}^{m} = A_{ia}^{m} \quad \text{iff} \quad I_{ia}^{cheat} = 0 \text{ and } I_{ia}^{miss} = 0$$

5.2 The likelihood function

The outcomes observed for individuals are the enrollment choices at each age, D_{ija}, whether the student passes the grade attended at age a I_{ia}^{Pass}, the grades completed for each age G_{ia}, and the observed achievement test scores in mathematics and Spanish at each age a, $\tilde{A}_{ia} = \{\tilde{A}_{ia}^{1}, \tilde{A}_{ia}^{2}\}$. Let a_i^f and a_i^l denote the first and the last ages at which we observe the individual i in the dataset. The state space elements $\tilde{\Omega}(a)$ (described above) include lagged test scores, whether passed last grade, grades completed, school-attendance decision in the previous period, as well as initial conditions $\Omega(0)$.

The timing within a model period affects how the likelihood is specified in terms of the information sets for the different probability events. In terms of timing, students (or their parents) first decide which type of primary school they attend. At the end of each school year, students take the tests and observe their scores (if they take the exam). They then observe whether they pass the grade or must repeat the current grade. After grade 6, they decide simultaneously whether to dropout or to enroll in a certain type of lower-secondary school for the next period.[25] After grades 7 and 8 they decide whether to dropout.

Given the vector of the observed test scores $\tilde{A}_i = \{\tilde{A}_{ia_f}^{1}, \tilde{A}_{ia_f}^{1}, ..., \tilde{A}_{ia_l}^{1}, \tilde{A}_{ia_l}^{1}\}$, the vector of the true test scores $A_i = \{A_{ia_f}^{1}, A_{ia_f}^{2}, ..., A_{ia_l}^{1}, A_{ia_l}^{2}\}$, and the vector of initial conditions and time-varying state-space elements $\tilde{\Omega}_i = \{\Omega_i(0), \Omega_i(a_f), ..., \Omega_i(a_l)\}$, the individual likelihood can be written as:

$$L_i(\Theta, \mu, \rho; \tilde{A}_i, A_i, \tilde{\Omega}_i) = \textstyle\sum_{l=1}^{4} \rho_l \Bigg\{ \underbrace{\prod_{j \in J_{ia_f}^{4}} \Pr(D_{ija_f} = 1 | \Omega_i(0), \mu_l)^{1(D_{ija_f}=1)}}_{\text{Primary-school choice at initial period}}$$

$$\textstyle\prod_{a=a_f}^{a_l-1} \Big\{ \underbrace{\Pr(I_{ia}^{Pass} = 1 | A_{ia}, D_{ij,a-1}, \mu_l)}_{\text{Prob of passing the current grade } g_a} \Pi_{j \in J_{ia}^{g}} \underbrace{\Pr(D_{i0a} = 1 | A_{i,a-1}, \mu_l)^{1(D_{i0a}=1)}}_{\text{Prob of dropout at grade } g_a}$$

$$\underbrace{\Big[\Pr(D_{ija} = 1 | A_{ia}, \mu_l) \phi(A_{i,a+1} | A_{ia}, D_{ija}, I_{ia}^{Pass}, \mu_l) \phi(A_{ia} | \tilde{A}_{ia}, D_{ij,a-1}, I_{ia-1}^{Pass})^{1(I_{ia-1}^{Miss}=0)} \Big]^{1(D_{ija}=1)} \Big\}^{1(I_{ia}^{Pass}=1)}}_{\text{Prob of observing test score } \tilde{A}_{i,a+1} \text{ when passing grade } g_a \text{in a school of type } j}$$

$$\underbrace{[\Pr(I_{ia}^{Pass} = 0 | A_{ia}, D_{ij,a-1}, \mu_l) \phi(A_{i,a+1} | A_{ia}, D_{ija}, I_{ia}^{Pass}, \mu_l) \phi(A_{ia} | \tilde{A}_{ia}, D_{ij,a-1}, I_{ia-1}^{Pass})^{1(I_{ia-1}^{Miss}=0)}]^{1(I_{ia}^{Pass}=0)}}_{\text{Prob of observing test score } \tilde{A}_{i,a+1} \text{ when repeating the grade in a type } j \text{ school}} \Bigg\}$$

where Θ defines the vector of model parameters and we suppress the dependence of all the probabilities on the state space $\tilde{\Omega}_i$ to simplify notation. Pr(.) represents the logit or multinomial logit probabilities of school-choice and retention decisions defined in equations 3 and 4. $\phi(.)$ represents the conditional density function of test scores derived from equation 1 and equation 5. J_{ia}^{g} is the available school-choice

[24] We choose this particular functional form so that $\zeta_{ia}^{m} > 0$ and $E(\zeta_{ia}^{m}) = 1$.

[25] Note that the lower-secondary school-choice decision is only made once at the end of grade 6 for each individual.

set at grade g_a defined in equation 2. The vector $\rho = \{\rho_1, ..., \rho_4\}$ denotes the vector of unobserved-type probabilities, where $\rho_l = \Pr(\mu_l = 1)$.

As described in the previous section, the true test scores are not observed in cases of cheating or when students are absent on the day of the test. In those cases, we integrate over the possible true test-score outcomes to obtain the individual likelihood function[26]:

$$L_i(\Theta, \mu, \rho; \tilde{A}_i, \tilde{\Omega}_i) = \int ... \int L_i(\Theta, \mu, \rho; \tilde{A}_i, A_i, \tilde{\Omega}_i) dA_{ia_i^1} ... dA_{ia_i^N}$$

where $\{a_i^1, a_i^2, ..., a_i^N\}$ are the ages that the test scores are either contaminated by cheating or missing for individual i.

We obtain standard errors of the parameter estimates from the inverse of the average of the product of the score matrices, where the derivatives of the log likelihood are evaluated numerically.[27]

5.3 Some simplifying assumptions

We impose three assumptions with regard to the choice problem that simplify the estimation and are basically consistent with the data. First, we assume students do not switch to a different type of school once enrolled. Therefore, the primary-school type is chosen once at the start of primary school and the lower-secondary school type is chosen once at the start of lower-secondary school. Second, we assume that individuals who dropout of school do not afterwards re-enroll. Third, because the fraction of students who repeat grades is fairly small, we restrict the value-added model coefficients to be invariant across ages, separately within primary school and lower-secondary school grades, for the small fractions of students who repeat grades (for whom the current test score and lagged score refer to the same grade).[28]

5.4 Evaluating the effects of *Prospera*-program participation

Prospera provides potential cash transfers for all children of beneficiary households who are enrolled in grades 3-12. Whether each child receives the transfers depends, however, on whether they regularly attend school (at least 85% of days). We consider the family's *Prospera* status as an initial condition of our model, so it is contained in the state space $\Omega(0)$. Once families are enrolled in *Prospera*, they rarely lose their eligibility. Also, even if one child is not attending school, the family may still receive transfers for other children and would still be considered to be participating. The transfers in grades 3-9 typically go to the mothers of the children. After the model parameters are estimated, we use the estimated model to simulate school-going and test-score outcomes for *Prospera* families' children had they not participated in *Prospera*. In this way, we are able to assess the grade-specific program impacts as well as the cumulative impacts of participating in *Prospera* for multiple years.

[26]This integration is performed numerically by taking 50 random draws of the shocks.

[27]This estimator is known as the BHHH estimator (Berndt et. al., 1974). To obtain the numerical derivatives needed to implement the estimator, we use a step-size parameter equal to 1% of the parameter estimates.

[28]That is, we restrict $\delta_{kj}^{m4I} = \delta_{kj}^{m5I}, \delta_{kj}^{m6I} = \delta_{kj}^{m7I} = \delta_{kj}^{m8I}, k = \{1, 2, 3\}$ in the value-added equation 1 and $\gamma_k^4 = \gamma_k^5 = \gamma_k^6, \gamma_k^7 = \gamma_k^8, k = \{1, 2, 3, 4\}$ in equation 4. But the intercept term δ_{0jl}^{mgI} and γ_{0l}^g are grade-specific.

6 Empirical results

6.1 Descriptive statistics

Table 1 shows summary statistics for the students in our sample, which consists of fourth graders in 2008. The columns show the means and standard deviations for children whose families are *Prospera*-program beneficiaries or non-beneficiaries. The average age of the children is around 9 and the gender split is 50-50. The parents of beneficiary students have much less schooling; 60% of the fathers and 62% of the mothers in beneficiary families have primary school (6 grades) or less in comparison to 24% and 29% for non-beneficiaries. Beneficiary fathers are more likely to work full-time, whereas mothers are less likely to work in the labor market at all. Beneficiary students are more likely to have either a mother or father who is not at home. 88% of non-beneficiary students live in urban areas in comparison to 39% for beneficiaries.

Prospera families tend to be larger, with 41% having seven or more household members in comparison to 25% for non-beneficiaries. Children from beneficiary families have fewer years of preschool education. They are also much less likely to have access to computers or the internet at home. 17% of beneficiary students have computers at home and 12% have access to the internet in comparison to 48% and 34% for non-beneficiaries.

In terms of languages spoken at home, 12% of children from beneficiary families speak indigenous languages at home (sometimes in combination with Spanish) in comparison to 2% of non-beneficiary children. There are also substantial heterogeneities in the regional distribution of families by beneficiary status, with a larger fraction of the children from *Prospera* families living in the South and a smaller fraction in the North.

Table 2 shows the average test scores and dropout rates by grade and by *Prospera*-beneficiary status. There are significant gaps in average mathematics and Spanish test scores in primary-school grades with the gaps declining over grades 4-6. For the Spanish test scores, the gaps persist throughout lower-secondary school grades but they are smaller than in primary school. Interestingly, the gaps in the mathematics test scores reverse in lower-secondary grades.

As seen in the last two columns of Table 2, the dropout rate is higher among *Prospera*-beneficiary children/youth at every grade. 26% of Prospera beneficiary youth dropout prior to grade 9 in comparison to 19% for non-beneficiaries. The higher dropout rates for the $P = 1$ group could in part explain the declining test-score gaps if dropouts come disproportionately from the lower parts of the test-score distributions. For this reason, it is important to control for selective dropout in evaluating achievement test-score program impacts.

Figure 1 shows the test-score distributions by grade and lower-secondary school types. The left column displays the mathematics test-score distributions and the right column the Spanish test-score distributions. Each row summarizes the test score for a different grade, from 6 to 9. There is a larger variance in mathematics test-score performance across school types than in Spanish performance. The initial mathematics score is substantially lower at grade 6 for students who go on to attend telesecondary schools. The gaps in mathematics scores start to reverse at grade 7 and the final mathematics score is substantially higher by grade 9. The general school type has the highest Spanish

Table 1: Summary Statistics †

	Prospera		Non-*Prospera*			*Prospera*		Non-*Prospera*	
	Mean	Std	Mean	Std		Mean	Std	Mean	Std
Age (at grade 4)	9.25	0.65	9.08	0.52	Number of household members				
Female	0.49	0.50	0.50	0.50	*≤4 people*	0.20	0.40	0.33	0.47
Education cat. (dad)					*5 people*	0.20	0.40	0.27	0.44
Below primary school	0.38	0.49	0.12	0.33	*6 people*	0.19	0.39	0.15	0.36
Primary school completed	0.22	0.42	0.12	0.33	*≥7 people*	0.41	0.49	0.25	0.43
Secondary or below	0.22	0.41	0.31	0.46	Number of preschool years				
College or above	0.05	0.22	0.35	0.48	*0 year*	0.13	0.34	0.08	0.27
Education cat. (mom)					*1 year*	0.09	0.29	0.04	0.20
Primary school	0.38	0.49	0.14	0.34	*2 years*	0.22	0.42	0.21	0.40
Primary school completed	0.24	0.43	0.15	0.36	*3 years*	0.28	0.45	0.40	0.49
Secondary or below	0.23	0.42	0.30	0.46	*4 years*	0.28	0.45	0.28	0.45
College or above	0.04	0.19	0.34	0.47	Internet at home	0.12	0.33	0.34	0.47
Working status (dad)					Computer at home	0.17	0.38	0.48	0.50
Full time	0.20	0.40	0.15	0.36	First language				
Not full time	0.67	0.47	0.75	0.43	*Spanish*	0.88	0.32	0.98	0.16
Working status (mom)					*Indigenous*	0.08	0.27	0.01	0.11
Housework	0.70	0.46	0.51	0.50	*Both*	0.04	0.19	0.01	0.11
Full time	0.11	0.31	0.25	0.44	Region				
Part time	0.08	0.28	0.17	0.37	*North*	0.11	0.31	0.28	0.45
Father at home	0.76	0.43	0.79	0.41	*North-center*	0.26	0.44	0.23	0.42
Mother at home	0.86	0.35	0.91	0.29	*Center*	0.37	0.48	0.41	0.49
Urban dummy	0.39	0.49	0.88	0.33	*South*	0.26	0.44	0.09	0.28
Obs.	29,964		118,266		Obs.	29,964		118,266	

†Source: Authors' calculations using *ENLACE* merged with student- and parent-context questionnaires.

Table 2: Average test scores and dropout rates by *Prospera* status (P)

	Mathematics score			Spanish score			dropout proportion†	
Prospera	$P = 0$	$P = 1$	Gap	$P = 0$	$P = 1$	Gap	$P = 0$	$P = 1$
Grade 4	534	479	11.3%	527	466	13.1%	-	-
Grade 5	539	495	9.0%	541	488	10.8%	-	-
Grade 6	564	525	7.5%	563	514	9.5%	-	-
Grade 7	508	498	2.0%	497	468	6.2%	0.05	0.09
Grade 8	537	540	-0.6%	508	485	4.7%	0.10	0.16
Grade 9	561	568	-1.4%	509	484	5.1%	0.19	0.26

†The dropout proportion in each row is measured prior to entering that grade. The sample includes students who enroll to at least 6th grade.

Figure 1: Mathematics and Spanish test scores distributions (cdfs) by grade and lower-secondary school types

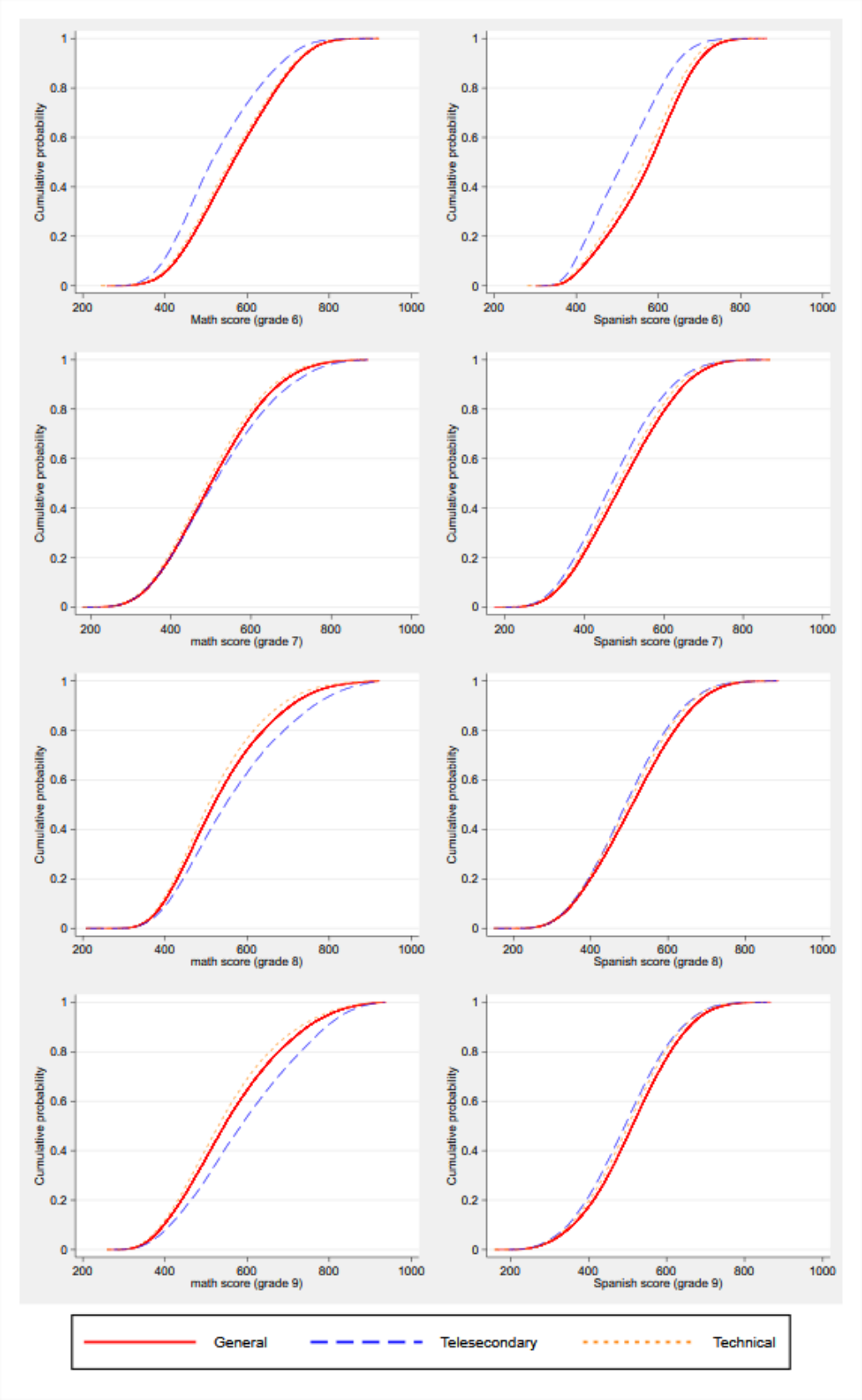


Table 3: Enrollment distribution by school type, grade and *Prospera* status (P)†

Primary school	General		Indigenous			
	$P = 0$	$P = 1$	$P = 0$	$P = 1$		
Grade 4	0.73	0.23	0.01	0.03		
Grade 5	0.73	0.23	0.01	0.03		
Grade 6	0.73	0.23	0.01	0.03		
Secondary school	General		Telesecondary		Technical	
	$P = 0$	$P = 1$	$P = 0$	$P = 1$	$P = 0$	$P = 1$
Grade 7	0.46	0.08	0.06	0.12	0.22	0.06
Grade 8	0.47	0.07	0.06	0.12	0.22	0.06
Grade 9	0.47	0.07	0.06	0.12	0.22	0.05

†This table displays the school type enrollment distribution by beneficiary status, where $P = 1$ denotes *Prospera* participation. The calculation is based on students who still remain in school. The rows add to 1.

Table 4: Percentage of retained students by school type, grades and *Prospera* status (P)†

Primary school	General		Indigenous			
	$P = 0$	$P = 1$	$P = 0$	$P = 1$		
4th year	1.1%	2.5%	2.4%	2.8%		
5th year	0.8%	1.6%	1.4%	2.5%		
6th year	0.1%	0.2%	0.0%	0.2%		
Lower-secondary school	General		Telesecondary		Technical	
	$P = 0$	$P = 1$	$P = 0$	$P = 1$	$P = 0$	$P = 1$
1st year	0.7%	0.5%	0.4%	0.4%	0.3%	0.4%
2nd year	0.8%	0.6%	0.3%	0.3%	0.5%	0.4%

†$P = 1$ denotes beneficiary and $P = 0$ non-beneficiary.

scores, but the advantage shrinks over grades. These graphs suggest that the test-score distributions in telesecondary schools show greater improvement across grades those in the other school types. However, the comparisons are not necessarily causal, because of the compositional differences in the students attending the different school types.

Table 3 shows the school-type enrollment distributions by beneficiary status, where $P = 1$ denotes *Prospera* participation. Only about 4% of students attend indigenous primary schools, and the vast majority of these are *Prospera* beneficiaries. At the lower-secondary school level, children in *Prospera* are more than twice as likely to attend telesecondary schools. Within telesecondary schools, about two-thirds of the students are participating in *Prospera*. The majority of children from non-beneficiary families are enrolled in general schools, whereas the largest fraction of children from beneficiary families are enrolled in telesecondary schools.

Table 4 shows the fraction of students who are retained at each grade level by *Prospera*-beneficiary status. Grade retention is more prevalent in primary-school grades, reaching 3.1% in fourth grade, and is higher for *Prospera* children. Once youth attend lower-secondary school, less than 1% are retained and there is no systematic pattern by beneficiary status.

Table 5 shows the transition patterns from primary school to attending one of the three types

Table 5: Primary and lower-secondary school transitions by *Prospera* status (P)†

		General	Tele	Tech	Dropout
General	$P = 1$	0.28	0.44	0.20	0.08
	$P = 0$	0.58	0.10	0.27	0.05
	Total	0.52	0.16	0.26	0.06
Indigenous	$P = 1$	0.08	0.57	0.21	0.14
	$P = 0$	0.15	0.48	0.20	0.16
	Total	0.11	0.54	0.21	0.14

†$P = 1$ denotes beneficiary and $P = 0$ non-beneficiary.

of lower-secondary schools or to dropping out, by *Prospera*-beneficiary status. 15% of children who attended indigenous primary schools dropout in comparison to 7% of children who attended general primary schools. For non-beneficiary children who attended general primary schools, 55% continue in general lower-secondary schools, 10% attend telesecondary schools, 28% attend technical schools, and the remainder (6%) drop out. The comparable school choice distribution for *Prospera* beneficiaries is 27% general, 43% telesecondary and 21% technical, with 9% dropping out. Thus, youth from beneficiary and non-beneficiary families attend different distributions of types of schools. Therefore, the effectiveness of different types of schools is potentially important to understanding the overall effectiveness of the *Prospera* program in promoting schooling attainment and academic achievement.

Table 6 provides information on the supply of local schools of different types at the individual level. In Mexico, multiple school sessions are often held in the same building, such as a morning and afternoon session. The different sessions may have different principals and teachers, so in the dataset they are regarded as different schools.[29] Primary schools tend to be small, with an average enrollment of less than 200, and consequently there are a large number of primary schools. Their small size partly reflects that the school systems do not usually provide transportation and students typically walk to school. Also, indigenous schools are mainly located in areas with greater indigenous populations. 58% of children do not have access to indigenous schools. At the lower-secondary level, there are fewer schools and they are larger in size. The last column of the table also shows the fractions of students in the sample who do not have access to certain types of schools, which is taken into account in estimating the school-choice model.

6.2 Parameter estimates

6.2.1 Sample-inclusion criteria

After two decades of operation, the *Prospera* program was widely known in Mexico. Because the transfers are relatively large (amounting to a 20% increase in family income on average), most families that are aware of their eligibility choose to participate. They may receive partial benefits if they send some but not all of their children to school. Program eligibility is not means-tested by income.

[29] In Appendix figure 6, we show one illustrative example of local primary-school sessions in Aguascalientes, a city in central Mexico. It has 316 school sessions distributed in 250 unique coordinates within 10 kilometers.

Table 6: Number of local schools of different types†

	Mean	Std	p10	p50	p75	Not available
Primary school (within 5 km)						
General	67	82	9	29	109	3.6%
Indigenous	5	6	1	3	6	75.5%
Secondary school (within 10 km)						
General	51	73	4	19	65	14.7%
Technical	11	12	2	6	17	15.7%
Telesecondary	14	12	4	11	20	7.5%

†The number is calculated based on the estimation sample. The sample-inclusion criteria are described in the next section. Columns 3-5 report selected quantiles. The last column gives the percentages of individuals for whom a given school type are not locally available.

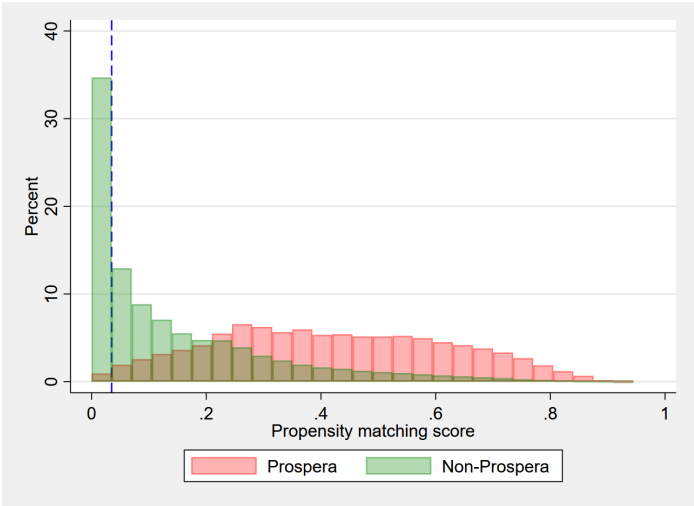
This is because income can be difficult to measure in a country with a significant informal sector and where many low-income individuals are engaged in agricultural work. The program eligibility criteria rather depend mainly on households' assets (such as car ownership), on characteristics of the houses themselves (such as whether they have dirt floors and how many rooms there are per person living the house) and on household demographics, such as numbers of children and numbers of dependents per worker. The detailed eligibility criteria vary regionally, have changed somewhat over time and are not disclosed to the public. Families who apply to the program typically fill out a survey to determine their eligibility and their answers on the survey may be checked through home visits.[30] Although the program benefits are relatively generous, there may be eligible families who do not apply because they are not aware of their eligibility or for other reasons.

As previously noted, the administrative data on children's families' *Prospera* participation statuses were linked with the test-score database. The families' participation statuses were measured when children were in grade 6. The vast majority of children enroll in grade 6, although there is a small fraction that drops-out prior to entering grade 6 who we could not include in our analysis sample.

In the program-evaluation literature, it is common practice to restrict the program impact analysis to individuals who meet program eligibility criteria so as to increase comparability between the treated group and the comparison group. Heckman et al. (1997) showed, in the context of evaluating a job-training program, that having a highly comparable comparison group is important to producing reliable non-experimental impact estimates that replicate experimental estimates. In the absence of detailed data on all of the precise *Prospera*-eligibility criteria that go into eligibility determination, we cannot restrict our analysis sample to children from eligible families. As an alternative strategy, however, we use the extensive information on housing characteristics and demographics that was gathered through the student and parent surveys to estimate a probit model of whether children come from households participating in *Prospera* and we use the estimated model to generate predicted probabilities, i.e. propensity scores. The estimated coefficients from this model are shown in Appendix A.4. The

[30]They usually remain eligible for at least three years. After that, they may be subject to recertification. If family circumstances substantially improve, they are usually transitioned to a less generous version of the program.

Figure 2: The propensity score distribution by *Prospera* status (P)†



†The red histogram represents the propensity score distribution for *Prospera* children/youth and the green histogram for Non-*Prospera* children/youth.

percentage correctly classified under the model is high - 89%.

Figure 2 plots the propensity-score distribution for children from *Prospera*-beneficiary and non-beneficiary households. As seen, a large fraction of non-beneficiaries fall in the first bin of the histogram, meaning that they have extremely low probabilities of participating in *Prospera*, likely because their characteristics make them ineligible. To increase comparability between the *Prospera* subsample and the comparison-group subsample, we exclude in our impact analysis the bottom 1% of the *Prospera* beneficiaries and the non-beneficiaries with similar or smaller propensity scores (the lowest bin of the histogram).[31] Our final analysis sample includes 105,256 children with 589,371 individual-period observations.

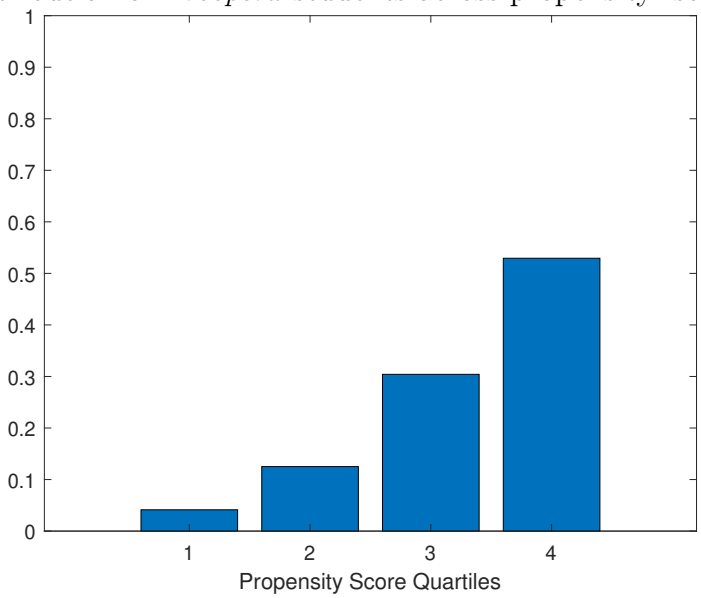
6.2.2 Evidence on *Prospera* impacts

Prospera may have heterogeneous impacts for students from different backgrounds. To capture this potential heterogeneity in a parsimonious way, we divide the analysis sample into quartiles based on the propensity scores and allow the *Prospera* impact to vary by quartile. Families in the highest quartile - i.e. with characteristics that make them most likely to be eligible for *Prospera* - tend to be the most disadvantaged. Figure 3 shows the distribution of *Prospera*-beneficiary children across quartiles. 83% of *Prospera* children/youth are concentrated in quartiles 3 and 4, 30% and 53% respectively, reflecting the fact that the program is targeted at the poorest families.

In estimation, we allow the coefficients associated with *Prospera* participation to be quartile-specific. In particular, we redefine $\{\delta_{2j}^{mgI}, \phi_{2j}^{g}, \gamma_{2}^{g}\}$ to be $\{\delta_{2j\eta}^{mgI}, \phi_{2j\eta}^{g}, \gamma_{2\eta}^{g}\}, \eta = \{1, 2, 3, 4\}$, in which η represents the quartile to which each individual belongs. This model nests a model where the *Prospera* impact is restricted to be the same across quartiles. However, our estimates indicate that impacts

[31]This trimming threshold excludes 299 children from *Prospera* families and 42,675 non-beneficiary children. This type of trimming is commonly done in the application of matching estimators as a way of imposing "common support." See e.g., Heckman et al. (1997).

Figure 3: Distribution of *Prospera* students across propensity- score quartiles



are heterogeneous, with the most disadvantaged children/youth experiencing the largest impacts on average.

6.2.3 Value-added models

Tables 7 and 8 show the estimated parameters of the value-added production functions for mathematics and Spanish. The tables display the coefficients related to lagged scores, *Prospera* impacts and gender. In addition, the specifications include other covariates, such as parents' education and numbers of siblings, to capture heterogeneous family inputs into the achievement process. Also, they include indicators for rural/urban residence and for region of residence to capture regional differences in school quality/infrastructure that may affect achievement. The full set of estimated parameters can be found in Appendix B.2. Each model includes lagged scores in both subjects, assuming that knowledge of Spanish might facilitate learning in mathematics and vice versa.[32] For example, children who are more proficient in one subject might be able to focus their efforts more on studying the other subject. The lagged parameters are highly statistically significant in all grades and in both subjects. As expected, the estimated lagged own-subject coefficients are larger than the other-subject coefficients. Boys have statistically significantly higher scores than girls in mathematics and lower scores in Spanish. The gender gaps in mathematics are larger in secondary school than in primary school.[33]

The table also reports the effect of being a *Prospera* beneficiary, allowing the program effects to differ by propensity-score quartile. These estimates represent one-year participation effects. Given that only a small portion of *Prospera* beneficiaries are in the low propensity-score quartiles (4% in quartile 1 and 13% in quartile 2), it is not surprising to see systematically larger standard errors

[32]Aucejo and James (2021) estimate production function models for mathematics and verbal test scores using UK data. They find cross-effects for verbal skills for learning math but not vice versa. We find cross-effects for both subjects.

[33]Bharadwaj et al. (2016) also find that boys perform better than girls in mathematics using Chilean test score data and that the gaps widen between fourth and eighth grades. Aucejo and James (2021) find a large female advantage in verbal skills.

for the estimates associated with these two quartiles, leading to their estimates to be insignificant. In contrast, we observe relatively more significant effects of *Prospera* participation in the top two quartiles, especially the one with the highest propensity-score (quartile 4), which captures *Prospera* beneficaries coming from the most disadvantaged backgrounds in terms of SES. In particular, *Prospera* has statistically significant impacts on mathematics for children/youth in quartile 4 in the lower-secondary school (grade 7, 8, 9), with a range from 0.05-0.14 standard deviations. In Spanish, the statistically significant effect sizes range from 0.03-0.05 standard deviations. For both mathematics and Spanish, the largest impacts are observed in 7th grade. Since quartile 4 contains the majority of the *Prospera* beneficaries (52.7%), the *Prospera* effects on this subgroup largely drive the average effects on the population.

6.3 School-choice/dropout model

Table 9 reports estimated coefficients from the school-choice/dropout model, where the omitted category is dropping out and working. The specification also includes other covariates not shown in the table as described in the table note. Family background and demographic variables capture heterogeneity in preferences for schooling and school types. Regional indicators capture regional differences in school quality and infrastructure that may affect school choices. As seen in the table, children with higher sixth-grade test scores are more likely to attend general or technical schools rather than telesecondary schools and are less likely to drop out. Being in a higher propensity-score quartile increases the likelihood of attending a telesecondary school. The estimated distance coefficients show that the distances that students need to travel to get to the nearest lower-secondary school of each type are major factors influencing their school enrollment decisions. As expected, greater distance decreases the likelihood of attending that type of school. Local school availability is also an important factor. The number of local general/technical/telesecondary lower-secondary schools increases the probability of attending that type of school.

Lastly, we imputed the wages that children could earn in the labor market, given their observed characteristics, including place of residence.[34] Children who have higher potential wage offers are less likely to attend lower-secondary school.

[34]The wages are derived from Census data and take into account the locality of residence as well as a set of observed child characteristics. See Appendix A for a description of the procedure, which controls for sample selection into working.

Table 7: Mathematics value-added model estimates†

	General		*Indigenous*			*General*			*Telesecondary*			*Technical*	
	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Mathematics score													
Lag mathematics	0.55	0.54	0.36	0.40	0.45	0.47	0.50	0.33	0.43	0.41	0.41	0.46	0.47
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lag Spanish	0.15	0.13	0.15	0.15	0.26	0.18	0.25	0.25	0.15	0.24	0.27	0.16	0.26
	(0.00)	(0.00)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prospera effect by propensity score quartile													
P*Q1 (4%)	0.62	-0.80	-1.05	2.89	2.76	2.49	-0.77	0.29	-2.86	2.25	1.73	2.81	2.34
	(3.00)	(2.88)	(21.81)	(16.49)	(4.14)	(3.56)	(4.50)	(12.52)	(10.99)	(11.83)	(5.65)	(5.17)	(7.38)
P*Q2 (13%)	0.31	-1.83	-2.61	-3.37	2.17	-2.20	2.59	2.45	2.31	1.92	2.02	1.87	2.64
	(1.71)	(1.64)	(9.99)	(8.39)	(2.75)	(2.58)	(3.03)	(5.42)	(4.34)	(5.28)	(3.49)	(3.09)	(3.90)
P*Q3 (30%)	-0.15	1.02	2.02	1.98	4.00	-1.01	4.36	6.17	6.77	6.09	4.34	3.23	3.39
	(1.21)	(1.14)	(5.40)	(4.90)	(2.14)	(1.97)	(2.29)	(3.11)	(2.64)	(3.13)	(2.56)	(2.17)	(2.73)
P*Q4 (53%)	-0.05	-1.05	-0.12	-3.33	13.57	7.50	6.97	12.39	6.63	9.32	12.54	5.75	7.74
	(1.10)	(1.05)	(3.71)	(3.23)	(2.24)	(1.97)	(2.45)	(2.41)	(2.06)	(2.45)	(2.36)	(2.10)	(2.58)
Gender													
Male	3.93	3.14	2.65	1.18	7.05	16.7	11.4	0.24	9.27	2.23	6.19	19.49	7.64
	(0.66)	(0.63)	(3.13)	(2.77)	(1.04)	(0.94)	(1.13)	(1.98)	(1.69)	(2.01)	(1.37)	(1.20)	(1.48)

†The percentages in the first column give the percentages of *Prospera* beneficiaries in each quartile. The model includes additional control variables for parents' schooling attainment, parents' working status, number of household members, child age and its square, gender, language spoken at home, internet access, computer access, number of years child attended preschool, urban-rural dummy, regional dummies (north, north-center, center, south) and unobserved types. The full set of estimated parameters can be found in Appendix B.2.

Table 8: Spanish value-added model estimates†

	General		*Indigenous*			*General*		*Telesecondary*			*Technical*		
	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Spanish score													
Lag mathematics	0.23	0.20	0.21	0.18	0.13	0.23	0.20	0.10	0.17	0.15	0.12	0.22	0.19
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lag Spanish	0.43	0.40	0.28	0.32	0.53	0.48	0.49	0.43	0.39	0.39	0.53	0.47	0.48
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prospera effect by propensity score quartile													
P*Q1 (4%)	0.02	-2.03	0.09	-3.73	0.06	-2.52	-2.81	-0.36	-0.72	-2.80	-2.21	0.54	-1.96
	(2.76)	(2.41)	(19.37)	(13.83)	(4.00)	(3.40)	(3.75)	(9.52)	(7.80)	(9.49)	(5.32)	(4.93)	(5.99)
P*Q2 (13%)	-1.41	-1.82	-2.86	0.13	1.34	-2.34	-0.99	1.38	0.51	-1.09	-0.17	1.72	-0.58
	(1.64)	(1.43)	(9.28)	(7.29)	(2.61)	(2.41)	(2.45)	(4.25)	(3.45)	(4.14)	(3.27)	(2.90)	(3.18)
P*Q3 (30%)	-2.76	-1.71	-2.29	-1.01	3.80	-1.62	-1.26	2.33	0.68	1.51	2.09	2.54	-1.35
	(1.16)	(0.99)	(4.85)	(4.24)	(2.07)	(1.88)	(1.96)	(2.50)	(2.11)	(2.44)	(2.39)	(2.14)	(2.32)
P*Q4 (53%)	-1.06	-1.48	0.95	-2.22	5.82	3.97	-1.58	5.43	1.89	1.05	5.76	3.23	0.87
	(1.09)	(0.92)	(3.30)	(2.73)	(2.13)	(2.03)	(2.14)	(1.98)	(1.68)	(1.92)	(2.30)	(2.14)	(2.26)
Gender													
Male	-16.30	-15.46	-7.46	-8.48	-25.80	-17.00	-12.23	-25.14	-19.8	-15.93	-26.24	-15.71	-13.72
	(0.62)	(0.55)	(2.79)	(2.32)	(0.98)	(0.88)	(0.94)	(1.61)	(1.36)	(1.57)	(1.25)	(1.13)	(1.23)

†The percentages in the first column give the percentages of *Prospera* beneficiaries in each quartile. The model includes additional control variables for parents' schooling attainment, parents' working status, number of household members, child age and its square, gender, language spoken at home, internet access, computer access, number of years child attended preschool, urban-rural dummy, regional dummies (north, north-center, center, south) and unobserved types. The full set of estimated parameters can be found in Appendix B.2.

Table 9: Estimated parameters for the lower-secondary school-choice/dropout model†

	General	Tele	Technical
mathematics (6th grade)	0.0015	0.0005	0.0019
	(0.0001)	(0.0001)	(0.0001)
Spanish (6th grade)	0.0036	0.0019	0.0031
	(0.0001)	(0.0001)	(0.0001)
Prospera effect by propensity score quartile			
P*Q1 (4%)	0.26	0.44	0.20
	(0.12)	(0.14)	(0.13)
P*Q2 (13%)	0.36	0.48	0.34
	(0.07)	(0.07)	(0.07)
P*Q3 (30%)	0.30	0.60	0.43
	(0.05)	(0.05)	(0.05)
P*Q4 (53%)	0.18	0.66	0.32
	(0.04)	(0.04)	(0.04)
Distance to general	-0.44	0.09	0.15
	(0.01)	(0.01)	(0.01)
Distance to telesecondary	0.04	-0.74	0.03
	(0.01)	(0.01)	(0.01)
Distance to technical	0.12	0.08	-0.51
	(0.01)	(0.00)	(0.01)
Number of general	0.0023	-0.0017	-0.0002
	(0.0003)	(0.0005)	(0.0002)
Number of telesecondary	-0.0018	0.0123	-0.0114
	(0.0011)	(0.0012)	(0.0012)
Number of technical	-0.0033	-0.0188	0.0033
	(0.0019)	(0.0027)	(0.0020)
Wage	-0.010	-0.044	-0.016
	(0.004)	(0.004)	(0.004)

†The percentages in the first column give the percentages of *Prospera* beneficiaries in each quartile. Additional control variables include parental schooling attainment, parental working status, whether mother at home, number of household members, age and its square, language spoken at home, internet access, computer access, number of years child attended preschool, urban-rural dummy, regional dummies (north, north-center, center, south) and unobserved types. The full set of estimated parameters can be found in Appendix B.2.

6.4 Grade-retention model

Table 11 shows estimated coefficients for the probability of being retained and also for the value-added model specifications that were estimated for retained children (who were taking the grade for the second time). The achievement test scores are not used in making retention decisions. However, one would expect them to be correlated with other school performance measures. As seen in the first and fourth columns, students with higher test scores are less likely to be retained, in both primary and secondary school. Also, the retention probability does not appear to depend on *Prospera*-participation status or to vary by propensity-score quartile. Gender is a very significant determinant of retention

Table 10: Estimated parameters for the probability of dropping out after grades 7 and 8†

	Grade 7	Grade 8
Lag mathematics	-0.0013	-0.0008
	(0.0001)	(0.0001)
Lag Spanish	-0.0030	-0.0034
	(0.0001)	(0.0001)
Prospera effect by propensity score quartile		
P*Q1 (4%)	0.09	0.20
	(0.12)	(0.09)
P*Q2 (13%)	0.15	-0.03
	(0.07)	(0.06)
P*Q3 (30%)	0.03	-0.07
	(0.05)	(0.04)
P*Q4 (53%)	-0.11	-0.16
	(0.04)	(0.04)
Distance to current lower-secondary school	0.03	0.03
	(0.01)	(0.01)
Telesecondary-school indicator	-0.07	0.10
	(0.04)	(0.04)
Technical-school indicator	0.04	0.07
	(0.03)	(0.03)
Wage	0.04	0.01
	(0.00)	(0.00)

†The percentages in the first column give the percentages of *Prospera* beneficiaries in each quartile. Additional control variables include parental schooling attainment, parental working status, whether mother is home, number of household members, age and its square, language spoken at home, internet access, computer access, number of years attended pre-school, urban-rural dummy, region dummies (north, north-center, center, south) and unobserved types. The full set of estimated parameters can be found in Appendix B.2.

Table 11: Estimated model parameters for retained students†

	Primary			Secondary		
	Prob Ret.	VA Math	VA Spanish	Prob Ret.	VA Math	VA Spanish
Lag mathematics	-0.0041	0.40	0.17	-0.0012	0.38	0.18
	(0.0001)	(0.03)	(0.02)	(0.0001)	(0.05)	(0.05)
Lag Spanish	0.0003	0.12	0.26	-0.0100	0.14	0.35
	(0.0001)	(0.03)	(0.03)	(0.0002)	(0.05)	(0.05)
Prospera effect by propensity score quartile						
P*Q1 (4%)	0.005	-1.12	0.52	-0.473	0.55	0.56
	(0.21)	(20.72)	(17.34)	(0.34)	(43.51)	(49.61)
P*Q2 (13%)	0.16	-0.68	0.42	-0.20	1.52	0.98
	(0.10)	(9.37)	(8.78)	(0.20)	(20.38)	(26.90)
P*Q3 (30%)	-0.002	1.35	1.90	-0.59	2.80	1.34
	(0.07)	(6.43)	(5.43)	(0.15)	(15.61)	(15.68)
P*Q4 (53%)	-0.07	1.98	2.62	-0.56	3.52	2.61
	(0.06)	(5.42)	(4.82)	(0.15)	(17.33)	(16.69)
Male	0.48	6.12	-12.45	0.67	8.49	-8.05
	(0.04)	(4.01)	(3.33)	(0.09)	(9.84)	(9.39)

†VA = value added. The probability of retention is estimated by a probit model. The percentages in the first column give the percentages of *Prospera* beneficiaries in each quartile. Additional control variables include parental schooling attainment, parental working status, whether mother present, number of household members, age and its square, language spoken at home, internet access, computer access, number of years child attended pre-school, urban-rural dummy, region dummies (north, north-center, center, south) and unobserved types. The full set of estimated parameters can be found in Appendix B.2.

with boys more likely to be retained than girls, a pattern widely found in developing countries (e.g., Grant and Behrman (2010)).

6.5 Test-score measurement equation

An unusual feature of our data is that they contain information on which students were flagged by SEP as likely copiers.[35] All tests, even low-stakes tests, can be affected by copying. Our test-score measurement equation allows the true test score to differ from the measured test score in the event of copying and also allows for heterogeneity in the gains from copying across grades and types of schools. We allow for this heterogeneity because the technology for classroom monitoring could vary across schools and depend on factors such as class size.

Within a value-added model, copying induces a one-sided measurement error in the dependent variable, the independent variable (lagged test scores) or both variables, but only for students who engage in copying. Table 12 shows the percentages of students suspected of copying, which ranges from a low of 1.6% in seventh grade in general schools to a high of 9.1% in eighth grade in telesecondary schools. At the primary-school level, indigenous schools have higher copying rates. Copying distorts average test scores by 20-100 points for copiers. Our estimation approach described earlier takes into

[35]See Appendix B for further discussion of the methods they use.

Table 12: Estimated test-score distortion from copying by lower-secondary school types

	Percentages	Math			Spanish		
		Raw	True	Diff	Raw	True	Diff
Grade 5							
General	4.1%	567	525	42	547	521	26
Indigenous	7.0%	539	468	72	510	461	49
Overall	4.3%	564	519	45	544	516	28
Grade 6							
General	3.8%	605	555	50	578	547	30
Indigenous	6.1%	566	503	62	536	490	46
Overall	3.9%	602	551	51	574	542	32
Grade 7							
General	1.6%	563	504	59	529	493	36
Telesecondary	3.9%	618	519	99	538	482	57
Technical	2.3%	583	496	88	536	488	49
Overall	2.4%	592	509	83	535	487	48
Grade 8							
General	3.3%	617	533	84	556	513	43
Telesecondary	9.4%	693	577	116	580	516	65
Technical	4.0%	613	528	84	553	511	42
Overall	5.1%	653	553	99	567	514	53
Grade 9							
General	2.5%	632	549	83	534	504	30
Telesecondary	5.4%	666	611	55	533	512	21
Technical	3.4%	617	553	64	531	505	26
Overall	3.5%	641	574	67	533	507	25

account possible test score distortion due to copying.

6.6 Model goodness of fit

Our model involves a fairly large number of parameters, in part because the specifications are flexible in the sense that we do not impose cross-grade parameter restrictions on the production functions or other relations. Our sample sizes are large and most of the parameters are precisely estimated; however, a model with many parameters raises possible concerns about over-fitting. To address such concerns, we performed an out-of-sample model validation test. In particular, we randomly split the sample into a 20% training sample used for model estimation and a 80% test sample used to evaluate the model's goodness-of-fit.[36] Tables 13 and 14 provide evidence on the model's goodness-of-fit based on the 80% test sample. Table 13 compares the average test scores by grade and by *Prospera*-beneficiary status, in the data and simulated under the model. With few exceptions, the data averages and the model-simulated averages are very close. The estimated model generates the observed pattern that

[36]The estimates based on the 20% subsample are only used for this model validation purpose. Otherwise, parameter estimation is based on the full sample, which is more efficient.

Table 13: Goodness of fit for average test scores by *Prospera* status (P)†

Prospera	Math				Spanish			
	$P = 0$		$P = 1$		$P = 0$		$P = 1$	
	Data	Sim	Data	Sim	Data	Sim	Data	Sim
Grade 5	520	519	492	493	519	520	487	489
Grade 6	548	549	522	527	543	544	512	516
Grade 7	491	493	493	505	478	480	465	475
Grade 8	519	519	531	535	489	488	481	485
Grade 9	544	544	565	567	491	489	483	487

†We simulate the test scores 100 times for each individual. We adjust for any copying in both the simulation and the data, so that both represent true scores.

Table 14: Goodness of fit to school-type distribution

Lower-secondary	General		Telesecondary		Technical		Dropout	
choice	Data	Sim	Data	Sim	Data	Sim	Data	Sim
Non beneficiary (P=0)								
Grade 7	0.50	0.49	0.16	0.16	0.28	0.28	0.07	0.07
Grade 8	0.47	0.47	0.15	0.15	0.26	0.26	0.12	0.12
Grade 9	0.42	0.42	0.13	0.13	0.23	0.24	0.23	0.21
Prospera beneficiary (P=1)								
Grade 7	0.25	0.24	0.46	0.46	0.20	0.19	0.09	0.12
Grade 8	0.23	0.23	0.43	0.43	0.18	0.17	0.16	0.17
Grade 9	0.20	0.20	0.38	0.37	0.16	0.16	0.26	0.27

the *Prospera* beneficiaries have lower average test scores in primary grades in both subjects. It also captures the pattern that the test-score gaps between beneficiaries and non-beneficiaries shrink in lower-secondary grades and even reverses in mathematics.

Table 14 shows the model fit to the school-type distribution. The proportions predicted by the model do not differ from the data by more than 0.03. The model captures the greater tendency for *Prospera*-beneficiary children to attend telesecondary lower-secondary schools as well as their higher dropout proportions. It also captures the observed dropout pattern by grade.

6.7 Estimated cumulative *Prospera*-program effects

As previously described, educational production functions typically assume that knowledge acquisition in mathematics and Spanish is a cumulative process. The value-added model specification allows lagged knowledge to have an effect on contemporaneous knowledge accumulation, so that the history of inputs into the learning process potentially matters. If *Prospera* participation increases knowledge in a particular grade, then this benefit can have a persistent effect on learning in future grades. That is, program participation can have direct effects on current test scores as well as indirect effects operating

through lagged test scores.

In Table 15, we use our estimated model of educational enrollment, attainment and achievement to simulate the effects of being a *Prospera* beneficiary over multiple grades, starting with grade 4. Our estimation procedure allows for *Prospera* effects that operate through all of the different channels of our school progression and achievement model. Columns labeled $P = 1$ show the outcomes for *Prospera*-beneficiary children/youth with their participation in the program. Columns labeled $\tilde{P} = 0$ show the simulated (counterfactual) outcomes were they not to participate in the program.

It is only possible to assess test-score impacts for children/youth who would attend school both with and without *Prospera*. Therefore, our reported program impacts on test scores in column "Diff" represent the lower bounds, as they do not include academic achievement gains for children/youth who in the absence of *Prospera* would not be attending the grade.[37] We estimate positive benefits to being a *Prospera* beneficiary in lower-secondary grades but essentially find little effect in grades 5 and 6. In lower-secondary school, the cumulative *Prospera* impact in mathematics increases with the grade level and reaches a high of 0.12 standard deviations by grade 9. In Spanish the cumulative gains are much smaller - about 0.04 standard deviations.[38]

We might expect the program effects to be larger in lower-secondary school than primary school because the transfer amounts that families receive for school attendance are substantially larger (in the fall semester of 2008 the transfers ranged from were 130 to 265 pesos for primary school and 405 to 495 (385 to 430) for females (males) in lower-secondary school (US1= 11 pesos in 2008). Also, older children typically have more demands on their time that compete with schoolwork than do younger children, such as taking care of younger siblings, housework, working for family businesses or working for pay after school. The *Prospera* cash transfers may reduce these outside demands, allowing them to focus more on schoolwork.[39]

The last three columns of Table 15 show the effects of being a *Prospera* beneficiary on the proportion of students who dropout prior to entering the grade. *Prospera* reduces the dropout proportion by 0.04. The effect occurs mainly at the transition between primary and lower-secondary school.

We next explore the heterogeneous *Prospera* impacts for students from different backgrounds in Figure 4 and Figure 5. Figure 4 shows the disaggregated effects of *Prospera* participation on test scores by propensity-scores quartiles. As we described in session 6.2.2, the propensity score is a summary statistic of student's family background, with quartile 1 denoting the most advantaged family and quartile 4 denoting the most disadvantaged family. Our estimates show larger impacts in later grades but smaller impacts in earlier grades, regardless of the quartiles. These patterns are consistent with the *Prospera* effects being cumulative with exposure to the *Prospera* program. Among the four quartiles, we observe the greatest estimated cumulative impacts for students in the highest propensity-score quartile, who are the ones from the most disadvantaged backgrounds. *Prospera* increases their test

[37] As we show in Figure 4 and Figure 5, these absent children are disproportionally from most disadvantaged family backgrounds and therefore tend to have larger program gains on average.

[38] Although the average effect on the Spanish score is not significant, it displays substantial heterogeneity among *Prospera* beneficiaries. We will come back to this point below.

[39] Because we only observe children for six years and we want to examine their performance through grade 9 with a model with one-year lags, we do not investigate whether *Prospera* participation affected academic achievement in earlier grades (1-4).

Table 15: Cumulative program impacts†

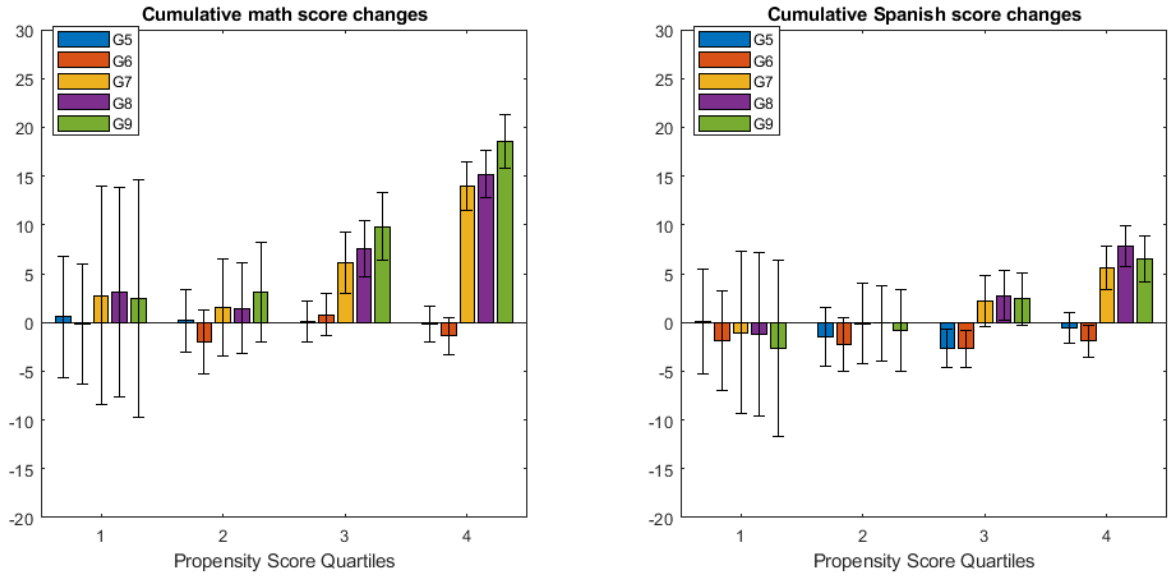
	Mathematics score				Spanish score				Dropout rate		
	$P = 1$	$\tilde{P} = 0$	Diff	S.E.	$P = 1$	$\tilde{P} = 0$	Diff	S.E.	$P = 1$	$\tilde{P} = 0$	Diff
Grade 5	493	493	-0.1	2.3	489	490	-1.3	2.1	-	-	-
Grade 6	528	528	-0.8	2.4	516	518	-2.1	2.1	-	-	-
Grade 7	505	496	9.3	3.4	475	472	3.5	2.9	0.115	0.153	-0.038
Grade 8	535	525	10.3	3.3	486	481	4.8	2.7	0.168	0.204	-0.036
Grade 9	567	554	12.9	3.7	488	484	3.7	3.0	0.265	0.305	-0.040

†We report the test scores for children/youth who would attend school both with and without *Prospera*. The columns "Diff" capture the test-score gain of these subgroups. The columns "S.E." report the standard errors of the test-score gains from the program. It is obtained by a parametric bootstrap with 500 replications. For each bootstrap iteration, we draw the model coefficients from their estimated distributions and re-simulate the cumulative impacts and derive the standard errors from the empirical distributions.

scores in mathematics by 0.2 standard deviations and their test scores in Spanish by 0.07 standard deviations, both are statistically significant. While the average impact of *Prospera* on Spanish is not significant, we find its impacts are clearly positive for the students in the top propensity-score quartile, which contains the majority (52.7%) of *Prospera* participants.

Figure 5 shows the cumulative *Prospera* impacts on schooling attainment and on dropouts by grade 9, disaggregated by propensity-score quartiles. The largest impacts are estimated for the highest quartile, which comprises students from the poorest families. Thus, the program has the greatest impact for the most disadvantaged. Table 16 shows cumulative *Prospera*-program impacts on test scores and on dropouts by gender. After grade 7, boys dropout at a slightly higher rate than girls. The pattern of test-score impacts is highly similar for girls and boys.

Figure 4: *Prospera* academic achievement effects by propensity-score quartiles†



†Confidence intervals are obtained by a parametric bootstrap with 500 replications. For each bootstrap iteration, we draw the model coefficients from their estimated distributions and resimulate the cumulative impacts and derive the confidence interval from the empirical distribution.

Figure 5: *Prospera* effects on educational attainment and dropout by propensity-score quartiles

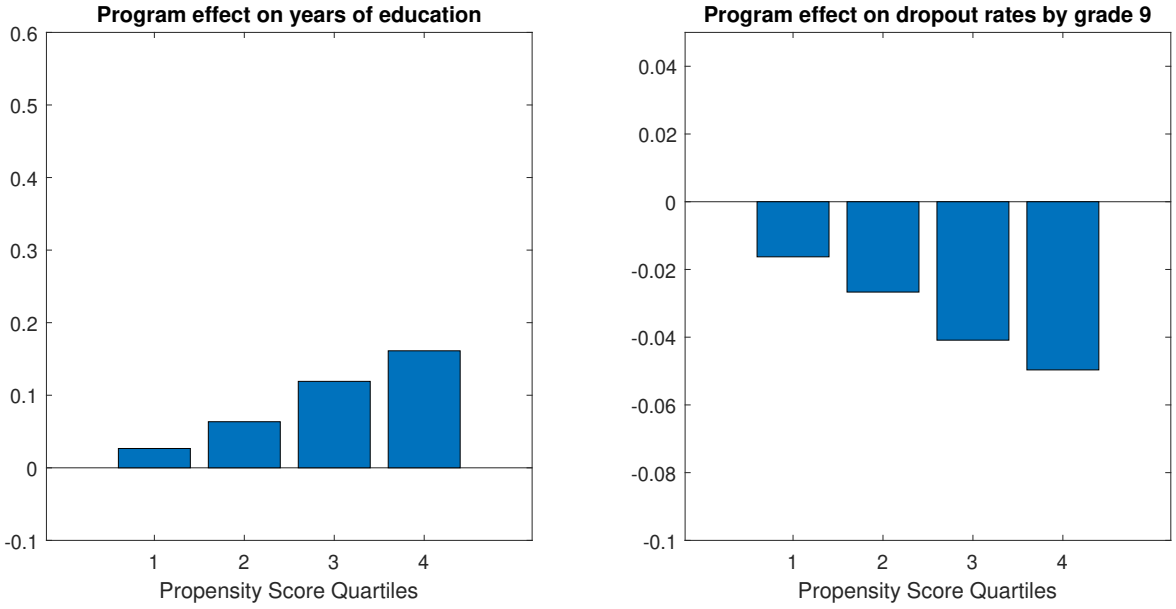


Table 16: Gender differences in cumulative *Prospera* effects (by grade 9)

	Female			Male		
	$P = 1$	$\tilde{P} = 0$	Diff	$P = 1$	$\tilde{P} = 0$	Diff
Mathematics						
Grade 5	494	494	0.0	491	491	0.0
Grade 6	527	528	-0.9	528	529	-0.7
Grade 7	507	497	9.6	503	494	8.9
Grade 8	533	522	10.8	537	527	9.8
Grade 9	568	554	13.4	566	554	12.2
Spanish						
Grade 5	503	504	-1.4	475	476	-1.2
Grade 6	529	532	-2.3	502	504	-2.1
Grade 7	493	490	3.5	457	453	3.4
Grade 8	504	499	4.9	468	463	4.6
Grade 9	502	498	4.0	472	469	3.4
Dropout rate						
Grade 7	0.12	0.16	-0.04	0.11	0.14	-0.03
Grade 8	0.16	0.21	-0.04	0.17	0.20	-0.03
Grade 9	0.24	0.29	-0.05	0.29	0.32	-0.03

6.8 How important is telesecondary schooling?

As previously described, children/youth from *Prospera*-beneficiary households are more likely to live in rural areas where telesecondary schools are available and they more often attend those schools. Our previous results indicate that students who attend these types of schools see significant improvement over the grades in their mathematics test scores.[40]

In this section, we use our estimated model to simulate what educational outcomes would look like were the telesecondary school option not available. The simulation takes into account that students might then have to travel further distances to get to different types of school or have to dropout if telesecondary schools were their only option. Table 17 shows the distribution of local lower-secondary school types in the data (baseline) and after removing telesecondary schools. 7.5% of students live in areas where the telesecondary schools are the only options locally available, so they would have to dropout of school without this option. Table 18 shows the simulated dropout proportion (by grade 9) for *Prospera*-beneficiary children/youth when the telesecondary schools are removed from choice sets. dropout increases dramatically from 0.22 to 0.58. Average educational attainment over the six years of our observation period (up to grade 9 for the children who were not retained) falls from 8.7 grades to 7.5 grades. Despite using different data sources and approaches, our results comparable with Navarro-Sola (2019). Using a difference-in-difference approach and Employment and Occupation National Survey (EONS) dataset, she finds that the construction of an additional telesecondary per 50 children encourages 10 individuals to enroll in junior secondary education, causing an average increase of one additional year of education among individuals that could have attended it.

Table 17: The distribution of school choice sets with and without the telesecondary option

	Baseline	No telesecondary
General, technical and telesecondary	0.705	N/A
General and telesecondary	0.072	N/A
General and technical	0.052	0.757
Telesecondary and technical	0.077	N/A
Only general	0.012	0.089
Only telesecondary	0.071	N/A
Only technical	0.007	0.080
No local schools	N/A	0.074

Table 18: Simulated dropout and educational attainment for *Prospera* telesecondary enrollees when the telesecondary option is removed

	With telesecondary	Without telesecondary
Dropout rate	0.22	0.58
Education attainment	8.70	7.46

[40]The copying rates were somewhat higher in telesecondary schools, but our estimates adjust for copying.

6.9 Quantifying the importance of the dynamic selection

The multi-equation modeling framework we implemented controlled for multiple sources of dynamic selection - due to dropout, school choice and grade retention - as well as for cheating and missing data. In the US context, value-added models are usually implemented without accounting for selection. Arguably, in Mexico, selection is a more important issue, given that school enrollment drops significantly in lower-secondary school grades. In addition, there are multiple types of schools along with school choice.[41]

To explore the importance of controlling for multiple sources of selection within our dynamic academic achievement model, we compare our baseline results with results obtained from a simpler value-added model. In particular, we estimate the following value-added regression grade-by-grade:

$$A_{ia}^{m} = \delta_0^{mg} + A_{i,a-1}\delta_1^{g} + \delta_2^{mg}P_i + Z_{ia}^{A}\delta_3^{mg} + \omega_{ia}^{mg}$$

Compared with equation 2, in this regression the contemporaneous *Prospera* program effect δ_2^{mg} is homogeneous across school types and we do not model school choice. Also, this model does not include permanent unobservable heterogeneity (types). The cumulative program effect can be calculated as:

$$\Delta_g^m = \begin{cases} \delta_2^{m5} & \text{if } g = 5 \\ \delta_2^{mg} + \delta_1^{g}\Delta_g & \text{if } g > 5 \end{cases}$$

where Δ_g^m represents the cumulative effect for subject m in grade g, $\Delta_g = [\Delta_g^1, \Delta_g^2]$ is a 2×1 vector of cumulative effects for both math and Spanish.

Table 19: Cumulative program impacts

	Math score		Spanish score	
	(1) Baseline	(2) Simple VA	(3) Baseline	(4) Simple VA
Grade 5	0.0	0.2	-1.3	-2.2
Grade 6	-0.9	-1.5	-2.2	-4.4
Grade 7	9.3	6.1	3.4	-0.5
Grade 8	10.2	7.4	4.6	1.0
Grade 9	12.4	9.7	3.5	0.1

†Simple VA = simple value added model.

Table 19 compares our baseline results for mathematics and Spanish test scores (Columns (1) and (3)) with results generated from the simpler model (Columns (2) and (4)). The cumulative program impacts derived from the simpler model are noticeably smaller, especially in the lower-secondary grades. As previously noted, not controlling for dropping-out will tend to downward bias the impact estimates if the program causes students at the margin of dropping-out to stay in school longer. Also,

[41] Cameron and Heckman (2001) deal with the problem of selection in US high schools, but they do not analyze test-score data.

our previous analysis showed that *Prospera* beneficiaries attended telesecondary schools in greater numbers and that these schools were particularly effective in teaching mathematics. These benefits are not captured in the simpler model, which does not allow heterogeneous impacts across different types of schools. Overall, we find the richer modeling framework is required to capture heterogeneous program impacts and to control for sources of selection bias.

7 Exploring school-quality and student-engagement differences by school type

Our analysis found substantial differences in the estimated production-function parameters for different school types. In this section, we explore differences in the characteristics of different school types, based on a separate school census database (called the 911 database). We also examine differences in reported student effort and student engagement across different types of schools.

7.1 School characteristics

The first two columns of Table 20 show average school characteristics for general and indigenous primary schools. Indigenous primary schools have on average 94 students in comparison to 174 students in general primary schools. The percentages of students who are disabled ranges from 1-2%. Despite having overall fewer students, the student-teacher ratio in indigenous schools is higher - 33 in comparison to 24. Another difference is that teachers in indigenous schools are more likely to have only an upper-secondary school degree (17% in comparison to 3%). At the same time, the fraction of teachers with an undergraduate or higher degree is 7 percentage points higher. Thus, teacher schooling attainment exhibits higher variance in indigenous schools.

The last three columns of Table 20 compare the average school characteristics in general, technical and telesecondary schools. Technical schools tend to be larger, with an average enrollment of 395. General schools have an average enrollment of 296 and telesecondary schools have an average enrollment of 75. Again, the proportion of disabled students across all types of schools is 1-2%. The student-teacher ratio is 14 in general schools, 19 in technical schools and 24 in telesecondary schools.[42] Thus, we see a general pattern of the smaller schools in rural areas having higher student-teacher ratios, which could either reflect that video learning is less teacher-intensive or teacher or resource shortages are more common in more remote areas. Comparing the teacher educational profiles across the different kinds of local secondary schools, we see that they are fairly similar. The main difference is that general-school teachers are more likely to have an undergraduate degree rather than a teaching college degree, compared to teachers in technical and telesecondary schools.

[42]These tabulations use regular teachers and exclude art and music teachers who often travel to teach at multiple schools.

Table 20: Mean lower-secondary school characteristics by school type†

	Primary		Secondary		
Characteristic	General	Indigenous	General	Technical	Telesecondary
Number of students	174	94	296	395	75
	(175)	(95)	(244)	(247)	(64)
Proportion disabled	0.02	0.01	0.01	0.02	0.01
	(0.06)	(0.07)	(0.04)	(0.05)	(0.03)
Student-teacher ratio	24	33	14	19	24
	(18)	(12)	(7)	(8)	(9)
Teachers with HS degree	0.03	0.17	0.03	0.02	0.03
	(0.09)	(0.30)	(0.07)	(0.06)	(0.16)
Teachers with teacher college	0.47	0.26	0.34	0.42	0.41
	(0.35)	(0.34)	(0.34)	(0.32)	(0.42)
Teachers with undergraduate degree	0.47	0.56	0.54	0.47	0.44
	(0.34)	(0.39)	(0.34)	(0.32)	(0.42)
Teachers with post-grad degree	0.03	0.01	0.07	0.08	0.11
	(0.10)	(0.07)	(0.12)	(0.11)	(0.23)

†Tabulations based on a school census dataset called the 911 data.

7.2 Student effort and engagement

In Table 21, we examine whether measures of student effort and student engagement in the classroom vary by type of school, including some controls for student-background characteristics (gender and lagged test scores). The effort and engagement variables are each reported in five ordered categories, so we estimate an ordered-probit model. The sample used in estimation are students who are beginning lower-secondary school in 2011 (i.e. 7th graders).[43] Students attending technical schools report studying the most hours. Students with higher lagged mathematics and Spanish scores report studying more hours, paying attention more often and participating in class more often. Interestingly, students in telesecondary schools report higher rates of paying attention and participating in class. There are substantial differences by gender in study time. Being female is associated with greater study effort and paying attention more frequently. However, girls report on average lower rates of participation in class.

8 Conclusions

Prior literature demonstrated substantial effects of the Mexican *PROGRESA/Oportunidades/Prospera* CCT program on educational attainment and schooling progression. Little was known about how the program affected children's academic achievement because the data to study that question were not available. Using newly available nationwide school-roster and test-score data, we develop and implement a model of school progression and academic achievements. The empirical framework features

[43]Our model was estimated using a sample of fourth graders in 2008. These students would be in 7th grade in 2011 if they stayed in school and progressed on time.

Table 21: How student engagement varies by school type in 7th grade†

	Hours study	Pay attention	Part. in Class
Technical	0.036*	0.043*	-0.035
Telesecondary	0.020	0.230*	0.353*
Lag mathematics	0.0005*	0.0008*	0.0010*
Lag Spanish	0.0004*	0.0013*	0.0010*
Female	0.186*	0.131*	-0.123*

†Estimates derived from ordered-probit models. The omitted category is general school. * denotes significant at <0.001 level.

value-added models for academic achievement, school-choice models that include the dropout option and incorporate local-labor-market work opportunities, measurement equations to allow for missing test score data or measurement error arising from copying, and grade-retention models. Our analysis incorporates rich observed heterogeneities in family and child characteristics and unobserved heterogeneity in the form of discrete types, which enter multiple model equations. The likelihood estimation approach explicitly controls for selective school enrollment/dropping-out and selection into different school types. Model parameters are estimated using multiple linked administrative and survey datasets as well as geocode data on school locations, used to characterize the individual school-choice sets.

The data show that children from *Prospera*-beneficiary households live in less urban areas and enroll in distance-learning schools (telesecondary) at much higher rates, so another goal of our analysis is to understand how academic achievement depends on the type of school attended. The question of whether a distance-learning modality is an effective way of teaching in comparison to fully in-person approaches is of considerable independent interest. Many countries face problems of how to provide access to high-quality education for students living in rural areas. There is scant evidence on the effectiveness of distance learning in such contexts.

We find the following key results. First, the *Prospera* program did not significantly impact test scores in grades 5 and 6 in either general or indigenous schools. However, there are positive and statistically significant impacts on test scores in grades 7, 8, 9 with the larger overall average impacts in mathematics (0.09-0.13 standard deviations) than in Spanish (0.03-0.05 standard deviations). The pattern of larger impacts at higher grade levels is perhaps to be expected given that the cash transfer amounts are significantly higher in lower-secondary than in primary grades. Second, there is a consistent pattern of larger impacts for children/youth in higher propensity-score quartiles, which shows that the program has the greatest effects on the most disadvantaged children. We also find that the *Prospera* program decreases school dropouts by 4 percentage points, at the sixth to seventh grade transition.

Third, the value-added parameter estimates indicate that lagged test scores are important determinants of current test scores, implying that *Prospera*-program effects on test scores accumulate over time. Our empirical findings show the importance of accounting for the dynamic nature of academic achievement production to quantify the program participation gains accruing over multiple years. We also find evidence of gender differences on test scores, with boys scoring higher on average on math-

ematics tests than girls and girls scoring higher on average on Spanish tests. The gender gaps in mathematics widen from primary to secondary grades.

Fourth, our analysis considers that a small proportion of students in each grade were identified as having cheated (copied). Although cheating rates overall are low, we find that they are somewhat higher in telesecondary schools and at higher grade levels (grades 8 and 9). We develop an approach to account for possible test-score distortion (one-sided measurement error) arising from copying, which can affect both the dependent and independent (lagged) variables in the value-added model. Even with the copying adjustment, the results show that telesecondary schools are an effective type of school, particularly in teaching mathematics. Children from disadvantaged backgrounds finish primary school with gaps in mathematics and attending telesecondary school helps to close these gaps. When we simulate the effects of removing the telesecondary school option for the children who attend these schools, their dropout rate prior to grade 9 increases from 21% to 58% and educational attainment decreases by 1.2 grades. Thus, telesecondary schools are important to *Prospera*'s success in improving educational outcomes for a relatively disadvantaged student population.

We also explored possible mechanisms to explain telesecondary schools' effectiveness. The average teacher educational qualifications are for the most part similar to other schools and the pupil-teacher ratios are somewhat higher. However, the youth who attend these schools report that they pay attention more often and participate in class more often. The video lectures and teaching materials used at the telescondary schools were created by highly qualified teachers, which may facilitate students' learning. Also, the video topics adhered closely to the curriculum that is tested on the national exams. Two recent studies (Fabregas (2019) and Navarro-Sola (2019)) estimate relatively large returns to telesecondary schooling on earnings, consistent with significant learning occurring in telesecondary schools.

Overall, our results indicate that *Prospera* was not only effective in increasing school enrollments but also that the program led to significant positive impacts on academic achievement in mathematics and Spanish, with the most disadvantaged children/youth experiencing the greatest benefit and with distant learning through the telesecondary schools playing a critical role.

A Appendices

A.1 Data sources

This appendix contains some additional information about the sample sizes and about the data elements in the different surveys that we use. We also describe how we obtain the GIS location data used in estimating the school choice model.

The student survey: Students answer questions related to their school and home life. They are asked about their own effort on school work in terms of how much attention they pay in class, whether they participate in class and how many hours they study each day. There are also questions about the home environment, for example, how many siblings they have.

The parent survey: For the parent survey, there are questions about the socioeconomic status and some questions about early childhood, such as whether the child attended preschool and whether parents read to the child when they were young. We use information on parents' education, work status, on the household size, housing characteristics and on household assets.

The geographic location data. : In the *ENLACE* data, each school has a unique identifier. Several years of data also contain geographic data for each school, including the state, municipality and locality where the school is located. Mexico has 31 states and a federal district, and within these districts there are 2,448 municipalities. In the *ENLACE* test score data, there are schools recorded in all states, close to 2,000 municipalities and over 20,000 localities. Given that the school IDs are constant over time, it is possible to use the years that contain geographic data, and create a database containing the majority of the schools and their respective locations. This database was merged with census data which allows us to link with information on the locality that the school is in with longitude and latitude coordinates. We used R software to obtain a distance measure of the distances between primary schools and between primary and secondary schools to determine the choice sets available to families given their location.

A.2 Sample restrictions

Our initial sample includes individuals who fill in surveys in year 2008 and collect their Prospera status in 2010, with a total size of 189,100 individuals.

1. Drop individuals with test scores that suggest that they did not do the exam (< 100) or those whose ages are outside the range $7 - 17$ or missing. This leaves a sample size of 186,572.

2. Drop individuals whose secondary school types are not coded as general, telesecondary or technical types. The leaves a sample size is 186,282.

3. Drop students who attend their primary school and secondary school in different states. The leaves a sample size of 183,992.

4. Drop students whose grade information or test scores are missing for more than one period, but keep the students who only miss their test scores for one period. The remaining sample size is 179,981.

5. Drop the observations whose primary school names are missing. This restriction leaves 173,580.

6. Drop the observations for which the distance information is missing. The remaining sample size is 162,936.

7. Drop individuals that miss any of the key variables (age, gender, retention, urban dummy, cheating factor, region, internet, computer, first language, dad at home, mom at home, number of household members). We allow some variables (parental education, household income, parental working status) to have missing values. The left sample size is 148,230.

8. Drop individuals whose propensity score is in the bottom 1% (pscore<0.038, trimming as described in the text). The remaining sample size is 105,256. (We drop 299 *Prospera* students and 42,675 non-Prospera students with this restriction.)

Our final sample has 105,256 different individuals and 589,371 individual-period observations.

A.3 Local wage imputation method

Our administrative test score database does not have students' wage information. We impute the potential wages that students could earn using data from the Mexican 2010 census obtained through the IPUMS site: https://international.ipums.org/international-action/sample_details/country/mx#tab_mx2010a. The census contains the age and gender, working status, school enrollment status, and the wages earned for children across Mexico. It also includes other information such as, the education level of their parents, family income and descriptive statistics about the home. Lastly, the municipality in which they live is also recorded. The full list of variables that we use appears in table 22. We exclude individuals whose age is <12 years old or >20 years old, as well as children/youth for whom the school attendance status is undefined. We further exclude students who have already finished high school (educational attainment levels $\geq$ 13).

We estimate a wage regression for the working children/youth sample. The dependent variable is hourly wage, which is calculated by monthly income (Inc) divided by 4 and divided by hours worked in the last week (Wkhs).[44] To account for selection into working in estimating the wage offer parameters, a Heckman (1979) selection model is estimated. The wage regression and labor force participation equations include age, a school attendance indicator, educational attainment, parents' education, missing indicators for parents' education, urban-rural dummies, north-south dummies, and municipality dummies. In addition, variables representing family socioeconomic levels, such as family income (household income) and home infrastructure (home electric access, home piped water access, home internet access, home computer access) are used as exclusion restrictions that affect selection into working but not the wage offers directly.

[44]We trimmed the hourly wage distribution at 99th quantile.

Table 22: Variables used from Census

Variable	Note	Name in database
Age	Age of subject	MX2010A AGE
Enroll	School enrollment dummy (1 = yes, 2 = no)	MX2010A_SCHOOL
Inc	Income of individual for the last month	MX2010A_INCOME
HInc	Household's income from work	MX2010A_INCHOME
Wkhs	Number of hours worked in the last week	MX2010A_HRSWORK
Edu	Educational attainment (in years)	MX2010A_EDATTAIN
Gender	Gender; Male = 1, Female = 2	MX2010A_SEX
Empl	Employment status	MX2010A_EMPSTAT
Pps	Position at work	MX2010A_CLASSWK
Edu_Mom	Mother's Educational attainment (in years)	MX2010A_EDATTAIN_MOM
Edu_Dad	Father's Educational attainment (in years)	MX2010A_EDATTAIN _POP
Electricity	Access to electricity	MX2010A_ELECTRIC
PipWater	Access to piped water	MX2010A_PIPEDWTR
Internet	Access to the internet	MX2010A_INTERNET
Comp	Access to computer	MX2010A_COMPUTR
State	State code	GEO1_MX2010
Mun	Municipality code	GEO2_MX2010
Urban	Urban-rural status; 1 = rural, 2 = urban	URBAN

Using the estimated coefficients from the probit estimation of the labor force participation equation, we form control functions for each student (the inverse Mills ratio λ). The second stage regression has hourly wages as the dependent variable, and regressions are done separately for girls and boys. The hourly wage specification includes municipality level fixed effects, which allows for substantial regional variation. The results are reported in Table 23. We use the regression estimates to impute hourly wage offers to the children/youth in our analysis sample, based on their observed characteristics.

Table 23: Wage regression with Heckman selection correction

	Boys		Girls	
Coefficient	Estimate	S.E.	Estimate	S.E.
Age	0.682***	(0.173)	0.593**	(0.239)
Enroll	5.109***	(1.346)	5.301***	(1.885)
Edu_mom_missing	−1.483	(1.517)	3.561	(2.438)
Edu_dad_missing	0.08	(1.317)	−0.051	(2.097)
Edu	0.394**	(0.175)	−0.459*	(0.279)
Edu_mom	−0.535***	(0.163)	0.389	(0.265)
Edu_dad	−0.190	(0.179)	−0.039	(0.291)
Urban	−0.096	(1.719)	5.753**	(2.442)
λ	0.023	(0.017)	−0.015	(0.029)
Age*Enroll	−0.252***	(0.081)	−0.292***	(0.111)
Age*Edu_mom_missing	0.08	(0.085)	−0.216	(0.135)
Age*Edu_dad_missing	0.006	(0.074)	0.001	(0.117)
Age*Edu	−0.014	(0.010)	0.042***	(0.015)
Age*Edu_mom	0.034***	(0.009)	−0.019	(0.015)
Age*Edu_dad	0.013	(0.010)	0.006	(0.016)
Age*Urban	0.038	(0.102)	−0.306**	(0.141)
Urban*Enroll	−2.634***	(0.786)	−3.666***	(1.104)
Urban*Edu_mom_missing	0.84	(0.943)	−2.197	(1.459)
Urban*Edu_dad_missing	0.081	(0.827)	0.105	(1.280)
Urban*Edu	−0.019	(0.110)	0.072	(0.169)
Urban*Edu_mom	0.229**	(0.097)	−0.203	(0.154)
Urban*Edu_dad	0.137	(0.106)	−0.021	(0.168)
North*Age	−0.585***	(0.156)	−0.302	(0.231)
North*Enroll	−4.866***	(1.170)	0.438	(1.743)
North*Edu_mom_missing	−0.161	(1.621)	−2.976	(2.625)
North*Edu_dad_missing	−0.092	(1.461)	−4.095*	(2.409)
North*Edu	−0.080	(0.187)	−0.301	(0.301)
North*Edu_mom	0.177	(0.159)	−0.504*	(0.259)
North*Edu_dad	0.113	(0.162)	−0.037	(0.260)
North*Urban	0.144	(0.124)	−0.341	(0.213)
Age*Enroll*Urban	0.128***	(0.047)	0.190***	(0.064)
Age*Edu_mom_missing*Urban	−0.012	(0.053)	0.162**	(0.081)
Age*Edu_dad_missing*Urban	−0.006	(0.047)	0.012	(0.072)
Age*Edu*Urban	0.0001	(0.006)	−0.007	(0.009)
Age*Edu_mom*Urban	−0.013**	(0.005)	0.013	(0.009)
Age*Edu_dad*Urban	−0.008	(0.006)	0.001	(0.009)
Age*Enroll*North	0.278***	(0.068)	−0.029	(0.099)
Age*Edu_mom_missing*North	0.021	(0.090)	0.159	(0.144)
Age*Edu_dad_missing*North	−0.002	(0.081)	0.21	(0.133)
Age*Edu*North	0.006	(0.010)	0.012	(0.016)
Age*Edu_mom*North	−0.007	(0.009)	0.028*	(0.014)
Age*Edu_dad*North	−0.007	(0.009)	0.001	(0.014)
Observation	174,905		76,978	
R^2	0.216		0.251	

Sample: Mexico 2010 Census. we exclude individuals whose age is beyond a suitable school age (< 20 or > 12), and whose school attendance status is undefined. We further exclude students who have already finished high schools (educational attainment levels ≥ 13). We defined a dummy variable $Edu\_mom\_missing = 1$ if Edu_mom is missing or unknown, a dummy variable $Edu\_dad\_missing = 1$ if Edu_dad is missing or unknown. And the variable North is also a binary variable whether the municipality is in the North or South region of Mexico (1 = North, 0 = South). The dependent variable is hourly wage, which is defined as the monthly income (Inc) divided by 4hours worked in the last week (Wkhs). The hourly wage is trimmed at the upper 99th quantile. λ is the inverse Mills ratio calculated from the first-stage probit regression. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

A.4 Propensity score model

As described in the text, we estimate a probit model for the probability that a child/youth comes from a *Prospera* beneficiary family. The precise eligibility criteria are not publicized but they are known to depend on household composition, household assets and demographic characteristics. Table A.2 shows the estimated coefficients from the model. The model includes missing indicators to account for individual item nonresponse. The child being female makes it more likely that the family participates. Also, having a larger household increases the probability f being a beneficiary. Having higher income (>30) makes it less likely to be a beneficiary. Higher education categories for the mother or father make it less likely that a family participates (the omitted category is less than primary education). If the mom works, the family is less likely to participate in *Prospera*. If the dad works more than 8 hours per day, the family is more likely to participate. Car ownership, having a refrigerator, owning a clothes washer or having sanitary facilities in the home makes all make it less likely that the family participates in *Prospera*. The percent correctly classified under the model is 89%.

Table 24: Propensity score model (probit)

coefficient	estimate	std. error
intercept	-0.762	0.022
female	-0.005	0.008
number of sibling	0.387	0.028
numsibmiss	-0.099	0.002
number total at home	-0.448	0.028
numhomemiss	-0.029	0.002
family income 15K-30K	-0.158	0.030
family income > 30K	0.284	0.041
famincmiss	-0.073	0.040
mom completed primary	-0.016	0.025
mom completed secondary	0.088	0.010
mom bachalaureate or tech	0.185	0.018
mom BA or more	0.455	0.017
dad completed primary	0.761	0.035
dad completed secondary	0.161	0.011
dad bachalaureate or tech	0.222	0.018
dad BA or more	0.495	0.017
mom ed miss	0.647	0.026
dad ed miss	0.276	0.035
cartruck	0.242	0.027
cartruckmiss	0.302	0.010
house has dirt floor	0.027	0.029
dirtfloormiss	-0.166	0.011
house drainage connected to public	0.041	0.025
drainagemiss	0.564	0.010
sanitary fac in home	0.217	0.027
sanitarymiss	0.170	0.014
mom works 4+ hrs/day	0.116	0.029
momworks < 4 or retired	0.274	0.012
dad works 4-8 hrs/day	0.213	0.015
dad works pensioned retired or not present	0.497	0.011
house has a refridgerator	0.453	0.011
fridgemiss	0.152	0.027
cook and sleep in the same room	0.303	0.020
cooksleepmiss	0.288	0.012
house has a clothes washer	0.207	0.030
clotheswashmiss	-0.114	0.013
Student spoke first indigenous language	-0.029	0.024
indiglangmiss	0.207	0.011

The model also includes state fixed effects. The omitted category for dad working is works 8+ hours per day. The omitted category for mom working is engaged in housework. The percent correctly classified under the model is 89%.

References

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of political economy*, 113(1):151–184.

Andersen, C. T., Reynolds, S. A., Behrman, J. R., Crookston, B. T., Dearden, K. A., Escobal, J., Mani, S., Sánchez, A., Stein, A. D., and Fernald, L. C. (2015). Participation in the juntos conditional cash transfer program in peru is associated with changes in child anthropometric status but not language development or school achievement. *The Journal of nutrition*, 145(10):2396–2405.

Angrist, J., Bettinger, E., Bloom, E., King, E., and Kremer, M. (2002). Vouchers for private schooling in colombia: Evidence from a randomized natural experiment. *American Economic Review*, 92(5):1535–1558.

Angrist, J., Bettinger, E., and Kremer, M. (2006). Long-term educational consequences of secondary school vouchers: Evidence from administrative records in colombia. *American economic review*, 96(3):847–862.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Attanasio, O. P., Meghir, C., and Santiago, A. (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.

Aucejo, E. and James, J. (2021). The path to college education: The role of math and verbal skills. *Journal of Political Economy*, 129(10):2905–2946.

Baird, S., Ferreira, F. H., Özler, B., and Woolcock, M. (2014). Conditional, unconditional and everything in between: a systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, 6(1):1–43.

Barham, T., Macours, K., and Maluccio, J. A. (2013). More schooling and more learning?: Effects of a three-year conditional cash transfer program in nicaragua after 10 years.

Behrman, J. R., Parker, S. W., and Todd, P. E. (2005a). Long-term impacts of the oportunidades conditional cash transfer program on rural youth in mexico. Technical report, IAI Discussion Papers.

Behrman, J. R., Parker, S. W., and Todd, P. E. (2009). Schooling impacts of conditional cash transfers on young children: Evidence from mexico. *Economic development and cultural change*, 57(3):439–477.

Behrman, J. R., Parker, S. W., and Todd, P. E. (2011). Do conditional cash transfers for schooling generate lasting benefits? a five-year followup of progresa/oportunidades. *Journal of Human Resources*, 46(1):93–122.

Behrman, J. R., Sengupta, P., and Todd, P. (2005b). Progressing through progresa: An impact assessment of a school subsidy experiment in rural mexico. *Economic development and cultural change*, 54(1):237–275.

Bharadwaj, P., De Giorgi, G., Hansen, D., and Neilson, C. A. (2016). The gender gap in mathematics: Evidence from chile. *Economic Development and Cultural Change*, 65(1):141–166.

Boardman, A. E. and Murnane, R. J. (1979). Using panel data to improve estimates of the determinants of educational achievement. *Sociology of Education*, 52(2):113–12.

Borgheson, E. and Vasey, G. (2021). The marginal returns to distance education: Evidence from mexico's telesecundarias. *manuscript*.

Bravo, D., Mukhopadhyay, S., and Todd, P. E. (2010). Effects of school reform on education and labor market performance: Evidence from chile's universal voucher system. *Quantitative economics*, 1(1):47–95.

Cameron, S. V. and Heckman, J. J. (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males. *Journal of Political economy*, 106(2):262–333.

Cameron, S. V. and Heckman, J. J. (2001). The dynamics of educational attainment for black, hispanic, and white males. *Journal of political Economy*, 109(3):455–499.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014a). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9):2593–2632.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014b). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9):2633–79.

Cunha, F. and Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of human resources*, 43(4):738–782.

Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1:697–812.

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

De Hoyos, R., Estrada, R., and Vargas, M. J. (2018). *Predicting individual wellbeing through test scores: evidence from a national assessment in Mexico*. The World Bank.

Epple, D., Jha, A., and Sieg, H. (2018). The superintendent's dilemma: Managing school district capacity as parents vote with their feet. *Quantitative Economics*, 9(1):483–520.

Fabregas, R. (2019). Broadcasting education: The long-term effects of mexico's telesecundarias. Technical report, Working Paper.

Figlio, D. N. and Rouse, C. E. (2006). Do accountability and voucher threats improve low-performing schools? *Journal of Public Economics*, 90(1-2):239–255.

Fiszbein, A. and Schady, N. R. (2009). *Conditional cash transfers: reducing present and future poverty*. World Bank Publications.

Gallego, F. and Hernando, A. (2009). On the determinants and implications of school choice: Structural estimates and simulations for chile. *Documento de Trabajo*, (343).

Glewwe, P. (2002). Schools and skills in developing countries: education policies and socioeconomic outcomes. *Journal of economic literature*, 40(2):436–482.

Grant, M. J. and Behrman, J. R. (2010). Gender gaps in educational attainment in less developed countries. *Population and Development Review*, 36(1):71–89.

Hadna, A. H. and Kartika, D. (2017). Evaluation of poverty alleviation policy: Can conditional cash transfers improve the academic performance of poor students in indonesia? *Cogent Social Sciences*, 3(1):1295548.

Hanushek, E. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, 14:351–388.

Hastings, J., Kane, T. J., and Staiger, D. O. (2009). Heterogeneous preferences and the efficacy of public school choice. *NBER Working Paper*, 2145:1–46.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society*, pages 271–320.

Heckman, J. J., Humphries, J. E., and Veramendi, G. (2016). Dynamic treatment effects. *Journal of econometrics*, 191(2):276–292.

Heckman, J. J., Humphries, J. E., and Veramendi, G. (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy*, 126(S1):S197–S246.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, 64(4):605–654.

Heckman, J. J. and Navarro, S. (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics*, 136(2):341–396.

Hsieh, C.-T. and Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from chile's voucher program. *Journal of public Economics*, 90(8-9):1477–1503.

Kane, T. J., McCaffrey, D. F., Miller, T., and Staiger, D. O. (2013). Have we identified effective teachers? validating measures of effective teaching using random assignment. In *Research Paper. MET Project. Bill & Melinda Gates Foundation*. Citeseer.

Kane, T. J. and Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. Technical report, National Bureau of Economic Research.

Keane, M. P., Todd, P. E., and Wolpin, K. I. (2011). The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. In *Handbook of labor economics*, volume 4, pages 331–461. Elsevier.

Leite, P. G., Narayan, A., and Skoufias, E. (2011). How do ex ante simulations compare with ex post evaluations? evidence from the impact of conditional cash transfer programs. *Evidence from the Impact of Conditional Cash Transfer Programs (June 1, 2011). World Bank Policy Research Working Paper*, (5705).

Manski, C. F. (1988). Identification of binary response models. *Journal of the American statistical Association*, 83(403):729–738.

McEwan, P. J. (2001). The effectiveness of public, catholic, and non-religious private schools in chile's voucher system. *Education economics*, 9(2):103–128.

Mroz, T. A. (1999). Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. *Journal of Econometrics*, 92(2):233–274.

Navarro-Sola, L. (2019). Secondary school expansion through televised lessons: The labor market returns of the mexican telesecundaria. Technical report, Working paper.

Neal, D. (1997). The effects of catholic secondary schooling on educational achievement. *Journal of Labor Economics*, 15(1, Part 1):98–123.

Neilson, C., Allende, C., and Gallego, F. (2019). Approximating the equilibrium effects of informed school choice.

Parker, S. W. and Todd, P. E. (2017). Conditional cash transfers: The case of progresa/oportunidades. *Journal of Economic Literature*, 55(3):866–915.

Parker, S. W. and Vogl, T. (2018). Do conditional cash transfers improve economic outcomes in the next generation? evidence from mexico. Technical report, National Bureau of Economic Research.

Paxson, C. and Schady, N. (2010). Does money matter? the effects of cash transfers on child development in rural ecuador. *Economic development and cultural change*, 59(1):187–229.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2):417–458.

Rouse, C. E. (1998). Private school vouchers and student achievement: An evaluation of the milwaukee parental choice program. *The Quarterly journal of economics*, 113(2):553–602.

Sapelli, C. and Vial, B. (2002). The performance of private and public schools in the chilean voucher system. *cuadernos de economía*, 39(118):423–454.

Schultz, T. P. (2004). School subsidies for the poor: evaluating the mexican progresa poverty program. *Journal of development Economics*, 74(1):199–250.

Simoes, A. A. and Sabates, R. (2014). The contribution of bolsa família to the educational achievement of economically disadvantaged children in brazil. *International Journal of Educational Development*, 39:141–156.

Snilstveit, B., Gallagher, E., Phillips, D., Vojtkova, M., Eyers, J., Skaldiou, D., Stevenson, J., Bhavsar, A., and Davies, P. (2017). Protocol: Interventions for improving learning outcomes and access to education in low-and middle-income countries: a systematic review. *Campbell Systematic Reviews*, 13(1):1–82.

Summers, A. A. and Wolfe, B. L. (1977). Do schools make a difference? *The American Economic Review*, 67(4):639–652.

Todd, P. E. and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485):F3–F33.

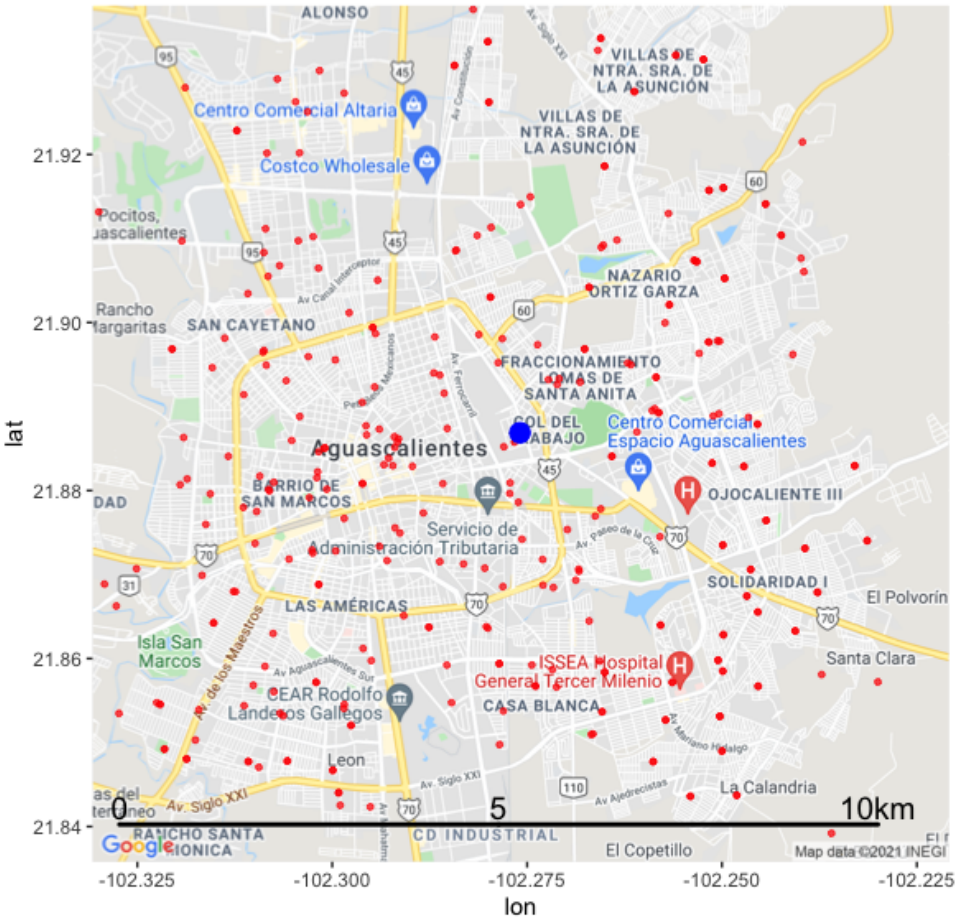
Todd, P. E. and Wolpin, K. I. (2006). Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American economic review*, 96(5):1384–1417.

Vasey, G. E. (2021). The impact of child labor on student enrollment, effort and achievement: Evidence from mexico. *manuscript*.

B Online Appendix (Not for publication)

B.1 An illustration of local primary school sessions

Figure 6: Local primary school sessions around Aguascalientes



This map shows the primary school sessions around Aguascalientes, which is a large city in central Mexico. It contains 316 school sessions with 250 unique coordinates, indicating multiple school sessions share the same physical teaching place.

B.2 Model estimates

Table B.1: Mathematics value-added estimates

	General		Indigenous		General			Telesecondary			Technical		
	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Lag Mathematics	0.55	0.54	0.36	0.40	0.45	0.47	0.50	0.33	0.43	0.41	0.41	0.46	0.47
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lag Spanish	0.15	0.13	0.15	0.15	0.26	0.18	0.25	0.25	0.15	0.24	0.27	0.16	0.26
	(0.00)	(0.00)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prospera score quartile													
P*Q1	0.62	-0.80	-1.05	2.89	2.76	2.49	-0.77	0.29	-2.86	2.25	1.73	2.81	2.34
	(3.00)	(2.88)	(21.81)	(16.49)	(4.14)	(3.56)	(4.50)	(12.52)	(10.99)	(11.83)	(5.65)	(5.17)	(7.38)
P*Q2	0.31	-1.83	-2.61	-3.37	2.17	-2.20	2.59	2.45	2.31	1.92	2.02	1.87	2.64
	(1.71)	(1.64)	(9.99)	(8.39)	(2.75)	(2.58)	(3.03)	(5.42)	(4.34)	(5.28)	(3.49)	(3.09)	(3.90)
P*Q3	-0.15	1.02	2.02	1.98	4.00	-1.01	4.36	6.17	6.77	6.09	4.34	3.23	3.39
	(1.21)	(1.14)	(5.40)	(4.90)	(2.14)	(1.97)	(2.29)	(3.11)	(2.64)	(3.13)	(2.56)	(2.17)	(2.73)
P*Q4	-0.05	-1.05	-0.12	-3.33	13.57	7.50	6.97	12.39	6.63	9.32	12.54	5.75	7.74
	(1.10)	(1.05)	(3.71)	(3.23)	(2.24)	(1.97)	(2.45)	(2.41)	(2.06)	(2.45)	(2.36)	(2.10)	(2.58)
Education cat. (dad)													
Below primary school	-4.38	-0.86	-5.15	-2.86	-8.10	-1.30	0.73	-15.11	7.69	-1.10	-3.63	1.41	-6.93
	(2.48)	(2.40)	(11.25)	(10.36)	(4.06)	(3.60)	(4.35)	(7.55)	(6.85)	(7.98)	(5.31)	(4.46)	(5.61)
Primary school completed	-2.98	1.08	-1.63	-5.52	-5.77	-2.41	2.40	-10.44	6.60	-1.64	-1.81	4.72	-6.84
	(2.50)	(2.42)	(11.45)	(10.55)	(4.07)	(3.59)	(4.32)	(7.66)	(6.96)	(8.08)	(5.37)	(4.49)	(5.63)
Secondary or below	-2.69	1.78	-3.52	2.10	-1.78	-0.74	3.67	-11.48	8.09	-0.76	-1.55	6.23	-5.90
	(2.45)	(2.37)	(11.84)	(10.71)	(3.96)	(3.50)	(4.22)	(7.69)	(6.97)	(8.06)	(5.27)	(4.38)	(5.46)
College or above	3.83	6.51	10.03	-2.25	4.69	3.62	6.51	-5.78	17.84	2.46	4.99	8.92	-1.91
	(2.57)	(2.48)	(13.29)	(12.15)	(4.09)	(3.58)	(4.36)	(8.62)	(7.63)	(8.83)	(5.46)	(4.53)	(5.66)
Working status (dad)													
Full time	4.50	1.26	-5.46	4.52	4.67	-0.19	0.02	11.64	4.00	11.71	2.18	-2.73	12.21
	(2.02)	(1.98)	(10.30)	(8.95)	(3.30)	(2.94)	(3.59)	(6.27)	(5.47)	(6.49)	(4.61)	(3.85)	(4.83)
Not full time	4.05	1.00	-3.65	1.08	4.99	-1.35	-0.89	8.06	0.28	10.85	1.29	-3.19	5.94
	(1.92)	(1.89)	(10.05)	(8.72)	(3.13)	(2.78)	(3.42)	(6.07)	(5.31)	(6.27)	(4.44)	(3.69)	(4.63)
Father present at home	0.22	1.01	5.69	5.92	3.49	2.77	2.52	5.36	-1.39	4.82	5.29	0.13	7.56
	(0.93)	(0.86)	(4.20)	(3.63)	(1.44)	(1.31)	(1.62)	(2.71)	(2.37)	(2.81)	(1.88)	(1.68)	(2.11)

Table B.1: Mathematics value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Education cat. (mom)													
Primary school	-1.98	-1.69	-7.57	11.90	3.28	-8.63	3.26	12.50	-7.43	17.05	3.37	-7.11	-2.40
	(3.22)	(3.03)	(12.17)	(12.05)	(5.25)	(4.61)	(5.76)	(9.80)	(7.96)	(10.11)	(6.60)	(5.41)	(6.98)
Primary school completed	-0.31	0.11	2.21	20.70	3.80	-7.75	2.88	14.61	-5.45	17.47	3.65	-10.85	-0.46
	(3.23)	(3.04)	(12.43)	(12.26)	(5.23)	(4.61)	(5.72)	(9.91)	(8.06)	(10.23)	(6.62)	(5.44)	(7.03)
Secondary or below	0.77	2.16	1.39	26.01	5.63	-5.43	1.38	15.25	-1.91	21.32	4.87	-10.85	-2.87
	(3.21)	(3.02)	(13.00)	(12.66)	(5.20)	(4.55)	(5.68)	(9.97)	(8.11)	(10.23)	(6.58)	(5.39)	(6.97)
College or above	7.40	5.07	0.96	38.31	12.40	-0.99	4.93	22.96	-0.80	28.54	11.74	-7.58	2.22
	(3.34)	(3.14)	(15.62)	(15.03)	(5.32)	(4.67)	(5.80)	(11.01)	(9.07)	(11.24)	(6.84)	(5.57)	(7.17)
Working status (mom)													
Housework	-6.73	-1.90	9.64	-8.48	-4.37	5.34	-5.45	-7.92	-6.05	-4.29	-6.53	10.03	-4.05
	(2.87)	(2.77)	(11.26)	(9.98)	(4.55)	(4.21)	(4.90)	(8.51)	(7.05)	(8.52)	(6.01)	(5.36)	(7.13)
Part time	-7.95	-2.47	14.33	-15.23	-7.88	5.53	-7.13	-10.73	-9.09	-4.67	-9.40	7.65	-8.68
	(2.91)	(2.81)	(11.64)	(10.38)	(4.60)	(4.24)	(4.94)	(8.84)	(7.33)	(8.84)	(6.07)	(5.40)	(7.20)
Full time	-6.57	-2.84	4.45	-25.15	-5.01	3.58	-5.36	-13.51	-9.31	-2.34	-7.42	8.24	-8.92
	(2.99)	(2.87)	(12.11)	(10.90)	(4.70)	(4.34)	(5.04)	(9.12)	(7.56)	(9.09)	(6.24)	(5.56)	(7.35)
Mother present at home	7.69	6.29	6.51	-8.68	10.10	3.42	7.89	5.94	5.07	-2.18	12.42	7.54	-5.42
	(1.17)	(1.10)	(4.66)	(3.93)	(1.86)	(1.72)	(2.18)	(3.22)	(2.79)	(3.38)	(2.43)	(2.21)	(2.73)
Number of people at home													
5 people	2.03	1.41	0.63	8.58	-0.78	-1.18	1.11	2.88	0.08	1.48	1.23	0.94	4.05
	(0.94)	(0.87)	(5.33)	(4.67)	(1.39)	(1.25)	(1.51)	(2.96)	(2.54)	(2.97)	(1.88)	(1.63)	(2.00)
6 people	0.55	0.62	-1.38	5.38	0.21	-1.69	2.29	-0.86	-0.34	-0.08	-0.40	0.88	3.14
	(1.03)	(0.96)	(5.02)	(4.42)	(1.56)	(1.39)	(1.71)	(3.11)	(2.63)	(3.10)	(2.07)	(1.79)	(2.24)
$\geq$ 7 people	0.33	-1.22	0.21	6.14	-4.10	-5.19	-0.53	-0.70	-2.01	3.52	-2.14	0.02	1.48
	(0.88)	(0.82)	(4.19)	(3.63)	(1.36)	(1.22)	(1.49)	(2.59)	(2.22)	(2.62)	(1.80)	(1.57)	(1.97)
Age	-4.01	-4.32	-2.09	-6.81	-4.65	-2.07	-5.47	-5.13	-6.40	-5.46	-3.60	-3.17	-3.63
	(0.74)	(0.73)	(3.47)	(3.13)	(1.16)	(1.00)	(1.23)	(2.38)	(2.03)	(2.44)	(1.50)	(1.29)	(1.64)
Age^2	-2.60	-2.06	-2.16	0.68	-2.99	-3.87	-2.39	-0.99	-0.79	-0.71	-1.97	-2.60	-4.13
	(0.47)	(0.41)	(1.85)	(1.51)	(0.78)	(0.74)	(0.97)	(1.30)	(1.17)	(1.66)	(0.94)	(0.90)	(1.25)
Gender	3.93	3.14	2.65	1.18	7.05	16.70	11.40	0.24	9.27	2.23	6.19	19.49	7.64

Table B.1: Mathematics value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
	(0.66)	(0.63)	(3.13)	(2.77)	(1.04)	(0.94)	(1.13)	(1.98)	(1.69)	(2.01)	(1.37)	(1.20)	(1.48)
First language spoken at home													
Indigenous	-3.46	-6.58	-2.68	-7.59	-8.21	-4.04	-6.26	10.60	5.78	28.55	16.29	14.90	14.81
	(2.39)	(2.30)	(3.80)	(3.31)	(4.62)	(4.21)	(5.32)	(3.55)	(3.24)	(4.06)	(3.49)	(2.89)	(3.81)
Both Spanish and indigenous	3.01	-1.44	10.56	3.76	-5.33	2.53	3.39	4.72	16.17	9.57	1.02	8.74	13.49
	(2.54)	(2.44)	(5.30)	(4.68)	(4.41)	(4.08)	(4.86)	(5.63)	(4.55)	(5.20)	(4.38)	(3.72)	(4.71)
Internet access	-5.16	-5.30	-8.87	-7.60	-6.00	-5.56	-6.04	-17.98	-2.49	-13.77	-7.48	-6.95	-4.42
	(0.99)	(0.94)	(5.44)	(4.69)	(1.48)	(1.33)	(1.61)	(3.46)	(2.91)	(3.58)	(1.97)	(1.73)	(2.12)
Computer access	2.69	2.79	-2.54	-0.13	5.08	5.41	4.11	-6.84	-0.06	-4.87	1.14	5.60	-1.66
	(0.90)	(0.85)	(4.83)	(4.13)	(1.35)	(1.20)	(1.46)	(3.00)	(2.52)	(3.02)	(1.77)	(1.56)	(1.88)
Number of pre-school years													
1 year	4.75	-2.39	2.59	2.68	-7.39	-4.58	-0.39	-5.75	3.36	-17.03	-1.20	-3.71	-7.45
	(2.52)	(2.36)	(9.99)	(8.81)	(4.29)	(3.65)	(4.68)	(6.79)	(5.70)	(7.07)	(5.01)	(4.70)	(6.03)
2 years	2.75	-0.64	12.08	1.89	-3.32	-6.50	-1.67	-9.33	1.88	-14.51	-2.88	-8.64	-1.44
	(2.30)	(2.13)	(9.73)	(8.68)	(3.80)	(3.16)	(4.07)	(6.30)	(5.23)	(6.43)	(4.53)	(4.26)	(5.51)
3 years	3.66	0.20	5.38	5.42	-2.03	-5.56	-0.40	-2.17	3.58	-10.90	2.22	-5.75	0.60
	(2.27)	(2.11)	(9.49)	(8.43)	(3.76)	(3.11)	(4.01)	(6.23)	(5.15)	(6.34)	(4.46)	(4.21)	(5.43)
4 years	4.68	-1.42	2.45	0.01	-3.44	-2.32	1.81	6.41	7.68	-6.69	0.38	-1.63	-2.92
	(2.30)	(2.14)	(9.09)	(8.16)	(3.80)	(3.16)	(4.05)	(6.18)	(5.13)	(6.33)	(4.53)	(4.26)	(5.50)
Urban dummy	-2.10	4.22	4.38	31.42	-1.13	-0.81	0.26	0.72	-0.60	0.21	-0.99	-0.47	3.09
	(0.83)	(0.81)	(4.60)	(3.82)	(1.62)	(1.44)	(1.74)	(2.45)	(2.11)	(2.46)	(1.82)	(1.60)	(1.95)
Regions													
North-center	-0.04	2.79	-10.15	-17.30	-2.51	8.72	-4.36	1.92	-10.34	0.91	-0.60	7.35	3.54
	(0.98)	(0.94)	(9.61)	(9.09)	(1.46)	(1.31)	(1.65)	(4.59)	(3.89)	(4.19)	(2.02)	(1.74)	(2.22)
Center	-0.53	-0.44	7.26	-8.19	-2.84	6.81	-0.74	18.11	7.43	8.48	2.45	13.16	6.41
	(0.94)	(0.91)	(8.10)	(7.47)	(1.35)	(1.21)	(1.46)	(4.51)	(3.82)	(4.13)	(1.93)	(1.66)	(2.06)
South	3.19	0.92	-2.51	-19.53	-1.15	-0.83	0.28	0.73	-0.63	0.21	-0.98	-0.47	3.10
	(1.18)	(1.10)	(8.29)	(7.54)	(1.75)	(1.64)	(1.87)	(4.96)	(4.23)	(4.69)	(2.08)	(1.87)	(2.27)
Unobserved types													
Type I	0.77	-0.75	2.64	-1.59	4.70	2.55	7.21	0.26	2.05	5.32	0.53	2.27	8.61

Table B.1: Mathematics value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
	(4.58)	(16.32)	(3.75)	(12.48)	(6.10)	(11.66)	(7.22)	(5.86)	(9.75)	(6.16)	(6.18)	(11.70)	(6.88)
Type II	0.83	-0.52	2.05	-0.39	1.38	1.12	3.35	1.33	1.13	1.85	0.55	0.96	0.80
	(5.04)	(22.27)	(3.89)	(15.77)	(6.59)	(14.05)	(7.88)	(6.24)	(11.42)	(6.47)	(6.30)	(13.32)	(6.97)
Type III	-0.98	-0.46	1.28	8.53	-1.26	-6.11	-0.19	-1.60	-3.60	-0.96	-0.91	-1.93	-1.15
	(5.09)	(20.14)	(3.95)	(15.02)	(6.80)	(13.86)	(7.85)	(6.13)	(11.23)	(7.04)	(6.46)	(12.97)	(8.04)
Intercept term	152	189	231	241	83	185	146	185	249	220	95	183	163
	(2.56)	(2.52)	(13.75)	(12.71)	(4.40)	(3.76)	(4.68)	(8.94)	(6.96)	(8.46)	(5.66)	(4.72)	(6.11)
Standard error (σ)	81	82	85	83	88	84	96	114	106	115	91	84	100
	(0.23)	(0.24)	(1.20)	(1.30)	(0.39)	(0.32)	(0.39)	(0.78)	(0.68)	(0.79)	(0.55)	(0.43)	(0.59)

Table B.2: Spanish value-added estimates

	General		Indigenous		General			Telesecondary			Technical		
	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Lag Mathematics	0.23	0.20	0.20	0.17	0.13	0.23	0.20	0.09	0.16	0.16	0.12	0.23	0.18
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lag Spanish	0.43	0.40	0.26	0.33	0.55	0.48	0.49	0.44	0.40	0.39	0.53	0.46	0.48
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prospera score quartile													
P*Q1	0.02	-2.04	0.09	-3.72	0.06	-2.53	-2.82	-0.37	-0.73	-2.79	-2.24	0.54	-1.97
	(2.76)	(2.41)	(19.22)	(13.75)	(3.99)	(3.40)	(3.75)	(9.50)	(7.81)	(9.38)	(5.33)	(4.92)	(5.99)
P*Q2	-1.42	-1.81	-2.86	0.13	1.34	-2.35	-1.01	1.36	0.50	-1.08	-0.18	1.73	-0.59
	(1.63)	(1.43)	(9.24)	(7.29)	(2.60)	(2.41)	(2.45)	(4.25)	(3.45)	(4.13)	(3.27)	(2.90)	(3.18)
P*Q3	-2.76	-1.69	-2.32	-0.98	3.79	-1.61	-1.30	2.38	0.67	1.49	2.12	2.55	-1.36
	(1.16)	(0.99)	(4.82)	(4.24)	(2.06)	(1.88)	(1.96)	(2.50)	(2.11)	(2.43)	(2.39)	(2.14)	(2.32)
P*Q4	-1.09	-1.49	0.93	-2.21	5.84	3.99	-1.58	5.41	1.86	1.04	5.77	3.21	0.90
	(1.09)	(0.92)	(3.28)	(2.73)	(2.13)	(2.03)	(2.14)	(1.98)	(1.68)	(1.92)	(2.30)	(2.14)	(2.26)
Education cat. (dad)													
Below primary school	-3.34	0.00	-9.55	5.96	-10.09	0.28	-1.51	-9.37	-5.70	-1.55	-5.32	1.07	-5.08
	(2.37)	(2.10)	(10.16)	(9.37)	(3.65)	(3.29)	(3.58)	(6.39)	(5.28)	(6.41)	(4.72)	(4.40)	(4.73)
Primary school completed	-2.54	1.98	-3.00	7.22	-9.71	0.42	0.45	-6.42	-7.00	-0.41	-4.70	0.17	-2.09
	(2.38)	(2.11)	(10.35)	(9.48)	(3.66)	(3.30)	(3.58)	(6.47)	(5.37)	(6.48)	(4.76)	(4.40)	(4.71)
Secondary or below	-0.84	3.00	-4.84	11.10	-7.14	1.50	2.69	-6.44	-3.96	3.05	-3.43	4.93	-1.34
	(2.34)	(2.07)	(10.59)	(9.63)	(3.55)	(3.20)	(3.47)	(6.48)	(5.37)	(6.47)	(4.63)	(4.30)	(4.58)
College or above	6.73	7.69	-0.69	14.61	2.64	7.29	4.76	0.65	3.05	8.19	5.54	9.46	2.97
	(2.45)	(2.16)	(11.71)	(10.87)	(3.67)	(3.30)	(3.58)	(7.25)	(5.91)	(7.07)	(4.80)	(4.44)	(4.72)
Working status (dad)													
Full time	1.34	-2.18	-5.55	-6.33	3.11	0.15	-1.06	5.13	-0.04	-1.41	3.53	-1.82	0.89
	(1.93)	(1.72)	(8.73)	(7.97)	(3.06)	(2.76)	(3.06)	(5.16)	(4.23)	(5.10)	(3.93)	(3.69)	(3.88)
Not full time	1.19	-2.25	-4.25	-7.94	4.86	-0.26	-2.24	2.18	-0.37	-1.21	1.70	0.58	1.11
	(1.84)	(1.64)	(8.50)	(7.84)	(2.91)	(2.60)	(2.92)	(4.98)	(4.10)	(4.94)	(3.77)	(3.53)	(3.68)
Father present at home	0.01	1.10	4.56	6.19	0.71	1.32	0.01	8.11	-2.33	-1.12	2.21	-0.04	4.81

Table B.2: Spanish value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
	(0.87)	(0.75)	(3.72)	(3.16)	(1.35)	(1.22)	(1.33)	(2.19)	(1.88)	(2.16)	(1.73)	(1.58)	(1.73)
Education cat. (mom)													
Primary school	-4.63	-3.26	5.09	-0.85	3.83	-12.33	1.72	10.68	6.11	8.29	2.82	-5.31	1.55
	(3.12)	(2.71)	(11.40)	(9.56)	(5.02)	(4.41)	(4.71)	(8.14)	(6.34)	(8.06)	(5.83)	(5.24)	(5.92)
Primary school completed	-2.94	-2.44	10.75	7.70	3.45	-9.11	1.51	12.20	8.76	8.90	2.59	-6.83	3.65
	(3.12)	(2.72)	(11.59)	(9.76)	(5.00)	(4.39)	(4.70)	(8.22)	(6.40)	(8.14)	(5.82)	(5.25)	(5.91)
Secondary or below	-1.54	0.15	14.80	12.41	5.45	-8.59	1.27	12.92	6.50	12.81	4.58	-4.80	5.97
	(3.11)	(2.71)	(11.97)	(10.05)	(4.96)	(4.35)	(4.65)	(8.26)	(6.45)	(8.15)	(5.78)	(5.20)	(5.85)
College or above	4.72	6.21	11.35	26.91	11.97	-0.70	5.94	21.99	18.92	18.61	9.78	0.65	7.45
	(3.21)	(2.80)	(14.02)	(12.20)	(5.07)	(4.45)	(4.74)	(9.14)	(7.15)	(8.91)	(5.99)	(5.38)	(5.98)
Working status (mom)													
Housework	-3.76	-1.00	9.84	-2.17	1.48	3.76	-2.00	-2.92	-6.45	0.38	-5.35	4.28	-5.33
	(2.79)	(2.42)	(10.64)	(8.84)	(4.49)	(4.07)	(4.22)	(7.10)	(5.65)	(6.76)	(5.46)	(4.95)	(5.51)
Part time	-5.02	-0.77	11.38	-7.29	0.36	4.50	-2.92	-0.75	-10.56	0.26	-9.60	2.56	-4.08
	(2.83)	(2.46)	(10.90)	(9.21)	(4.52)	(4.10)	(4.24)	(7.38)	(5.85)	(7.03)	(5.53)	(4.99)	(5.56)
Full time	-3.76	-1.79	1.50	-10.10	-1.10	3.01	-2.27	-4.96	-11.18	-0.29	-7.84	1.68	-3.39
	(2.89)	(2.51)	(11.40)	(9.51)	(4.62)	(4.18)	(4.32)	(7.56)	(6.01)	(7.16)	(5.62)	(5.10)	(5.68)
Mother present at home	9.48	8.54	3.95	0.02	12.61	5.50	11.46	7.79	12.50	6.56	12.48	9.29	7.10
	(1.12)	(0.95)	(4.21)	(3.50)	(1.80)	(1.60)	(1.78)	(2.65)	(2.26)	(2.60)	(2.31)	(2.11)	(2.26)
Number of people at home													
5 people	-0.57	-0.27	-1.11	7.48	-2.21	-3.32	0.27	1.18	1.27	4.25	-1.42	1.97	-1.27
	(0.87)	(0.76)	(4.76)	(3.94)	(1.30)	(1.16)	(1.25)	(2.40)	(2.00)	(2.32)	(1.70)	(1.53)	(1.65)
6 people	-1.58	-1.59	-0.16	4.31	-0.11	-4.06	-1.19	-0.09	-0.58	1.05	-3.02	-0.10	-0.14
	(0.97)	(0.82)	(4.38)	(3.72)	(1.47)	(1.30)	(1.42)	(2.54)	(2.09)	(2.41)	(1.93)	(1.70)	(1.86)
$\geq$ 7 people	-3.36	-3.76	-1.39	6.33	-6.71	-6.51	-2.73	-0.84	-0.50	1.58	-6.74	-3.30	-3.85
	(0.83)	(0.71)	(3.70)	(3.07)	(1.27)	(1.14)	(1.23)	(2.11)	(1.78)	(2.06)	(1.65)	(1.48)	(1.62)
Age	-2.96	-3.27	-2.41	-2.55	-3.34	-2.42	-6.22	-5.05	-3.52	-5.19	-3.42	-0.27	-4.96
	(0.69)	(0.62)	(3.11)	(2.58)	(1.07)	(0.94)	(1.00)	(1.87)	(1.58)	(1.84)	(1.43)	(1.25)	(1.36)
Age^2	-2.38	-2.27	-0.92	-1.52	-3.14	-3.42	-3.45	-0.24	-1.76	-1.87	-2.14	-4.43	-5.01
	(0.45)	(0.35)	(1.68)	(1.25)	(0.70)	(0.68)	(0.80)	(1.05)	(0.92)	(1.27)	(0.90)	(0.83)	(1.02)

Table B.2: Spanish value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Gender	-16.30	-15.46	-7.46	-8.48	-25.80	-17.00	-12.23	-25.14	-19.80	-15.93	-26.24	-15.71	-13.72
	(0.62)	(0.55)	(2.79)	(2.32)	(0.98)	(0.88)	(0.94)	(1.61)	(1.36)	(1.57)	(1.25)	(1.13)	(1.23)
First language spoken at home													
Indigenous	-8.67	-8.43	-8.55	-12.96	4.38	-12.70	-11.25	0.57	-6.79	2.98	7.92	5.30	-20.21
	(2.39)	(2.03)	(3.41)	(2.79)	(4.49)	(4.13)	(4.96)	(2.92)	(2.65)	(3.04)	(3.68)	(3.16)	(3.46)
Both Spanish and indigenous	2.24	-4.64	7.51	-0.44	-2.32	-3.73	-6.84	2.99	3.50	0.22	5.25	3.91	0.61
	(2.38)	(2.08)	(4.67)	(3.88)	(4.23)	(3.86)	(4.36)	(4.57)	(3.61)	(4.11)	(4.56)	(3.85)	(3.87)
Internet access	-3.55	-4.57	-1.91	-12.01	-4.87	-3.91	-3.67	-17.61	-8.26	-14.38	-6.50	-7.63	-1.29
	(0.93)	(0.81)	(4.95)	(4.07)	(1.39)	(1.23)	(1.32)	(2.78)	(2.34)	(2.74)	(1.83)	(1.61)	(1.78)
Computer access	3.19	4.12	-5.22	-0.73	4.86	4.36	3.74	-6.70	0.09	-1.14	3.00	3.48	-1.01
	(0.84)	(0.73)	(4.45)	(3.64)	(1.27)	(1.13)	(1.22)	(2.41)	(2.00)	(2.34)	(1.64)	(1.45)	(1.59)
Number of pre-school years													
1 year	4.16	-3.54	-1.18	-0.90	-6.56	-3.36	-3.84	-5.64	9.04	-3.46	-1.58	-3.96	-7.97
	(2.39)	(2.07)	(9.21)	(7.69)	(4.13)	(3.59)	(3.84)	(5.48)	(4.61)	(5.50)	(5.01)	(4.49)	(4.91)
2 years	3.03	-0.91	1.55	4.24	-5.70	-1.69	2.70	-5.40	8.95	-4.21	0.61	-3.78	-0.92
	(2.16)	(1.87)	(8.95)	(7.47)	(3.64)	(3.11)	(3.34)	(5.06)	(4.23)	(5.05)	(4.50)	(3.98)	(4.32)
3 years	4.05	1.01	1.91	5.35	-4.24	-0.70	2.20	-0.76	10.14	-2.69	4.34	-2.54	0.41
	(2.13)	(1.85)	(8.84)	(7.30)	(3.58)	(3.07)	(3.28)	(4.99)	(4.16)	(4.97)	(4.44)	(3.92)	(4.27)
4 years	5.38	0.34	-1.69	3.76	-3.33	3.11	3.26	6.33	8.84	-0.30	2.06	-2.15	-0.30
	(2.17)	(1.87)	(8.54)	(7.10)	(3.63)	(3.11)	(3.34)	(4.95)	(4.15)	(4.97)	(4.51)	(3.98)	(4.33)
Urban dummy	1.40	6.11	4.47	19.22	0.46	0.06	0.00	-1.32	-1.03	-1.62	0.15	-0.86	0.09
	(0.79)	(0.70)	(3.80)	(3.15)	(1.56)	(1.43)	(1.48)	(1.96)	(1.66)	(1.91)	(1.72)	(1.57)	(1.67)
Regions													
North-center	-0.89	-2.69	-13.20	-5.50	4.71	5.37	2.21	3.66	-1.71	3.08	-0.55	4.69	8.74
	(0.91)	(0.81)	(8.64)	(7.32)	(1.36)	(1.22)	(1.34)	(3.62)	(3.00)	(3.42)	(1.78)	(1.60)	(1.80)
Center	-0.45	-1.09	7.97	-0.66	6.12	7.19	4.87	18.87	10.09	13.35	4.74	5.34	9.64
	(0.88)	(0.78)	(7.10)	(6.15)	(1.25)	(1.10)	(1.21)	(3.56)	(2.96)	(3.36)	(1.71)	(1.54)	(1.69)
South	-1.25	-4.34	-15.96	-22.91	0.46	0.05	0.00	-1.31	-1.06	-1.64	0.14	-0.86	0.10
	(1.13)	(0.96)	(7.30)	(6.25)	(1.73)	(1.57)	(1.62)	(3.96)	(3.34)	(3.75)	(1.96)	(1.83)	(1.94)
Unobserved types													

Table B.2: Spanish value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Type I	0.34	-1.21	-0.86	-2.69	5.70	2.34	3.99	-0.76	2.23	0.93	0.76	0.85	2.18
	(4.53)	(14.53)	(3.18)	(11.28)	(5.97)	(9.45)	(7.43)	(5.61)	(7.95)	(6.52)	(5.29)	(9.04)	(6.24)
Type II	0.41	-1.28	0.13	-1.53	0.78	1.44	2.44	-0.09	1.84	-1.17	-0.62	-0.52	-0.90
	(4.99)	(19.08)	(3.28)	(13.80)	(6.46)	(10.99)	(8.14)	(5.98)	(9.41)	(6.95)	(5.46)	(10.05)	(6.62)
Type III	0.14	0.59	-1.25	8.12	0.31	-1.67	0.01	-0.85	-0.41	-0.75	0.44	0.44	0.99
	(5.18)	(18.17)	(3.38)	(13.24)	(6.74)	(11.03)	(8.16)	(5.79)	(8.87)	(7.00)	(5.76)	(9.74)	(7.35)
Intercept term	186	229	247	264	110	144	144	188	207	202	123	145	154
	(2.34)	(2.13)	(11.79)	(10.00)	(4.17)	(3.50)	(3.68)	(7.10)	(5.47)	(6.43)	(5.20)	(4.51)	(4.91)
Standard error (σ)	76	70	76	70	83	79	80	93	86	90	84	80	82
	(0.22)	(0.21)	(1.00)	(1.15)	(0.42)	(0.33)	(0.33)	(0.60)	(0.51)	(0.55)	(0.54)	(0.42)	(0.46)

Table B.3: Primary and secondary school choices

	Primary school	Secondary school		
	Indigenous	General	Telesecondary	Technical
Lag mathematics		0.0015	0.0005	0.0019
		(0.0001)	(0.0001)	(0.0001)
Lag Spanish		0.0036	0.0019	0.0031
		(0.0001)	(0.0001)	(0.0001)
Prospera score quartile				
P*Q1	0.53	0.26	0.44	0.20
	(0.21)	(0.12)	(0.14)	(0.13)
P*Q2	0.08	0.36	0.48	0.34
	(0.10)	(0.07)	(0.07)	(0.07)
P*Q3	0.04	0.30	0.60	0.43
	(0.06)	(0.05)	(0.05)	(0.05)
P*Q4	0.16	0.18	0.66	0.32
	(0.04)	(0.04)	(0.04)	(0.04)
Distance to general school (primary)	-0.03			
	(0.00)			
Distance to indigenous school	0.08			
	(0.00)			
Distance to general school (Secondary)		-0.44	0.09	0.15
		(0.01)	(0.01)	(0.01)
Distance to telesecondary school		0.04	-0.74	0.03
		(0.01)	(0.01)	(0.01)
Distance to technical school		0.12	0.08	-0.51
		(0.01)	(0.00)	(0.01)
Number of general schools		0.0023	-0.0017	-0.0002
		(0.0003)	(0.0005)	(0.0002)
Number of telesecondary schools		-0.0018	0.0123	-0.0114
		(0.0011)	(0.0012)	(0.0012)
Number of technical schools		-0.0033	-0.0188	0.0033
		(0.0019)	(0.0027)	(0.0020)
Imputed wages		-0.01	-0.04	-0.02
		(0.00)	(0.00)	(0.00)
Education cat. (dad)				
Below primary school	0.08	-0.13	0.10	-0.16
	(0.13)	(0.09)	(0.10)	(0.09)
Primary school completed	0.02	0.02	0.18	0.07
	(0.13)	(0.09)	(0.10)	(0.10)
Secondary or below	-0.24	0.35	0.32	0.37
	(0.14)	(0.09)	(0.10)	(0.09)
College or above	-0.33	0.44	0.11	0.49
	(0.15)	(0.10)	(0.11)	(0.10)

Table B.3: Primary and secondary school choices

	Indigenous	General	Telesecondary	Technical
Working status (dad)				
Full time	0.21	-0.10	-0.19	-0.12
	(0.11)	(0.07)	(0.08)	(0.08)
Not full time	0.03	0.14	-0.03	0.16
	(0.11)	(0.07)	(0.08)	(0.07)
Father present at home	-0.01	0.08	0.12	0.08
	(0.05)	(0.03)	(0.04)	(0.03)
Education cat. (mom)				
Primary school	0.09	-0.24	0.07	-0.37
	(0.16)	(0.12)	(0.12)	(0.12)
Primary school completed	0.04	0.04	0.24	-0.11
	(0.16)	(0.12)	(0.12)	(0.12)
Secondary or below	-0.23	0.31	0.34	0.16
	(0.16)	(0.12)	(0.12)	(0.12)
College or above	-0.22	0.53	0.19	0.26
	(0.18)	(0.13)	(0.14)	(0.13)
Working status (mom)				
Housework	-0.17	-0.19	-0.30	0.02
	(0.14)	(0.10)	(0.11)	(0.11)
Part time	0.00	-0.18	-0.27	-0.01
	(0.14)	(0.11)	(0.11)	(0.11)
Full time	0.05	-0.20	-0.24	-0.07
	(0.15)	(0.11)	(0.12)	(0.11)
Mother present at home	-0.34	0.14	0.04	0.20
	(0.05)	(0.04)	(0.04)	(0.04)
Number of people at home				
5 people	-0.21	0.03	0.04	0.06
	(0.06)	(0.04)	(0.04)	(0.04)
6 people	0.00	-0.10	-0.02	-0.08
	(0.06)	(0.04)	(0.04)	(0.04)
$\geq$ 7 people	-0.06	-0.23	-0.07	-0.22
	(0.05)	(0.03)	(0.04)	(0.03)
Age	-0.14	-0.76	-0.64	-0.75
	(0.04)	(0.03)	(0.03)	(0.03)
Age^2	0.03	-0.08	-0.03	-0.11
	(0.02)	(0.02)	(0.02)	(0.02)
Gender	0.03	0.30	0.31	0.27
	(0.04)	(0.03)	(0.03)	(0.03)
First language spoken at home				
Indigenous	1.59	-0.72	-0.19	-0.65
	(0.06)	(0.07)	(0.05)	(0.07)

Table B.3: Primary and secondary school choices

	Indigenous	General	Telesecondary	Technical
Both Spanish and indigenous	1.00	-0.25	-0.19	-0.32
	(0.07)	(0.09)	(0.08)	(0.08)
Internet access	0.00	-0.05	-0.17	-0.12
	(0.06)	(0.04)	(0.04)	(0.04)
Computer access	-0.11	0.39	0.11	0.34
	(0.06)	(0.04)	(0.04)	(0.04)
Number of pre-school years				
1 year	-0.18	-0.36	-0.12	-0.42
	(0.12)	(0.08)	(0.09)	(0.09)
2 years	-0.10	0.08	0.19	0.03
	(0.11)	(0.08)	(0.08)	(0.08)
3 years	-0.08	0.37	0.41	0.40
	(0.11)	(0.08)	(0.08)	(0.08)
4 years	0.18	0.39	0.46	0.40
	(0.11)	(0.08)	(0.08)	(0.08)
Urban dummy	-1.15	0.07	-0.26	-0.09
	(0.04)	(0.04)	(0.03)	(0.04)
Regions				
North-center	-0.47	-0.37	-0.57	-0.20
	(0.12)	(0.05)	(0.06)	(0.05)
Center	0.08	-0.25	-0.29	-0.18
	(0.09)	(0.05)	(0.07)	(0.06)
South	0.02	-0.27	-0.86	-0.19
	(0.10)	(0.07)	(0.08)	(0.07)
Type I	0.61	-0.27	0.00	-0.22
	(0.17)	(0.15)	(0.16)	(0.16)
Type II	0.21	-0.06	0.06	-0.06
	(0.20)	(0.17)	(0.18)	(0.17)
Type III	-0.27	0.05	-0.10	0.01
	(0.22)	(0.16)	(0.18)	(0.17)
Intercept	-0.94	-0.58	0.72	-0.53
	(0.13)	(0.11)	(0.12)	(0.11)

Table B.4: Retention estimates (choice and value-added)

	Choice	Mathematics	Spanish	Choice	mathematics	Spanish
Lag mathematics	-0.0041	0.40	0.17	-0.0012	0.38	0.18
	(0.0001)	(0.03)	(0.02)	(0.0001)	(0.05)	(0.05)
Lag Spanish	0.0003	0.12	0.26	-0.0100	0.14	0.35
	(0.0001)	(0.03)	(0.03)	(0.0002)	(0.05)	(0.05)
Prospera score quartile						
P*Q1	0.005	-1.12	0.52	-0.473	0.55	0.56
	(0.21)	(20.72)	(17.34)	(0.34)	(43.51)	(49.61)
P*Q2	0.16	-0.68	0.42	-0.20	1.52	0.98
	(0.10)	(9.37)	(8.78)	(0.20)	(20.38)	(26.90)
P*Q3	-0.002	1.35	1.90	-0.59	2.80	1.34
	(0.07)	(6.43)	(5.43)	(0.15)	(15.61)	(15.68)
P*Q4	-0.07	1.98	2.62	-0.56	3.52	2.61
	(0.06)	(5.42)	(4.82)	(0.15)	(17.33)	(16.69)
Education cat. (dad)						
Below primary school	-0.21	5.44	3.74	-0.37	12.21	1.95
	(0.14)	(13.89)	(11.05)	(0.26)	(30.60)	(30.74)
Primary school completed	-0.06	6.70	4.58	-0.36	-1.70	-16.23
	(0.14)	(14.07)	(11.39)	(0.26)	(31.14)	(31.54)
Secondary or below	0.00	7.09	7.70	-0.09	9.38	-12.79
	(0.14)	(14.02)	(11.19)	(0.25)	(29.83)	(30.08)
College or above	-0.08	10.04	8.22	-0.11	4.54	3.25
	(0.15)	(15.10)	(12.24)	(0.26)	(32.25)	(31.47)
Working status (dad)						
Full time	-0.13	1.58	-5.34	-0.27	16.64	17.20
	(0.11)	(11.37)	(9.70)	(0.20)	(23.70)	(22.55)
Not full time	0.01	1.46	-6.11	-0.24	23.69	25.19
	(0.11)	(10.95)	(9.24)	(0.19)	(22.09)	(21.21)
Father present at home	-0.11	-1.94	1.04	-0.01	-4.39	-19.69
	(0.05)	(4.52)	(3.88)	(0.09)	(10.08)	(9.92)
Education cat. (mom)						
Primary school	-0.03	0.28	0.13	-0.25	-2.86	7.10
	(0.19)	(16.94)	(15.53)	(0.33)	(34.62)	(35.68)
Primary school completed	0.15	0.62	4.52	-0.07	-5.41	13.73
	(0.19)	(17.03)	(15.80)	(0.33)	(34.73)	(35.52)
Secondary or below	0.24	6.31	2.38	0.25	2.02	11.03
	(0.19)	(17.19)	(15.80)	(0.32)	(33.59)	(34.88)
College or above	0.44	15.17	13.28	0.42	8.82	24.22
	(0.20)	(18.15)	(16.54)	(0.33)	(35.01)	(35.93)
Working status (mom)						
Housework	-0.09	-15.58	-1.12	0.30	-0.88	-3.79
	(0.17)	(14.89)	(14.51)	(0.35)	(40.94)	(41.22)

Table B.4: Retention estimates (choice and value-added)

	Choice	Mathematics	Spanish	Choice	mathematics	Spanish
Part time	-0.08	-14.78	4.70	0.40	-5.14	-6.73
	(0.17)	(15.04)	(14.83)	(0.35)	(41.01)	(41.07)
Full time	-0.17	-16.09	2.23	0.35	-2.44	-5.41
	(0.17)	(15.39)	(15.12)	(0.35)	(42.38)	(42.29)
Mother present at home	-0.07	9.20	2.66	-0.01	16.07	20.32
	(0.06)	(5.28)	(4.65)	(0.11)	(12.57)	(11.81)
Number of people at home						
5 people	-0.04	5.61	1.71	0.01	3.64	-2.42
	(0.06)	(5.64)	(4.77)	(0.10)	(11.53)	(10.75)
6 people	-0.15	8.82	4.76	-0.13	-13.31	-10.77
	(0.06)	(5.85)	(4.99)	(0.12)	(12.93)	(12.74)
$\geq$ 7 people	-0.10	12.17	4.26	-0.11	-5.20	2.93
	(0.05)	(4.80)	(4.11)	(0.09)	(10.67)	(10.25)
Age	3.44	-2.49	-0.75	3.95	-0.36	8.54
	(0.07)	(8.40)	(7.05)	(0.12)	(18.28)	(16.44)
Age^2	-0.68	0.44	-0.91	-1.03	-0.97	-4.24
	(0.02)	(2.42)	(2.05)	(0.05)	(6.27)	(5.49)
Gender	0.48	6.12	-12.45	0.67	8.49	-8.05
	(0.04)	(4.01)	(3.33)	(0.09)	(9.84)	(9.39)
First language spoken at home						
Indigenous	-0.02	-31.39	-17.64	-0.81	43.86	20.12
	(0.08)	(7.66)	(6.86)	(0.26)	(23.89)	(30.51)
Both Spanish and indigenous	-0.20	-20.36	-16.36	-0.44	-30.70	-14.48
	(0.11)	(11.30)	(8.66)	(0.31)	(36.41)	(39.00)
Internet access	0.29	-10.50	-8.39	0.03	-3.21	-7.87
	(0.05)	(5.21)	(4.50)	(0.10)	(11.96)	(10.44)
Computer access	0.01	0.31	6.42	0.01	-0.39	2.07
	(0.05)	(4.99)	(4.40)	(0.10)	(11.32)	(10.15)
Number of pre-school years						
1 year	-0.24	-4.29	-6.54	-0.33	-23.98	-47.31
	(0.13)	(13.73)	(10.85)	(0.29)	(38.15)	(34.43)
2 years	0.26	-4.53	-11.67	0.05	-18.11	-25.12
	(0.12)	(12.70)	(9.77)	(0.26)	(35.73)	(30.16)
3 years	0.37	-0.76	-6.09	0.27	-25.90	-24.42
	(0.12)	(12.53)	(9.71)	(0.25)	(35.34)	(29.72)
4 years	0.13	2.35	-5.77	-0.04	-36.48	-34.37
	(0.12)	(12.54)	(9.68)	(0.26)	(35.75)	(30.07)
Urban dummy	-0.05	9.61	8.99	0.46	4.91	9.08
	(0.05)	(4.54)	(3.91)	(0.12)	(14.83)	(14.86)
Regions						
North-center	-0.18	-6.52	2.98	-0.20	2.14	-2.20

Table B.4: Retention estimates (choice and value-added)

	Choice	Mathematics	Spanish	Choice	mathematics	Spanish
	(0.07)	(6.74)	(5.81)	(0.12)	(13.03)	(12.83)
Center	0.50	3.25	7.73	0.37	8.56	14.38
	(0.06)	(6.14)	(5.30)	(0.10)	(11.39)	(10.55)
South	0.23	-3.10	4.11	-0.06	12.37	14.10
	(0.07)	(6.92)	(6.11)	(0.13)	(14.22)	(13.71)
Indigenous school	-0.01	-8.50	-14.65			
	(0.09)	(7.84)	(7.34)			
Grade 5	-1.31	-5.88	-52.05			
	(0.05)	(4.11)	(3.90)			
Grade 6	-3.47	-44.17	-7.72			
	(0.12)	(11.06)	(9.83)			
Telesecondary school				-0.11	41.60	26.37
				(0.13)	(14.25)	(13.79)
Technical school				-0.02	-2.46	-4.98
				(0.09)	(9.99)	(9.62)
Grade 8				1.00	-50.89	16.87
				(0.09)	(10.27)	(9.07)
Unobserved types						
Type I	0.13	-0.44	0.50	-0.07	2.01	-1.90
	(0.24)	(20.30)	(17.87)	(0.61)	(81.85)	(68.49)
Type II	0.10	-0.62	-0.83	-0.06	-2.92	-1.98
	(0.28)	(22.52)	(21.75)	(0.70)	(94.20)	(78.59)
Type III	0.00	1.49	3.26	0.15	-1.66	1.10
	(0.27)	(26.73)	(22.64)	(0.54)	(71.98)	(60.83)
Intercept term	-3.78	231.86	295.40	-3.05	257.80	172.89
	(0.12)	(16.46)	(13.34)	(0.25)	(36.23)	(35.31)
Standard error (σ)		90.60	79.60		100.41	92.98
		(1.43)	(1.25)		(3.40)	(3.02)

Table B.5: Dropout during secondary school

Dropout period	Grade 7	Grade 8
Lag mathematics	-0.0013	-0.0008
	(0.0001)	(0.0001)
Lag Spanish	-0.0030	-0.0034
	(0.0001)	(0.0001)
Prospera score quartile		
P*Q1	0.09	0.20
	(0.12)	(0.09)
P*Q2	0.15	-0.03
	(0.07)	(0.06)
P*Q3	0.03	-0.07
	(0.05)	(0.04)
P*Q4	-0.11	-0.16
	(0.04)	(0.04)
Education cat. (dad)		
Below primary school	0.27	0.03
	(0.10)	(0.08)
Primary school completed	0.22	-0.07
	(0.10)	(0.08)
Secondary or below	0.06	-0.03
	(0.10)	(0.08)
College or above	-0.14	-0.18
	(0.11)	(0.09)
Working status (dad)		
Full time	-0.15	-0.07
	(0.08)	(0.07)
Not full time	-0.24	-0.11
	(0.07)	(0.06)
Father present at home	-0.12	-0.15
	(0.03)	(0.03)
Education cat. (mom)		
Primary school	0.01	0.05
	(0.12)	(0.11)
Primary school completed	-0.05	0.08
	(0.12)	(0.11)
Secondary or below	-0.19	0.08
	(0.12)	(0.11)
College or above	-0.42	-0.08
	(0.13)	(0.12)
Working status (mom)		
Housework	0.11	-0.19
	(0.11)	(0.10)

Table B.5: Dropout during secondary school

Dropout period	Grade 7	Grade 8
Part time	0.31	0.05
	(0.11)	(0.10)
Full time	0.18	-0.03
	(0.12)	(0.10)
Mother present at home	-0.14	-0.12
	(0.04)	(0.04)
Number of people at home		
5 people	-0.002	-0.05
	(0.04)	(0.03)
6 people	0.001	-0.02
	(0.04)	(0.04)
$\geq$ 7 people	0.14	-0.01
	(0.04)	(0.03)
Age	0.69	1.29
	(0.04)	(0.03)
Age^2	0.04	-0.07
	(0.02)	(0.02)
Gender	0.00	0.15
	(0.03)	(0.02)
First language spoken at home		
Indigenous	-0.08	0.08
	(0.07)	(0.06)
Both Spanish and indigenous	-0.20	-0.05
	(0.09)	(0.08)
Internet access	0.08	0.14
	(0.04)	(0.03)
Computer access	-0.22	-0.14
	(0.04)	(0.03)
Number of pre-school years		
1 year	-0.02	0.0002
	(0.09)	(0.08)
2 years	-0.11	0.08
	(0.08)	(0.07)
3 years	-0.29	-0.01
	(0.08)	(0.07)
4 years	-0.17	-0.06
	(0.08)	(0.07)
Urban dummy	0.23	0.24
	(0.04)	(0.03)
Regions		
North-center	0.32	0.01

Table B.5: Dropout during secondary school

Dropout period	Grade 7	Grade 8
	(0.05)	(0.04)
Center	0.20	0.21
	(0.05)	(0.04)
South	1.02	0.39
	(0.06)	(0.06)
Distance to the current secondary school	0.03	0.03
	(0.01)	(0.01)
Telesecondary school dummy	-0.07	0.10
	(0.04)	(0.04)
Technical school dummy	0.04	0.07
	(0.03)	(0.03)
Imputed wages	0.04	0.01
	(0.00)	(0.00)
Unobserved types	0.03	0.05
Type I	(0.15)	(0.12)
	-0.03	0.04
Type II	(0.15)	(0.13)
	-0.06	0.06
Type III	(0.15)	(0.12)
	-1.60	-0.66
Intercept term	(0.12)	(0.10)
	(0.12)	(13.34)

Table B.6: Coefficients associated with cheating effect

	Math		Spanish	
	Value	S.D.	Value	S.D.
Grades are not retained				
General				
G5	40	1.4	25	1.4
G6	76	6.6	53	5.1
Indigenous				
G5	45	1.5	27	1.2
G6	62	7.2	47	4.9
General				
G7	58	3.5	37	3.3
G8	95	5.4	59	4.4
G9	88	4.4	50	4.3
Telesecondary				
G7	79	2.0	42	1.9
G8	100	3.1	62	2.4
G9	79	2.5	41	2.5
Technical				
G7	75	3.4	30	3.0
G8	45	3.8	20	3.3
G9	57	3.3	26	3.0
Grades are retained				
Primary	75	14.3	58	12.5
Secondary	137	34.8	104	26.0

Table B.7: The measurement-model estimates

	Mathematics		Spanish	
	Value	S.D.	Value	S.D.
For non-retained students				
Grade 5				
General	40	(1.4)	25	(1.4)
Indigenous	76	(6.6)	53	(5.1)
Grade 6				
General	45	(1.5)	27	(1.2)
Indigenous	62	(7.2)	47	(4.9)
Grade 7				
General	58	(3.5)	37	(3.3)
Telesecondary	95	(5.4)	59	(4.4)
Technical	88	(4.4)	50	(4.3)
Grade 8				
General	79	(2.0)	42	(1.9)
Telesecondary	100	(3.1)	62	(2.4)
Technical	79	(2.5)	41	(2.5)
Grade 9				
General	75	(3.4)	30	(3.0)
Telesecondary	45	(3.8)	20	(3.3)
Technical	57	(3.3)	26	(3.0)
For retained students				
Primary school	75	(14.3)	58	(12.5)
Secondary school	137	(34.8)	104	(26.0)

Table B.8: Type distribution

Type	Fraction	Standard errors
I	0.332	0.025
II	0.129	0.024
III	0.201	0.025
IV	0.328	-

Note: Notice that these four types only have three degrees of freedom. The fraction of Type IV can be uniquely pinned down by the fraction of other three types 1-0.332-0.129-0.201 = 0.328. Therefore, we have standard errors only for the first three types.