Prospering through *Prospera*: A Dynamic Model of CCT Impacts on Educational Attainment and Achievement in Mexico*

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

December 28, 2022

Abstract

This paper develops and estimates a dynamic model, which integrates value-added and school-choice models, to evaluate grade-by-grade and the cumulative impacts of the Mexican *Prospera* conditional cash transfer (CCT) program on educational achievement. The empirical application advances the previous literature by estimating policy impacts on learning, accounting for dynamic selective school attendance and incorporating both observed and unobserved heterogeneities. A dynamic framework is critical for estimating cumulative learning effects because lagged achievement is an important determinant of current achievement. The model is estimated using rich nationwide Mexican administrative data on schooling progression and math and Spanish test scores in grades 4-9 along with student and family survey data. The estimates show significant CCT learning impacts, inter alia, particularly for students from poorer households. School-choice decisions, especially whether to attend telesecondary schools, play crucial roles.

*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 Goktas, 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 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 helpful suggestions.

1 Introduction

Conditional cash transfer (CCT) programs aim to alleviate current poverty by transfers to poor families and reduce future poverty by making these transfers conditional on investments in the human capital of children and youth. In 1998-2000, a large-scale randomized evaluation of the Mexican *PROGRESA* CCT program demonstrated substantial impacts on schooling enrollment and attainment, child work and family income (Parker and Todd, 2017). These findings contributed to a large scaling-up in Mexico and an impressive adoption of similar programs in more than 60 countries on five continents (Fiszbein and Schady, 2009).

This paper contributes to the literature by developing and estimating a dynamic model of academic achievement and school progression and using the model to analyze the understudied, substantive question of whether CCTs improve learning. Several studies, using various methods including the original experiment, matching and structural dynamic models, examined the impacts of *PROGRESA/Oportunidades/Prospera* on school enrollment and, in some cases, on longer-term schooling attainment (e.g., Schultz (2004), Behrman et al. (2005b), Behrman et al. (2005a), Behrman et al. (2009), Todd and Wolpin (2006), Attanasio et al. (2012) and Parker and Vogl (2018)). This literature demonstrated positive program impacts on school enrollment and schooling attainment. However, a longstanding concern has been whether and to what extent increased school enrollment translates into higher academic achievement, presumably important to affect significantly earnings potential and other outcomes of interest. Most prior studies did not analyze academic achievement impacts because the original data did not include achievement test scores.[1]

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 schooling attainment but also on academic achievement in mathematics and Spanish. Nationwide standardized 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 dynamically affects their school enrollment, school choice, grade progression and academic achievements over time.

A major challenge in evaluating the effect of a CCT program on academic achievement is controlling for selection bias because nonrandom selection occurs at multiple stages. First, participation in the

[1]In 2003, Woodcock-Johnson tests in mathematics and Spanish were 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, because the original control group was enrolled in the program 1.5 years after the original treatment group, the schooling differences between them were relatively small, at about 0.2 years of additional schooling.

Prospera CCT program is selective. Students from high-poverty backgrounds and rural areas are much more likely enrolled in CCTs compared with other students. Second, school enrollments/dropouts are selective. In primary school (up through grade 6), enrollment rates are high and dropouts are rare. However, dropouts increase substantially after grade 6 when students choose whether to enroll in lower-secondary school.[2] Because the ENLACE standardized tests are administered in schools, test scores are only observed for enrolled children. The 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. These two selection mechanisms are interrelated; 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.[3]

Our goal in this paper is to examine how the *Prospera* program affects schooling and academic achievement, accounting for selective program participation and school enrollment. To this end, we develop and estimate a dynamic model of students' school progression that incorporates decision-making in each grade (4-9) with regard to enrollment, school choice, and dropping-out as well as grade- and subject-varying value-added functions for academic achievement. Our framework combines value-added academic achievement models with school-choice models and links equations across ages/grades, incorporating both observed and unobserved heterogeneity. Valued-added models typically specify a relationship among academic achievement, key learning inputs in the current period, and lagged achievement, which is a sufficient statistic for past learning inputs under some assumptions about coefficients of past learning inputs following geometric patterns (Summers and Wolfe, 1977; Boardman and Murnane, 1979; Hanushek, 1979; Todd and Wolpin, 2003; Cunha et al., 2006, 2010).[4] School-choice models generally focus on the decision of what type of school to attend (Neal, 1997; McEwan, 2001; Altonji et al., 2005; Sapelli and Vial, 2002; Gallego and Hernando, 2009).[5] Value-added models and school-choice models are usually estimated in isolation, although there are a few papers that combine

[2]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.

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

[4]There is some debate about whether value-added models with teacher fixed effects should be used to measure teacher effectiveness (see, e.g., Kane and Staiger (2008), Kane et al. (2013), Chetty et al. (2014a), Chetty et al. (2014b)). Our focus is rather on using these models to capture the cumulative student learning process.

[5]Some studies use school-choice models to estimate school-voucher effects (Rouse (1998), Figlio and Rouse (2006), Hsieh and Urquiola (2006), Bravo et al. (2010), Angrist et al. (2002), and Angrist et al. (2006)), to study parents' preferences for school quality and to analyze the welfare effects of school policies (Epple et al. (2018), Hastings et al. (2009) and Neilson et al. (2019)).

them (e.g., Hastings et al. (2012), Allende et al. (2019), and Schellenberg and Walters (2020).) Our model additionally considers dropout decisions, grade retention, learning dynamics, and cumulative program effects.

The model begins when students finish fourth grade and continues through the ninth grade, which is the end of lower-secondary schooling in Mexico.[6] Students differ in terms of their family backgrounds and their fourth-grade knowledge in math and Spanish as measured by standardized test scores. 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 each grade, students can progress to the next grade, repeat the same grade or dropout. Conditional on progressing to lower-secondary school (at the end of grade 6), students/parents make a one-time choice of a lower-secondary school type, from up to three public school options–general, telesecondary or technical schools–all of which are academically oriented.[7]

In each period, we model skill accumulation in math and Spanish using value-added production functions, with coefficients varying by grades, school types, and grade-retention status. The dynamic-panel specification captures the notion that skill accumulation at one stage affects skill attainment at other stages, which has been shown to be essential to characterizing human-capital-skill formation processes (e.g. Cunha et al. (2006, 2010)). When students/parents choose from among different school types, they essentially choose a learning technology. The same inputs (including *Prospera* participation) may generate substantially different outcome trajectories depending on the school type chosen.

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 progression and academic achievements. It also accounts for the selection arising from the *Prospera* enrollment process. We directly control for the observed heterogeneity between *Prospera* and non-*Prospera* students using a rich set of family demographics and further allow for selection on unobserved factors.[8] In particular, we model the unobserved heterogeneity as discrete multinomial types, as in Heckman and Singer (1984) and Cameron and Heckman (1998). These types enter multiple model equations and, in doing so, allow for correlated error structures. We allow the unobserved type distribution to vary by *Prospera* status.

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

[7] Technical schools differ from general schools by including vocational/technical educational curricular components. Telesecondary schools are distance-learning schools that largely serve rural communities and that enroll almost 20% of lower-secondary school students. *Prospera*-beneficiary family children attend telesecondary schools in greater proportions than average. Section two provides more detail on how the school types differ.

[8] As discussed in Todd and Wolpin (2003) and Rivkin et al. (2005) for value-added models and in numerous other studies of schooling (e.g., Behrman et al. (1980); Behrman and Rosenzweig (1999); Altonji et al. (2005); Rothstein (2009)), it is important to control for unobserved inherent student abilities, personality traits, or motivation, that matter for children's achievement growth.

The data also contain information on a small percentage of students suspected to have copied answers on the multiple-choice standardized tests. Copying induces one-sided measurement errors in the dependent variables, the lagged independent variables, or both, which we take into account in estimation. Model parameters are estimated by maximum likelihood where the outcomes at different ages/grades are school enrollment, school choices, mathematics and Spanish test scores, dropping-out, and grade retention.

We use the estimated model to evaluate how *Prospera*-beneficiary status affects schooling progression and academic achievements in different grades. In particular, we simulate school-choice decisions, school-enrollment decisions and academic achievements with and without the *Prospera* program, for children from different family backgrounds. There are multiple channels through which *Prospera* participation can affect these outcomes. First, past participation may increase lagged achievement, which can facilitate present learning. For example, greater comprehension of sixth-grade mathematics can facilitate learning and comprehension of the seventh-grade curriculum.[9] Second, contemporaneous program participation can affect directly learning if the program encourages regular school attendance, student engagement and study efforts. There are two reasons why we might expect the *Prospera* 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 may reduce the pressure on children/youth to work in labor markets while in school and thereby allow for greater focus on schoolwork (Skoufias and Parker (2001)).

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.[10]

Our results contribute to the understudied yet substantive question of whether CCTs improve learning. Fiszbein and Schady (2009) and Baird et al. (2014) systematically reviewed the previous CCT programs and concluded that the effects of CCTs in various countries on achievement tests

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

[10] While initial studies suggested larger impacts of *PROGRESA* on the secondary school enrollment of girls (Schultz (2004)), studies of *PROGRESA*'s impacts on completed schooling based on the original evaluation sample have suggested similar overall effects for boys and girls through lower secondary school (Behrman et al. (2005b) and Todd and Wolpin (2006)) although Parker and Vogl (2018) suggest higher impacts on completed schooling of girls based on nationwide data on the population residing in very poor municipalities.

were disappointingly "small, at best." One caveat is that these conclusions are mostly drawn from relatively short evaluation periods and based on relatively small sample sizes.[11] When evaluating CCT programs on a longer horizon, some papers found significant effects on academic achievement. For example, 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 schooling attainment and learning for boys ten years later. They found a half-grade increase in schooling and substantial gains (approximately 0.25 standard deviations) in mathematics and language achievement scores. By comparing two cohorts (2007 and 2013), 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. Our results can reconcile the mixed findings in the literature by highlighting the accumulative feature of the CCT program effects.

Our analysis also shows that telesecondary schools are important determinants of *Prospera*-program impacts. Distance-learning schools are often the only option for students living in rural areas. We find that telesecondary schools are similar in effectiveness and, in some cases, more effective than the regular public schools for children/youth who attend them. When we use the estimated model to analyze the effect of removing the telesecondary option from the choice set, we find that the dropout rate would be substantially higher and average schooling attainment 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 schooling attainments for local students.

We also find that failing to control for dynamic selection would lead to underestimation of *Prospera's* impact. In particular, we estimate a simpler value-added model grade-by-grade, 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 three sources that cause these downward biases. First, the program causes students at the margin of dropping-out to stay in school longer and failure to control for this changing composition of students would lead to a downward bias in the impact estimates. Second, the simpler model does not allow heterogeneous impacts across different types of schools and, therefore, does not capture that telesecondary schools are particularly effective for *Prospera* beneficiaries. Lastly, the simpler model ignores the negative selection of unobserved types, which also leads to an underestimation of program impacts. We find a richer modeling framework is required to capture heterogeneous program impacts and to control for multiple sources of selection bias.

[11] Due to the data limitation, 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.

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 describes the model and section four the estimation approach. Section five presents the model empirical results. Section six presents the estimated cumulative *Prospera* program effects. 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 concludes.

2 Background

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.[12] 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. 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 has 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.[13]

[12] The National Institute for Assessment of Education (INEE) monitored standards during our period of study.
[13] Based on the authors' tabulations.

2.2 *ENLACE* test score data and additional survey data

From 2006 to 2013, the SEP applied the *Evaluación Nacional de Logro Académico en Centros Escolares*, called the *ENLACE*. The test evaluated student performance in mathematics, Spanish and a rotating subject for all third-to-ninth graders in private and public schools at the end of each academic year. The test is directly based on the curriculum SEP (2010) and intended to be a low-stakes assessment that would be informative about learning outcomes to SEP and to parents. Beginning in 2008, *ENLACE* was also given to students in their final year of upper-secondary school (grade 12). [14] It is supposed to have no bearing on a student's GPA, graduation, or admissions to higher education. The exams were designed to have a mean of 500 and a standard deviation of 100 in their first year of implementation, and subsequent test years were calibrated to allow measurement of changes in learning over time.(See SEP (2010)). 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. In addition to test scores, the ENLACE data also contain information on the age, gender, *Prospera* status, school attendance, school ID, and school type for each student. We examine a cohort of students who were in grade 4 in 2007/08 for whom ENLACE test scores are available for six years (i.e. from fourth grades through ninth grades).

We link the ENLACE data with information on student, parent, and school characteristics from a random sample of schools. These surveys provide detailed information on parental education, monthly family income, home infrastructure, number of siblings, and other household characteristics. The combination of the ENLACE data with the student/parent/school surveys produces an exceptionally rich data set that is representative of the population of Mexican school children. The detailed survey information available on housing characteristics is useful for reliably approximating the *Prospera* program eligibility criteria, as will be discussed below.

In addition, we gather geocode information (latitude and longitude) for each primary and secondary school. We use such information to characterize the number of local primary schools (within 5 km radius) and number of local secondary schools (within 10 km radius). We then further calculate the distance (in kilometers) between the primary school and the nearest secondary school of each type. Some further details are provided in appendix A.

[14] The ENLACE exams have been used as a means of evaluating educational interventions by several papers (Avitabile and De Hoyos (2018); De Hoyos Navarro et al. (2019); De Hoyos et al. (2017)). Scores on these exams have been shown to have predictive power on important life outcomes including university enrollment and wages (De Hoyos et al. (2018)).

2.3 Descriptive statistics

Table 1 shows summary statistics for the students in our sample. 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 49% are female. The parents of beneficiary students have much less schooling; 65% of the fathers and 67% of the mothers in beneficiary families have primary school (6 grades) or less in comparison to 23% and 26% for non-beneficiaries. Beneficiary fathers are less likely to work full-time, whereas mothers are less likely to work in the labor market at all. Beneficiary students are slightly more likely to have either a mother or father who is not at home. 91% of non-beneficiary students live in urban areas in comparison to 43% for beneficiaries.

Prospera families tend to be larger, with 38% having six or more household members in comparison to 21% for non-beneficiaries. Children from beneficiary families have fewer years of preschool education on average. They are also much less likely to have access to computers or the internet at home. 15% of beneficiary students have computers at home and 10% have access to the internet in comparison to 50% and 34% for non-beneficiaries. In terms of languages spoken at home, 6% of children from beneficiary families speak indigenous languages (sometimes in combination with Spanish) in comparison to 1% of non-beneficiary children. Also, a larger fraction of the children from *Prospera* families live in the South and a smaller fraction in the North.

Figure 1 shows the average test scores by grade and by *Prospera*-beneficiary status, where $P = 1$ denotes that the family participates in *Prospera*. In fourth grade, there are considerable gaps in average mathematics and Spanish test scores. However, the gaps narrow in lower-secondary school grades. By grades 8 and 9, *Prospera* and *non-Prospera* students have similar math test score distributions but there is still a gap for Spanish test scores. Gaps in secondary school grades are smaller than in primary school. We further compare the test score distributions by grade and by lower secondary school types in Appendix Figure B2. The figures show that students in telesecondary schools show greater test score improvements across grades than in other school types. However, the comparisons are not necessarily causal due to compositional differences in the students attending different school types and differential dropout rates across grades and school types. Our model will control for these sources of selection.

Table 2 shows the school-type enrollment distributions by *Prospera* beneficiary status (P). Only about 1% of non-*Prospera beneficiary* children/youth attend indigenous primary schools, while 10% *Prospera* beneficiaries attend indigenous primary schools. In secondary school, the majority of children from non-beneficiary families are enrolled in general schools, whereas about 40% of beneficiary families are enrolled in telesecondary schools. In contrast, only 12% of non-*Prospera beneficiary* children/youth attend telesecondary school in grade 7. The last column of Table 2 reports the fraction of dropouts at every grade during lower-secondary school. 24% of Prospera beneficiary youth dropout prior to

Table 1: Summary Statistics

	Prospera		Non-*Prospera*			*Prospera*		Non-*Prospera*	
	Mean	Std	Mean	Std		Mean	Std	Mean	Std
Age (at grade 4)	9.22	0.64	9.05	0.49	Number of household members				
Female	0.49	0.50	0.49	0.50	*≤3 people*	0.21	0.41	0.37	0.48
Education cat. (dad)					*4 people*	0.22	0.41	0.28	0.45
Below primary school	0.41	0.49	0.11	0.31	*5 people*	0.19	0.39	0.15	0.35
Primary school completed	0.24	0.43	0.12	0.32	*≥6 people*	0.38	0.49	0.21	0.41
Secondary or below	0.26	0.44	0.33	0.47	Number of preschool years				
College or above	0.07	0.25	0.43	0.50	*0 year*	0.02	0.16	0.01	0.12
Education cat. (mom)					*1 year*	0.11	0.31	0.04	0.19
Primary school	0.42	0.49	0.12	0.32	*2 years*	0.26	0.44	0.21	0.41
Primary school completed	0.25	0.44	0.14	0.35	*3 years*	0.32	0.47	0.43	0.50
Secondary or below	0.27	0.44	0.32	0.47	*4 years*	0.29	0.45	0.30	0.46
College or above	0.05	0.22	0.42	0.49	Internet at home	0.10	0.30	0.34	0.47
Working status (dad)					Computer at home	0.15	0.36	0.50	0.50
Part time	0.22	0.42	0.16	0.36	First language				
Full time	0.75	0.44	0.81	0.39	*Spanish*	0.91	0.29	0.98	0.12
Working status (mom)					*Indigenous*	0.06	0.24	0.01	0.08
Housework	0.78	0.42	0.52	0.50	*Both*	0.03	0.18	0.01	0.09
Part time	0.12	0.33	0.28	0.45	Region				
Full time	0.09	0.29	0.19	0.39	*North*	0.11	0.31	0.30	0.46
Father at home	0.84	0.36	0.85	0.35	*North-center*	0.26	0.44	0.25	0.43
Mother at home	0.96	0.20	0.98	0.14	*Center*	0.38	0.48	0.37	0.48
Urban dummy	0.43	0.50	0.91	0.29	*South*	0.26	0.44	0.09	0.28
Obs.	39,098		120,344		Obs.	39,098		120,344	

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

grade 9 in comparison to 20% 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 tails of the test-score distributions. For this reason, it is important to control for selective dropout in evaluating achievement test-score program impacts.

Table 3 shows the percentage of students retained at each grade by *Prospera*-beneficiary status. Grade retention is more prevalent in primary-school grades and is higher for *Prospera* children. In lower-secondary school, about 1% are retained and there is no systematic pattern by beneficiary status. Table 4 shows the transition patterns from the two types of primary school (general and indigenous) to attend one of the three types of lower-secondary schools or to dropping-out, by *Prospera*-beneficiary status. 16% of children who attended indigenous primary schools drop-out in comparison to 11% of children who attended general primary schools. For non-beneficiary children who attended general primary schools, 61% continue in general lower-secondary schools, 7% attend telesecondary schools, 28% attend technical schools, and the remainder (4%) drop-out. The comparable school-choice distribution for *Prospera* beneficiaries is 30% general, 42% telesecondary and 21% technical, with 7%

Figure 1: Mathematics and Spanish test scores distributions (CDFs) by grade and *Prospera* status

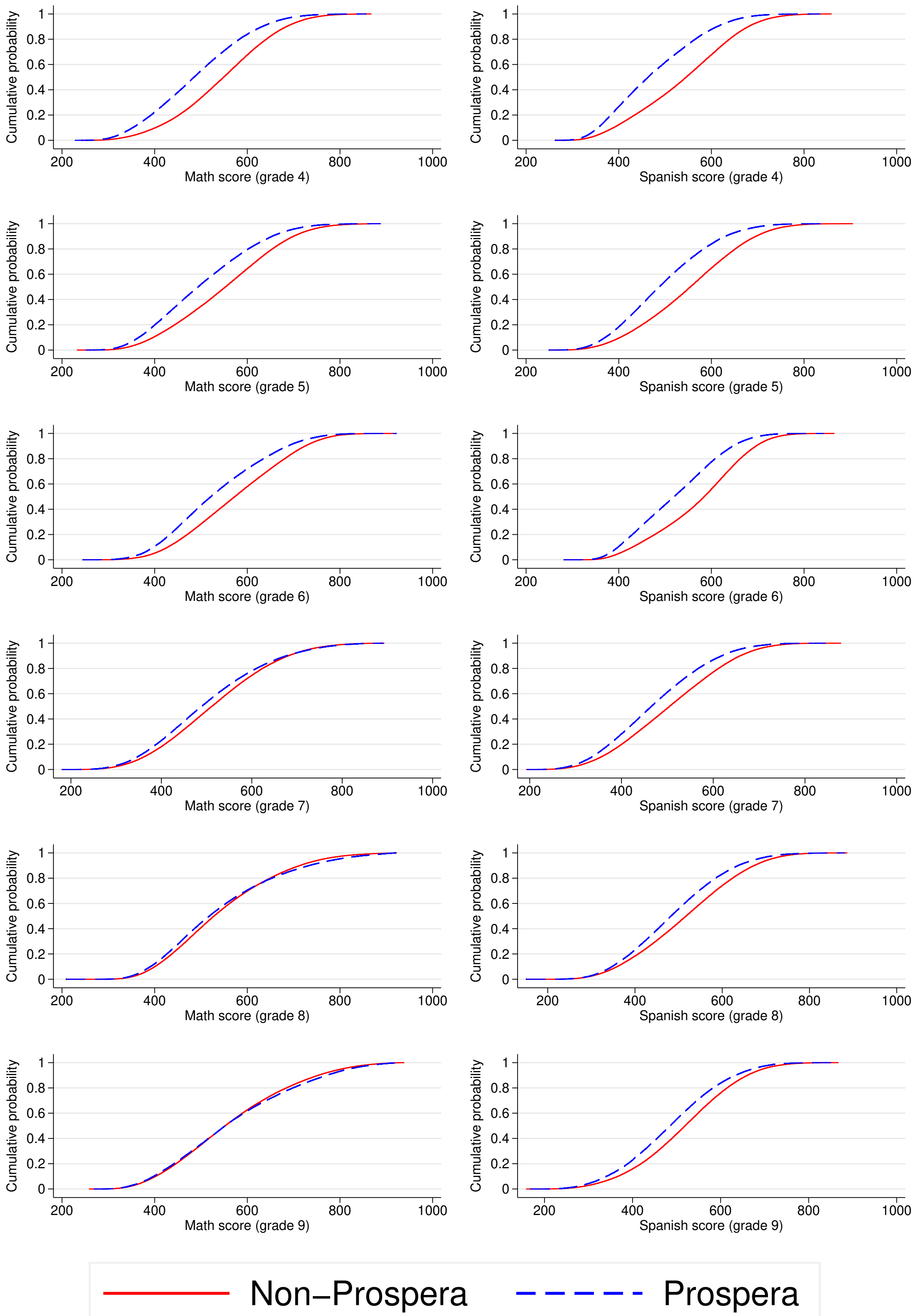


Table 2: Enrollment distribution by school type, grade and *Prospera* status (P)

Primary school	General	Indigenous		
Beneficiary (P=1)				
Grade 4	0.90	0.10		
Grade 5	0.89	0.11		
Grade 6	0.90	0.10		
Non beneficiary (P=0)				
Grade 4	0.99	0.01		
Grade 5	0.99	0.01		
Grade 6	0.99	0.01		
Secondary school	General	Telesecondary	Technical	Dropout
Beneficiary (P=1)				
Grade 7	0.27	0.44	0.21	0.08
Grade 8	0.26	0.43	0.21	0.10
Grade 9	0.22	0.37	0.17	0.24
Non beneficiary (P=0)				
Grade 7	0.53	0.12	0.29	0.06
Grade 8	0.52	0.12	0.29	0.08
Grade 9	0.45	0.10	0.25	0.20

Note: This table displays the school-type enrollment distribution by beneficiary status and grades, where $P = 1$ denotes *Prospera* participation. The dropout only matters for lower-secondary school and is measured prior to entering that grade and is cumulative. Each row totals add to 1.

drop-out. Thus, youth from beneficiary and non-beneficiary families enroll in different types of schools, which is potentially important to understanding the overall effectiveness of the *Prospera* program in promoting schooling attainment and academic achievement.

In Appendix Table B1, we provide 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 considered to be different schools.[15] 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, 77% of children do not have access to indigenous schools, which are typically located in areas with significant indigenous populations. At the lower-secondary level, there are fewer schools and they are larger.

[15] In Appendix Figure B1, 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 3: Percentage of retained students
by school type, grade and *Prospera* status (P)

Primary school	General		Indigenous			
	$P=0$	$P=1$	$P=0$	$P=1$		
4th year	1.0%	2.5%	1.5%	3.2%		
5th year	0.8%	1.7%	1.7%	3.0%		
6th year	0.1%	0.3%	0.0%	0.3%		
Lower-secondary school	**General**		**Telesecondary**		**Technical**	
	$P=0$	$P=1$	$P=0$	$P=1$	$P=0$	$P=1$
1st year	1.2%	1.1%	1.8%	0.8%	0.8%	1.1%
2nd year	1.2%	1.0%	0.6%	0.5%	0.7%	0.7%

Table 4: Primary and lower-secondary school transitions
by *Prospera* status (P)

		General	Tele	Tech	Dropout
General	$P=1$	0.30	0.42	0.21	0.07
	$P=0$	0.61	0.07	0.28	0.04
	Total	0.54	0.15	0.26	0.05
Indigenous	$P=1$	0.09	0.59	0.21	0.11
	$P=0$	0.25	0.35	0.24	0.16
	Total	0.12	0.55	0.21	0.12

3 Model

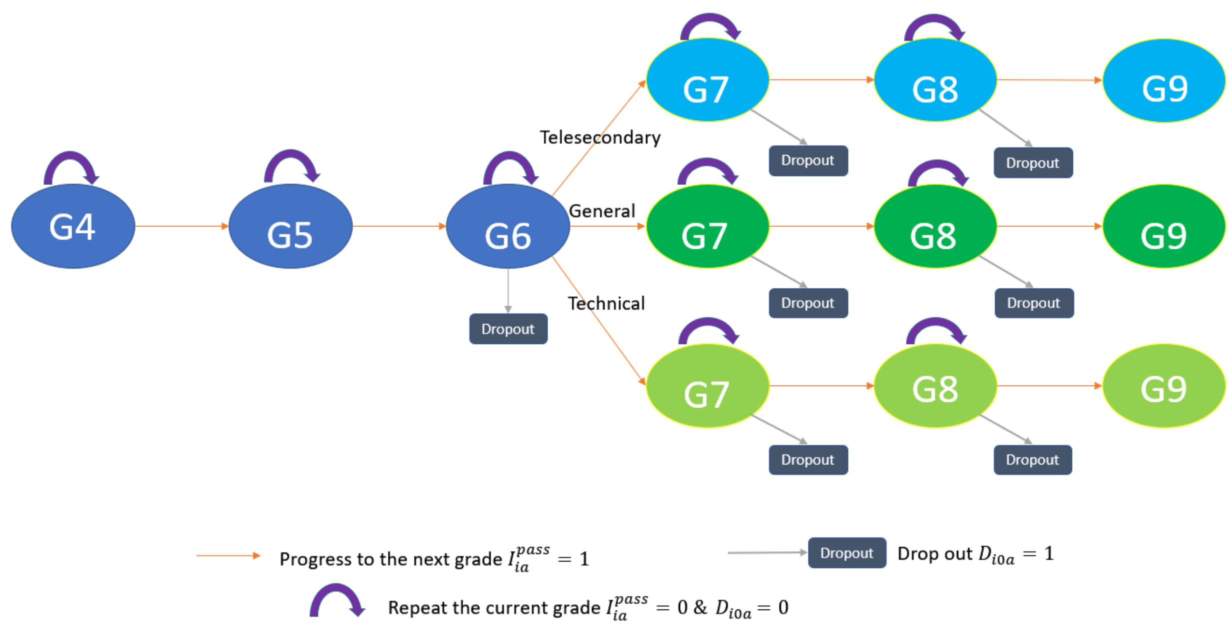
Our modeling framework combines a model of school attendance decisions at different types of schools with dynamically linked models of academic achievement in mathematics and Spanish at different ages/grades. We specify test-score gains from year-to-year using a value-added framework that relates current achievement to lagged achievement, family and school inputs into the learning process, and student unobserved heterogeneity (e.g. arising from ability or preferences). Our framework also allows for dropping-out of school and for grade retention.

Our empirical framework could be considered 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, likely reflecting the decisions of students, parents, and school administrators. As discussed below, the outside option in the school choice model is to drop out.

3.1 General environment and sequential outcomes

Individuals are indexed by i, $i = 1, .., n$ and each model period corresponds to one school year. In the initial period (a_f, corresponding to the age at grade 4), students/parents can choose to attend one of

Figure 2: Potential sequential outcomes from grade 4 to 9



two types of primary schools: general ($j = 1$) or indigenous (bilingual) ($j = 4$), depending on the types locally available (within 5 km). At the end of grade 6, students simultaneously make school enrollment decisions and school-type choices from up to three options: general ($j = 1$), telesecondary ($j = 2$), or technical ($j = 3$), depending on the types locally available (within 10 km). We can summarize the choice set J_{ia}^{g} at different grades as:

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

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.[16]

Let $D_{ija} = 1$ if the individual i 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 in which she is enrolled at age a, else $I_{ia}^{Pass} = 0$. We assume the passing outcome I_{ia}^{Pass} is realized at the end of the school year, prior to the other decisions being made. The potential sequential outcomes from grade 4 to grade 9 are illustrated in figure 2:

As seen in figure 2, until grade 6 the possible outcomes are whether a student is retained in the current grade ($I_{ia}^{Pass} = 0$) or progresses to the next grade ($I_{ia}^{Pass} = 1$), conditioning on their primary school

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

types when entering into the model at grade 4, their family background and their *Prospera*-beneficiary status. Upon passing grade 6 ($I_{ia}^{Pass} = 1$), students simultaneously make enrollment decisions D_{i0a} and school choices D_{ija}, depending on locally available school types.[17] Once enrolled in a secondary school, students decide in each period whether to drop out or stay in school. In each grade, there is a probability of passing or having to repeat the grade. 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:

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

3.2 Accounting for selective program participation

As previously noted, our aim is to use our estimated model to assess the impacts of *Prospera* participation on schooling progression and academic achievement, where we treat *Prospera* beneficiary status as a family characteristic. $P_i = 1$ denotes that a child/youth comes from a *Prospera*-beneficiary family, else $P_i = 0$.[18] The vast majority of families who are eligible to participate opt to do so, for the following reasons (i) the transfers that families receive under the program are large, accounting for a 20% increase in family income on average; (ii) the program provides transfers in grades 3-6 when school attendance is nearly universal; and (iii) the program has been in operation since 1997 and is well-known (see Parker and Todd (2017). However, there may still be some nonparticipating families, particularly ones that are unaware of their eligibility. Also, families may receive partial benefits if they send some but not all of their children to school. We consider families to be participating if they are listed in the administrative data as being enrolled. As was seen in Table 1, the average *Prospera* student is not directly comparable to an average *non-Prospera* student. Beneficiary students come from higher poverty backgrounds and are more likely to live in rural areas.

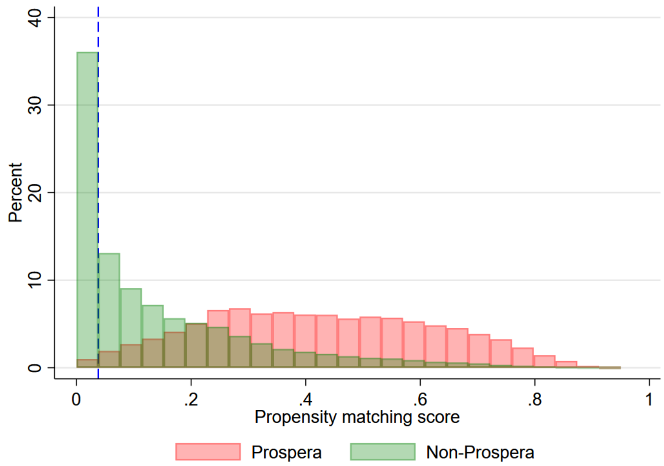
Program selection in this context is largely based on observed characteristics, because the general eligibility criteria are known and most eligible families opt to participate. Although the survey data we use were not collected for the purpose of ascertaining Prospera eligibility, the data are rich and contain information on most of the eligibility determinants (see below).[19] We therefore estimate a probit model for the probability that each family is eligible for and participates in the *Prospera* program given the available information. Program eligibility is not means-tested by income, because income can be difficult to measure in a country with a significant informal sector and where many low-

[17]Because school enrollment is very high during primary school, we assume students do not drop-out during primary grades. Therefore, they do not make choices about continuing in school until the end of grade 6.

[18]The administrative data on participation in *Prospera*-was linked with the test-score database when children were in grade 6. The overwhelming majority of children enroll in grade 6, although there is a small fraction that drops-out prior to entering grade 6 that we do not include in our analysis sample.

[19]The precise eligibility criteria are not made public, but some of the authors of this paper were involved in the design of the criteria.

Figure 3: The propensity score distribution by *Prospera* status (P)



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

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, piped water 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.[20] We use information on housing characteristics and demographics that was gathered through the student and parent surveys to calculate a household's probability of being eligible and participating in the program (a propensity score). The estimated coefficients from this model are shown in Appendix A.3. The percentage of individuals correctly classified as being beneficiaries or not under the estimated model is high (90%).

Figure 3 plots the propensity-score distributions for children from *Prospera*-beneficiary households (in red) and non-beneficiary households (in green). As seen in the figure, a large fraction of non-beneficiaries fall in the first histogram bin, meaning that they have extremely low probabilities of participating in *Prospera*, generally because their characteristics make them ineligible. To increase comparability between the *Prospera* subsample and the comparison-group subsample, we impose a common support restriction and exclude in our impact analysis *Prospera* beneficiaries and the non-beneficiaries with propensity scores below the 1% quantile (the lowest bin of the histogram).This trimming threshold excludes 383 children from *Prospera* families and 50,798 non-beneficiary children.[21]

[20]The detailed eligibility criteria 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.

[21]This type of trimming is common in the application of matching estimators as a way of imposing "common support." 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 addition, our model also includes observed covariates to further capture differences between *Prospera* and non-*Prospera* households (such as parents' education). We allow for the possibility that children/youth from *Prospera* beneficiary families may differ in unobservable ways. In particular, we include unobserved heterogeneity by allowing for four discrete multinomial types (see, e.g. Heckman and Singer (1984) and Cunha and Heckman (2008)) that enter into all the model equations. Models with discrete unobserved types are sometimes called latent class models.[22] Let $\mu_{il} = 1$ if individual i is of type l, =0 else, where $l \in \{1, ..., L\}$ and $L = 4$. The probability of type l may depend on the individual's *Prospera* status but is assumed to be independent of other initial conditions. We use the notation $\rho_l^P \equiv \Pr(\mu_{il} = 1|P)$ to denote the fraction of type l among students with *Prospera* status P, with $\sum_l \rho_l^P = 1, P = \{0, 1\}$.

3.3 Other model components

We next describe how we specify different model components including the value-added academic achievement model, the school-choice/dropout model, and the grade-repetition process.

Value-added model: Achievements in mathematics and Spanish evolve over time with attendance in different grades and 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 $I_{i,a-1}^{Pass} = 1$ or is repeating the grade $I_{i,a-1}^{Pass} = 0$.[23] Let Z_{ia}^A denote the vector of observed characteristics of the youth and of the family that enter the achievement production function.

$$A_{ia}^m = \delta_{0jl}^{mgI} + A_{i,a-1}\delta_{1j}^{gI} + \delta_{2j}^{mgI} P_i + Z_{ia}^A \delta_{3j}^{mgI} + \omega_{ija}^{mgI}. \tag{2}$$

In this equation, δ_{0jl}^{mgI} is the type-specific intercept that allows for unobserved heterogeneity (l denotes the type). $A_{i,a-1} = \left\{A_{i,a-1}^1, A_{i,a-1}^2\right\}$ is a 2×1 vector including both the mathematics score and the Spanish score from the previous period $a - 1$. The lagged test score terms are assumed to be sufficient statistics for the impacts of past inputs in the learning process. Our specification also allows cross effects between Spanish and math. For example, better Spanish skills may enhance student's understanding in their math classes, implying a positive effect of past Spanish score on the current math score. δ_{2j}^{mgI} captures the impact of the *Prospera* program. We assume that the error terms ω_{ija}^{mgI},

[22] Alternatively, we could impose a continuous distribution for the unobserved heterogeneity, for example, a mixture of normal distributions. 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, he finds that the discrete-type method performs better.

[23] 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.

conditional on the unobserved types, are iid and normally distributed.

Most of the literature considers learning technology to be exogenous. By combining a school-choice model with value-added models that vary by school type, we allow students/parents to select a learning technology.

School-choice model: We next specify how individuals make their schooling choice D_{ija} from the available options, J^g_{ia}(depending on his/her grade and geographic location, defined in equation 1). Assuming a random-utility model with Type 1 extreme-value errors (taste heterogeneity), we obtain that the probability of choosing option j_{ia} is multinomial logistic:

$$\Pr(D_{ija}=1|\tilde{\Omega}(a),\mu_l)=\begin{cases}\frac{\exp(\mu^g_{0jl}+A_{ia-1}\phi^g_{1j}+P_i\phi^g_{2j}+Z^D_{ia}\phi^g_{3j}+w_{ia}\phi_{4j}+S_{ija}\phi_{5j})}{\sum_{j'\in J^g_a}\exp(\mu^g_{0j'l}+A_{ia-1}\phi^g_{1j'}+P_i\phi^g_{2j'}+Z^D_{ia}\phi^g_{3j'}+w_{ia}\phi_{4j'}+S_{ij'a}\phi_{5j'})} & \text{if } j_{ia}\in J^g_a\\ 0 & \text{if } j_{ia}\notin J^g_a\end{cases} \quad (3)$$

where μ^g_{0jl} is a type-specific intercept (l denotes the type). ϕ^g_{1j} captures the effect of test scores on schooling choices. ϕ^g_{2j} captures the impact of *Prospera* on schooling choices. $Z^D_{ia}\in\tilde{\Omega}(a)$ includes demographic and family-background characteristics. There are three additional variables that enter the school-choice equations but not other parts of the model: (i) the imputed hourly wage w_{ia}; (ii) the distance to the closest school of each type S^1_{ija}; and (iii) the local supply of schools of each type S^2_{ija}.

Because our test-score databases do not contain information on wages for children/youth who 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.[24] The imputed wage, w_{ia}, captures the opportunity costs of being enrolled. The other two variables in S_{ija}, the distance to the closest school (S^1_{ija}) and the total number of schools for each type (S^2_{ija}), capture the impact of local supplies of different types of schools on schooling choices.

Our assumption is that these three variables significantly affect the school-choice decisions but do not directly enter the test-score equations, so that the variables can be considered exclusion restrictions. This assumption could be invalid, for example, if higher wages provide incentives for students to work part-time while enrolled in school. Another potential threat to validity is that the travel distance may directly affect the commuting time required to attend schools. Both channels may negatively impact academic performance through fatigue or a reduced ability to concentrate on studying. To examine the empirical relevance of such concerns, we also estimated a specification in which the variables $\{S_{ija},w_{ija}\}$ are added to the value-added equation . We examined (i) whether the coefficients associated with *Prospera*-beneficiary status change and (ii) whether the estimated coefficients associated with the variables $\{S_{ija},w_{ija}\}$ are significantly different from 0. We did not find evidence of direct impacts of

[24]The imputation procedure also includes a selection correction to account for selective labor-force participation. For details, see appendix A.2

$\{S_{ija}, w_{ija}\}$ on test scores, conditional on the other model covariates.

Grade retention: 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, $Z^I_{ia} \in \tilde{\Omega}(a)$:

$$Pr(I^{Pass}_{ia} = 1 | \tilde{\Omega}(a), \mu_l) = \Phi(\gamma^g_{0l} + A_{ia}\gamma^g_1 + \gamma^g_2 P_i + j_{ia}\gamma^g_3 + Z^I_{ia}\gamma^g_4). \tag{4}$$

The coefficients of the passing probability are grade-specific. γ^g_{0l} is a unobserved type-specific intercept.

3.4 Treatment effect heterogeneity

Prospera may have heterogeneous impacts for students from different backgrounds. Inspired by the marginal-treatment-effect (MTE) literature, we divide the analysis sample into quartiles based on the *Prospera*-eligible propensity scores and allow the *Prospera* impact to vary by quartile.[25] 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 4 shows the distribution of *Prospera*-beneficiary children across quartiles. 83% of *Prospera* children/youth are concentrated in quartiles 3 and 4 (30% and 53%), reflecting the fact that the program is targeted at the poorest families.

In estimation, we allow coefficients associated with *Prospera* participation to be quartile-specific. In particular, we redefine $\{\delta^{mgI}_{2j}, \phi^g_{2j}, \gamma^g_2\}$ to be $\{\delta^{mgI}_{2j\eta}, \phi^g_{2j\eta}, \gamma^g_{2\eta}\}, \eta = \{1, 2, 3, 4\}$, in which η represents the quartile to which each individual belongs. Our estimates (reported below) indicate that impacts are heterogeneous, with the most disadvantaged children/youth experiencing the largest impacts.

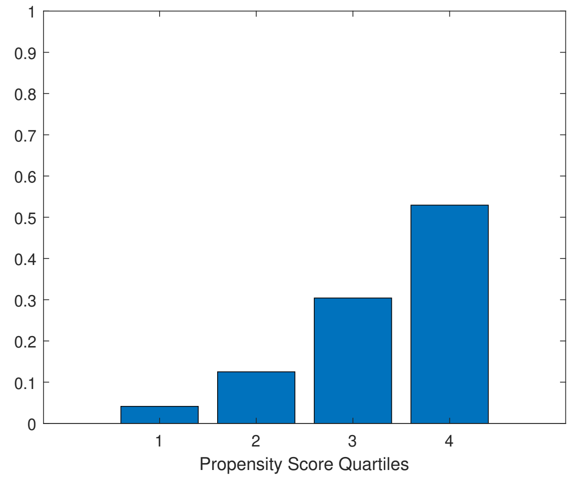
3.5 Model interpretation

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 do not specify the utility functions used by agents to make these decisions.[26] In

[25] We capture potential heterogeneity more parsimoniously than a standard approach in the literature, which is to estimate the treatment effect nonparametrically as a function of the propensity-matching score. (e.g. Heckman and Vytlacil (2001, 2005, 2007)) However, most literature implements MTE in a static set-up whereas our model is dynamic.

[26] 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. For discussion of how to approximate the decision rules, see, e.g. Keane et al. (2011); Heckman and Navarro (2007); Heckman et al. (2018).

Figure 4: Distribution of *Prospera* students across propensity-score quartiles



our case, the school-choice probability represents the difference between the value functions associated with alternative schooling-work options.

There are some advantages of adopting this kind of "quasi-structural" approach. First, if multiple individuals are involved in the school-choice decisions (students, parents, school administrators) then our approach avoids having to model the decision processes of all of these agents and how they interact. Second, estimating a structural dynamic schooling model with academic achievement is complicated, because the state space would need to include two continuous lagged test scores (which usually requires 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 the 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 transfers given only upon completing certain grades in the future. This limits the range of counterfactual policies that can be considered. In the analysis reported below, we consider the effects of two time-invariant policy changes, eliminating *Prospera* and removing telesecondary schools from the choice set.

3.6 Identification

In our model, prior to grade 6, there is no dropout decision and there is only the binary choice over two types of elementary schools. Manski (1988) and Heckman and Navarro (2007) consider semiparametric identification of binary-choice index models in static and dynamic settings.[27] 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. At the end of

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

grade 6, students make a decision about what type of secondary school to enter or whether to drop out. The decision is observed for everyone in our sample (recall that dropout in primary school is rare and our sample includes children who attend through sixth grade). Our multinomial logistic model with discrete unobserved types falls within the more general class of mixed logit models. Fox et al. (2012) show nonparametric identification in such models.

At the end of grades 7 and 8, students only decide whether to dropout, with the dropout decisions only being observed for students who did not dropout in prior grades. Heckman and Navarro (2007) establish semiparametric identification of a class of dynamic discrete-choice models. They consider a model of optimal stopping time of schooling as their running example.[28] Their framework allows for general forms of duration dependence and unobserved heterogeneity. Let γ^t denote the parameters of the period t index equations and let F_ν^t denote the distribution of the period t unobserved shocks. As discussed in Heckman and Navarro (2007), the set (γ^t, F_ν^t) is identified relative to all other sets of parameters $(\gamma^{*t}, F_\nu^{*t})$ if there exists a sequence of past choices such that the probability of observing a choice under the true parameters differs from that under alternative parameter values.[29]. Their identification proof (Theorem 1) follows an identification-in-the-limit strategy that conditions on large values of the indices associated with preceding choice probabilities. In our application, it means that the dropout probability function at any grade can be identified using the subset of individuals who attain that grade level with probability close to 1. Although their proof does not require conventional exclusion restrictions, they note that the assumptions of their Theorem 1 will be satisfied if there are transition-specific exclusion restrictions (that satisfy certain rank and support conditions). In our context, the hourly wage w_{ia} that is imputed from the census data and that varies with age, educational attainment, and region of residence constitutes an exclusion restriction.

The value-added achievement model is a standard dynamic-panel-data model, of the kind discussed in Arellano and Bond (1991) with two modifications. First, the model incorporates discrete unobserved types to control for unobserved variables. Second, there is the potential for dropout after grades 6, 7, and 8, and the test score outcomes are only available for students who enroll in school. Heckman and Navarro (2007) also consider the identification of continuous outcomes and counterfactual outcomes in a stopping-time model, again using an identification-in-the-limit strategy based on sets of students with characteristics such that they advance to the next grade with a probability close to 1. Our permanent-transitory specification of the error term in the value-added equation is analogous to the one-factor error structure that they present. They note that one needs to have at least three outcome measures to identify a one-factor structure. In our case, we have 12 continuous outcomes (Math and

[28]They call this model a time-to-treatment model, where the treatment is stopping school.

[29]Their theorem allows for general error structures F_ν^t that can be correlated over time for each individual but are assumed to be independent across individuals and of regressors in the initial time period. Our permanent-transitory error structure is a special case. (see footnote #16 in Heckman and Navarro (2007))

Spanish, from grades 4 to grade 9). They establish nonparametric identifiability of both the factor and error term distributions.[30] Given our multi-equation setting, we implement a parametric model.

4 Estimation

4.1 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 often 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.[31]

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 randomness of the cheating effect.[32] When students do not cheat ($I_{ia}^{cheat} = 0$) and the test scores are not missing ($I_{ia}^{miss} = 0$), under this specification, the observed test score equals the true test score:

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

[30] Heckman and Navarro (2007) assume the factors are independent of covariates. We assume the type distribution can vary by *Prospera* status (an initial condition) and that the conditional mean type effect is zero. The Heckman and Navarro (2007) identification strategy can still be applied under this slightly more general formulation by stratifying the data by *Prospera* status.

[31] 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).

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

4.2 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 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*.[33] We use notation $\Omega_i(0)$ to denote the set of these initial state variables. 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.

Given the vector of the observed test scores $\tilde{A}_i = \{\tilde{A}_{ia_f}^1, \tilde{A}_{ia_f}^2, ..., \tilde{A}_{ia_l}^1, \tilde{A}_{ia_l}^2\}$, 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:

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

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

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

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

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 2 and equation 5. J_{ia}^g is the available school-choice set at grade g_a defined in equation 3.3. The vector $\rho^P = \{\rho_1^P, ..., \rho_4^P\}$ denotes the vector of unobserved-type probabilities conditioning on *Prospera* status P, where $\rho_l^P = \Pr(\mu_l = 1 | P)$.

[33] 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.

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[34]:

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

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.[35]

4.3 Some simplifying assumptions

We impose three assumptions with regard to the choice problem that simplify the estimation and are largely 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.[36]

4.4 Evaluating the effects of *Prospera*-program participation

Prospera potentially provides cash transfers for all children of beneficiary households who are enrolled in grades 3-12. For children in grades 3-9, the transfers typically go to the mothers. Whether a transfer is received for each child depends, however, on whether that child regularly attends school (at least 85% of school days). We consider the family's *Prospera* status as a time-invariant characteristic, so it is contained in the initial state space $\Omega(0)$. Once families are enrolled in the program, they rarely lose their eligibility. Even if one child is not attending school, the family may still receive transfers for other children and is still considered to be participating. We use the estimated schooling 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.

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

[35] 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.

[36] 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 2 and $\gamma_k^4 = \gamma_k^5 = \gamma_k^6, \gamma_k^7 = \gamma_k^8, k = \{1,2,3,4\}$ in equation 4 in the periods when $I_{ia}^{pass} = 0$. But the intercept term δ_{0jl}^{mgI} and γ_{0l}^g are grade-specific.

5 Model estimates

5.1 Value-added models

Tables 5 and 6 provide the estimates for the value-added production functions for mathematics and Spanish. The tables show the coefficients associated with the lagged scores, the *Prospera* participation indicator 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 production process. Also, they include indicators for rural/urban residence and for region of residence to capture regional differences in school quality/infrastructure. The full set of estimated parameters is in Appendix C. Each model includes lagged scores in both subjects, assuming that knowledge of Spanish might facilitate learning in mathematics and vice versa.[37] 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 coefficient estimates. Boys have 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.[38]

The previous tables report the one-year effect of being a *Prospera* beneficiary, allowing the effects to vary by propensity-score quartile. 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 for the estimates associated with these two quartiles, leading to their coefficient estimates being insignificant. 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 includes the *Prospera* beneficiaries 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. Because quartile 4 contains the majority of the *Prospera* beneficaries (52.7%), the *Prospera* impacts on this subgroup primarily determine the average population impacts.

[37] Aucejo and James (2021) estimate production functions 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.

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

Table 5: Mathematics value-added model estimates

	General		*Indigenous*			*General*			*Telesecondary*			*Technical*	
	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Math score													
Lag math	0.56	0.55	0.41	0.45	0.45	0.48	0.49	0.35	0.42	0.41	0.42	0.49	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	0.16	0.13	0.11	0.12	0.27	0.17	0.25	0.26	0.16	0.25	0.27	0.16	0.27
	(0.01)	(0.00)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prospera effect by propensity score quartile													
P*Q1 (4%)	0.62	-0.79	-1.05	2.89	2.74	2.49	-0.74	0.33	-2.84	2.24	1.70	2.81	2.36
	(3.23)	(3.15)	(31.28)	(40.96)	(4.35)	(4.03)	(4.75)	(12.34)	(11.99)	(12.73)	(6.57)	(5.55)	(6.56)
P*Q2 (13%)	0.33	-1.83	-2.61	-3.38	2.18	-2.19	2.59	2.45	2.31	1.93	2.05	1.90	2.64
	(2.07)	(1.99)	(15.63)	(13.96)	(3.17)	(2.92)	(3.27)	(6.78)	(5.26)	(6.24)	(4.27)	(3.71)	(4.29)
P*Q3 (30%)	-0.14	1.02	2.05	2.01	3.99	-1.00	4.37	6.16	6.76	6.09	4.34	3.22	3.41
	(1.68)	(1.60)	(9.15)	(8.27)	(2.73)	(2.53)	(2.75)	(4.68)	(3.80)	(4.36)	(3.51)	(3.05)	(3.37)
P*Q4 (53%)	-0.05	-1.05	-0.11	-3.34	13.57	7.50	6.97	12.40	6.62	9.33	12.53	5.75	7.74
	(1.61)	(1.55)	(6.43)	(5.80)	(2.72)	(2.50)	(2.76)	(4.41)	(3.55)	(4.13)	(3.33)	(2.95)	(3.24)
Gender													
Male	2.99	2.47	1.97	-0.41	7.16	16.65	10.37	1.02	10.22	2.01	4.39	17.45	9.20
	(0.66)	(0.65)	(3.40)	(3.06)	(1.04)	(0.93)	(1.09)	(2.03)	(1.71)	(2.02)	(1.39)	(1.22)	(1.46)

Note: Standard errors are shown in parentheses. 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 C.

5.2 School-choice/dropout model

Table 7 reports estimated coefficients from the school-choice/dropout model after grade 6, where the omitted category is dropping-out and working.[39] The specification also includes other covariates as described in the table footnote. Family background and demographic variables capture heterogeneities 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 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. Male students are significantly more likely to enter secondary school than female students. The estimated distance coefficients show that the distances that students need to travel 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.[40] Children who have higher potential

Table 6: 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 math	0.23	0.20	0.25	0.19	0.14	0.23	0.19	0.09	0.17	0.16	0.13	0.23	0.18
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lag Spanish	0.43	0.40	0.27	0.32	0.53	0.47	0.48	0.45	0.41	0.39	0.53	0.48	0.49
	(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%)	0.02	-2.03	0.09	-3.72	0.06	-2.53	-2.81	-0.38	-0.72	-2.81	-2.22	0.53	-1.95
	(2.99)	(2.72)	(23.69)	(37.37)	(4.29)	(3.78)	(3.99)	(9.65)	(9.07)	(9.99)	(5.80)	(5.01)	(5.60)
P*Q2 (13%)	-1.42	-1.81	-2.86	0.13	1.34	-2.35	-1.01	1.37	0.51	-1.08	-0.16	1.75	-0.59
	(1.99)	(1.67)	(13.25)	(11.30)	(3.11)	(2.75)	(2.72)	(5.25)	(4.21)	(4.96)	(3.88)	(3.42)	(3.59)
P*Q3 (30%)	-2.77	-1.69	-2.30	-0.99	3.79	-1.59	-1.26	2.33	0.68	1.50	2.09	2.55	-1.35
	(1.63)	(1.36)	(8.01)	(6.91)	(2.73)	(2.38)	(2.34)	(3.74)	(3.02)	(3.42)	(3.29)	(2.85)	(2.83)
P*Q4 (53%)	-1.07	-1.48	0.95	-2.22	5.83	3.99	-1.60	5.42	1.89	1.03	5.76	3.21	0.89
	(1.58)	(1.30)	(5.95)	(4.75)	(2.73)	(2.44)	(2.38)	(3.53)	(2.88)	(3.22)	(3.19)	(2.84)	(2.76)
Gender													
Male	-16.70	-15.51	-8.74	-12.29	-26.17	-16.97	-12.87	-24.89	-18.66	-15.26	-27.15	-16.74	-13.64
	(0.62)	(0.56)	(3.05)	(2.61)	(0.98)	(0.86)	(0.91)	(1.64)	(1.37)	(1.57)	(1.25)	(1.12)	(1.19)

Note: See note to Table 5.

higher hourly wages are less likely to attend lower-secondary school.

In grades 7 and 8, students decide only whether to drop out and Table 8 shows how being a *Prospera* beneficiary affects this decision. The estimates show a statistically significant negative effect on dropout for youth from lower SES households (i.e. those with propensity scores in higher quartiles), with larger impacts for seventh than for eighth grade. As expected, youth with higher test scores are less likely to drop out. Although males enter secondary school at slightly higher rates than girls, they are significantly more likely to drop out at grades 7 and 8. Greater distance to school significantly increases the dropout probability. Students attending telesecondary are slightly less likely to drop out after 7th grade but more likely to drop out after eighth grade. Lastly, higher local wages increase the probabilities of dropping out.

5.3 Grade-retention model

Table 9 shows estimated coefficients for the probability of being retained (second and fifth columns) and also for the value-added model specifications that were estimated for retained children (taking the grade for the second time). The achievement test scores are not used in retention decisions, but one would expect them to be correlated with school performance. As seen in the second and fifth columns, students with higher test scores are less likely to be retained, in both primary and secondary school.

[39] In estimation, the choice set for individuals varies depending on the set of schools available to them.

[40] 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.2 for a description of the procedure, which controls for sample selection into working.

Table 7: Estimated parameters for the lower-secondary school-choice/dropout model (after grade 6)

	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%)	0.35	0.35	0.33
	(0.17)	(0.19)	(0.17)
P*Q2 (13%)	0.44	0.60	0.42
	(0.10)	(0.10)	(0.10)
P*Q3 (30%)	0.41	0.79	0.45
	(0.08)	(0.07)	(0.08)
P*Q4 (53%)	0.20	0.71	0.30
	(0.07)	(0.07)	(0.07)
Male	0.27	0.27	0.26
	(0.03)	(0.03)	(0.03)
Distance to closest general	-0.45	0.08	0.14
	(0.01)	(0.01)	(0.01)
Distance to closest telesecondary	0.05	-0.77	0.05
	(0.01)	(0.01)	(0.01)
Distance to closest technical	0.11	0.08	-0.50
	(0.01)	(0.01)	(0.01)
Number of general	0.0029	-0.0014	-0.0005
	(0.0002)	(0.0004)	(0.0002)
Number of telesecondary	-0.0032	0.0118	-0.0119
	(0.0013)	(0.0015)	(0.0014)
Number of technical	-0.0053	-0.0177	0.0037
	(0.0020)	(0.0031)	(0.0020)
Wage	-0.01	-0.03	-0.02
	(0.00)	(0.00)	(0.00)

Note: See note to Table 5.

Also, the retention probability does not appear to depend on *Prospera*-participation status or to vary by quartile. Gender is a significant determinant of retention with boys more likely to be retained than girls, a pattern widely found in developing countries (e.g., Grant and Behrman (2010)).

5.4 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. 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

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

	Grade 7	Grade 8
Lag mathematics	-0.0014	-0.0008
	(0.0002)	(0.0001)
Lag Spanish	-0.0025	-0.0037
	(0.0002)	(0.0001)
Prospera effect by propensity-score quartile		
P*Q1 (4%)	0.01	0.16
	(0.21)	(0.10)
P*Q2 (13%)	-0.36	0.02
	(0.13)	(0.06)
P*Q3 (30%)	-0.32	-0.07
	(0.10)	(0.05)
P*Q4 (53%)	-0.33	-0.12
	(0.09)	(0.05)
Male	0.28	0.15
	(0.05)	(0.02)
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)

Note: See note to Table 5

Table 9: Estimated model parameters for retained students

	Primary			Secondary		
	Prob Ret.	VA Math	VA Spanish	Prob Ret.	VA Math	VA Spanish
Lag mathematics	-0.0053	0.19	0.39	-0.0016	0.39	0.17
	(0.0001)	(0.03)	(0.03)	(0.0001)	(0.03)	(0.03)
Lag Spanish	0.0001	0.24	0.13	-0.0105	0.13	0.30
	(0.0001)	(0.03)	(0.04)	(0.0002)	(0.04)	(0.04)
Prospera effect by propensity-score quartile						
P*Q1 (4%)	0.39	0.53	0.54	-0.16	0.54	0.55
	(0.21)	(18.83)	(31.53)	(0.23)	(31.53)	(26.83)
P*Q2 (13%)	0.05	0.39	1.53	-0.14	1.53	0.99
	(0.12)	(11.00)	(15.94)	(0.14)	(15.94)	(14.37)
P*Q3 (30%)	0.01	1.91	2.81	-0.34	2.81	1.35
	(0.09)	(8.11)	(11.50)	(0.10)	(11.50)	(11.70)
P*Q4 (53%)	-0.03	2.63	3.53	-0.55	3.53	2.64
	(0.08)	(7.55)	(12.14)	(0.10)	(12.14)	(11.90)
Male	0.42	-11.62	20.27	0.81	20.27	-7.70
	(0.04)	(3.70)	(6.79)	(0.05)	(6.79)	(6.17)

Note: VA = value added. The probability of retention is estimated by a probit model. The notes from Table 5 also apply here.

variable, a right-side variable (lagged test scores) or both variables, but only for students who engage in copying. Table 10 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 9.4% in eighth grade in telesecondary schools. At the primary-school level, indigenous schools have higher copying rates. Copying distorts average test scores by 17-115 points for copiers. Our estimation approach described earlier takes into account possible test-score distortion due to copying.

Table 10: Estimated test-score distortion from copying, by school types

		Percentages	Math			Spanish		
			Raw	True	Diff	Raw	True	Diff
Grade 5	General	4.1%	573	531	42	553	526	26
	Indigenous	7.0%	543	476	67	519	469	50
	Overall	4.3%	571	527	44	550	522	28
Grade 6	General	3.8%	609	559	49	580	550	30
	Indigenous	6.1%	574	517	56	544	505	39
	Overall	3.9%	606	556	50	577	547	30
Grade 7	General	1.6%	565	507	58	527	496	32
	Telesecondary	3.9%	621	524	97	541	486	55
	Technical	2.3%	576	502	74	537	493	44
	Overall	2.4%	591	513	78	536	491	44
Grade 8	General	3.3%	615	534	81	555	514	41
	Telesecondary	9.4%	700	585	115	588	522	66
	Technical	4.0%	615	533	82	555	514	40
	Overall	5.1%	656	558	97	571	518	53
Grade 9	General	2.5%	627	553	73	532	507	26
	Telesecondary	5.4%	669	614	55	532	515	17
	Technical	3.4%	624	554	70	537	512	25
	Overall	3.5%	642	577	65	534	511	22

Note: The percentages in the third column give the percentages of students suspected of copying in either math or Spanish tests.

5.5 The importance of unobserved types

Table 11 shows the unobserved type distribution, which is allowed to vary by *Prospera*-beneficiary status. Non-*Prospera* students are more likely to be type II whereas the *Prospera* students are more likely to be type IV. Table 12 examines the effects of types in the determination of outcomes by simulating the ninth grade test score and dropout outcomes when all *Prospera* beneficiaries are assumed to be a particular type. Type I has the highest academic achievement (both in Spanish and math), followed by type II and III. Type IV has the lowest academic achievement on average. Thus, we find the low types are disproportionately more likely to be enrolled in the *Prospera* program, indicating the presence of negative selection on unobservables in addition to that accounted for by the observables. If we simulate average program impacts on 9th grade test scores assuming that the unobserved type distribution for *Prospera* beneficiaries is the same as the distribution for non-beneficiaries (column labeled "Avg P=0"), we find that the positive effects of participating in *Prospera* are understated for both math (by 9 points) and Spanish (by 6 points). It also slightly under-states the reduction in dropout out. Thus, our results suggest that selection on unobservables is an important feature of the data.

Table 11: The unobserved type distributions

	Type proportions			
	Type I	Type II	Type III	Type IV
Non-Prospera (P=0)	0.05	0.46	0.19	0.30
Prospera (P=1)	0.19	0.23	0.17	0.40

Table 12: Test-score and dropout outcomes by unobserved types

Outcomes	Type I	Type II	Type III	Type IV	Avg $P = 1$	Avg $P = 0$	Diff
Spanish	579	569	554	544	559	550	-9
Mathematics	492	486	483	475	482	476	-6
Dropout	0.234	0.218	0.234	0.225	0.224	0.225	0.001

Note: Columns "Type I" to "Type IV" display the simulated test score and dropout rate outcomes when all *Prospera* students are assumed to be a particular type. The column "Avg P=1" reports the outcomes when *Prospera* students follow their own unobserved type distribution. The column "Avg $P = 0$" reports the alternative outcomes when *Prospera* students are assigned the unobserved type distribution for non-*Prospera* students. "Diff" reports the difference between "Avg $P = 0$" and "Avg $P = 1$".

5.6 Model goodness of fit

Our model involves a fairly large number of parameters, in part because the specifications and we do not impose cross-grade parameter restrictions on the production functions. Students learn different curricula in different grades, which is why we allow the production function for achievement to vary in an unrestricted way by grade. Our sample sizes are large and most 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. We randomly split the sample into a 30% training sample for model estimation and a 70% test sample to evaluate the model's goodness-of-fit.[41] Appendix Table D1 and Appendix Table D2 provide evidence on the model's goodness-of-fit

[41] The estimates based on the 30% subsample are only used for this model-validation purpose. Otherwise, parameter estimation is based on the full sample, which is more efficient.

based on the 70% test sample. Table D1 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 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 reverse in mathematics.

Appendix Table D2 shows the model fit to the school-type distribution. The predicted proportions do not differ from the data by more than 0.04. 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 Assessing cumulative *Prospera*-program effects

6.1 The 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 13, 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.[42]

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.[43] We estimate positive benefits of being a *Prospera* beneficiary in lower-secondary grades but essentially find little effect in grades 5 and 6. In

[42]The simulation keeps the distribution of unobserved types for *Prospera* beneficiaries.

[43]As we show in Figure 5 and Figure 6, absent children are disproportionally from most-disadvantaged family backgrounds and therefore tend to have larger program gains on average.

Table 13: 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	500	500	-0.1	2.3	495	496	-1.3	2.1	-	-	-
Grade 6	535	536	-0.9	2.4	523	525	-2.2	2.1	-	-	-
Grade 7	500	491	9.1	3.4	473	470	3.0	2.9	0.120	0.166	-0.046
Grade 8	538	528	10.5	3.3	488	483	4.9	2.7	0.134	0.183	-0.049
Grade 9	570	557	13.3	3.7	490	487	3.7	3.0	0.236	0.284	-0.048

Note: We report the test scores for children/youth who would attend school both with and without *Prospera*. Standard errors are 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. 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.

lower-secondary school, the cumulative *Prospera* impact in mathematics increases with the grade level and reaches a high of 0.13 standard deviations by grade 9. In Spanish, the cumulative gains are much smaller - about 0.04 standard deviations.[44]

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.[45] 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.

The last three columns of Table 13 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 5 and Figure 6. Figure 5 shows the disaggregated effects of *Prospera* participation on test scores by propensity-scores quartiles. As described in session 3.4, the propensity score is a summary statistic of student's family background, with quartile 1 denoting the most-advantaged families and quartile 4 denoting the most-disadvantaged families. 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,

[44] 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.

[45] in the fall semester of 2008 the transfers ranged from 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).

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 scores in mathematics by 0.2 standard deviations and their test scores in Spanish by 0.07 standard deviations, both are statistically significant. Although 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 6 shows the cumulative *Prospera* impacts on schooling attainment and on dropouts by grade 9, by propensity-score quartiles. The largest impacts are estimated for quartile four, comprised of students from the most-disadvantaged families. In Appendix Table D3, we show the cumulative *Prospera*-program impacts on test scores and on dropouts by gender, which shows very similar impacts for girls and boys.

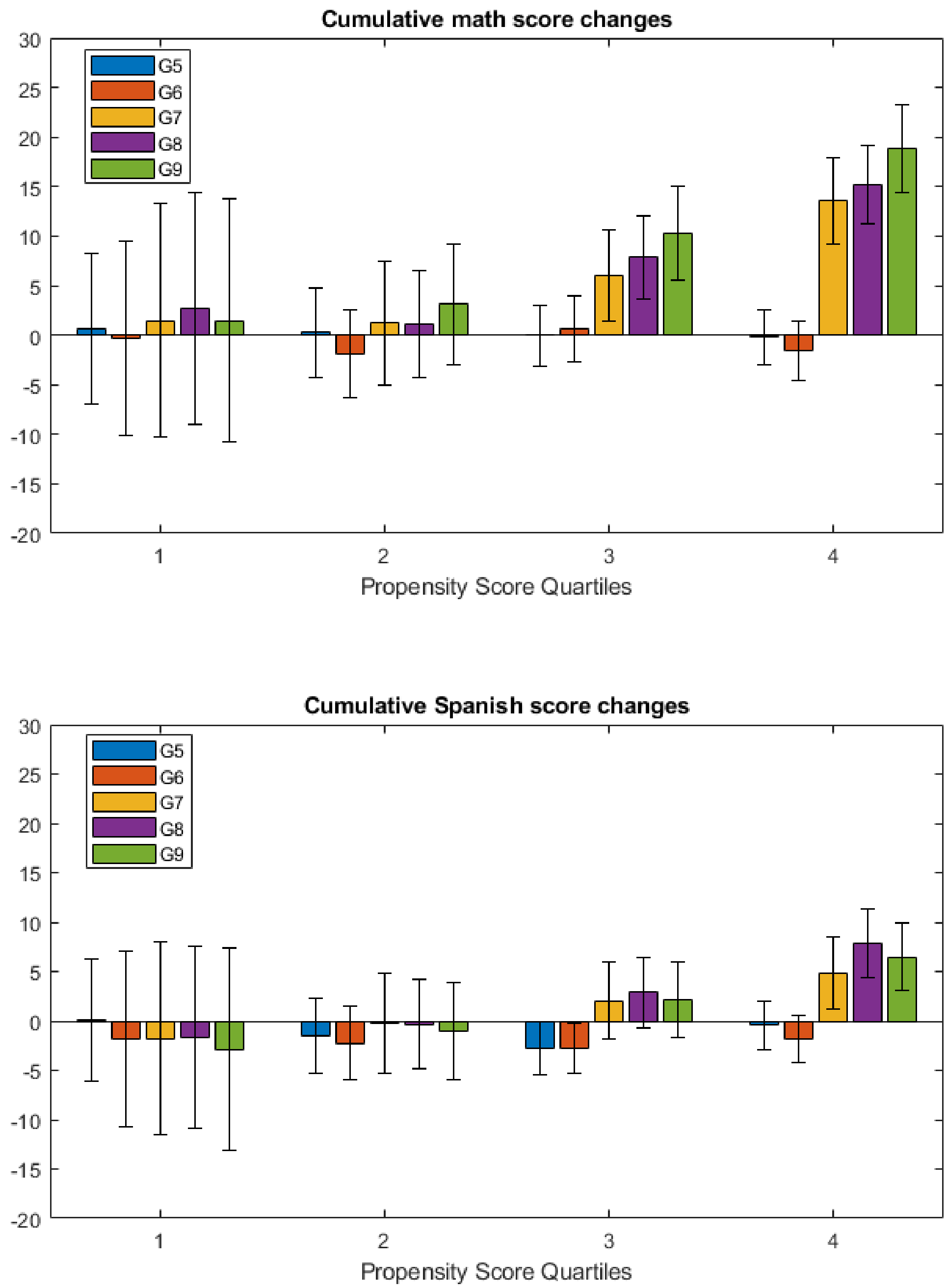
6.2 How do program effects vary by school type?

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 this type of school. Figure B2 indicated that students attending telesecondary schools show greater test-score improvement than students attending other school types. However, such comparisons are not necessarily causal because of the compositional differences in the students attending the different school types.

To examine how the effect of *Prospera* participation varies by school type, we simulate schooling outcomes and academic-achievement outcomes with and without participating in *Prospera*, restricting all the beneficiary students to attend one type of school. We repeat this exercise for the three different types of lower-secondary schools (general/telesecondary/technical). The results are reported in table 14. After controlling for compositional differences, we find that the program effects for telesecondary schools are very close to the estimated effects for the other two types of schools. This means the substantial differences observed in figure B2 across the types of schools are mainly due to telesecondary schools admitting a more disadvantaged student population. When further examining the program effects across different propensity-scores quartiles, we find students from quartile 3 benefit much more from *Prospera* by choosing the telesecondary option rather than the other two types of schools. The program increases their math score by 0.10 standard deviations in telesecondary schools but only 0.05 (0.07) standard deviations in general (technical) schools. For Spanish, the program effect is significantly positive for quartile 3 students in telesecondary school but is insignificantly different from zero in the other two types of schools.

We next evaluate the importance of telesecondary school as a determinant of *Prospera* impacts on school enrollment. In particular, we use the estimated model to simulate what educational outcomes

Figure 5: *Prospera* academic achievement effects by propensity-score quartiles



Note: Confidence intervals are obtained by a parametric bootstrap with 500 replications.

Figure 6: *Prospera* effects on schooling grade attainment and dropout by propensity-score quartiles

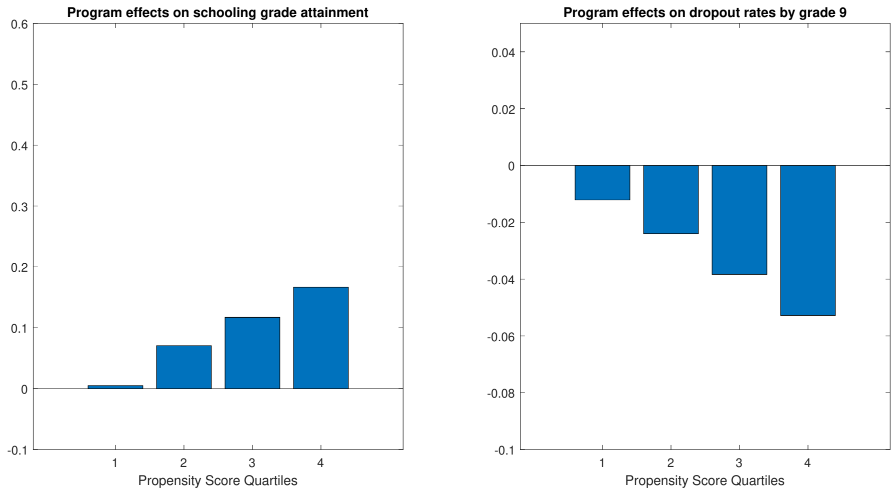


Table 14: Comparing program effects at grade 9 across different school types

	Assuming all schools to be one type		
	General	Telesecondary	Technical
Mathematics at grade 9			
Q1 (4%)	0.2	0.6	3.4
Q2 (13%)	1.1	3.3	3.9
Q3 (30%)	5.0	10.4	6.8
Q4 (53%)	16.3	15.7	15.5
Average	10.0	11.8	10.7
Spanish at grade 9			
Q1 (4%)	-3.4	-3.8	-1.8
Q2 (13%)	-2.4	-0.3	0.3
Q3 (30%)	-0.9	3.8	1.4
Q4 (53%)	5.6	5.2	7.0
Average	2.1	3.6	4.0

Note: The table reports the program effect on both math and Spanish test scores by grade 9 across school types and propensity score quantiles when assuming all lower-secondary schools to be one type.

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 15 shows the distribution of local lower-secondary school types in the data (baseline) and after removing telesecondary schools. 6.9% of students live in areas where the telesecondary schools are the only options locally available. Table 16 shows the simulated dropout proportion (by grade 9) for current *Prospera* telesecondary enrollees (at grade 7) when the telesecondary schools are removed from choice sets. The dropout proportion increases dramatically from 0.13 to 0.54. Average educational attainment over the six years of our observation period (up to grade 9 for the children who were not retained) falls from 8.9 grades to 7.5 grades. Despite using different data sources and approaches, our results are in line with evidence on the importance of telesecondary schools reported in Navarro-Sola (2019). Using a difference-in-difference approach and Employment and Occupation National Survey (EONS) dataset, she showed that the construction of an additional telesecondary per 50 children would encourage 10 individuals to enroll in lower-secondary education, causing an average increase of one additional grade of education among individuals that could have attended it.

Table 15: The school-choice distribution with and without the telesecondary option

	Baseline	No telesecondary
General, technical and telesecondary	0.699	N/A
General and telesecondary	0.070	N/A
General and technical	0.061	0.760
Telesecondary and technical	0.078	N/A
Only general	0.012	0.090
Only telesecondary	0.069	N/A
Only technical	0.008	0.078
No local schools	N/A	0.072

Table 16: Simulated dropout and educational attainment for *Prospera* telesecondary enrollees when the telesecondary option is removed

	With telesecondary	Without telesecondary
Dropout rate	0.13	0.54
Grades attained	8.85	7.53

Note: The simulation is based on the *Prospera* beneficiaries who are currently enrolled in telesecondary school at grade 7. The outcomes are measured by grade 9.

6.3 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 often 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 and that parents can select from among available schools.[46]

To explore the importance of controlling for multiple sources of selection, we compare our baseline results with results obtained from a simpler value-added model that we estimate 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 3.3, in this regression the contemporaneous *Prospera* effect δ_2^{mg} is homogeneous across school types and we do not model school choice. Also, this model does not include permanent unobserved heterogeneity (types). The cumulative program effect can be calculated as:

$$\Delta_g^m = \begin{cases} \delta_2^{m5} & \text{if } g = 5 \\ \delta_2^{mg} + \delta_1^{g}\Delta_{g-1} & \text{if } g > 5 \end{cases}$$

where Δ_g^m represents the cumulative effect for subject m in grade g, $\Delta_{g-1} = [\Delta_{g-1}^1, \Delta_{g-1}^2]$ is a 2×1 vector of cumulative effects for both math and Spanish in grade $g-1$. Table 17 compares our baseline

Table 17: Cumulative program impacts

	Math score		Spanish score	
	(1) Baseline	(2) Simple VA	(3) Baseline	(4) Simple VA
Grade 5	-0.1	-0.1	-1.3	-2.3
Grade 6	-0.9	-1.9	-2.2	-4.6
Grade 7	9.1	5.3	3.0	-1.0
Grade 8	10.5	6.3	4.9	0.2
Grade 9	13.3	8.7	3.7	-0.5

Note: Simple VA = simple value added model.

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

[46] Cameron and Heckman (2001) consider the problem of selection in modeling grade progression in US high schools, but they do not analyze test-score data.

students at the margin of dropping-out to stay in school longer. Also, 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. Lastly, the simpler model also ignores the negative selection on unobserved types, which also leads to the under-estimation on the program impacts. In summary, we find that the richer modeling framework is required to capture heterogeneous program impacts and to control for sources of selection bias to avoid the underestimation on the program impacts.

7 Exploring school-quality and student-engagement differences

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. The analysis reveals that school types do differ on average in observable quality dimensions, such as class size and teacher qualifications, and also that students report different levels of engagement, particularly in telesecondary schools.

7.1 School characteristics

The first two columns of Table 18 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 18 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.[47] Thus, we

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

Table 18: Mean lower-secondary school characteristics by school type

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

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

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 that 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.

7.2 Student effort and engagement

In Table 19, we examine whether 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).[48] 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

schools.

[48]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 19: 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.

substantial differences by gender in study time. Being female is associated with greater study effort and paying attention more frequently. However, girls report lower rates of class participation.

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 comprehensive 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 modeling framework goes beyond the previous literature by integrating value-added models for academic achievement, school-choice models that include the dropout option, incorporating local-labor-market work opportunities, grade retention, and measurement equations to allow for missing test score data or measurement error arising from copying. 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.

Our key results include the following. 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 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. We also find that the *Prospera* program decreases school dropouts by 4 percentage points, at the sixth to seventh grade transition. Second, there is a consistent pattern of larger impacts for the most-disadvantaged children/youth, those in higher *Prospera*-participation propensity-score quartiles.

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 gains accruing over multiple years. Fourth, we also find evidence of gender differences on test scores, with boys scoring higher on average on mathematics tests than girls and girls scoring higher on average on Spanish tests. The gender gaps in mathematics widen from primary to secondary grades.

Fifth, 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 distortions (one-sided measurement error) arising from copying, which can affect both the dependent and right-side (lagged) variables in the value-added model.

Sixth, our analysis shows that telesecondary schools play an important role in the *Prospera*-program effect. Even with the copying adjustment, we find that telesecondary schools are at least similar in effectiveness and, in some case, more effective than the regular public school in teaching mathematics and Spanish. 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 13% to 54% and educational attainment decreases by 1.4 grades. Thus, telesecondary schools are important to *Prospera*'s success in improving educational outcomes for a relatively disadvantaged student population.

Seventh, 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 critical roles. They also illustrate how our extended approach integrating value-added and school-choice models with multiple unobserved types can lead to estimates that in some respects, including program impacts, are substantially different than those obtained from simpler approaches.

References

Allende, C., Gallego, F., Neilson, C., et al. (2019). Approximating the equilibrium effects of informed school choice. Technical report.

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.

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. *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. *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.

Avitabile, C. and De Hoyos, R. (2018). The heterogeneous effect of information on student performance: evidence from a randomized control trial in Mexico. *Jour. Development Economics*, 135:318–348.

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. *InterAmerican Development Bank*.

Behrman, J. R., Hrubec, Z., Taubman, P., and Wales, T. J. (1980). *Socioeconomic Success: A Study of the Effects of Genetic Endowments, Family Environment, and Schooling*, volume 128. Amsterdam; New York: North-Holland Publishing Company.

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. and Rosenzweig, M. R. (1999). "Ability" biases in schooling returns and twins: a test and new estimates. *Economics of Education Review*, 18(2):159–167.

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.

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.

De Hoyos, R., Garcia-Moreno, V. A., and Patrinos, H. A. (2017). The impact of an accountability intervention with diagnostic feedback: Evidence from Mexico. *Economics of Education Review*, 58:123–140.

De Hoyos Navarro, R. E., Attanasio, O., and Meghir, C. (2019). Can scholarships increase high school graduation rates? evidence from a randomized control trial in Mexico. *Evidence from a Randomized Control Trial in Mexico (April 22, 2019). World Bank Policy Research Working Paper*, (8826).

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.

Fox, J. T., il Kim, K., Ryan, S. P., and Bajari, P. (2012). The random coefficients logit model is identified. *Journal of Econometrics*, 166(2):204–212.

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.

Hastings, J. S., Neilson, C. A., and Zimmerman, S. D. (2012). The effect of school choice on intrinsic motivation and academic outcomes. Technical report, National Bureau of Economic Research.

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. *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.

Heckman, J. J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52:271–320.

Heckman, J. J. and Vytlacil, E. (2001). Policy-relevant treatment effects. *American Economic Review*, 91(2):107–111.

Heckman, J. J. and Vytlacil, E. (2005). Structural equations, treatment effects, and econometric policy evaluation 1. *Econometrica*, 73(3):669–738.

Heckman, J. J. and Vytlacil, E. J. (2007). Econometric evaluation of social programs, part i: Causal models, structural models and econometric policy evaluation. *Handbook of Econometrics*, 6:4779–4874.

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.

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

Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education Finance and Policy*, 4(4):537–571.

Rouse, C. E. (1998). Private school vouchers and student achievement: An evaluation of the milwaukee parental choice program. *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.

Schellenberg, J. and Walters, C. R. (2020). Do parents value school effectiveness? *American Economic Review*, 110(5):1502–39.

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

SEP (2010). Enlace: Manual tecnico 2010. Technical report, Working Paper.

Skoufias, E. and Parker, S. W. (2001). Conditional cash transfers and their impact on child work and schooling: Evidence from the progresa program in mexico. *Economia*, 2(1):45–96.

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? *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. *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.

A 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.1 Sample restrictions

Our initial sample includes individuals who fill in surveys in year 2008 and collect their *Prospera* status in 2010, with a total of 216,575 individuals. We then:

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 214,268.

2. Drop individuals whose lower-secondary school types are not coded as general, telesecondary or technical types. This leaves a sample size of 213,914.

3. Drop students who attend their primary school and secondary school in different states. This leaves a sample size of 211,328.

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 207,252.

5. Drop the observations whose primary school names are missing. This restriction leaves 206,406.

6. Drop the observations for which the distance information is missing. The remaining sample size is 200,308.

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 remaining sample size is 159,442.

8. Drop individuals whose propensity score is in the bottom 1% (pscore<0.031, trimming as described in the text). The remaining sample size is 105,256. (We drop 383 *Prospera* students and 50,798 non-Prospera students with this restriction.)

Our final sample has 108,261 different individuals and 617,919 individual-period observations.

A.2 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 2010 Mexican 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 schooling 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 A1. 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 ≥ 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).[49] 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.

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

Table A1: 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 A2. We use the regression estimates to impute hourly wage offers to the children/youth in our analysis sample, based on their observed characteristics.

Table A2: 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 school (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 4 times hours worked in the last week. 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.3 *Prospera*-beneficiary propensity-score model

We estimate a probit model for the probability that a child/youth comes from a *Prospera*-beneficiary family. The precise eligibility criteria are not made public and they vary somewhat by geographic region. Eligibility is not income-based but is rather based on housing characteristics and demographics that are highly correlated with poverty. From the program administers, we ascertained that the following characteristics are often used in determining eligibility: measures of overcrowding in the household, a demographic dependency index, sex of the household head, whether the household has access to social security (IMSS), number of children age 0-11 years, education of the household head, whether the house has a bathroom with water, type of floor in the home, whether the home has a gas stove, whether the house has a refrigerator, whether the house has a washing machine, whether the family has a vehicle and an indicator of rural/urban/semi-urban status. To estimate a probit model for the probability of being a *Prospera* beneficiary, we use variables from the context surveys that are equal to or are close proxies for the above known eligibility determinants. The probit model was estimated using 216,747 observations and the percentage correctly classified under the model is 90%.

Table A3 shows the estimated coefficients from the model. The child being female makes it more likely that the family participates in *Prospera*. Also, having more siblings or a larger household increases the probability of being a beneficiary. Having higher income (>1500 pesos) 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; but if the dad works more than 8 hours per day, the family is more likely to participate. A higher ratio of number of persons to number of rooms in the home increases the probability of being a participant. Car ownership, having drainage connected to a public system, having a refrigerator, owning a clothes washer, having garbage collection service, having internet, having a tv, and having sanitary facilities in the home makes all make it less likely that the family participates in *Prospera*. Owning the home, having a dirt floor, having electricity (which is nearly universal), cooking and sleeping in the same room, and speaking an indigenous language, ceteris paribus, increases the participation probability. We allowed for item non-response by including indicator variables in the probit regression, but we excluded any observations for which both mother's and father's education information was missing. Thus, individuals were required to have parental survey information, in which the parental education information was gathered.

Table A3: *Prospera*-participant propensity-score model (probit)

Coefficient	Estimate	Std. Error
intercept	0.642	0.037
female	-0.016	0.008
one sibling	0.086	0.021
2-3 siblings	0.324	0.020
4-5 siblings	0.515	0.021
6+ siblings	0.675	0.023
num siblings missing	0.503	0.034
father only present	0.029	0.024
mother only present	0.010	0.011
number total at home	0.013	0.002
num home missing.	0.108	0.036
family monthly income between 1500 and 29999	-0.165	0.009
family monthly income between 3000 and 7499	-0.385	0.011
family monthly income between 7500 and 14999	-0.517	0.021
family monthly income between 15000 and 30000	-0.470	0.037
family monthly income >30000	-0.103	0.037
famincmiss	-0.115	0.023
mom completed primary	-0.065	0.010
mom completed secondary	-0.072	0.011
mom bachalaureate or tech	-0.405	0.015
mom BA or more	-0.635	0.031
dad completed primary	-0.139	0.010
dad completed secondary	-0.282	0.011
dad bachalaureate or tech	-0.547	0.015
dad BA or more	-0.698	0.023
mom ed miss	-0.125	0.047
dad ed miss	-0.286	0.027
mom works 4+ hours daily	-0.258	0.009
mom works <4 hours or retired/looking for work	-0.199	0.017
dad works <8 hours daily	0.112	0.012
dad does not work or is retired	-0.005	0.013
mom work info missing	-0.178	0.036
dad work info missing	0.023	0.021
ratio of number in home/number of rooms	0.027	0.003
missing ratio	0.053	0.040
whether own home	0.177	0.010
missing home own info	0.173	0.030
whether have a car or truck	-0.264	0.009
cartruckmiss	-0.030	0.027
house has dirt floor	0.112	0.010
dirtfloormiss	-0.071	0.024
house drainage connected to public	-0.519	0.009
drainagemiss	-0.229	0.026
sanitary fac in home	-0.127	0.013
sanitarymiss	-0.113	0.028
electric power in the home	0.102	0.022
electric power miss	0.032	0.031
house has a refridgerator	-0.232	0.011
refridgerator miss	-0.142	0.027
cook and sleep in the same room	0.102	0.012
cooksleepmiss	0.086	0.035
house has a clothes washer	-0.147	0.010
clotheswashmiss	-0.037	0.029
speak indigenous language	0.343	0.014
indiglangmiss	0.045	0.028
covered by IMSS social security	-0.138	0.009
IMSS info missing	0.069	0.011
household has a stove	-0.283	0.015
stove info missing	-0.163	0.024
household has internet	-0.272	0.016
internet missing	0.055	0.026
household has tv	-0.042	0.020
tv missing	-0.208	0.029
household has garbage collection	-0.111	0.009
garbage collection missing	-0.028	0.028

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%.

B Additional descriptive statistics

Table B1 provides information on the supply of local schools of different types at the individual level. The last column indicates the proportion of students for which that type of school is not an available option. See further discussion in the text.

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 considered to be different schools. Figure B1 shows 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.

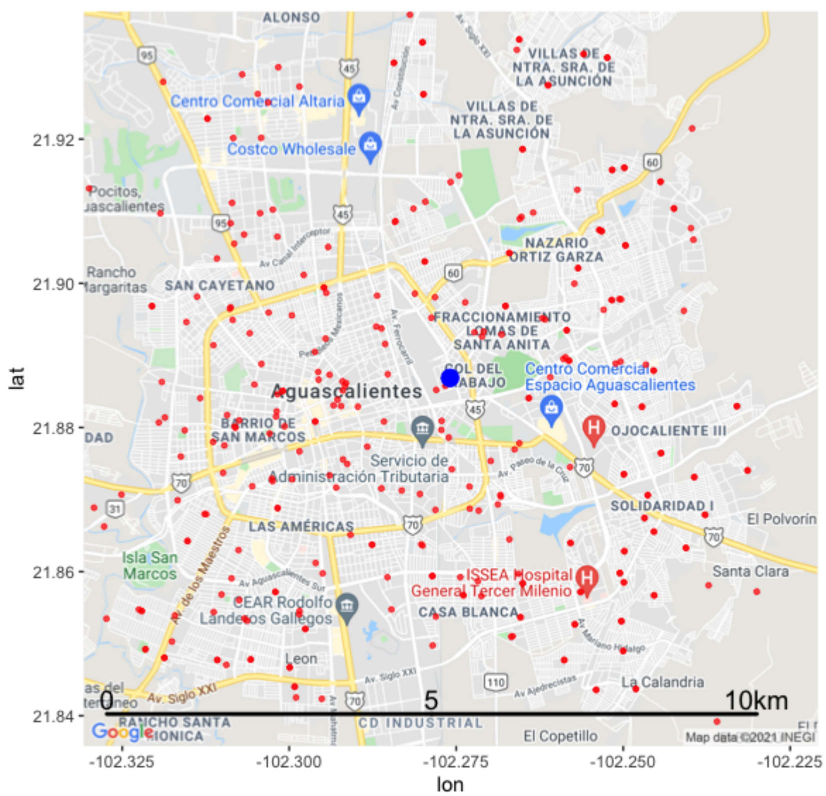
Table B1: Number of local schools of different types

	Mean	Std	p10	p50	p75	Not available
Primary school (within 5 km)						
General	65	80	9	26	104	3.4%
Indigenous	5	8	1	3	6	76.9%
Secondary school (within 10 km)						
General	48	69	4	18	63	14.6%
Telesecondary	13	12	4	10	20	8.4%
Technical	10	12	2	6	16	15.9%

Note: Columns 3-5 report selected percentiles. The last column gives the percentages of individuals for whom a given school type are not locally available.

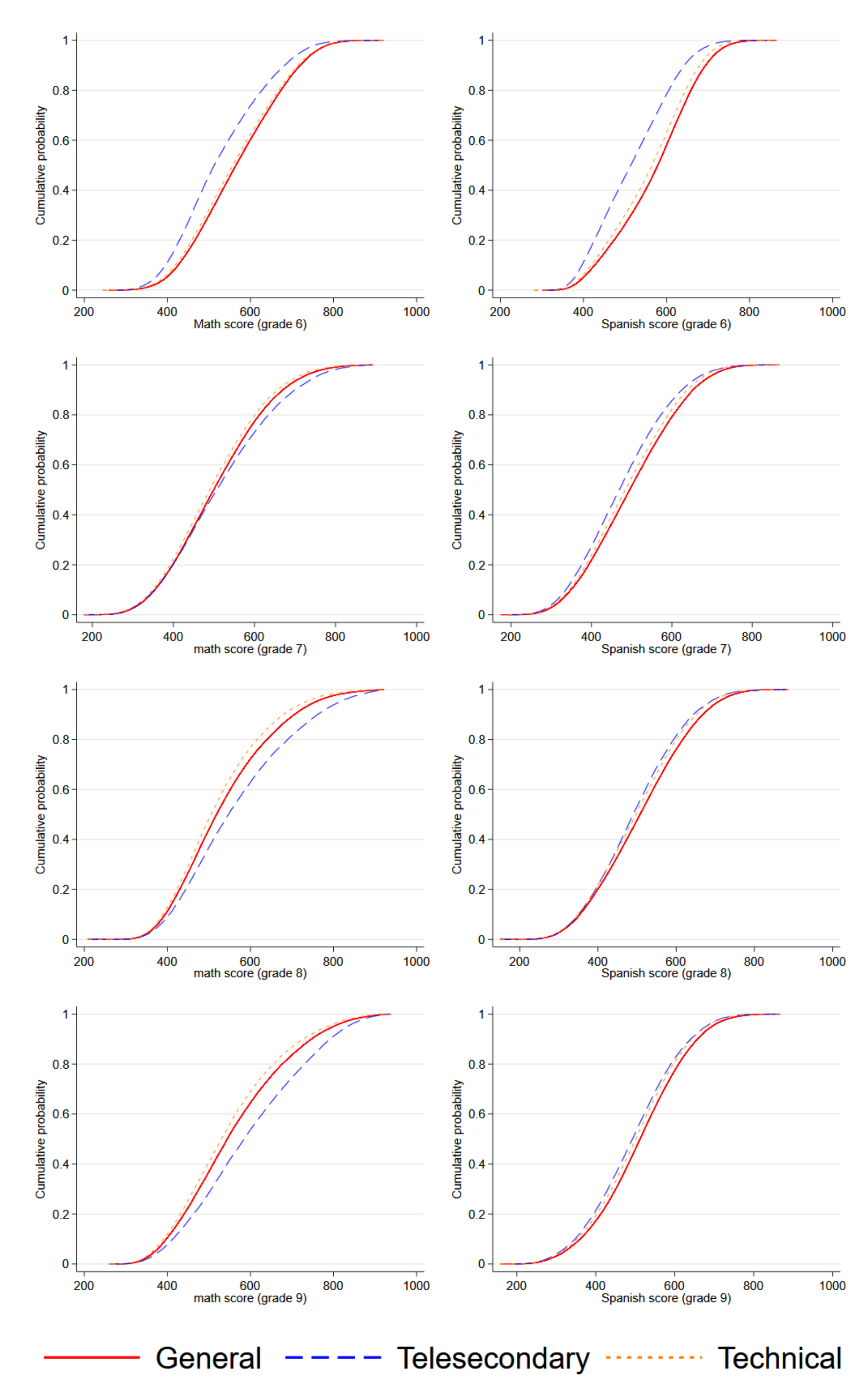
Figure B2 shows the test-score distributions by grade and lower-secondary school types. The left column displays the mathematics test-score distributions and the right column the Spanish test-score distributions. Each row shows test scores for a different grade, from 6 to 9. There is a larger variance 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 following year. The gaps in mathematics scores start to reverse at grade 7 and the final score is substantially higher for *Prospera* students by grade 9. The general school type has the highest Spanish scores, but the advantage shrinks over grades. These graphs suggest that students in telesecondary schools show greater test score improvements 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 and because of differential dropout rates across grades and by school types.

Figure B1: Local primary school sessions around Aguascalientes



Aguascalientes is a city with about 1.4 million inhavitants in 2020 in central Mexico. It contains 316 school sessions with 250 unique coordinates, indicating multiple school sessions share the same physical teaching place.

Figure B2: Mathematics and Spanish test scores distributions (CDFs) by grade and lower-secondary school types



C Model estimates

We show the full model estimates in this section. Tables C1 and C2 show the estimated parameters of the value-added production functions for mathematics and Spanish. Table C3 reports estimated coefficients from primary and lower-secondary school choices. Table C4 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). Table C5 reports estimated coefficients from the dropout decision during lower-secondary school choices, where the omitted category is dropping out and working. Lastly, Table C6 shows the coefficients associated with the cheating equation. See further discussion in the text.

Table C1: Mathematics value-added estimates

	General		Indigenous		General			Telesecondary			Technical		
	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Lag Mathematics	0.56	0.55	0.41	0.45	0.45	0.48	0.49	0.35	0.42	0.41	0.42	0.49	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	0.16	0.13	0.11	0.12	0.27	0.17	0.25	0.26	0.16	0.25	0.27	0.16	0.27
	(0.01)	(0.00)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prospera score quartile													
P*Q1 (4%)	0.62	-0.79	-1.05	2.89	2.74	2.49	-0.74	0.33	-2.84	2.24	1.70	2.81	2.36
	(3.23)	(3.15)	(31.28)	(40.96)	(4.35)	(4.03)	(4.75)	(12.34)	(11.99)	(12.73)	(6.57)	(5.55)	(6.56)
P*Q2 (13%)	0.33	-1.83	-2.61	-3.38	2.18	-2.19	2.59	2.45	2.31	1.93	2.05	1.90	2.64
	(2.07)	(1.99)	(15.63)	(13.96)	(3.17)	(2.92)	(3.27)	(6.78)	(5.26)	(6.24)	(4.27)	(3.71)	(4.29)
P*Q3 (30%)	-0.14	1.02	2.05	2.01	3.99	-1.00	4.37	6.16	6.76	6.09	4.34	3.22	3.41
	(1.68)	(1.60)	(9.15)	(8.27)	(2.73)	(2.53)	(2.75)	(4.68)	(3.80)	(4.36)	(3.51)	(3.05)	(3.37)
P*Q4 (53%)	-0.05	-1.05	-0.11	-3.34	13.57	7.50	6.97	12.40	6.62	9.33	12.53	5.75	7.74
	(1.61)	(1.55)	(6.43)	(5.80)	(2.72)	(2.50)	(2.76)	(4.41)	(3.55)	(4.13)	(3.33)	(2.95)	(3.24)
Education cat. (dad)													
Below primary school	-4.60	-0.10	-15.93	-9.54	-5.79	-0.30	-1.48	-13.93	2.60	4.53	2.16	-6.35	-1.78
	(2.64)	(2.60)	(14.79)	(14.12)	(4.19)	(3.76)	(4.60)	(8.47)	(7.53)	(8.56)	(5.74)	(4.73)	(5.95)
Primary school completed	-4.01	2.29	-11.75	-9.41	-3.91	-0.36	-0.74	-10.30	4.02	4.55	3.68	-2.19	0.68
	(2.66)	(2.62)	(14.93)	(14.25)	(4.19)	(3.76)	(4.58)	(8.57)	(7.61)	(8.64)	(5.77)	(4.74)	(5.94)
Secondary or below	-2.64	2.78	-16.61	-7.66	-0.09	1.76	-0.31	-9.37	3.62	8.99	4.11	0.05	0.70
	(2.61)	(2.57)	(15.25)	(14.36)	(4.09)	(3.67)	(4.49)	(8.57)	(7.61)	(8.62)	(5.67)	(4.62)	(5.82)
College or above	3.53	8.78	-5.96	-7.40	6.52	7.39	3.35	-4.56	8.89	5.53	8.62	2.36	6.17
	(2.70)	(2.66)	(16.48)	(15.23)	(4.19)	(3.74)	(4.59)	(9.48)	(8.29)	(9.41)	(5.82)	(4.77)	(5.96)
Working status (dad)													
Full time	6.21	2.98	-7.77	-11.09	2.86	-1.81	2.05	7.13	3.13	10.90	-1.91	-1.98	14.19
	(2.03)	(2.03)	(10.73)	(9.47)	(3.23)	(2.91)	(3.50)	(6.46)	(5.33)	(6.33)	(4.49)	(3.86)	(4.68)
Not full time	4.91	1.77	-5.10	-13.15	3.22	-2.37	0.26	4.84	0.09	8.46	-2.41	-3.27	8.80

Table C1: Mathematics value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
	(1.96)	(1.96)	(10.45)	(9.22)	(3.10)	(2.79)	(3.39)	(6.28)	(5.17)	(6.14)	(4.35)	(3.72)	(4.50)
Father present at home	0.36	-0.65	4.19	3.71	2.92	-0.49	-0.42	0.61	3.19	1.34	-0.54	-0.66	-1.57
	(1.00)	(0.96)	(4.88)	(4.57)	(1.55)	(1.37)	(1.63)	(3.03)	(2.65)	(3.15)	(2.09)	(1.85)	(2.27)
Education cat. (mom)													
Primary school	1.99	10.48	-9.56	6.47	4.59	8.56	-2.78	11.25	-16.70	6.23	9.05	-9.43	3.01
	(4.66)	(4.26)	(21.84)	(21.90)	(8.41)	(6.39)	(8.12)	(14.90)	(11.90)	(14.64)	(10.50)	(7.61)	(9.72)
Primary school completed	3.44	12.81	-6.10	14.34	6.32	8.60	-3.67	14.24	-13.60	6.30	12.65	-13.52	4.28
	(4.66)	(4.26)	(22.07)	(22.06)	(8.40)	(6.39)	(8.10)	(14.95)	(11.93)	(14.68)	(10.51)	(7.61)	(9.75)
Secondary or below	4.85	14.29	-4.36	17.62	6.21	10.45	-3.82	15.47	-10.72	9.26	13.71	-12.55	5.24
	(4.65)	(4.25)	(22.38)	(22.27)	(8.38)	(6.35)	(8.06)	(14.98)	(11.96)	(14.69)	(10.46)	(7.58)	(9.71)
College or above	10.71	16.53	-1.54	27.17	12.87	13.94	-0.35	20.37	-5.18	13.06	18.09	-8.74	8.74
	(4.73)	(4.33)	(23.87)	(23.75)	(8.44)	(6.42)	(8.14)	(15.75)	(12.67)	(15.41)	(10.57)	(7.71)	(9.83)
Working status (mom)													
Housework	1.05	0.79	4.30	-3.45	2.71	7.67	-1.42	-13.64	4.75	4.62	13.68	8.56	0.02
	(3.36)	(3.28)	(14.89)	(11.80)	(5.53)	(4.96)	(6.08)	(9.93)	(8.75)	(10.58)	(7.58)	(6.39)	(8.42)
Part time	-0.05	0.30	7.16	-10.71	0.60	6.78	-3.12	-16.63	0.67	2.75	10.45	5.57	-5.21
	(3.43)	(3.35)	(15.37)	(12.40)	(5.61)	(5.02)	(6.15)	(10.33)	(9.05)	(10.92)	(7.69)	(6.48)	(8.52)
Full time	1.87	0.81	1.83	-14.65	1.20	6.90	-1.51	-17.82	1.74	9.56	11.16	7.06	-3.55
	(3.47)	(3.38)	(15.96)	(12.93)	(5.66)	(5.08)	(6.20)	(10.53)	(9.21)	(11.10)	(7.77)	(6.56)	(8.62)
Mother present at home	0.84	-0.18	-6.63	8.25	7.84	1.43	-1.00	-1.06	-6.07	-3.47	5.39	-0.07	-0.07
	(1.92)	(1.81)	(6.71)	(6.03)	(3.13)	(2.78)	(3.38)	(4.85)	(4.13)	(4.84)	(4.10)	(3.76)	(4.35)
Number of people at home													
4 people	1.06	0.94	4.56	9.25	0.66	0.45	3.74	3.22	0.41	2.42	0.04	0.48	1.38
	(0.90)	(0.87)	(5.72)	(5.06)	(1.36)	(1.21)	(1.43)	(2.96)	(2.48)	(2.93)	(1.85)	(1.60)	(1.91)
5 people	0.50	1.01	2.32	4.77	-0.27	2.03	2.79	0.22	-1.98	2.03	-1.70	-0.92	3.66
	(1.00)	(0.97)	(5.43)	(4.82)	(1.53)	(1.35)	(1.62)	(3.12)	(2.63)	(3.09)	(2.10)	(1.79)	(2.17)
$\geq$ 6 people	-0.33	-0.05	1.13	0.32	-2.38	-2.31	1.32	-1.40	-1.59	5.35	-3.54	-1.79	0.95

Table C1: Mathematics value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
	(0.88)	(0.85)	(4.43)	(3.92)	(1.36)	(1.22)	(1.46)	(2.64)	(2.23)	(2.63)	(1.85)	(1.62)	(1.96)
Age	-3.72	-4.84	-2.26	-3.81	-3.69	-2.78	-5.65	-6.22	-6.08	-5.43	-3.24	-2.55	-4.56
	(0.77)	(0.77)	(3.91)	(3.65)	(1.16)	(1.01)	(1.19)	(2.43)	(2.02)	(2.41)	(1.57)	(1.35)	(1.65)
Age^2	-2.31	-1.77	-2.17	-0.61	-2.94	-3.05	-2.39	-1.45	-1.13	-1.53	-3.06	-1.70	-3.42
	(0.50)	(0.45)	(2.12)	(1.88)	(0.82)	(0.73)	(0.95)	(1.45)	(1.21)	(1.67)	(1.02)	(0.95)	(1.29)
Male	2.99	2.47	1.97	-0.41	7.16	16.65	10.37	1.02	10.22	2.01	4.39	17.45	9.20
	(0.66)	(0.65)	(3.40)	(3.06)	(1.04)	(0.93)	(1.09)	(2.03)	(1.71)	(2.02)	(1.39)	(1.22)	(1.46)
First language spoken at home													
Indigenous	-6.29	-5.02	-6.32	-7.86	-9.67	-8.37	-8.50	4.49	3.10	17.43	8.96	14.75	14.37
	(2.66)	(2.54)	(4.20)	(3.68)	(5.11)	(4.50)	(5.71)	(3.81)	(3.40)	(4.35)	(3.92)	(3.28)	(4.21)
Both Spanish and indigenous	2.30	-4.34	5.18	6.51	-5.66	4.58	4.26	5.24	6.78	6.57	2.85	1.36	4.41
	(2.65)	(2.57)	(5.74)	(5.18)	(4.81)	(4.32)	(4.94)	(6.05)	(4.82)	(5.36)	(4.88)	(4.32)	(5.35)
Internet access	-5.91	-7.11	-3.02	-11.90	-7.21	-4.22	-9.56	-19.12	-1.80	-14.32	-8.24	-6.48	-4.75
	(1.05)	(1.03)	(6.81)	(5.73)	(1.57)	(1.40)	(1.65)	(3.94)	(3.17)	(4.00)	(2.16)	(1.90)	(2.22)
Computer access	1.10	4.62	-10.23	1.09	4.55	5.17	5.00	-0.90	1.01	-1.04	0.51	4.81	3.21
	(0.90)	(0.88)	(6.01)	(4.94)	(1.33)	(1.18)	(1.39)	(3.21)	(2.68)	(3.22)	(1.84)	(1.61)	(1.86)
Number of pre-school years													
1 year	5.61	1.99	-2.59	15.64	-0.60	-10.67	-1.15	-1.22	6.86	-8.66	-3.12	-1.44	-1.87
	(2.51)	(2.47)	(11.01)	(10.27)	(4.04)	(3.46)	(4.59)	(7.18)	(6.09)	(7.51)	(5.23)	(4.52)	(5.90)
2 years	5.33	3.79	9.51	13.37	1.07	-9.26	-1.07	-5.19	2.30	-1.88	-1.21	-4.12	-0.11
	(2.31)	(2.27)	(10.62)	(10.05)	(3.62)	(2.98)	(4.06)	(6.75)	(5.74)	(6.97)	(4.80)	(4.03)	(5.42)
3 years	5.71	4.01	2.05	11.78	2.59	-6.90	-0.26	2.03	4.99	0.32	4.04	0.88	2.59
	(2.29)	(2.25)	(10.53)	(9.89)	(3.56)	(2.92)	(4.00)	(6.69)	(5.68)	(6.92)	(4.74)	(3.99)	(5.37)
4 years	6.70	2.85	-2.62	7.98	0.64	-4.68	2.54	8.40	10.29	2.24	2.77	2.69	-1.06
	(2.32)	(2.28)	(10.18)	(9.67)	(3.62)	(2.99)	(4.06)	(6.67)	(5.67)	(6.91)	(4.81)	(4.04)	(5.43)
Urban dummy	-0.42	2.79	11.08	28.14	-0.44	1.19	-2.43	-1.77	-0.12	-0.06	-0.07	-0.03	0.66
	(0.83)	(0.82)	(4.52)	(3.91)	(1.56)	(1.39)	(1.63)	(2.52)	(2.17)	(2.54)	(1.84)	(1.60)	(1.88)

Table C1: Mathematics value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Regions													
North-center	2.18	2.99	-9.04	-15.85	0.25	6.58	-0.94	0.02	-6.88	0.60	1.42	7.04	3.91
	(0.97)	(0.95)	(9.64)	(9.83)	(1.43)	(1.29)	(1.58)	(4.85)	(4.04)	(4.38)	(2.02)	(1.75)	(2.17)
Center	0.51	-0.76	9.54	-2.44	-1.28	6.89	2.71	16.75	11.04	8.72	3.61	12.31	6.30
	(0.96)	(0.95)	(8.15)	(8.04)	(1.38)	(1.22)	(1.45)	(4.81)	(3.98)	(4.33)	(2.03)	(1.74)	(2.07)
South	2.20	3.90	-2.09	-6.22	-0.42	1.19	-2.44	-1.78	-0.12	-0.06	-0.07	-0.03	0.65
	(1.16)	(1.12)	(8.56)	(8.20)	(1.71)	(1.57)	(1.79)	(5.18)	(4.34)	(4.77)	(2.16)	(1.91)	(2.22)
Unobserved types													
Type I	1.52	5.29	-1.52	-3.11	9.37	0.49	1.04	5.06	4.08	4.56	14.39	10.63	17.21
	(2.69)	(2.32)	(9.14)	(7.64)	(4.10)	(4.14)	(3.99)	(6.38)	(5.45)	(6.15)	(4.40)	(4.39)	(4.35)
Type II	1.70	4.13	-1.01	-0.76	2.82	2.72	1.18	2.28	2.28	1.92	6.74	3.70	1.61
	(5.28)	(4.33)	(22.07)	(18.54)	(6.96)	(7.22)	(6.95)	(14.96)	(13.09)	(14.13)	(9.39)	(9.40)	(8.50)
Type III	-1.99	2.58	-0.91	17.11	-2.51	-3.19	-1.77	-12.23	-7.27	-3.85	-0.34	-1.92	-2.29
	(2.32)	(2.07)	(12.02)	(9.94)	(3.38)	(3.23)	(3.41)	(8.02)	(6.38)	(7.35)	(4.70)	(4.18)	(4.15)
Intercept term	136	174	251	233	67	175	159	187	253	209	69	195	142
	(6.69)	(6.38)	(31.34)	(30.57)	(11.71)	(8.93)	(11.47)	(21.55)	(17.06)	(21.65)	(15.32)	(11.11)	(15.31)
Standard error (σ)	81	82	85	83	89	85	96	114	106	116	92	85	98
	(0.23)	(0.25)	(1.29)	(1.49)	(0.46)	(0.33)	(0.38)	(0.92)	(0.69)	(0.79)	(0.64)	(0.46)	(0.65)

Table C2: 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	0.25	0.19	0.14	0.23	0.19	0.09	0.17	0.16	0.13	0.23	0.18
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lag Spanish	0.43	0.40	0.27	0.32	0.53	0.47	0.48	0.45	0.41	0.39	0.53	0.48	0.49
	(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 (4%)	0.02	-2.03	0.09	-3.72	0.06	-2.53	-2.81	-0.38	-0.72	-2.81	-2.22	0.53	-1.95
	(2.99)	(2.72)	(23.69)	(37.37)	(4.29)	(3.78)	(3.99)	(9.65)	(9.07)	(9.99)	(5.80)	(5.01)	(5.60)
P*Q2 (13%)	-1.42	-1.81	-2.86	0.13	1.34	-2.35	-1.01	1.37	0.51	-1.08	-0.16	1.75	-0.59
	(1.99)	(1.67)	(13.25)	(11.30)	(3.11)	(2.75)	(2.72)	(5.25)	(4.21)	(4.96)	(3.88)	(3.42)	(3.59)
P*Q3 (30%)	-2.77	-1.69	-2.30	-0.99	3.79	-1.59	-1.26	2.33	0.68	1.50	2.09	2.55	-1.35
	(1.63)	(1.36)	(8.01)	(6.91)	(2.73)	(2.38)	(2.34)	(3.74)	(3.02)	(3.42)	(3.29)	(2.85)	(2.83)
P*Q4 (53%)	-1.07	-1.48	0.95	-2.22	5.83	3.99	-1.60	5.42	1.89	1.03	5.76	3.21	0.89
	(1.58)	(1.30)	(5.95)	(4.75)	(2.73)	(2.44)	(2.38)	(3.53)	(2.88)	(3.22)	(3.19)	(2.84)	(2.76)
Education cat. (dad)													
Below primary school	-3.81	0.17	-11.46	3.26	-6.67	-0.67	-4.80	-12.36	-1.85	3.68	-3.18	0.06	1.56
	(2.47)	(2.20)	(14.25)	(13.66)	(3.80)	(3.50)	(3.73)	(6.75)	(5.94)	(6.91)	(5.09)	(4.48)	(4.89)
Primary school completed	-2.83	2.37	-8.09	3.70	-5.78	-0.02	-4.20	-10.14	-1.71	4.07	-3.99	-0.40	3.83
	(2.48)	(2.21)	(14.43)	(13.78)	(3.80)	(3.49)	(3.71)	(6.83)	(6.00)	(6.97)	(5.10)	(4.49)	(4.88)
Secondary or below	-0.33	3.15	-10.33	6.97	-3.97	0.84	-2.67	-8.37	-1.22	11.74	-1.57	4.36	3.68
	(2.44)	(2.17)	(14.58)	(13.89)	(3.69)	(3.41)	(3.63)	(6.83)	(6.00)	(6.96)	(4.99)	(4.37)	(4.78)
College or above	5.96	8.77	-6.57	9.32	4.37	5.83	-0.46	-2.74	1.99	10.11	6.00	10.40	5.81
	(2.53)	(2.25)	(15.54)	(14.87)	(3.78)	(3.49)	(3.71)	(7.57)	(6.50)	(7.57)	(5.13)	(4.49)	(4.88)
Working status (dad)													
Full time	4.12	-1.33	-5.29	-14.13	2.50	0.38	0.50	5.22	2.11	2.73	3.06	-5.11	2.16
	(1.94)	(1.74)	(8.99)	(8.63)	(3.06)	(2.76)	(2.95)	(5.33)	(4.28)	(5.09)	(3.91)	(3.64)	(3.68)

Table C2: Spanish value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
Not full time	3.58	-1.51	-2.45	-15.21	4.21	0.67	-0.28	4.72	2.40	2.48	1.40	-2.40	3.07
	(1.87)	(1.68)	(8.74)	(8.44)	(2.94)	(2.65)	(2.85)	(5.18)	(4.15)	(4.94)	(3.78)	(3.51)	(3.55)
Father present at home	0.14	-0.98	1.62	3.19	1.32	-1.29	0.01	1.06	0.81	-1.35	-0.31	0.00	-1.20
	(0.95)	(0.83)	(4.37)	(3.71)	(1.43)	(1.28)	(1.36)	(2.51)	(2.14)	(2.41)	(1.89)	(1.69)	(1.84)
Education cat. (mom)													
Primary school	-4.80	-0.06	3.15	-2.04	3.90	-3.69	-3.07	-0.46	-9.54	4.52	17.29	-5.77	-7.55
	(4.31)	(3.71)	(19.00)	(14.26)	(7.23)	(5.85)	(6.47)	(11.77)	(9.10)	(11.49)	(8.68)	(7.07)	(7.91)
Primary school completed	-3.04	1.21	4.27	4.58	4.10	-1.57	-2.24	0.67	-6.14	4.57	17.85	-5.86	-5.13
	(4.32)	(3.71)	(19.17)	(14.43)	(7.22)	(5.83)	(6.46)	(11.80)	(9.14)	(11.53)	(8.67)	(7.06)	(7.89)
Secondary or below	-1.38	3.95	9.18	5.48	5.87	-0.18	-1.78	2.40	-7.69	6.50	19.25	-5.64	-1.68
	(4.31)	(3.70)	(19.39)	(14.68)	(7.19)	(5.81)	(6.43)	(11.82)	(9.16)	(11.54)	(8.63)	(7.03)	(7.84)
College or above	3.68	8.76	7.99	16.17	12.78	5.38	3.36	8.26	1.43	10.51	22.83	2.65	-1.00
	(4.38)	(3.77)	(20.74)	(16.24)	(7.25)	(5.87)	(6.49)	(12.49)	(9.67)	(12.12)	(8.74)	(7.14)	(7.92)
Working status (mom)													
Housework	1.95	0.34	8.36	3.39	6.66	13.15	-1.17	-6.24	-1.82	5.19	13.36	1.45	-0.14
	(3.29)	(2.84)	(12.92)	(10.91)	(5.40)	(4.72)	(4.87)	(8.25)	(6.70)	(8.55)	(6.92)	(5.78)	(6.19)
Part time	0.04	0.05	8.38	-5.77	6.32	12.15	-1.63	-3.64	-6.96	3.52	10.71	0.95	0.82
	(3.35)	(2.89)	(13.43)	(11.35)	(5.46)	(4.77)	(4.93)	(8.57)	(6.95)	(8.84)	(7.01)	(5.86)	(6.29)
Full time	2.79	0.22	-2.01	-3.58	3.29	13.04	-1.26	-7.89	-6.55	4.45	9.43	0.05	-0.20
	(3.38)	(2.92)	(13.89)	(11.70)	(5.51)	(4.82)	(4.97)	(8.69)	(7.06)	(8.92)	(7.08)	(5.93)	(6.35)
Mother present at home	2.20	3.67	-7.95	6.46	10.14	3.02	4.88	-0.89	0.14	3.81	4.33	1.20	4.30
	(1.85)	(1.56)	(6.11)	(5.30)	(3.05)	(2.69)	(2.84)	(3.98)	(3.45)	(3.84)	(3.78)	(3.44)	(3.65)
Number of people at home													
4 people	-0.78	-0.90	4.12	9.51	-0.39	-1.94	0.97	1.83	2.43	3.19	-1.49	1.75	-0.03
	(0.85)	(0.75)	(5.13)	(4.38)	(1.28)	(1.12)	(1.18)	(2.38)	(1.97)	(2.29)	(1.66)	(1.46)	(1.56)
5 people	-1.11	-2.08	2.18	3.10	-0.60	-1.98	-0.25	1.09	-0.61	0.59	-1.73	-2.19	1.09
	(0.95)	(0.82)	(4.88)	(4.15)	(1.46)	(1.27)	(1.35)	(2.53)	(2.11)	(2.40)	(1.90)	(1.66)	(1.79)

Table C2: Spanish value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
$\geq$ 6 people	-2.88	-3.18	0.75	0.81	-4.91	-4.65	-2.14	-0.01	1.67	-0.57	-6.05	-3.40	-2.90
	(0.83)	(0.73)	(4.01)	(3.39)	(1.28)	(1.13)	(1.20)	(2.16)	(1.80)	(2.07)	(1.68)	(1.49)	(1.59)
Age	-2.51	-3.46	-1.72	-1.41	-1.98	-2.73	-6.67	-5.82	-4.37	-5.14	-1.98	-0.50	-5.47
	(0.72)	(0.65)	(3.54)	(2.94)	(1.09)	(0.94)	(0.98)	(1.91)	(1.62)	(1.82)	(1.47)	(1.26)	(1.35)
Age^2	-2.61	-2.28	-1.29	-1.58	-2.94	-3.45	-3.99	-0.65	-1.09	-2.22	-2.48	-4.35	-5.92
	(0.47)	(0.38)	(1.99)	(1.48)	(0.77)	(0.69)	(0.79)	(1.15)	(0.98)	(1.28)	(1.00)	(0.85)	(1.06)
Male	-16.70	-15.51	-8.74	-12.29	-26.17	-16.97	-12.87	-24.89	-18.66	-15.26	-27.15	-16.74	-13.64
	(0.62)	(0.56)	(3.05)	(2.61)	(0.98)	(0.86)	(0.91)	(1.64)	(1.37)	(1.57)	(1.25)	(1.12)	(1.19)
First language spoken at home													
Indigenous	-11.99	-8.66	-9.04	-13.12	0.74	-10.75	-12.70	-5.67	-7.90	-6.33	7.02	6.88	-17.95
	(2.65)	(2.24)	(3.82)	(3.14)	(5.12)	(4.42)	(5.47)	(3.20)	(2.82)	(3.30)	(4.11)	(3.49)	(3.66)
Both Spanish and indigenous	-0.86	-6.85	7.27	3.14	-4.11	-0.57	-3.74	3.53	-4.37	-0.29	7.89	3.81	0.95
	(2.54)	(2.23)	(5.14)	(4.37)	(4.68)	(4.13)	(4.32)	(4.87)	(3.90)	(4.40)	(4.77)	(4.27)	(4.02)
Internet access	-4.84	-5.10	0.62	-9.21	-4.17	-3.87	-6.64	-15.42	-9.26	-11.85	-4.93	-7.67	-4.98
	(0.99)	(0.87)	(5.94)	(5.10)	(1.48)	(1.27)	(1.36)	(3.14)	(2.60)	(3.00)	(1.95)	(1.71)	(1.85)
Computer access	1.62	3.19	-5.57	-6.05	3.34	1.97	4.31	-4.10	-0.82	-2.13	1.77	2.42	1.98
	(0.85)	(0.74)	(5.39)	(4.46)	(1.25)	(1.09)	(1.17)	(2.60)	(2.16)	(2.46)	(1.65)	(1.46)	(1.55)
Number of pre-school years													
1 year	3.94	-3.41	-0.51	4.72	0.55	-4.31	-6.51	-4.79	5.61	0.11	1.45	-3.04	2.42
	(2.39)	(2.08)	(10.62)	(8.77)	(3.94)	(3.43)	(3.67)	(5.73)	(4.82)	(5.73)	(4.94)	(4.43)	(4.81)
2 years	4.87	-0.13	1.93	8.20	0.50	-2.00	-0.54	-7.52	5.33	5.22	8.30	-1.11	5.07
	(2.18)	(1.89)	(10.27)	(8.51)	(3.48)	(2.98)	(3.24)	(5.36)	(4.50)	(5.32)	(4.47)	(3.92)	(4.31)
3 years	4.80	0.83	1.69	7.71	2.54	-0.47	-1.07	-3.00	7.36	5.58	11.53	1.06	6.39
	(2.16)	(1.87)	(10.22)	(8.40)	(3.43)	(2.94)	(3.18)	(5.31)	(4.44)	(5.27)	(4.43)	(3.88)	(4.26)
4 years	6.18	0.44	-2.24	6.48	1.93	1.41	0.36	2.23	6.12	6.27	9.46	-0.05	6.30
	(2.19)	(1.90)	(9.93)	(8.19)	(3.49)	(3.00)	(3.24)	(5.30)	(4.44)	(5.27)	(4.50)	(3.95)	(4.32)
Urban dummy	2.06	5.17	10.72	20.55	1.72	1.67	0.49	0.89	1.21	-2.90	-1.23	-1.34	-2.64

Table C2: Spanish value-added estimates

	G5	G6	G5	G6	G7	G8	G9	G7	G8	G9	G7	G8	G9
	(0.79)	(0.70)	(3.94)	(3.28)	(1.54)	(1.38)	(1.39)	(2.03)	(1.70)	(1.97)	(1.71)	(1.52)	(1.59)
Regions													
North-center	1.40	-2.79	-9.79	3.73	5.36	5.90	3.93	0.78	1.09	0.75	-0.88	4.58	7.08
	(0.90)	(0.81)	(9.13)	(8.06)	(1.35)	(1.18)	(1.29)	(3.78)	(3.12)	(3.51)	(1.76)	(1.55)	(1.73)
Center	0.93	-1.29	9.52	11.58	6.45	8.02	5.73	17.27	12.67	11.45	5.25	5.15	8.18
	(0.89)	(0.81)	(7.54)	(6.92)	(1.30)	(1.12)	(1.21)	(3.76)	(3.08)	(3.46)	(1.77)	(1.55)	(1.69)
South	-0.42	-2.14	-13.14	-9.00	1.72	1.67	0.50	0.90	1.21	-2.89	-1.23	-1.34	-2.65
	(1.11)	(0.97)	(7.96)	(7.14)	(1.68)	(1.48)	(1.52)	(4.11)	(3.42)	(3.80)	(1.98)	(1.80)	(1.88)
Unobserved types													
Type I	0.68	-1.72	-2.44	-5.39	11.37	-1.51	1.52	4.69	4.45	1.73	8.01	1.90	4.39
	(2.65)	(2.01)	(8.54)	(6.63)	(4.20)	(3.96)	(3.53)	(5.23)	(4.30)	(4.74)	(4.49)	(4.44)	(3.83)
Type II	0.77	0.25	-2.56	-3.04	1.55	-0.18	-1.19	2.87	3.70	-1.01	4.90	-2.34	-1.80
	(5.33)	(3.75)	(21.79)	(15.45)	(7.00)	(6.57)	(6.01)	(12.65)	(10.18)	(10.66)	(9.43)	(8.91)	(7.38)
Type III	0.28	-2.56	1.17	16.27	0.65	-1.67	0.90	-3.32	-0.77	0.89	0.02	-1.53	1.97
	(2.28)	(1.75)	(11.47)	(8.82)	(3.31)	(2.97)	(2.82)	(6.42)	(5.09)	(5.72)	(4.61)	(3.93)	(3.73)
Intercept term	181	232	257	256	94	133	160	209	215	191	88	156	151
	(5.90)	(4.98)	(26.13)	(22.23)	(10.35)	(8.83)	(9.07)	(16.08)	(13.06)	(16.22)	(13.39)	(10.87)	(11.71)
Standard error (σ)	76	70	77	71	84	79	80	93	85	90	85	80	82
	(0.22)	(0.21)	(1.14)	(1.40)	(0.54)	(0.35)	(0.33)	(0.66)	(0.53)	(0.57)	(0.61)	(0.42)	(0.49)

Table C3: Primary and lower-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 (4%)	-0.21	0.35	0.35	0.33
	(0.38)	(0.17)	(0.19)	(0.17)
P*Q2 (13%)	-0.24	0.44	0.60	0.42
	(0.20)	(0.10)	(0.10)	(0.10)
P*Q3 (30%)	-0.01	0.41	0.79	0.45
	(0.13)	(0.08)	(0.07)	(0.08)
P*Q4 (53%)	0.53	0.20	0.71	0.30
	(0.11)	(0.07)	(0.07)	(0.07)
Distance to general school (primary)	-0.03			
	(0.00)			
Distance to indigenous school	0.06			
	(0.00)			
Distance to general school (Secondary)		-0.45	0.08	0.14
		(0.01)	(0.01)	(0.01)
Distance to telesecondary school		0.05	-0.77	0.05
		(0.01)	(0.01)	(0.01)
Distance to technical school		0.11	0.08	-0.50
		(0.01)	(0.01)	(0.01)
Number of general schools		0.0029	-0.0014	-0.0005
		(0.0002)	(0.0004)	(0.0002)
Number of telesecondary schools		-0.0032	0.0118	-0.0119
		(0.0013)	(0.0015)	(0.0014)
Number of technical schools		-0.0053	-0.0177	0.0037
		(0.0020)	(0.0031)	(0.0020)
Imputed wages		-0.01	-0.03	-0.02
		(0.00)	(0.00)	(0.00)
Education cat. (dad)				
Below primary school	0.06	-0.28	-0.17	-0.31
	(0.22)	(0.11)	(0.12)	(0.11)
Primary school completed	-0.01	-0.06	-0.06	-0.06
	(0.22)	(0.11)	(0.12)	(0.11)
Secondary or below	-0.20	0.28	0.07	0.25

Table C3: Primary and secondary school choices

	Indigenous	General	Telesecondary	Technical
	(0.22)	(0.11)	(0.12)	(0.11)
College or above	-0.06	0.37	-0.14	0.38
	(0.24)	(0.12)	(0.13)	(0.12)
Working status (dad)				
Full time	0.27	-0.10	-0.22	-0.16
	(0.17)	(0.08)	(0.09)	(0.09)
Not full time	0.08	0.15	-0.08	0.04
	(0.17)	(0.08)	(0.09)	(0.08)
Father present at home	-0.02	0.13	0.15	0.21
	(0.08)	(0.04)	(0.04)	(0.04)
Education cat. (mom)				
Primary school	-0.15	-0.22	-0.32	-0.36
	(0.38)	(0.19)	(0.20)	(0.20)
Primary school completed	-0.27	0.03	-0.19	-0.11
	(0.39)	(0.19)	(0.20)	(0.20)
Secondary or below	-0.50	0.35	-0.05	0.20
	(0.39)	(0.19)	(0.20)	(0.20)
College or above	-0.44	0.56	-0.20	0.34
	(0.41)	(0.19)	(0.21)	(0.20)
Working status (mom)				
Housework	-0.14	0.00	0.06	0.15
	(0.25)	(0.14)	(0.14)	(0.14)
Part time	0.04	0.01	-0.04	0.10
	(0.26)	(0.14)	(0.15)	(0.14)
Full time	0.08	0.03	0.07	0.10
	(0.27)	(0.14)	(0.15)	(0.15)
Mother present at home	-0.25	0.13	0.15	0.22
	(0.11)	(0.07)	(0.07)	(0.07)
Number of people at home				
5 people	-0.19	0.07	0.05	0.10
	(0.08)	(0.04)	(0.05)	(0.04)
6 people	-0.06	-0.14	-0.11	-0.15
	(0.08)	(0.04)	(0.05)	(0.05)
$\geq$ 7 people	0.02	-0.27	-0.19	-0.29
	(0.07)	(0.04)	(0.04)	(0.04)
Age	-0.13	-0.83	-0.71	-0.84
	(0.06)	(0.04)	(0.04)	(0.04)
Age^2	0.02	-0.08	-0.03	-0.10

Table C3: Primary and secondary school choices

	Indigenous	General	Telesecondary	Technical
	(0.04)	(0.02)	(0.02)	(0.02)
Male	0.01	0.27	0.27	0.26
	(0.05)	(0.03)	(0.03)	(0.03)
First language spoken at home				
Indigenous	1.74	-0.69	-0.16	-0.61
	(0.11)	(0.09)	(0.07)	(0.08)
Both Spanish and indigenous	1.22	-0.30	-0.38	-0.40
	(0.11)	(0.10)	(0.09)	(0.10)
Internet access	0.06	-0.24	-0.31	-0.25
	(0.10)	(0.05)	(0.05)	(0.05)
Computer access	-0.11	0.31	0.05	0.33
	(0.09)	(0.04)	(0.05)	(0.04)
Number of pre-school years				
1 year	-0.21	-0.59	-0.26	-0.51
	(0.19)	(0.09)	(0.10)	(0.10)
2 years	-0.12	-0.13	0.02	-0.06
	(0.18)	(0.09)	(0.09)	(0.09)
3 years	-0.15	0.27	0.30	0.37
	(0.18)	(0.09)	(0.09)	(0.09)
4 years	0.10	0.20	0.21	0.27
	(0.17)	(0.09)	(0.09)	(0.09)
Urban dummy	-0.83	0.11	-0.16	-0.04
	(0.07)	(0.04)	(0.04)	(0.04)
Regions				
North-center	-0.13	-0.48	-0.48	-0.32
	(0.16)	(0.06)	(0.07)	(0.06)
Center	0.22	-0.31	-0.25	-0.25
	(0.13)	(0.06)	(0.07)	(0.06)
South	-0.35	-0.11	-0.42	-0.07
	(0.14)	(0.07)	(0.09)	(0.08)
Type I	1.53	-0.64	-0.01	-0.55
	(0.17)	(0.10)	(0.10)	(0.10)
Type II	0.52	-0.14	0.12	-0.16
	(0.28)	(0.22)	(0.23)	(0.23)
Type III	-0.68	0.13	-0.26	0.03
	(0.19)	(0.10)	(0.11)	(0.11)
Intercept	-1.06	-0.25	1.05	-0.40

Table C3: Primary and secondary school choices

	Indigenous	General	Telesecondary	Technical
	(0.52)	(0.27)	(0.29)	(0.29)

Table C4: Retention estimates (choice and value-added)

	Choice	Mathematics	Spanish	Choice	mathematics	Spanish
Lag mathematics	-0.0053	0.19	0.39	-0.0016	0.39	0.17
	(0.0001)	(0.03)	(0.03)	(0.0001)	(0.03)	(0.03)
Lag Spanish	0.0001	0.24	0.13	-0.0105	0.13	0.30
	(0.0001)	(0.03)	(0.04)	(0.0002)	(0.04)	(0.04)
Prospera score quartile						
P*Q1 (4%)	0.39	0.53	0.54	-0.16	0.54	0.55
	(0.21)	(18.83)	(31.53)	(0.23)	(31.53)	(26.83)
P*Q2 (13%)	0.05	0.39	1.53	-0.14	1.53	0.99
	(0.12)	(11.00)	(15.94)	(0.14)	(15.94)	(14.37)
P*Q3 (30%)	0.01	1.91	2.81	-0.34	2.81	1.35
	(0.09)	(8.11)	(11.50)	(0.10)	(11.50)	(11.70)
P*Q4 (53%)	-0.03	2.63	3.53	-0.55	3.53	2.64
	(0.08)	(7.55)	(12.14)	(0.10)	(12.14)	(11.90)
Education cat. (dad)						
Below primary school	-0.42	9.77	-1.53	-0.36	-1.53	-13.29
	(0.15)	(12.22)	(18.44)	(0.16)	(18.44)	(17.90)
Primary school completed	-0.27	8.01	-12.37	-0.42	-12.37	-23.88
	(0.15)	(12.49)	(18.98)	(0.16)	(18.98)	(18.07)
Secondary or below	-0.10	16.86	6.21	-0.14	6.21	-17.26
	(0.15)	(12.33)	(18.04)	(0.16)	(18.04)	(17.08)
College or above	-0.38	14.03	11.87	-0.11	11.87	0.11
	(0.16)	(13.54)	(19.42)	(0.17)	(19.42)	(18.29)
Working status (dad)						
Full time	-0.12	-23.41	1.90	-0.13	1.90	4.39
	(0.11)	(10.79)	(17.19)	(0.13)	(17.19)	(15.03)
Not full time	-0.07	-19.21	7.45	-0.22	7.45	-3.09
	(0.11)	(10.39)	(16.29)	(0.13)	(16.29)	(14.24)
Father present at home	0.11	-5.36	-6.09	-0.06	-6.09	-8.38
	(0.05)	(4.86)	(7.88)	(0.06)	(7.88)	(7.38)
Education cat. (mom)						
Primary school	-0.46	1.95	8.49	-0.19	8.49	-3.10
	(0.26)	(24.17)	(57.66)	(0.37)	(57.66)	(57.56)
Primary school completed	-0.25	2.32	-9.28	0.05	-9.28	0.71
	(0.26)	(24.23)	(57.70)	(0.37)	(57.70)	(57.63)
Secondary or below	-0.20	1.00	4.53	0.28	4.53	-0.97
	(0.26)	(24.17)	(57.42)	(0.36)	(57.42)	(57.38)
College or above	0.00	7.76	13.50	0.52	13.50	14.84
	(0.27)	(24.85)	(57.86)	(0.37)	(57.86)	(57.70)

Table C4: Retention estimates (choice and value-added)

	Choice	Mathematics	Spanish	Choice	mathematics	Spanish
Working status (mom)						
Housework	0.09	12.95	45.37	0.33	45.37	46.15
	(0.17)	(16.97)	(32.03)	(0.27)	(32.03)	(42.89)
Part time	0.02	16.43	40.16	0.44	40.16	35.10
	(0.18)	(17.35)	(32.05)	(0.27)	(32.05)	(43.08)
Full time	-0.01	16.91	55.22	0.32	55.22	49.88
	(0.18)	(17.60)	(32.65)	(0.28)	(32.65)	(43.32)
Mother present at home	0.27	9.96	18.73	-0.05	18.73	1.38
	(0.10)	(9.38)	(14.34)	(0.12)	(14.34)	(13.31)
Number of people at home						
4 people	-0.06	7.50	7.71	-0.09	7.71	-6.01
	(0.06)	(5.03)	(8.04)	(0.07)	(8.04)	(7.47)
5 people	-0.12	5.78	-8.74	-0.17	-8.74	-16.20
	(0.06)	(5.55)	(9.03)	(0.07)	(9.03)	(8.48)
$\geq$ 6 people	-0.17	6.13	5.64	-0.18	5.64	1.06
	(0.05)	(4.48)	(7.24)	(0.06)	(7.24)	(6.96)
Age	3.94	4.96	18.32	4.03	18.32	15.86
	(0.06)	(8.36)	(12.36)	(0.09)	(12.36)	(12.44)
Age^2	-0.86	-2.01	-5.71	-1.02	-5.71	-5.18
	(0.02)	(2.49)	(3.89)	(0.03)	(3.89)	(4.01)
Male	0.42	-11.62	20.27	0.81	20.27	-7.70
	(0.04)	(3.70)	(6.79)	(0.05)	(6.79)	(6.17)
First language spoken at home						
Indigenous	-0.05	-15.35	33.27	-0.46	33.27	28.48
	(0.09)	(8.01)	(17.82)	(0.16)	(17.82)	(17.13)
Both Spanish and indigenous	-0.26	-21.25	-31.36	-0.09	-31.36	-21.20
	(0.12)	(11.00)	(21.72)	(0.17)	(21.72)	(23.03)
Internet access	0.27	-4.59	-18.19	-0.05	-18.19	-9.94
	(0.06)	(5.38)	(8.93)	(0.07)	(8.93)	(7.98)
Computer access	-0.02	1.96	2.98	0.04	2.98	-4.96
	(0.06)	(5.04)	(8.23)	(0.07)	(8.23)	(7.25)
Number of pre-school years						
1 year	0.00	-0.50	-33.28	-0.34	-33.28	-32.86
	(0.12)	(11.85)	(21.99)	(0.18)	(21.99)	(19.27)
2 years	0.21	-9.03	-23.31	-0.06	-23.31	-26.28
	(0.12)	(11.13)	(20.00)	(0.17)	(20.00)	(16.92)
3 years	0.40	3.56	-29.93	0.13	-29.93	-30.16
	(0.12)	(11.05)	(19.75)	(0.16)	(19.75)	(16.60)

Table C4: Retention estimates (choice and value-added)

	Choice	Mathematics	Spanish	Choice	mathematics	Spanish
4 years	0.06	1.24	-37.12	-0.23	-37.12	-35.24
	(0.12)	(11.07)	(19.93)	(0.17)	(19.93)	(17.10)
Urban dummy	0.06	14.26	-14.59	0.19	-14.59	-5.78
	(0.05)	(4.30)	(9.70)	(0.08)	(9.70)	(9.90)
Regions						
North-center	0.02	1.86	0.58	-0.12	0.58	8.17
	(0.07)	(6.44)	(9.40)	(0.07)	(9.40)	(8.48)
Center	0.57	7.64	17.99	0.39	17.99	27.00
	(0.06)	(5.99)	(8.31)	(0.07)	(8.31)	(7.57)
South	0.46	2.54	0.92	-0.06	0.92	16.15
	(0.07)	(6.67)	(10.15)	(0.09)	(10.15)	(9.40)
Indigenous school	0.09	-8.96	22.71			
	(0.10)	(9.14)	(9.84)			
Grade 5	0.50	-0.61	-3.44			
	(0.04)	(3.73)	(7.36)			
Grade 6	0.00	13.13	6.07			
	(0.08)	(7.49)	(5.99)			
Telesecondary school				-0.74	22.71	8.96
				(0.08)	(9.84)	(9.70)
Technical school				-0.28	-3.44	4.78
				(0.06)	(7.36)	(6.88)
Grade 8				-0.92	6.07	-0.66
				(0.05)	(5.99)	(5.69)
Unobserved types						
Type I	-0.17	-0.69	-2.67	0.10	-2.67	2.56
	(0.13)	(12.68)	(12.77)	(0.12)	(12.77)	(12.65)
Type II	-0.11	1.11	3.94	0.09	3.94	2.67
	(0.24)	(26.84)	(16.22)	(0.18)	(16.22)	(18.56)
Type III	0.00	4.33	-2.18	0.20	-2.18	1.43
	(0.11)	(11.41)	(10.00)	(0.10)	(10.00)	(9.84)
Intercept term	-4.50	253	186	-1.08	186	240
	(0.35)	(30)	(71)	(0.47)	(71)	(63)
Standard error (σ)		78	100		100	94
		(1.28)	(2.15)		(2.15)	(2.00)

Table C5: Dropout during lower-secondary school

Dropout period	Grade 7	Grade 8
Lag mathematics	-0.0014	-0.0008
	(0.0002)	(0.0001)
Lag Spanish	-0.0025	-0.0037
	(0.0002)	(0.0001)
Prospera score quartile		
P*Q1 (4%)	0.01	0.16
	(0.21)	(0.10)
P*Q2 (13%)	-0.36	0.02
	(0.13)	(0.06)
P*Q3 (30%)	-0.32	-0.07
	(0.10)	(0.05)
P*Q4 (53%)	-0.33	-0.12
	(0.09)	(0.05)
Education cat. (dad)		
Below primary school	0.15	-0.09
	(0.16)	(0.08)
Primary school completed	0.11	-0.19
	(0.16)	(0.09)
Secondary or below	-0.07	-0.17
	(0.15)	(0.08)
College or above	-0.30	-0.34
	(0.17)	(0.09)
Working status (dad)		
Full time	-0.09	0.00
	(0.12)	(0.07)
Not full time	-0.13	-0.04
	(0.11)	(0.06)
Father present at home	-0.15	-0.17
	(0.06)	(0.03)
Education cat. (mom)		
Primary school	0.10	-0.04
	(0.29)	(0.16)
Primary school completed	0.12	-0.03
	(0.29)	(0.16)
Secondary or below	-0.05	-0.05
	(0.29)	(0.16)
College or above	-0.25	-0.18
	(0.30)	(0.16)

Table C5: Dropout during secondary school

Dropout period	Grade 7	Grade 8
Working status (mom)		
Housework	-0.39	-0.08
	(0.18)	(0.12)
Part time	-0.24	0.11
	(0.19)	(0.12)
Full time	-0.34	0.03
	(0.19)	(0.12)
Mother present at home	-0.07	-0.03
	(0.10)	(0.06)
Number of people at home		
5 people	0.00	-0.04
	(0.06)	(0.03)
6 people	-0.03	0.00
	(0.07)	(0.03)
$\geq$ 7 people	0.13	0.00
	(0.06)	(0.03)
Age	1.10	1.25
	(0.05)	(0.03)
Age^2	0.01	-0.05
	(0.03)	(0.02)
Male	0.28	0.15
	(0.05)	(0.02)
First language spoken at home		
Indigenous	0.00	0.08
	(0.13)	(0.07)
Both Spanish and indigenous	-0.13	0.03
	(0.17)	(0.08)
Internet access	0.18	0.22
	(0.07)	(0.03)
Computer access	-0.10	-0.12
	(0.07)	(0.03)
Number of pre-school years		
1 year	-0.12	-0.02
	(0.14)	(0.08)
2 years	-0.16	-0.04
	(0.13)	(0.07)
3 years	-0.33	-0.11
	(0.13)	(0.07)

Table C5: Dropout during secondary school

Dropout period	Grade 7	Grade 8
4 years	-0.29	-0.14
	(0.13)	(0.07)
Urban dummy	0.17	0.22
	(0.06)	(0.03)
Regions		
North-center	0.32	0.01
	(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		
Type I	0.03	0.05
	(0.15)	(0.12)
Type II	-0.03	0.04
	(0.15)	(0.13)
Type III	-0.06	0.06
	(0.15)	(0.12)
Intercept term	-1.60	-0.66
	(0.12)	(0.10)

Table C6: Coefficients associated with cheating equation

	Mathematics		Spanish	
	Value	S.D.	Value	S.D.
For non-retained students				
Grade 5				
General	39	1.4	25	1.4
Indigenous	71	7.0	54	6.4
Grade 6				
General	44	1.5	27	1.2
Indigenous	54	6.9	39	5.4
Grade 7				
General	57	3.3	32	3.1
Telesecondary	92	5.4	56	4.3
Technical	74	3.8	44	3.9
Grade 8				
General	76	2.0	39	1.9
Telesecondary	98	3.1	63	2.4
Technical	77	2.3	39	2.2
Grade 9				
General	66	3.0	25	2.7
Telesecondary	44	3.9	17	3.1
Technical	63	3.1	24	2.7
For retained students				
Primary school	62	14.4	50	12.2
Secondary school	122	20.2	84	16.9

D Additional estimation results

Tables D1 and D2 provide evidence on the model's goodness-of-fit based on the 70% test sample. Table D1 compares the average test scores by grade and by *Prospera*-beneficiary status, in the data and simulated under the model. Table D2 shows the model fit to the school-type distribution. See further discussion in the text. Table D3, we show the cumulative Prospera-program impacts on test scores and on dropouts by gender, which shows very similar impacts for girls and boys.

Table D1: Goodness of fit for average test scores by *Prospera* status (P)

	Math				Spanish			
Prospera	$P=0$		$P=1$		$P=0$		$P=1$	
	Data	Sim	Data	Sim	Data	Sim	Data	Sim
Grade 5	528	526	500	500	527	527	494	495
Grade 6	556	551	530	535	551	549	519	523
Grade 7	496	488	496	501	483	477	468	473
Grade 8	522	517	532	537	493	487	482	486
Grade 9	545	545	566	570	495	494	485	489

Note: 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 D2: 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	0.53	0.52	0.12	0.12	0.29	0.29	0.06	0.07
Grade 8	0.52	0.51	0.12	0.12	0.29	0.29	0.08	0.08
Grade 9	0.45	0.45	0.10	0.11	0.25	0.26	0.20	0.18
Prospera beneficiary (P=1)								
Grade 7	0.27	0.27	0.44	0.42	0.21	0.20	0.08	0.12
Grade 8	0.26	0.26	0.43	0.41	0.21	0.19	0.10	0.14
Grade 9	0.22	0.23	0.37	0.37	0.17	0.17	0.24	0.24

Table D3: 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	503	503	0.0	497	497	-0.1
Grade 6	536	537	-0.9	535	536	-0.9
Grade 7	501	492	9.3	499	490	9.0
Grade 8	535	524	10.9	542	532	10.2
Grade 9	568	554	13.6	572	559	13.0
Spanish						
Grade 5	509	510	-1.3	480	481	-1.2
Grade 6	536	538	-2.2	510	512	-2.1
Grade 7	492	489	3.0	454	451	3.1
Grade 8	504	499	4.9	470	466	4.8
Grade 9	503	499	3.8	476	472	3.7
Dropout rate						
Grade 7	0.12	0.17	-0.05	0.12	0.16	-0.05
Grade 8	0.13	0.18	-0.05	0.14	0.18	-0.05
Grade 9	0.22	0.26	-0.05	0.26	0.31	-0.05