# The Marginal Returns to Distance Education: Evidence from Mexico's Telesecundarias*

Emilio Borghesan† and Gabrielle Vasey‡

This paper analyzes a large-scale and long-running distance education program in Mexico. We estimate marginal treatment effects (MTEs) for learning in math and Spanish in telescundarias relative to traditional Mexican secondary schools using an empirical framework that allows for unobserved sorting on gains. The estimated MTEs reveal that school choice is not random and that the average student experiences significant improvements in both math and Spanish after just one year of attendance in telesecundarias. We find that the existing policy reduces educational inequality, and our policy-relevant treatment effects show that expanding telesecundarias would yield significant improvements in academic performance.

**JEL Codes:** I21, I24, O15.

**Keywords:** Education Policy, Achievement Gap, Education and Inequality

---

*We thank Jere Behrman, Susan Parker, Uta Schönberg, and Petra Todd for their help and guidance, as well as seminar participants at the University of Pennsylvania, LACEA/LAMES, and the Young Economists Symposium for helpful comments. We are also deeply thankful to Mira Potter-Schwartz, Elizaveta Brover, and Erika Trevino Alvarado for providing excellent research assistance. The Mexican Secretariat of Public Education (SEP) has generously provided data for this analysis. We thank Hector Robles Vaquez for preparing databases and for assistance working with these data. 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".

†University of Pennsylvania; email: borghesa@sas.upenn.edu

‡Postdoctoral Research Fellow at Evidence for Policy Design at Center for International Development, Harvard Kennedy School; email: gabrielle.vasey@gmail.com

## I. Introduction

Distance education elicits strong reactions from parents, policymakers, and teachers alike (Almagor 2021). Recent studies of the effects of pandemic-related school closures on academic achievement in developed countries show that more time away from in-person school is associated with less learning (Engzell, Frey, and Verhagen 2021; Maldonado and De Witte 2021). The effectiveness of distance education depends on two factors: how it is implemented and the quality of the counterfactual alternative. Distance education can be implemented as an online version of in-person schooling or it can be scaled, with a small number of highly effective communicators delivering lectures to multiple school districts. In addition, remote learning is more likely to be effective where the quality of in-person alternatives is low. These two factors mean that distance education may be of particular value in developing countries, where rural-urban divides and high rates of teacher absenteeism deprive many students of high-quality instruction (Duflo, Hanna, and Ryan 2012). Yet, large-scale implementation of a single remote curriculum is rare and there has been little research in economics evaluating the effectiveness of such programs.

In this paper, we study the effects of a long-running distance education program on student learning in Mexico. Since 1968, Mexico has undertaken an ambitious effort to provide a distance education option to secondary school students.[1] These students decide between traditional schools, with subject-specific teachers delivering their own lectures in person, and telesecundarias, brick and mortar establishments where students watch televised lectures and work on standardized assignments under the supervision of a single adult monitor. If shown to be effective, telesecundarias could provide a model for operating distance education at scale that differs considerably from the pandemic approach to schooling.

We analyze the effectiveness of telesecundarias relative to traditional public schools in Mexico using data on low-stakes end-of-year standardized tests in math and Spanish—the Evaluación National de Logro Académico en Centros Escolares, or ENLACE.[2] We examine students who were in the sixth grade in 2007/08 and who

[1] Secondary schools in Mexico enroll students in grades seven through nine and are akin to middle schools in the United States, while upper secondary schools enroll students in grades ten through twelve. Throughout the paper, we use the terms "secondary" and "upper secondary" to be consistent with the Mexican educational system.

[2] Numerous studies have used the ENLACE exams as a means of evaluating educational interventions (De Hoyos, Garcia-Moreno, and Patrinos 2017; Dustan, De Janvry, and Sadoulet 2017; Estrada and Gignoux 2017; Avitabile and De Hoyos 2018; Estrada 2019; Michaelsen and Salardi 2020). Scores on these exams have been shown to predict important life outcomes including university enrollment and wages (de Hoyos, Estrada, and Vargas 2021).

Table 1
Mean Scores on Exams, by School Type

| | Math | | | | Spanish | | | |
|---|---|---|---|---|---|---|---|---|
| | All | Traditional | Telesecundaria | Difference | All | Traditional | Telesecundaria | Difference |
| Grade 6 | 521.4 | 531.0 | 484.2 | -46.7 | 516.3 | 526.7 | 476.1 | -50.5 |
| Grade 7 | 500.1 | 501.5 | 494.8 | -6.70 | 499.7 | 503.6 | 484.7 | -18.9 |
| Grade 8 | 508.5 | 505.0 | 522.1 | 17.1 | 477.2 | 479.2 | 469.4 | -9.80 |
| Grade 9 | 524.5 | 518.7 | 546.9 | 28.2 | 499.6 | 501.2 | 493.6 | -7.60 |

The table shows mean scores on the math and Spanish ENLACE exams for public school students in Mexico who advanced to secondary school. Students in grade 6 attend primary school. Starting in grade 7, public school students attend either traditional secondary schools or telesecundarias. The standard deviations on the math ENLACE exams are 120.7, 100.7, 104.0, and 119.3 in grades 6, 7, 8, and 9, respectively. The standard deviations on the Spanish ENLACE exams are 107.3, 100.0, 109.2, and 100.8 in grades 6, 7, 8, and 9, respectively.

advanced to secondary school the following year. Table 1 presents mean scores on the ENLACE exams for this sample of public school students for grades six through nine. The table shows that in grade six, the year prior to enrollment in secondary school, students who will attend telesecundarias the following year score $46.7$ points lower in math and $50.5$ points lower in Spanish, on average, than their peers who will eventually attend traditional secondary schools. These gaps in the sixth grade, of $0.387$ and $0.471$ standard deviations, respectively, diminish considerably after students enroll in secondary school. For each year of enrollment, the performance gap between students at traditional schools and telesecundarias narrows in Spanish, and in math students at telesecundarias overtake their peers at traditional schools by grade eight and are performing $28.2$ points better by grade nine.

The marked improvement of telesecundaria students relative to their traditional school peers is neither caused by differential rates of dropout, as we restrict our sample to students who advance to secondary school, nor is it likely to be an artifact caused by differential rates of cheating or exam-sitting (see Appendix A). Nevertheless, since secondary school type is not randomly assigned, the extent to which this improvement represents a causal effect of telesecundarias or whether it stems from selection based on unobservable characteristics is unclear.

To assess the causal effect of telesecundarias on student learning, this paper combines a value-added model with the marginal treatment effects (MTE) framework, introduced by Björklund and Moffitt (1987) and generalized in Heckman and Vytlacil (1999, 2001b, 2005, 2007). We measure learning in two subjects, math and Spanish, by value added, the intertemporal difference in scores on the ENLACE exams administered the year before and one year after the start of secondary school. We focus on the transition from primary school to secondary school (grades six to seven), as this is when students face a choice regarding the type of school to attend. We define the relative effectiveness of telesecundarias as the difference in value added between the two school types between grades six and seven. Focusing on value added allows us to isolate how much student knowledge in seventh grade is directly attributable to the type of secondary school attended.

We use the MTE framework to jointly model student's choice of school and their academic outcomes at each school. This framework, which allows for the possibility that the choice of secondary school may be correlated with unobserved factors affecting performance on the exams, is a flexible semiparametric selection model that can estimate treatment effects for policies as they are currently implemented as well as evaluate a wide range of counterfactual policies. It is important to conduct the MTE analysis that we do in this paper instead of using traditional instrumental variables (IV) methods for several reasons. The estimates that are obtained by IV with a continuous instrument can be represented as a particular weighted average of the MTE and correspond to some treatment effect, but they do not usually identify typical treatment parameters of interest in the presence of sorting on gains. We use our estimated MTEs to calculate these parameters – the average effect of treatment (ATE), the average effect of treatment on the treated (TT), and the average effect of treatment on the untreated (ATU) – as well as analyze additional counterfactual policies that expand access to telesecundarias. Finally, we use the MTE to conduct distributional analysis of the effects of telesecundaris on the rural-urban test score gap and the gap in test scores between students who speak an indigenous language at home and those who speak Spanish. The results in this paper highlight the flexibility and usefulness of the MTE framework as a tool for policy evaluation.

Identification of the MTE requires an instrumental variable that affects the decision of which school to attend but does not directly affect outcomes at each school. The instrument we use in this paper is a measure of relative distance, which we compute as the difference between two distance measures. The first measure is the distance between each student's primary school and the nearest telesecundaria, while the second measure is the distance between the same primary school and the nearest traditional secondary school. The resulting instrument measures how many kilometers farther the telescundaria is compared to the traditional school.[3,4]

This instrumental variable, relative distance, is highly predictive of attendance

[3]A large body of empirical work uses a measure of distance to school as an instrument. See Card (1995), Kane and Rouse (1995), Kling (2001), Currie and Moretti (2003), Cameron and Taber (2004), Carneiro, Heckman, and Vytlacil (2011), and Carneiro, Lokshin, and Umapathi (2017).

[4]We eliminate students from the sample who lack both schooling options within a 15 km radius.

in telesecundarias: A one kilometer reduction in this measure, meaning that the telesecundaria becomes relatively closer, causes the student's probability of attending a telesecundaria to increase by 3.4% on average. Cameron and Taber (2004) and Carneiro and Heckman (2002) have raised concerns that distance to secondary school is correlated with student ability in the United States. We discuss why endogeneity of this sort is less likely in Mexico than in the US, and we conduct several robustness checks that suggest there is little dependence between the instrument and unobserved determinants of academic performance. Importantly, we rule out the possibility that fatigue resulting from walking to school worsens academic performance and affects our result.

We find that telesecundarias are highly beneficial: The average treatment effect (ATE) of telesecundaria attendance relative to attendance in traditional schools is a $36.7$ point increase in math scores (equivalent to $0.364$ standard deviations on the exam) and a $23.0$ point increase in Spanish scores (equivalent to $0.230$ standard deviations) after just one year of attendance.[5] These ATEs conceal considerable heterogeneity in who benefits from telesecundarias. Some students see gains of over $50$ points, while others experience no benefit. Our findings suggest that much of the improvement of telesecundaria students documented in Table 1 represents causal effects of telesecundarias.

Our analysis uncovers a pattern of negative sorting on gains. Students who are more likely to attend telesecundarias benefit less than students who are less likely to attend. This pattern is statistically significant: Our estimated MTEs for math and Spanish reject the hypothesis of no sorting on gains. Two theories can explain the existence of negative sorting. In the first theory, telesecundarias' negative reputation may induce parents to have biased beliefs regarding school quality (Cornelissen et al. 2018, Ainsworth et al. 2020). Another theory is that families may base their decisions on factors other than academic achievement, such as educational attainment.[6] We present evidence, consistent with this theory, that students who are most likely to attend telesecundarias experience some of the largest gains in the likelihood of advancing to the final year of secondary school.

We use our estimated MTEs to evaluate counterfactual policies that expand ac-

[5] As both the mean and standard deviation of test scores in the seventh grade may be influenced by the treatment effect, we conduct our analysis using raw scores on the exams. The standard deviations on the seventh grade ENLACE exams ($100.7$ for math and $100.0$ for Spanish) are close enough to $100$ to allow for an easy translation between points and standard deviations.

[6] Abdulkadiroğlu et al. (2020) show that parents in New York City do not value school effectiveness conditional on peer quality. As telesecundaria students are considerably more disadvantaged than traditional school students, this may also explain the pattern of reverse sorting.

cess to telesecundarias. The first policy we consider is a dramatic increase in telesecundaria availability that reduces the instrument, relative distance, by five kilometers (km) for everyone in the sample. The second policy is a school building program that constructs a telesecundaria adjacent to the eighteen percent of Mexican primary schools without one within five km. We find that the first policy raises math (Spanish) scores by $35.5$ ($23.8$) points, while the second raises scores by $21.8$ ($16.4$) points. The effects of the two policies differ and neither corresponds to the estimates obtained by a Two-Stage Least Squares regression that uses distance as an instrument ($25.1$ and $14.6$ points for math and Spanish, respectively), highlighting the importance of adopting a framework allowing for heterogeneous treatment effects and self-selection as we do in this paper. These policy-relevant treatment effects show that further expansions of this already-widespread program would cause gains in test scores.

We additionally show how the MTE framework can be used to conduct distributional analysis, and we estimate the causal effects of telesecundarias on two prominent divisions in Mexico: the gap in test scores between rural and urban students and between students who speak Spanish at home and those who speak an indigenous language. Our distributional treatment effects show that the rural-urban test score gap would be 142% larger in math and 43% larger in Spanish if all students were to attend traditional schools. Similarly, the Spanish-Indigenous language test score gaps would be 40% larger in math and 16% larger in Spanish. These distributional treatment effects reveal that, not only do telesecundarias improve educational outcomes in Mexico, they do so in a way that reduces persistent educational inequities across different subgroups of the population.

Our study has several main contributions. First, we combine the value-added and MTE frameworks to evaluate the quality of distinct educational alternatives, and we are the first to document that telesecundarias have significant causal effects on academic achievement relative to traditional schools. In addition, we know of no other papers in the education literature in which the marginal treatment effect can be nonparametrically identified and precisely estimated over its entire support. We compute all treatment parameters and the effects of counterfactual policies using our semiparametric estimates of the MTE.[7] These estimates are precise both because of the large size of the sample (over $120,000$ observations) and the predictive power of the instrument, which induces significant variation in the probability of attending a telesecundaria.

[7] We estimate a marginal treatment effect, $MTE(X, U_D)$ that is parametric in $X$ and nonparametric in $U_D$. We use the term "nonparametric identification" to refer to identification of $MTE(X, U_D)$ with respect to $U_D$ and "semiparametric estimate" to refer to estimates obtained via the semiparametric function, $MTE(X, U_D)$.

The semiparametric approach provides assurance that our estimated treatment parameters are not biased by a subpopulation with markedly different treatment effects who are especially likely (or unlikely) to attend telesecundarias.[8] Finally, we show how the MTE framework can be used to conduct distributional analysis, and we find that telesecundarias have significant causal effects on reducing educational inequality in Mexico.

This paper proceeds as follows. Section II describes the relevant literature to which our paper contributes, while section III provides information on secondary schooling in Mexico. Section IV describes the model and empirical strategy, section V describes the data, and section VI presents our main empirical results. Section VII presents estimates of counterfactual and distributional treatment effects and explores potential mechanisms behind telesecundarias' effectiveness. Section VIII concludes.

## II. Related Literature

Our paper contributes to three interrelated branches of the literature on education in economics. The first strand evaluates the effectiveness of new educational technologies, while the second uses the MTE framework to evaluate educational policies, and the third analyzes telesecundarias directly.

Bianchi, Lu, and Song (2020) analyze an educational technology in western China that is similar to telesecundarias. Exploiting the differential rollout of a large-scale distance education initiative across space and time in a difference-in-differences (DiD) design, they find that exposure to computer-aided learning raises math and Chinese skills by $0.18$ and $0.23$ standard deviations, respectively. Relative to their paper, we examine the effect on educational outcomes after a year of attendance rather than seven to ten years later, we allow for the choice of school to be nonrandom, we demonstrate the extent of heterogeneity in educational outcomes for students with various probabilities of enrolling in distance education, and we compute counterfactual and distributional treatment effects.

Several recent papers evaluate experimental interventions that combine new educational technologies with traditional learning methods in developing countries. Muralidharan, Singh, and Ganimian (2019) analyze a program that provides a personal-

[8] In this paper, the common support of propensity scores for treated and control units is the full $[0, 1]$ interval. Whenever this condition is not satisfied, treatment parameters that require full support, like ATE, TT, and ATU, cannot be nonparametrically identified. This means that estimates of these parameters based on a parametric MTE may be biased by a subpopulation outside the region of common support which has treatment effects different from those inside the region of common support.

ized computer-aided learning software to students in Delhi, India for ninety minutes each day after school. They find that the technology, which identifies each students' skill level in math and Hindi and then asks skill-appropriate questions, causes dramatic gains in both math and Hindi proficiency, of $0.60$ and $0.39$ standard deviations, respectively, on tests administered after ninety school days of exposure.

Beg et al. (2019) use data from two randomized controlled trials (RCTs) in Pakistan and find that the combination of high quality videos with in-person teaching raises student performance on standardized tests. In another RCT, Johnston and Ksoll (2017) find that a similar program in Ghana, which broadcasts live instruction from experts into rural schools, has a positive impact on test scores. Given the positive findings on distance education, it is perhaps surprising that this low-cost alternative is not prevalent in developing countries and rural areas around the world. Our paper validates these findings from RCTs by conducting analysis of a longstanding distance education policy using observational data and a framework that allows for nonrandom school choice.

Earlier work on the interaction of technology and education is surveyed in three excellent reviews by Bulman and Fairlie (2016), Escueta et al. (2017), and Rodriguez-Segura (2020). Bulman and Fairlie (2016) distinguish between educational interventions that provide supplements versus alternatives to the existing educational curriculum. Most educational interventions offer some learning technology to a treatment group without an offsetting reduction in resources, a practice that favors finding a positive effect of the technology. It is therefore not surprising that we estimate treatment effects for telesecundarias that are of a similar magnitude as those in Beg et al. (2019) and Johnston and Ksoll (2017) but are smaller than the supplemental intervention analyzed in Muralidharan, Singh, and Ganimian (2019).

A second strand of literature related to our paper evaluates the effects of schooling using the MTE framework. Carneiro, Heckman, and Vytlacil (2011) analyze the decision to attend college in the United States. They estimate positive treatment effects of college attendance on income and uncover a pattern of positive sorting on gains, whereby the return to the marginal student induced to attend college by policies that expand college access is significantly lower than the return to the average student already attending college. Carneiro, Lokshin, and Umapathi (2017) find similar results for the returns to attending secondary school in Indonesia using similar methods and with distance to the nearest secondary school as the instrument. Cornelissen et al. (2018) analyze the decision of parents to enroll their children in daycare in Germany, and, contrary to the two previous papers, find a pattern of negative sorting on gains.

Students who are not enrolled in Germany's universal child care program would experience increases in their readiness for primary school had they attended child care, while those currently attending experience little benefit. An advantage of our paper relative to this strand of literature is that we are able to nonparametrically identify the MTE over its full support, and we estimate all our policy effects without strong assumptions on the joint distribution of unobservables or on the shape of the MTE.[9] In addition, we evaluate competing educational technologies by comparing value added, and we extend the usefulness of the MTE framework by using it to conduct analysis of the distributional consequences of policies.

Although telesecundarias are widespread and have existed for over fifty years in Mexico, little research on their effectiveness exists. Recently, two papers have used DiD designs to estimate the effect of proximity to telesecundarias on educational attainment and labor market outcomes (Fabregas 2020, Navarro-Sola 2019). Both papers use data from the Mexican Census and thus lack information on student test scores and the type of school attended. Even without this information, both papers find that telesecundarias raise educational attainment and future income, although the estimates in Navarro-Sola (2019) are substantially larger than those in Fabregas (2020). Behrman, Parker, and Todd (2020) and Behrman et al. (2021) look at the impact of conditional cash transfers on schooling trajectories and find that cash transfers primarily raise schooling and educational achievement through increasing enrollment in telesecundarias. Our contribution to the literature on telesecundarias is through demonstrating their significant causal effects on student learning. Our results suggest that the causal effects of telesecundarias on the labor market found in earlier works stem not only from the extensive margin of educational attainment but also from the intensive margin of academic performance conditional on staying in school.

## III. Secondary Schooling in Mexico

Throughout the twentieth century, Mexico struggled to attract qualified teachers to rural areas to instruct the millions of school-age children living there. Introduced in 1968, telesecundarias were seen as a solution to this problem. Telesecundarias are physical structures with four key scholastic components. The first is the television program. Every subject begins with students watching fifteen minutes of a pre-recorded televised lecture. Lectures for each subject are recorded in Mexico City by high-quality

[9] Sapelli and Vial (2002) conduct a similar analysis of the effects of private schools on test scores in Chile. They estimate treatment effects using a generalized Roy Model, which assumes that unobservables have a multivariate normal distribution, without estimating the MTE curve directly.

instructors, so-called *telemaestros*, who are selected for their communication skills and undergo extensive training. Considerable effort is made to ensure that the videos are of high quality: The typical fifteen-minute televised lecture undergoes twenty days of production time and costs $30,000-$50,000 (Calderoni 1998).

Following the video lectures, a single teacher leads students in a 35-minute lesson on the same subject. The teacher does not specialize in a particular subject: Students have the same teacher for all courses. This teacher follows a guide designed for telesecundaria teachers that is filled with suggestions for each subject. The 35 minutes are spent in myriad ways, with the teacher leading question and answer sessions, engaging students in group activities, and giving assignments for students to do on their own.

The third educational resource is an encyclopedia-like book that contains the essential information in each subject taught during that year. These books are similar to textbooks in traditional secondary schools and are used by students as references while doing their assignments. The final component of telesecundaria education is the learning guide. Like a workbook, learning guides are filled with questions that students can answer individually as well as suggestions for group activities that reinforce learning. Class time is frequently devoted to doing assignments in the learning guides.

The four main educational components – televised instruction, in-person teaching, reference texts, and learning guides – are designed to be complementary. Students evidently see them that way: Ethnographic research indicates that students see each component as reinforcing the knowledge acquired through the televised lectures (Estrada 2003).

Telesecundarias were first introduced in rural areas and predominate in Mexico's poorer South. While they have expanded into suburban and urban areas, students from the South and from rural areas are still over-represented (see Table B-2). The typical telesecundaria is a purpose-built structure with between three and nine classrooms, a library, restrooms, a science lab, and a playground. Students in telesecundarias and traditional schools attend school for the same number of hours per week (30) and days per year (200). A reform in 1993 mandated schooling through grade nine and resulted in increases in both the construction of new telesecundarias and telesecundaria enrollment. We study the cohort of students who were in grade six in 2007/08, after secondary schooling became compulsory.

## IV. Model

### A. *Student Achievement*

We apply the potential outcomes framework of Rubin (1974) to a value-added model of learning. Students can either attend a telesecundaria or a traditional school.[10] We define the random variable $D$, where $D = 1$ denotes attendance in a telesecundaria and $D = 0$ denotes attendance in a traditional school. We study the effects of telesecundaria attendance on two outcomes: math and Spanish test scores in seventh grade. For each course $C \in \{\text{Math}, \text{Spanish}\}$, the potential outcomes $Y_0^C$ and $Y_1^C$ correspond to the grade seven test score a student would achieve had she enrolled in a traditional or telesecundaria school, respectively. For ease of notation, we will omit the $C$ superscripts. The same causal model will be used for each of the two outcomes.

We model test score outcomes and school choice according to the selection model in Heckman and Vytlacil (2005):

$$Y_1 = X\beta_1 + U_1 \ , \tag{1}$$

$$Y_0 = X\beta_0 + U_0 \ , \tag{2}$$

$$D = \mathbb{1}(\mathbf{Z}\gamma > V) \ , \tag{3}$$

where $X$ is a vector of observable characteristics influencing outcomes, $Z$ is a vector of observable characteristics influencing the choice of secondary school, and $(U_1, U_0, V)$ are unobserved by the econometrician. Students are assumed to know $(U_1, U_0, V)$ and may act upon them. Equations (1) and (2) are value-added equations: $X$ contains the previous year's test scores in both math and Spanish.[11]

The effect of attending a telesecundaria relative to a traditional school on a student's test scores is given by $Y_1 - Y_0$, and the average effect for individuals with a

[10] Mexico has three public secondary school types: General, Technical, and Telesecundaria. We consider the choice between a traditional school (General/Technical) and a telesecundaria. Table B-1 in Appendix B reveals that General and Technical school have similar distributions of observable household characteristics and student test scores in the seventh grade, so we believe that it is reasonable to consider them as a single alternative for the purposes of evaluating learning in math and Spanish between the sixth and seventh grades. However, the ensuing analysis goes through without modification if General and Technical schools are treated as separate alternatives as long as the results are re-interpreted as the causal effect of telesecundaria education relative to the next best alternative. We exclude the eight percent of students who attend private school from the analysis. Students who drop out are also omitted (six percent of the sample).

[11] The value-added framework admits a causal interpretation when the effects of time-varying investments on test scores decline geometrically with the time between when the investment was made and when the test was taken (Boardman and Murnane 1979, Todd and Wolpin 2003).

specific set of observable characteristics is $ATE(X) = \mathbb{E}[Y_1 - Y_0|X = x]$. The fundamental challenge in estimating any sort of treatment effect is that the econometrician only observes one of the two potential outcomes, $Y = Y_0 + D(Y_1 - Y_0)$.

The instrument, $Z$, is an exclusion restriction that must simultaneously induce variation in the choice of school conditional on the covariates, $X$, and have no direct effect on the outcome variables. When $Z$ is a sufficiently predictive instrument, as in this paper, it can shift the probability of attending a telesecundaria continuously between 0 and 1. Such a shift would cause even the most unlikely student to attend a telesecundaria, thereby permitting a comparison of $Y_0$ and $Y_1$ for all students.

The marginal treatment effect (MTE), introduced by Björklund and Moffitt (1987) and extended in Heckman and Vytlacil (1999, 2001b, 2005, 2007), is the average treatment effect for an individual at a particular margin of "resistance to treatment." $V$, in equation (3), represents this resistance to treatment. An individual with a higher $V$ is, on the basis of unobservables, less likely to attend a telesecundaria. As in Heckman and Vytlacil (2005), we apply the following useful transformation to equation (3) to obtain

$$\begin{aligned} D &= \mathbb{1}(Z\gamma > V) \\ &= \mathbb{1}(F_V(Z\gamma) > F_V(V)) \\ &= \mathbb{1}(P(Z) > U_D)\ , \end{aligned}$$

where $P(Z)$ is the propensity score and $U_D \sim U[0,1]$. Following the transformation, the MTE can be written as

$$MTE(x, u_D) = \mathbb{E}[Y_1 - Y_0|X = x, U_D = u_D]\ . \tag{4}$$

The marginal treatment effect measures the average difference in outcomes at telesecundarias relative to traditional schools for individuals with observable characteristics $x$ and latent resistance to treatment $u_D$.

*B. Identification*

The MTE is identified under the following assumptions, stated in Heckman and Vytlacil (2005) and modified slightly to fit the notation presented here:

**(A-1)** $Z$ is a nondegenerate random variable conditional on $X$.

**(A-2)** $(U_0, U_1, U_D) \perp\!\!\!\perp Z|X$.

**(A-3)** $U_D$ is absolutely continuous with respect to the Lebesque measure.

**(A-4)** $\mathbb{E}|Y_1|$ and $\mathbb{E}|Y_0|$ are finite.

**(A-5)** $1 > \mathbb{P}(D = 1|X) > 0$ .

Assumption (A-1) ensures that the instrument influences attendance in telesecundarias conditional on covariates, $X$, while (A-2) assumes that the instrument is exogenous in the sense that it is independent of unobservable variables in the selection and outcome equations conditional on $X$. Assumptions (A-3) - (A-5) are technical assumptions that are satisfied in our setting. Under these assumptions, $MTE(X, U_D)$ is nonparametrically identified by the Local Instrumental Variables (LIV) estimand (Heckman and Vytlacil 2001a):

$$\frac{\partial \mathbb{E}[Y|X = x, P(z) = p]}{\partial p}\Big|_{p=u_D} = MTE(x, u_D) \ . \tag{5}$$

The LIV estimand shows that the marginal treatment effect at each value of the latent resistance to treatment, $U_D$, is identified by individuals who are indifferent between being treated and not, because when $p = U_D$, the individual is on the knife edge between participating and not participating.

*C. Parameters of Interest*

A large class of parameters corresponding to the effect of telesecundaria attendance on schooling outcomes can be written as weighted averages of the marginal treatment effect. In this paper we are interested in $ATE(X)$, as well as the average effect of treatment on the treated, $TT(X) = \mathbb{E}[Y_1 - Y_0 \mid X = x, D = 1]$, the average effect of treatment on the untreated, $TUT(X) = \mathbb{E}[Y_1 - Y_0 \mid X = x, D = 0]$, and various treatment effects corresponding to the effects of never-before implemented policies. These Policy-Relevant Treatment Effects, first defined in Heckman and Vytlacil (2001b), are defined for a shift from a pre-existing policy $a$ to a new policy $a'$ and provide a normalized effect of the policy change:

$$PRTE_{a',a}(X) = \frac{\mathbb{E}[Y_{a'} - Y_a \mid X = x]}{\mathbb{P}(D_{a'} = 1 \mid X = x) - \mathbb{P}(D_a = 1 \mid X = x)} \ . \tag{6}$$

Any treatment parameter, including $ATE$, $TT$, $TUT$, and $PRTE_{a',a}$ for policies $a$ and $a'$, can be computed by integrating the MTE with respect to the distribution of $U_D$

induced by the treatment parameter under consideration:

$$ATE = \int MTE(X, U_D) dF(X, U_D) \ ,$$

$$TT = \int MTE(X, U_D) dF_{U_D, X|D=1}(x, u_D \mid D = 1) \ , \tag{7}$$

$$TUT = \int MTE(X, U_D) dF_{U_D, X|D=0}(x, u_D \mid D = 0) \ , \tag{8}$$

$$PRTE_{a',a} = \int MTE(X, U_D) dF_{U_D, X|D_a=0, D_{a'}=1}(x, u_D \mid D_a = 0, D_{a'} = 1) \ . \tag{9}$$

In the next section, we discuss the methods we use to integrate the MTE to obtain these treatment parameters.

*D. Estimation*

Assumptions (A-1) - (A-5) require that $MTE(X, U_D)$ be estimated separately for each $X$. When $X$ is high-dimensional, as in our setting, $MTE(X, U_D)$ is only identified by the support of $P(Z)$ given X. Even if the unconditional support of $P(Z)$ is the entire unit interval, $supp(P(Z)) \mid X = x)$ may consist of only a few points. Since this will be too few to estimate $MTE(X, U_D)$ with any degree of precision, we strengthen assumption (A-2) to (A-2)′:

**(A-2)′** $(X, Z) \perp\!\!\!\perp (U_1, U_0, U_D)$

Assumption (A-2)′ is standard in the literature estimating selection models. It has two consequences. The first is that, together with the linear framework in equations (1) and (2), it yields an MTE function that is additively separable in $X$ and $U_D$ so that

$$MTE(X, U_D) = X(\beta_1 - \beta_0) + K(U_D) \ .$$

An additional consequence of assumption (A-2)′ is that $MTE(X, U_D)$ can now be identified on the unconditional support of $P(Z)$ rather than $supp(P(Z) \mid X)$. The cost of the assumption is that it restricts the pattern of selection on unobservables – given by the shape of $MTE(X, \cdot)$ – to be the same across individuals with different observable characteristics, $X$. It rules out the possibility that $MTE(X, \cdot)$ has a different slope depending on the value of $X$ (level shifts can be accommodated). Sensitivity analysis in Appendix E conditions on subsamples defined by different variables in $X$. Reassuringly, the shape of selection and the estimated treatment parameters do not vary much across the different subsamples, suggesting that assumption (A-2)′ is not overly strong in our setting.

We estimate the MTE using two methods: a fully parametric approach that specifies the joint distribution of unobservables and a semiparametric approach that leaves the joint distribution of unobservables unspecified and estimates $\frac{\partial \mathbb{E}[Y|X=x,P(z)=p]}{\partial p}$ using local polynomial modeling. The parametric approach is straightforward: We refer the reader to Appendix C for details. The semiparametric approach, the LIV estimator of Heckman and Vytlacil 2001a, involves estimating $\frac{\partial \mathbb{E}[Y|X=x,P(z)=p]}{\partial p}$ using the partially linear model estimator of Robinson (1988). To understand this method note that assumptions (A-1), (A-2)′, (A-3) - (A-5), together with the assumption that the outcome models in (1) and (2) are linear, yields a conditional expectation function for $Y$ that is linear in $X$ and $XP$ and nonlinear in the propensity score, $P$:

$$\begin{aligned}\mathbb{E}[Y|X=x,P(z)=p] =& \mathbb{E}[Y_0 + D(Y_1 - Y_0)|X=x,P(z)=p] \\ =& X\beta_0 + \mathbb{E}[DX(\beta_1-\beta_0)|X=x,P(z)=p] + \\ & \mathbb{E}[U_0 + D(U_1 - U_0)|X=x,P(z)=p] \\ =& X\beta_0 + PX(\beta_1-\beta_0) + K(P)\ ,\end{aligned}$$

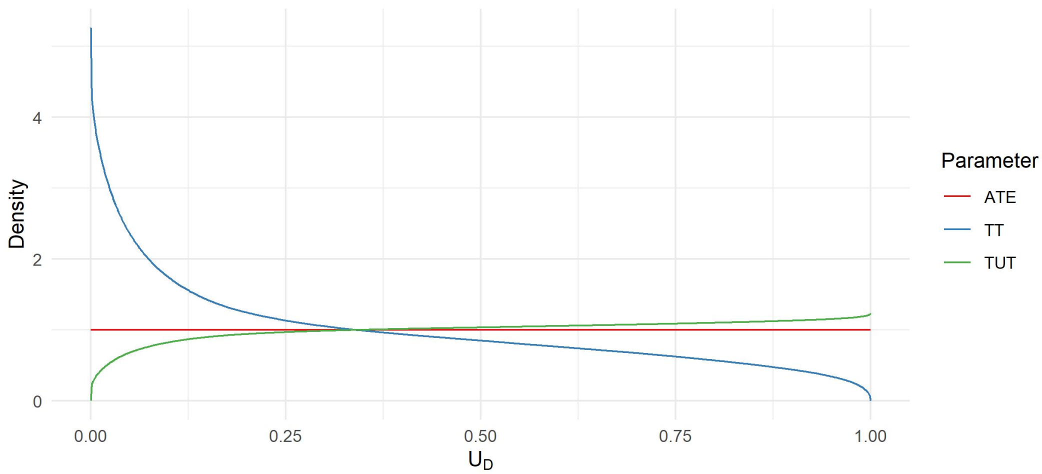
where $K(P)$ is some nonparametric function of the propensity score. This form for the conditional expectation means that the marginal treatment effect is given by

$$MTE(X,U_D) = X(\beta_1-\beta_0) + \frac{\partial K(P)}{\partial P}\Big|_{p=U_D}\ . \tag{10}$$

The Robinson (1988) semiparametric estimator of (10) entails two steps. First, the estimated propensity score, $P$, is partialed out of the other variables by running nonparametric regressions of $Y$, $X$, and $PX$ on $P$. Then the residualized $Y$ is regressed linearly on the residualized $X$ and $PX$ to obtain estimates of $\beta_0$ and $\beta_1-\beta_0$. In the second step, the derivative of the conditional expectation of $\hat{Y} \equiv Y - X\hat{\beta}_0 - XP(\hat{\beta}_1-\hat{\beta}_0)$ with respect of $P$ is estimated nonparametrically to obtain an estimate of $\frac{\partial K(P)}{\partial P}$.

All nonparametric regressions are estimated using local polynomial regression. Following the recommendations in Fan and Gijbels (1996), we use local linear regression to estimate the conditional expectations in the first stage and local quadratic regression to estimate $\frac{\partial K(P)}{\partial P}$ in the second stage. A single bandwidth, chosen using the plug-in estimator of Fan and Gijbels (1996), is used for all nonparametric regressions. This bandwidth, which aims to minimize the Integrated Mean Square Error (IMSE) in the final nonparametric regression, depends negatively on the function's curvature (second derivative) and on the density of the data, and positively on the conditional

Figure 1
Treatment Parameter Weights




The figure shows the distribution of weighting functions used to construct three standard treatment parameters. The average treatment effect (ATE) integrates the MTE with respect to the unit uniform distribution. The average effect of treatment on the treated (TT) integrates the MTE with respect to the distribution of $U_D$ conditional on attendance in telesecundarias, $f_{U_D,X|D=1}(x, u_D \mid D = 1)$, while the average effect of treatment on the untreated (TUT) integrates the MTE with respect to the distribution of $U_D$ conditional on attendance in traditional schools, $f_{U_D,X|D=0}(x, u_D \mid D = 0)$.

variance of the outcome variable. The plug-in method selects a bandwidth of 0.317.

We estimate treatment parameters by integrating the semiparametric $MTE(X, U_D)$ with respect to the appropriate distributions in equations (7) - (9) using the simulation method introduced in Carneiro, Lokshin, and Umapathi (2017). The simulation approach, which is only valid under assumption (A-2)′, involves creating an equally-spaced grid for $U_D$ for each individual and averaging $MTE(X, U_D)$ for the values of $U_D$ on the grid that are less than that individual's propensity score, $P(Z)$, for TT, greater than that individual's propensity score for TUT, and between $P(Z_a)$ and $P(Z_{a'})$ for $PRTE_{a',a}$. Figure 1 displays the densities used to compute ATE, TT, and TUT, plotted as a function of $U_D$. The figure shows that ATE uniformly samples individuals with all levels of $U_D$, while TT oversamples individuals with low $U_D$, and TUT oversamples individuals with high $U_D$.

## V. Data

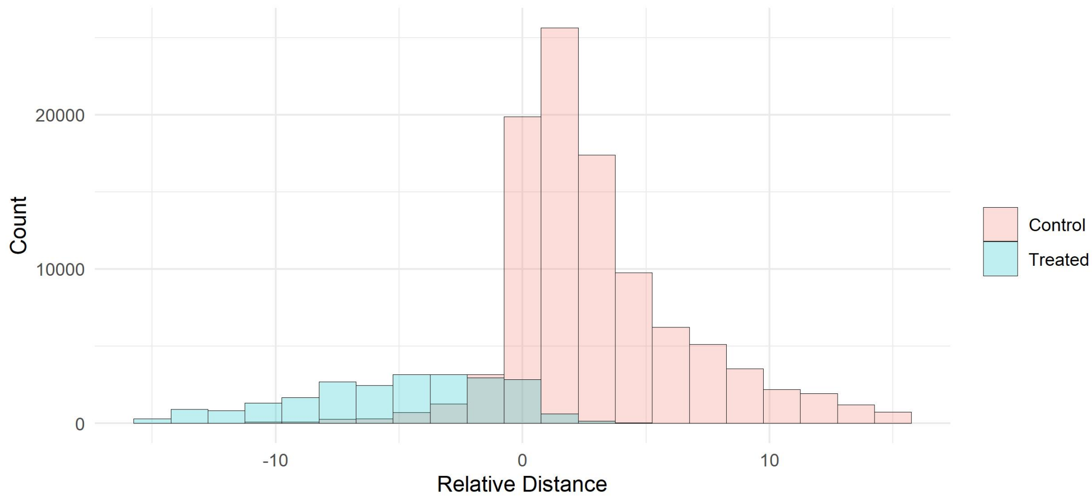
We examine the cohort of students who were in the sixth grade in 2007/08, and combine data on them from three different sources. The first is an administrative data set comprising the universe of student scores on the ENLACE exams in math and Spanish. The ENLACE exams, administered to Mexican schoolchildren in grades three through nine between 2006 and 2013, test end-of-year academic knowledge in math, Spanish, and a rotating third subject. They are low-stakes exams, with 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, but exam scores were not re-standardized in subsequent years. They are scored using item response theory, which allows improvements in test scores to be interpreted as learning gains. In addition to test scores, the ENLACE data contains information on the age, gender, conditional cash transfer status, school attendance, school ID, and school type for each student.

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 secondary school surveys produces an exceptionally rich data set that is representative of the population of Mexican schoolchildren.

We combine these two data sources with information on the latitude and longitude of each primary and secondary school and calculate the distance (in kilometers) between the primary school each student attends and the nearest secondary school of each type. We subtract the distance to the nearest traditional school from the distance to the nearest telesecundaria to obtain a measure of relative distance. This relative distance measure serves as an exclusion restriction that affects the choice of school, $D$, but does not affect outcomes, $Y_1$ and $Y_0$, directly. While it might be preferable to measure distance from the student's actual home (rather than primary school) to each secondary school, this would rely on information that is not present in any of our sources. In the next section, we show that the instrument we construct is highly predictive of attendance in telesecundarias.

Figure 2 presents the distribution of the instrument by school attendance. A negative value on the $x$-axis indicates that a telesecundaria is closer, while a positive value indicates that a traditional school is closer. The figure reveals that students mostly attend the school that is closer, but when the two schools are equally close (relative

Figure 2
Relative Distance to Telesecundaria




The figure plots a histogram of the instrumental variable by treatment status. Control units refer to students in traditional schools, while treated units refer to students in telesecundarias. The instrument is the difference between two measures of distance. The first is the distance from the student's primary school to the nearest telesecundaria, while the second is the distance from the student's primary school to the nearest traditional school. Relative distance is negative whenever telesecundarias are closer.

distance $= 0$), many more students attend traditional schools.

We impose minimal sample selection criteria. First, we eliminate the roughly six percent of students who drop out between the sixth and seventh grades.[12] We also eliminate the eight percent of students who attend private schools. There is little overlap in the distributions of family income for students who attend private schools versus telescundarias, so we do not consider private school to be a feasible alternative for the vast majority of Mexican schoolchildren who are considering attending a telesecundaria. We remove students from three states – Guerrero, Michoacan, and Oaxaca – that have limited data on ENLACE exam scores. Administration of the ENLACE exams was opposed by the teachers unions in these states.

Finally, we omit from our analysis students whose relative distance measure lies outside the middle 99% of the distribution and students who attend a secondary school more than fifteen km from their primary school. This is done for two reasons. First, we want to consider students who have a choice set consisting of two feasible alternatives, and so we drop students with only one nearby school. Second, movements

[12] Section VII. discusses how omitting dropouts influences the interpretation of treatment effects corresponding to counterfactual policies that increase access to telesecundarias.

Table 2
Variables Used in Estimation

| Variable | Definition |
|---|---|
| $Y$ | Math score in 7th grade, Spanish score in 7th grade. |
| $X$ | **Parent:** Mother's education, family income, Prospera status, rural residence, residence in Northern state, Spanish-speaking household, number of books in the home, whether the family has access to a computer.<br>**Child:** Math score in 6th grade, Spanish score in 6th grade, age, sex, number of siblings. |
| $Z\backslash X$ | Relative distance between the nearest telesecundaria and the nearest traditional school. |

to a distant school district, if they are done with knowledge of $(U_1, U_0)$, could violate assumption (A-2)′. We end up with a sample of $122,689$ students.

Table 2 lists the outcome variables ($Y$), covariates ($X$), and the instrumental variable ($Z\backslash X$) that we use in our empirical analysis. Apart from the outcome variables, sixth grade test scores, age, number of siblings, and the instrument, all variables are categorical.

Table B-2 in Appendix B presents summary statistics for these variables by school type. The table reveals that students who attend telesecundarias are disadvataged according to a wide range of metrics relative to students who attend traditional schools. They are disproportionately beneficiaries of the conditional cash transfer program Prospera, they come from poorer household with less educated mothers, and they fare worse academically in the year prior to secondary school.[13] Nevertheless, they make up over half of the gap in Spanish and nearly the entire gap in math relative to their peers at traditional secondary schools after just a year of secondary school.

*A. Is Relative Distance a Valid Instrument?*

Identification of the marginal treatment effect requires that the instrument satisfy assumptions (A-1) and (A-2)′. The first assumption, that relative distance predict attendance in telesecundarias conditional on observable covariates, $X$, is easily verified. Table 3 displays the average marginal effects of each variable in the propensity score model on the probability of attending a telesecundaria (estimated via probit). The av-

[13]The conditional cash transfer program in Mexico began in 1997 and has been has been called PROGRESA, Oportunidades, and Prospera. The main educational component of the program is a cash transfer that families receive if their child is enrolled in school. Parker and Todd (2017) provide a review of the literature on the effects of conditional cash transfer in Mexico and conclude that it has been effective in increasing school enrollment, reducing grade retention, and increasing educational attainment.

erage marginal effect of relative distance is an increase of 3.4% in the probability of attendance per kilometer, and it is highly significant.

The second assumption, that the instrument be independent of unobservable variables in the outcome and selection equations $(U_1, U_0, V)$, is more difficult to verify. In what follows, we discuss potential threats to instrumental exogeneity.

Since our specification controls for lagged test scores in addition to family and child characteristics, any threats to instrumental exogeneity must be caused by a correlation between relative distance and unobserved *time-varying* determinants of student outcomes that occur between the sixth and seventh grades. Such a correlation could occur if parents knew their child's realizations of $U_1$ and $U_0$ and moved to be closer to the school with the higher unobserved outcome. As discussed in the previous section, we drop from the sample any children who attend a secondary school more than fifteen kilometers from their primary school in an effort to eliminate this threat to exogeneity.

Another threat to identification could occur if distance traveled to school directly caused worse academic performance, perhaps through fatigue or a reduced ability to concentrate. To alleviate this concern, we analyze an augmented specification that controls for the distance actually traveled to secondary school in outcome equations (1) and (2). In this case, identification of the MTE is achieved by variation in the distance to the school that is not attended. The results of this augmented specification, presented in Appendix E, differ minimally from our main results.

Finally, the MTE can only be nonparametrically identified where the propensity score has positive support for both treated and untreated individuals. The region of common support in this paper, depicted in Figure 3, is the full $[0, 1]$ interval, owing to both the large sample size and the strong instrument. Each decile of the propensity score's distribution contains over $350$ treated and control individuals, permitting precise estimation of the MTE at both the interior and boundaries of its support.

## VI. Empirical Results

### *A. Main Findings*

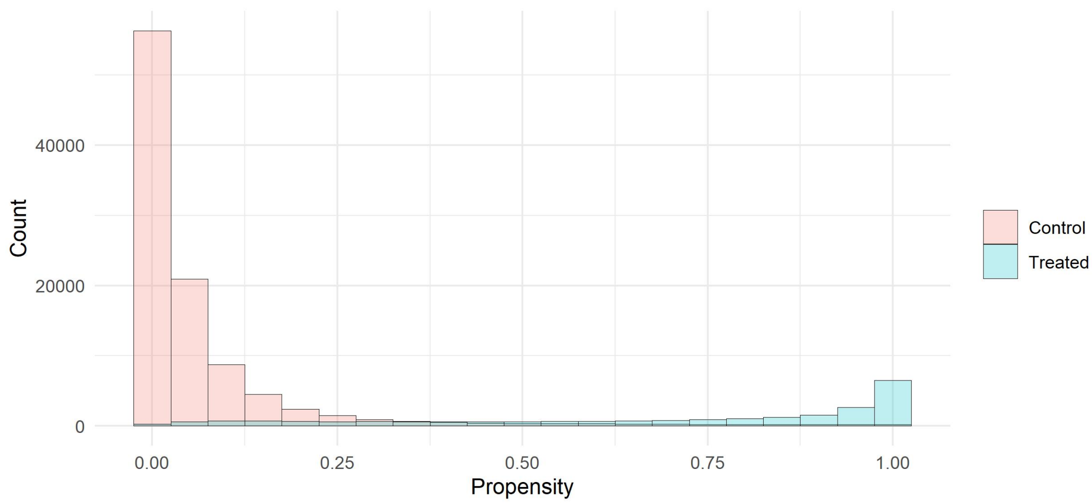
This section presents the main findings of the paper. Figure 4 shows the estimated MTEs for seventh grade math scores evaluated at mean values of $X$. The parametric and semiparametric MTEs are plotted side-by-side with 90% confidence bands in grey. Figure 5 repeats the analysis with Spanish as the outcome variable. In both figures, the horizontal axis measures the latent variable, $U_D$, while the vertical axis measures

Table 3
Propensity Score Model

| | Average Derivative | Standard Error |
|---|---|---|
| Relative Distance | -0.0341 | (0.0004) |
| Math Score (6th Grade) | -0.0060 | (0.0010) |
| Spanish Score (6th Grade) | -0.0110 | (0.0010) |
| Age | 0.0148 | (0.0010) |
| Siblings | 0.0041 | (0.0003) |
| Female | 0.0035 | (0.0014) |
| Prospera | 0.0269 | (0.0015) |
| Computer | -0.0221 | (0.0017) |
| Rural Residence | 0.0189 | (0.0017) |
| Northern State | -0.0452 | (0.0028) |
| Speaks Language other than Spanish at Home | 0.0058 | (0.0024) |
| Family Income : Low | -0.0114 | (0.0017) |
| Family Income : Medium | -0.0164 | (0.0020) |
| Family Income : High | -0.0185 | (0.0027) |
| Mother's Education : Middle | -0.0160 | (0.0017) |
| Mother's Education : Secondary | -0.0298 | (0.0024) |
| Mother's Education : Post-Secondary | -0.0091 | (0.0046) |
| Books in the Home : 20 | -0.0103 | (0.0016) |
| Books in the Home : 50 | -0.0177 | (0.0022) |
| Books in the Home : $\geq$100 | -0.0109 | (0.0029) |

The table shows the average marginal effects of each variable in the propensity score model for telesecundaria attendance. The instrument, relative distance, is computed as the difference between two distance measures. The first is the distance from the student's primary school to the nearest telesecundaria, and the second is the distance from the student's primary school to the nearest traditional school. Relative distance is negative whenever telesecundarias are closer. All other variables are included in the outcome models for seventh grade test scores. The omitted category in each of Family Income, Mother's Education, and Books in the Home is the lowest one. Computer is a binary variable that equals one if the student has access to a computer at home. Standard errors are calculated via 250 bootstrap replications.

Figure 3
Estimated Propensity Score by Treatment Status

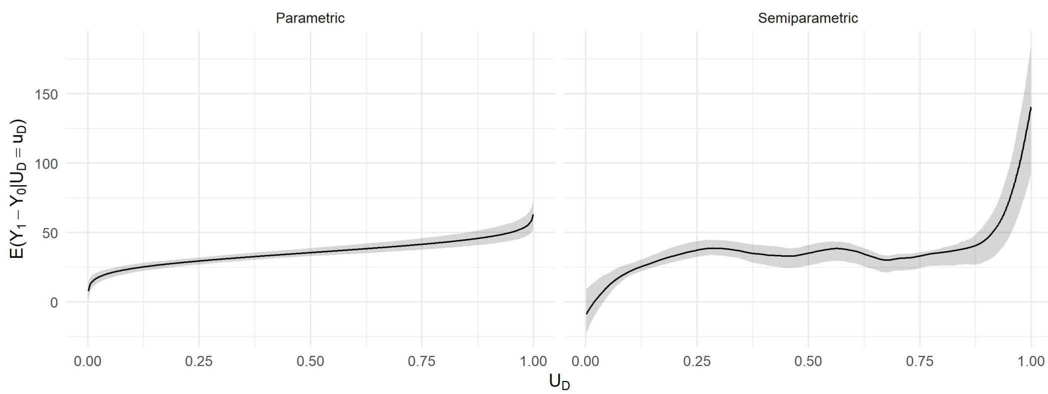



The figure plots histograms of estimated propensity scores by treatment status. Control refers to students in traditional schools, while treated refers to students in telesecundarias. The propensity score models telesecundaria attendance as a function the child's sixth grade math and Spanish scores, age, sex, number of siblings, the mother's education, family income, number of books in the home, family access to a computer, Prospera status, rural residence, residence in a Northern state, whether the family speaks Spanish at home, and the relative distance between the nearest telesecundaria and nearest traditional school. Relative distance is the difference between the distance from the student's primary school to the nearest telesecundaria and the distance from the student's primary school to the nearest traditional school. The propensity score model is estimated via probit. The figure shows that the common support of propensity scores across treated and control units is the full $[0, 1]$ interval.

the expected benefit to attending a telesecundaria relative to a traditional school for students with that level of $U_D$, $\mathbb{E}[Y_1 - Y_0 | X = \overline{x}, U_D = u_D]$. The MTEs in both figures are precisely estimated.

The semiparametric figures reveal a pattern of negative sorting on gains. Students who, on the basis of unobservables, are most likely to attend telesecundarias (low $U_D$) experience value added that is indistinguishable from zero for both math and Spanish. As $U_D$ increases, students are less likely to attend telesecundarias but their benefits from attendance increase significantly. This increase is monotonic for math, but nonmonotonic for Spanish. The Spanish MTE reaches a local minimum near $U_D = 0.35$. From this point onward, as $U_D$ increases, average benefits decrease in Spanish and stay roughly constant in math until about $U_D = 0.75$, after which there are large (but noisy) gains to telesecundaria attendance in both math and Spanish. Both MTEs

Figure 4
Marginal Treatment Effect: Math

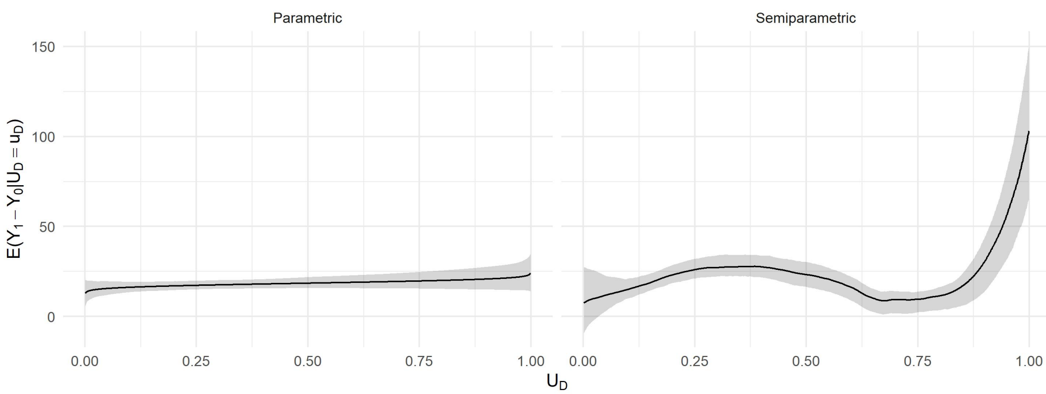



The dependent variable in the outcome equation is the raw score on the seventh grade ENLACE math exam. The outcome equations include controls for sixth grade math and Spanish scores, the age, sex, and number of siblings of the child, family income, mother's education, the number of books in the home, whether the family has access to a computer at home, and dummies for rural residence, residence in a Northern state, speaking a language other than Spanish at home, and participation in the conditional cash transfer program Prospera. The school choice model includes the same controls and also includes the relative distance between the nearest telesecundaria and nearest traditional secondary school as an exclusion restriction. The school choice model is estimated via probit. The parametric MTE is estimated using a two-step Least Squares method that controls for selection on unobservables. The semiparametric MTE is estimated using Local Quadratic Regression and an Epanechnikov kernel with a bandwidth of 0.317. The bandwidth is chosen to minimize the Integrated Mean Square Error in the final stage of estimation. Both MTEs are evaluated at the mean value of the covariates, $X = \overline{x}$. Ninety percent confidence intervals are computed via nonparametric bootstrap with 250 draws.

show that the treatment effects are nonnegative for all values of $U_D$ and larger for those values of $U_D$ that make students less likely to attend.

Tables 4 and 5 present estimates of standard treatment parameters for math and Spanish, respectively. All treatment parameters are positive, underscoring the findings from the MTE curve that telesecundaria attendance is beneficial for a large majority of students. Standard errors reveal that the treatment parameters are precisely estimated and are significant at conventional levels of significance. The semiparametric estimate of the ATE for math indicates that a randomly selected student would be expected to perform $36.7$ points better in mathematics in the seventh grade had she attended a telesecundaria instead of a traditional school. As the standard deviation on the 7th grade math ENLACE exam is $100.7$, this represents an increase of $0.364$ standard deviations. The estimates of TT ($27.5$ points) are smaller than those of

Figure 5
Marginal Treatment Effect: Spanish




The dependent variable in the outcome equation is the raw score on the seventh grade ENLACE Spanish exam. The outcome equations include controls for sixth grade math and Spanish scores, the age, sex, and number of siblings of the child, family income, mother's education, the number of books in the home, whether the family has access to a computer at home, and dummies for rural residence, residence in a Northern state, speaking a language other than Spanish at home, and participation in the conditional cash transfer program Prospera. The school choice model includes the same controls and also includes the relative distance between the nearest telesecundaria and nearest traditional secondary school as an exclusion restriction. The school choice model is estimated via probit. The parametric MTE is estimated using a two-step Least Squares method that controls for selection on unobservables. The semiparametric MTE is estimated using Local Quadratic Regression and an Epanechnikov kernel with a bandwidth of 0.317. The bandwidth is chosen to minimize the Integrated Mean Square Error in the final stage of estimation. Both MTEs are evaluated at the mean value of the covariates, $X = \overline{x}$. Ninety percent confidence intervals are computed via nonparametric bootstrap with 250 draws.

TUT ($38.8$ points), reinforcing the finding of negative sorting on gains for math. These numbers indicate that the average student currently attending a telesecundaria gains $0.273$ standard deviations while the average student currently attending a traditional school would gain $0.385$ standard deviations on the math ENLACE exam if she were to attend a telesecundaria. The semiparametric estimates of ATE, TT, and TUT for Spanish are $23.0$, $17.3$, and $24.3$ points, which represent $0.230$, $0.173$, and $0.243$ standard deviations, respectively. These estimates, while smaller than those for math, are significantly different from zero. In addition, while the Spanish MTE is nonmonotonic, the finding that TT<ATE<TUT indicates that the pattern of sorting on gains is primarily negative. Appendix D shows that this negative pattern is statistically significant and offers two possible explanations that are consistent with the data and with perceptions of schools in Mexico. Sensitivity analysis in Appendix E shows that our

Table 4
Estimated Treatment Effects: Math

| | Parametric | | Semiparametric | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| Average Treatment Effect | 35.5 | (1.79) | 36.7 | (2.47) |
| Treatment on the Treated | 30.5 | (1.48) | 27.5 | (1.88) |
| Treatment on the Untreated | 36.7 | (1.91) | 38.8 | (2.81) |

The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on raw scores on the seventh grade math ENLACE exam. The three treatment parameters are obtained by integrating the MTE with respect to the densities displayed in Figure 1. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the semiparametric MTE. Standard errors are obtained by 250 bootstrap replications.

results are unlikely to be biased by unobserved confounders.

The semiparametric estimates of treatment parameters for Spanish are noticeably higher than the parametric estimates. The parametric approach forces the MTE curve to be monotonic and therefore does a worse job of fitting the MTE for Spanish, in which there is more evidence of nonmonotonicity, than for math. Parametric estimates for math are quite similar to the semiparametric estimates, owing to the fact that math's parametric MTE curve is well-centered between the maxima and minima of the semiparametric MTE curve. Even when parametric estimates differ little from the semiparametric estimates, this does not detract from the usefulness of conducting the semiparametric analysis we do here. Rather, estimating treatment effects semiparametrically is often the only way of evaluating whether parametric assumptions bias our inference on the effectiveness of telesecundarias.

*B. Impacts on Enrollment*

Our analysis shows that telesecundarias have significant causal effects on student learning between the sixth and seventh grades. One might naturally expect that they also have positive effects on other outcomes that are correlated with test scores. Navarro-Sola (2019) shows that the expansion of telesecundarias caused significant increases in labor market earnings and an increase in formal labor force participation at the expense of employment in the agricultural and informal sectors. While we do not have labor market data, we do follow students through the ninth grade, which is the final year of secondary school and the final year of compulsory education. Since

Table 5
Estimated Treatment Effects: Spanish

| | Parametric | | Semiparametric | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| Average Treatment Effect | 18.5 | (1.80) | 23.0 | (2.66) |
| Treatment on the Treated | 17.5 | (1.46) | 17.3 | (2.08) |
| Treatment on the Untreated | 18.7 | (1.93) | 24.3 | (3.00) |

The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on raw scores on the seventh grade Spanish ENLACE exam. The three treatment parameters are obtained by integrating the MTE with respect to the densities displayed in Figure 1. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the semiparametric MTE. Standard errors are obtained by 250 bootstrap replications.

telesecundarias have positive effects on student learning, it is reasonable to expect that they would also reduce rates of dropout, and this is exactly what we find.

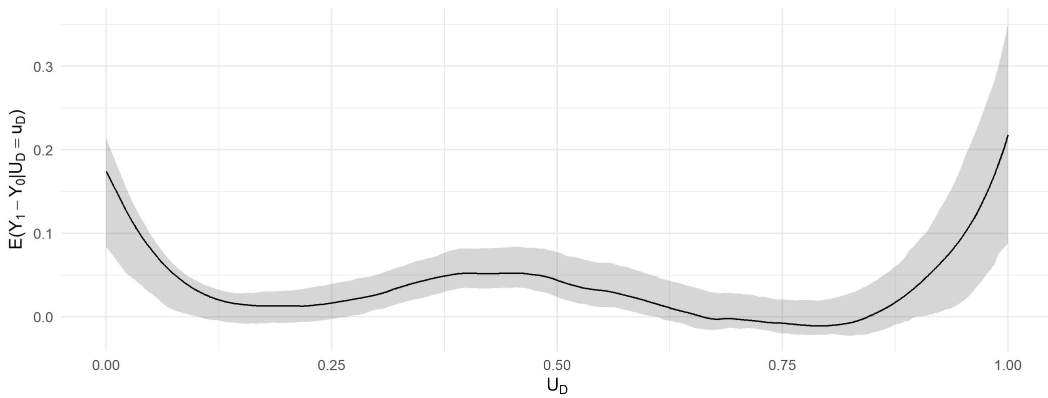
We replicate our main analysis with a dependent variable that equals one if the student attends school in the ninth grade and zero otherwise. Our sample conditions on students who enroll in secondary school, so this analysis answers the following question: Conditional on enrolling in secondary school, does enrollment in telesecundarias cause an increase in the probability of attending school in the ninth grade relative to attendance in traditional secondary schools?

We answer this question by estimating a semiparametric MTE and integrating it to compute treatment parameters.[14] Figure 6 displays the estimated MTE for this analysis. The bandwidth, 0.317, is the same as in the main analysis. There are two important findings from the figure. First, the MTE displays a pattern of nonmonotic selection: Students who are the most likely to attend telesecundarias, on the basis of having low values of $U_D$, experience larger causal effects on the probability of enrolling in the ninth grade than do students who have a moderate likelihood of attending telesecundarias. Students whose unobservables make them least likely to attend telesecundarias also experience increases in the probability of attending the ninth grade. Second, the figure reveals that the average causal effect of telesecundaria attendance on ninth grade enrollment is nonnegative for nearly all values of $U_D$.

Table 6 presents estimates of ATE, TT, and TUT for the effect of telesecundaria

[14]The parametric model is misspecified in the case of a binary outcome.

Figure 6
Marginal Treatment Effect: Ninth Grade Enrollment




The dependent variable in the outcome equation is a binary variable that equals one if the individual enrolls in the ninth grade and zero otherwise. The outcome equations include controls for sixth grade math and Spanish scores, the age, sex, and number of siblings of the child, family income, mother's education, the number of books in the home, whether the family has access to a computer at home, and dummies for rural residence, residence in a Northern state, speaking a language other than Spanish at home, and participation in the conditional cash transfer program Prospera. The school choice model includes the same controls and also includes the relative distance between the nearest telesecundaria and nearest traditional secondary school as an exclusion restriction. The school choice model is estimated via probit. The MTE is estimated semiparametrically using Local Quadratic Regression and an Epanechnikov kernel with a bandwidth of 0.317. The MTE is evaluated at the mean value of the covariates, $X = \overline{x}$. Ninety percent confidence intervals are computed via nonparametric bootstrap with 250 draws.

attendance on ninth grade enrollment. These are computed by integrating the semiparametric MTE with respect to the relevant densities in Figure 1. The average treatment effect of telesecundaria attendance on ninth grade completion is $3.6$ percentage points. The average effect is larger for the treated than the untreated, $7.2$ versus $2.8$ percentage points. The mean of the depedendent variable, which represents the probability of enrolling in the ninth grade conditional on attending secondary school, is 0.915. The MTE curve in Figure 6 and the treatment parameters in Table 6 show that, among students who enroll in any type of secondary school, telesecundarias have a positive causal effect on educational attainment. This is consistent with a story in which telesecundarias cause improvements in academic performance, which in turn makes dropout less likely.

Table 6
Estimated Treatment Effects: Ninth Grade Enrollment

| | Estimate | Standard Error |
|---|---|---|
| Average Treatment Effect | 0.036 | (0.0080) |
| Treatment on the Treated | 0.072 | (0.0076) |
| Treatment on the Untreated | 0.028 | (0.0091) |
| Mean : 9th Grade Enrollment | 0.915 | |

The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on ninth grade enrollment. The three treatment parameters are obtained by integrating the semiparametric MTE for ninth grade enrollment with respect to the densities displayed in Figure 1. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the MTE. Mean 9th grade enrollment is estimated conditional on attending secondary school for at least one year. Standard errors are obtained by 250 bootstrap replications.

## VII. Counterfactuals

The finding of a pattern of negative sorting on gains suggests that policies that can expand access to telesecundarias will raise test scores in Mexico in relation to the status quo. We therefore evaluate the effects of two counterfactual policies that are likely to induce more children with high latent resistance to treatment (high $U_D$) to switch from traditional secondary schools to telesecundarias.

The effects of counterfactual policies can be evaluated by integrating the MTE with respect to the probability distribution induced by the proposed policy. A baseline policy, $a$, is characterized by a particular distribution of the instrument, $Z^a$. The move from policy $a$ to a new policy $a'$ corresponds to a shift in the distribution of the instrument from $F_{Z^a}$ to $F_{Z^{a'}}$. This shift induces some students to attend telesecundarias who would not otherwise attend. The treatment effect for students induced to switch attendance from traditional schools to telesecundarias as a result of the policy is given $PRTE_{a',a}$, which is positive if these students learn more at telesecundarias than at traditional schools.

*A. Increasing Access to Telesecundarias*

We consider a class of counterfactual policies that expand access to telesecundarias so that the relative distance to them decreases for all students, $Z^{a'} \leq Z^{a}$. As our sample omits individuals who drop out between the sixth and seventh grades, we will not be able to say anything about the distribution of test scores for students who are induced to attend telesecundarias instead of dropping out as a result of the counterfactual policy. Our PRTE estimates apply only to the population of students who were already attending secondary school under the baseline policy in 2008. We measure only the achievement effects, not the enrollment effects, of each policy.[15] We acknowledge that enrollment effects may be significant, although they are bounded above by the 6% of students who dropped out between the sixth and seventh grades in 2008 (Table B-1).

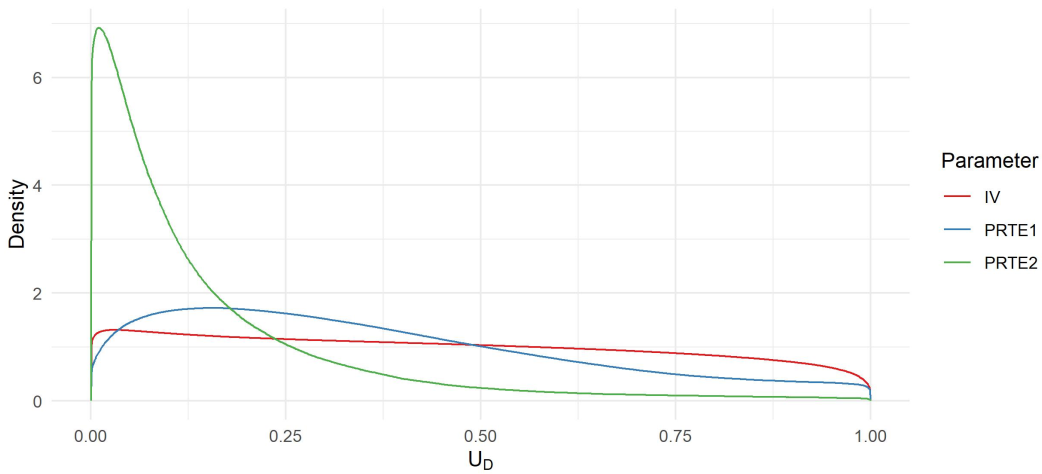
We consider two counterfactual policies. The first is a hypothetical policy that reduces the relative distance to telesecundarias by five km for every student. This policy, which would entail moving traditional schools farther away for students whose nearest telesecundaria is under five km, is infeasible but provides an instructive example of the gains to a policy that can drastically raise telesecundaria attendance.

The second counterfactual is a feasible school-building policy that constructs a telesecundaria directly adjacent to the eighteen percent of primary schools that have no telesecundaria within a five km radius. This has the effect of reducing the distance between primary schools and telesecundarias to zero for all students who formerly had only a distant telesecundaria.

Figure 7 displays the probability distributions of $U_D$ corresponding to these two treatment parameters. The first counterfactual induces a probability distribution unlike any of the other treatment parameters: The distribution is considerably less skewed than TT and has the most mass around $U_D = 0.30$. As a result, it oversamples individuals with relatively large values of $\mathbb{E}[Y_1 - Y_0 \mid U_D]$ and produces a large positive value for PRTE. The distribution corresponding to the second counterfactual is closer to the distribution for TT. Relative to TT, it oversamples individuals with low $U_D$ and undersamples those with high $U_D$. The figure also shows the weights corresponding to a Two-Stage Least Squares regression that uses relative distance as an instrument for telesecundaria attendance. The IV weights do not correspond to either of the PRTE

[15] A revealed preference argument demonstrates that no students will transition from dropping out to attendance in traditional schools as a result of the policy. The counterfactual policy, however, may induce students to transition from dropout to telesecundarias. The treatment effects we estimate therefore pertain to the set of students attending secondary school under the status quo.

Figure 7
Counterfactual Treatment Parameter Weights




The figure shows the distribution of weighting functions used to construct estimates of Policy-Relevant Treatment Effects (PRTEs) for two policies discussed in section VII. as well as the weights induced by traditional IV, which uses relative distance as an instrument for telesecundaria attendance. PRTE1 corresponds to the weights induced by a counterfactual policy that reduces relative distance between telesecundarias and traditional secondary schools by five km. PRTE2 corresponds to the weights induced by a counterfactual policy that constructs telesecundarias adjacent to all primary schools that do not have a telesecundaria within a five km radius. The IV weights are positive everywhere but do not correspond to any of the counterfactuals under consideration.

weights, nor do they correspond to the weights for ATE, TT, or TUT, highlighting the importance of using the MTE to conduct counterfactual analysis.

Table 7 presents estimates of the PRTEs for both math and Spanish alongside the IV estimate from a Two-Stage Least Squares regression with relative distance to telesecundarias as the excluded instrument. All parameters are precisely estimated and statistically significant at conventional levels of significance. We find that the first policy causes a 35.5 point increase in mean math scores and a 23.8 point increase in mean Spanish scores in the seventh grade. The second policy also causes improvements, but they are smaller, 21.3 points for math and 16.4 points for Spanish, owing to the less dramatic nature of the policy. The IV estimate for math, 25.1 points, lies in the middle of the treatment effects of the two counterfactual policies, while the IV estimate for Spanish, 14.6 points, falls below the two PRTEs.

The results underscore the positive academic gains that can result from policies that make telesecundarias more accessible. In addition, the considerable variation across our many estimated treatment effects validates our approach of estimating

Table 7
Counterfactual Treatment Effects

| | PRTE1 | PRTE2 | IV |
|---|---|---|---|
| Math | 35.5 | 21.3 | 25.1 |
| | (1.96) | (3.13) | (1.42) |
| Spanish | 23.8 | 16.4 | 14.6 |
| | (1.99) | (3.91) | (1.44) |

The table displays treatment parameters corresponding to two counterfactual policies discussed in section VII. PRTE1 corresponds to a counterfactual policy that reduces relative distance between telesecundarias and traditional secondary schools by five km. PRTE2 corresponds a counterfactual policy that constructs telesecundarias adjacent to all primary schools with no telesecundaria within a five km radius. PRTE1 and PRTE2 are calculated using the semiparametric MTE combined with the simulation method of Carneiro, Lokshin, and Umapathi (2017). Standard errors for these two treatment parameters are displayed in parentheses and are obtained by 250 bootstrap replications. The IV estimates are obtained by a Two-Stage Least Squares regression that uses relative distance as an instrument for telesecundaria attendance.

MTE curves for math and Spanish and using them to forecast specific policies rather than relying on IV for policy evaluation.

*B. Distributional Treatment Effects*

Many societies are characterized by deep cleavages, of geography, race, or language. In what follows we show how the MTE can be used to evaluate the distributional consequences of policies for a particular division. We apply this approach to estimate the effect of telesecundarias on two important divisions in Mexico: the test score gaps between rural and urban students and between those who speak Spanish at home and those who speak another language.

Let the random variable $G \in \{0, 1\}$ denote the subdivision to which an individual belongs (rural versus urban, for example). We estimate the effect of telesecundarias on the test score gap for the division defined by $G$ by considering a counterfactual policy that eliminates telesecundarias. The causal effect of telesecundarias on the test score gap, $TE^G$, is the difference between the gap under the status quo policy, $a$, and the new policy, $a'$:

$$TE^G = \mathbb{E}[Y^a \mid G=1] - \mathbb{E}[Y^a \mid G=0] - \{\mathbb{E}[Y^{a'} \mid G=1] - \mathbb{E}[Y^{a'} \mid G=0]\} \,. \tag{11}$$

We can write $TE^G$ as a function of the average treatment effect on the treated and the probability of obtaining treatment for each separate subgroup as follows:

$$\begin{aligned}
TE^G =& \mathbb{E}[Y_0 + D^a(Y_1 - Y_0) \mid G=1] - \mathbb{E}[Y_0 + D^a(Y_1 - Y_0) \mid G=0] - \\
& \{\mathbb{E}[Y_0 + D^{a'}(Y_1 - Y_0) \mid G=1] - \mathbb{E}[Y_0 + D^{a'}(Y_1 - Y_0) \mid G=0]\} \\
=& \mathbb{E}[D^a(Y_1 - Y_0) \mid G=1] - \mathbb{E}[D^a(Y_1 - Y_0) \mid G=0] - \\
& \{\mathbb{E}[D^{a'}(Y_1 - Y_0) \mid G=1] - \mathbb{E}[D^{a'}(Y_1 - Y_0) \mid G=0]\} \\
=& \mathbb{E}[D^a(Y_1 - Y_0) \mid G=1] - \mathbb{E}[D^a(Y_1 - Y_0) \mid G=0] \\
=& \mathbb{E}[Y_1 - Y_0 \mid G=1, D^a=1]\mathbb{P}(D^a=1 \mid G=1) - \\
& \mathbb{E}[Y_1 - Y_0 \mid G=0, D^a=1]\mathbb{P}(D^a=1 \mid G=0) \,,
\end{aligned} \tag{12}$$

where the third quality follows because $D^{a'} = 0$ for everyone when telesecundarias are eliminated, and the fourth equality follows from the law of iterated expectations. Equation (12) makes clear that telesecundarias can reduce the test score gap in two ways: the first through unequal treatment effects across different values of $G$, and the second through differential treatment probabilities.

We estimate distributional treatment effects for two prominent divisions in Mexico, between rural and urban residents and between students who speak Spanish at home versus those who speak an indigenous language. Table 8 shows both the size of the gaps and the distributional treatment effects for these two divisions. All distributional treatment effects are computed by integrating the MTE curves for math and Spanish with respect to the densities of $(X, U_D)$ conditional on $G$, $f(X, U_D \mid G)$. The table shows that the urban-rural test score gap in math is $10.8$ points, or $0.107$ standard deviations, but that it would be $15.3$ points larger (142%) if telesecundarias were eliminated. The corresponding gap for seventh grade Spanish scores is $18.7$ points, or $0.187$ standard deviations, and it would be $8.0$ points larger (43%) in the absence of telesecundarias. These differences are statistically significant.

The gap in test scores between students who speak Spanish at home versus those who speak an indigenous language is $15.5$ points in math and $22.7$ points in Spanish. We find that telesecundarias also cause statistically significant reductions in these test score gaps. The gaps would be $6.2$ points (40%) larger in math and $3.7$ points (16%)

Table 8
Distributional Treatment Effects

| | Math | | Spanish | |
|---|---|---|---|---|
| | Gap | Treatment Effect | Gap | Treatment Effect |
| Urban - Rural | 10.8 | 15.3 | 18.7 | 8.0 |
| | | (1.10) | | (1.07) |
| Spanish - Indigenous | 15.5 | 6.2 | 22.7 | 3.7 |
| | | (1.20) | | (0.40) |

The table shows the effects of telesecundarias on reducing the test score gaps between urban and rural residents and between students who speak Spanish at home and those who speak an indigenous language at home. The dependent variables are raw scores on the seventh grade math and Spanish ENLACE exams. A positive treatment effect indicates that the test score gap would be larger in the absence of telesecundarias. This analysis conditions on the set of students who currently attend secondary school. It does not consider how the elimination of telesecundarias might cause differential rates of dropout in telesecundarias relative to traditional schools. The treatment effects are obtained by integrating the semiparametric MTE with respect to the density of $(X, U_D)$ conditional on the urban, rural, Spanish-speaking, and indigenous-speaking subsamples, respectively. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the MTE. Standard errors, in parentheses, are obtained by 250 bootstrap replications.

larger in Spanish if telesecundarias were eliminated. Hence, not only do telesecundarias improve educational outcomes for the students who attend them, they do so in an equitable way by reducing the rural-urban and Spanish-Indigenous test score gaps.

*C. Mechanisms*

Telesecundarias offer a bundle of characteristics – televised lectures, in-person teaching, and standardized reference texts and learning guides – and it would be a mistake to attribute all the positive gains documented in this paper to the lectures alone. In addition, this paper has documented the effectiveness of telesecundarias relative to a *specific alternative*, that of traditional secondary education in Mexico. In this section we explore probable mechanisms that could explain why telesecundarias are an effective learning technology in Mexico.

One set of potential explanations for the effectiveness of telesecundarias stems from the challenges of providing high quality traditional education in many parts of Mexico. An initial justification for introducing telesecundarias was the existence of a shortage of qualified secondary school teachers who were willing to work in remotes areas, a shortage that persists to this day (Navarro-Sola 2019). This means that the feasible traditional alternatives for many students who attend telesecundarias will likely have teachers who are poorly trained in comparison with the *telemaestros*.

An additional reason why the telesecundaria model may be effective in Mexico is that traditional secondary schools suffer from considerable teacher absenteeism. Absent teachers often mean that students have no one to teach them a particular lesson that day.[16] Our data, which contain the responses of principals to a wide range of questions concerning their schools, reveals that absenteeism is a regular occurrence. Overall, only 56% of principals state that absenteeism is rare or never a problem, but the incidence of absenteeism varies markedly by the type of secondary school. Figure 8 shows the distribution of responses to the principal survey for the schools in our data. The figure shows that principals at telesecundarias are much less likely to consider teacher absenteeism a problem at their schools in comparison with principals at traditional secondary schools. One cause of telesecundarias' effectiveness may be that their students spend more time exposed to educational content and less time with no supervision.[17]

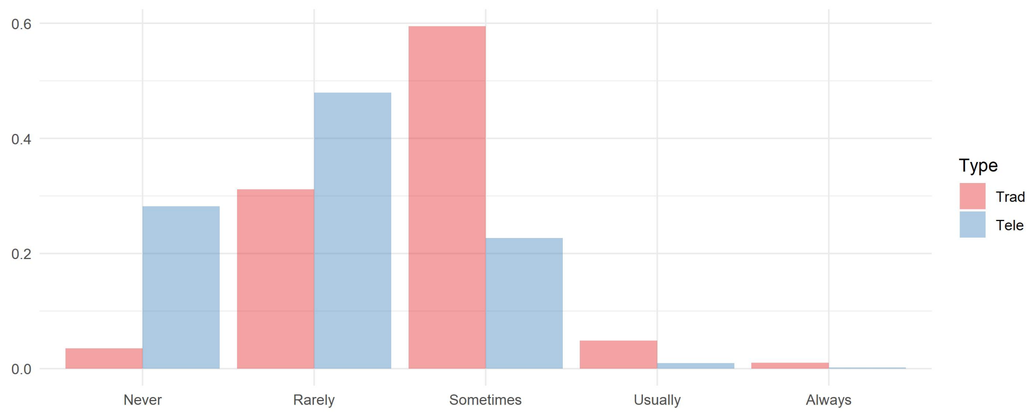
Apart from reasons stemming from the Mexican context, telesecundarias may also be effective due to intrinsic features of the program. The lectures are of uniformly high quality. The teaching materials and student learning guides are also uniform across schools, which can be an advantage wherever there is high variation in local teacher quality and teacher preparedness.[18] Lastly, because the lessons are delivered via satellite according to a pre-set schedule, students at telesecundarias are exposed to the entire curriculum planned for that school year. Classrooms at schools across the world often lag behind the anticipated pace of the curriculum, so that by the end of the year some students have not been exposed to material that the state expects them to know. Some of the gains of telesecundaria students relative to traditional students on the ENLACE exams may simply be due to exposure to more of the material tested

[16] Chaudhury et al. (2006) document high rates of teacher absenteeism across a range of developing countries. They explain that, in contrast to high-income nations, there are rarely substitutes to replace absent teachers in low-income nations, and students typically go home, wander aimlessly around the school, or join another class (which may be in a different subject or a different grade) when their teacher is absent.

[17] Duflo, Hanna, and Ryan (2012) find that reductions in teacher absenteeism have large effects on student learning.

[18] Variation in teacher quality is large across the developing world, with UNESCO (2016) reporting that the percentage of secondary school teachers who lack basic training is above 80% in 30 of 73 countries reporting data after 2012. Bau and Das (2020) document that primary school teachers in rural Pakistan have nearly as high a standard deviation on an exam designed for primary school students as the students themselves (0.87 versus 1.00) and that a nonneglible fraction of teachers score below the top students.

Figure 8
Absenteeism in Secondary Schools




The figure reports the fraction of responses by secondary school principals to the question "How often is teacher absenteeism a problem at your school?" Responses are disaggregated by school type, red for traditional secondary schools and blue for telesecundarias.

on the exam.

There is some evidence that the effectiveness of telesecundarias varies with student ability. All students benefit from the high quality of the lectures, the standardized teaching materials, and on-time exposure to the full curriculum. There may, however, be costs of telesecundarias that vary with student ability. A cost for high ability students may be that they are unable to ask more advanced questions that could enrich their understanding and that a specialized teacher could answer. Low ability students may instead need more personalized forms of instruction that do not fit into the telesecundaria lesson plans but that a teacher with experience in the field may be able to impart. Consistent with these explanations, we find that students who enter secondary school with average scores in the sixth grade are particularly likely to benefit from telesecundarias. Table 9 presents estimated treatment effects from our semiparametric model applied to three separate samples determined by conditioning on each tercile of sixth grade math scores. The table shows that students in the middle tercile of sixth grade scores experience the greatest gains in both math and Spanish from attending telesecundarias. While the treatment effects for all groups are positive, the information in the table suggests that telesecundarias are a learning technology that may be best aimed at the median student.

Table 9
Average Treatment Effects by Tercile of Sixth Grade Scores

| | Math | | Spanish | |
|---|---|---|---|---|
| | ATE | Standard Error | ATE | Standard Error |
| Bottom Third | 35.4 | (4.12) | 23.8 | (3.61) |
| Middle Third | 46.7 | (4.89) | 30.9 | (3.55) |
| Top Third | 41.8 | (6.06) | 25.2 | (5.25) |

The table shows treatment parameters corresponding to the effect of telesecundaria attendance on raw scores on the seventh grade math and Spanish ENLACE exams. The treatment effects for both math and Spanish condition on the tercile of sixth grade ENLACE scores in math. The average treatment effect (ATE) is obtained by averaging $MTE(X, U_D)$ over the values of $X$ for individuals in each subsample and the unit uniform distribution for $U_D$. Standard errors, in parentheses, are obtained through 250 bootstrap replications.

## VIII. Conclusion

In this paper we evaluate the effectiveness of a long-running distance learning program in Mexico. Our empirical approach, which combines value-added modeling with the MTE framework, simultaneously allows for rich heterogeneity in treatment effects and lets us isolate the degree of student learning that is directly attributable to attendance in telesecundarias. We find evidence of considerable heterogeneity in value added in math and Spanish caused by attending telesecundarias but that nearly all students benefit. The gains are large and correspond to a $0.365$ standard deviation increase in math scores and a $0.229$ standard deviation increase in Spanish scores after a single year of attendance. Our counterfactual simulations suggest that telesecundaria expansions would yield positive academic dividends, with the magnitude of these policy effects varying with the degree to which the policies make telesecundarias more accessible. Furthermore, we show that, by providing a high-quality alternative to students with initially low academic skills, telescundarias reduce educational inequality in Mexico. Since telesecundarias are cheaper than traditional schools, the benefits of sending a child to a telesecundaria instead of a traditional school would exceed its costs regardless of the assumptions used in any cost-benefit analysis.

While COVID-19 caused many school districts to hastily move existing lesson plans online and students to Zoom in from home, telesecundarias show that distance education can take different forms than the pandemic model. Based on the findings presented in this paper, we believe that the telesecundaria model can be an effective learning technology, especially when a few preconditions are met. In areas with low and variable teaching quality and high rates of teacher absenteeism, a combination of high quality video instruction, in-person supervision, and standardized teaching materials may help students, particularly those with abilities near the median of the distribution, learn more. Such a technology, which could be offered at a low per-unit cost by either governments or private entities, likely causes additional benefits not analyzed in this paper. We've provided evidence that, by raising knowledge, telesecundarias also raise educational attainment. Standard theories of human capital, as in Becker (2009) and Ben-Porath (1967), predict that this would have causal effects on earnings later in life. Finally, by raising the level of knowledge in isolated regions of Mexico, telesecundarias may help to raise a new generation of teachers that can alleviate the shortage this program was initially designed to address.

## References


Abdulkadiroğlu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters. 2020. "Do Parents Value School Effectiveness?" *American Economic Review* 110 (5): 1502–39.

Acta Educativa, Universidad Abierta. 2016, November 21. La Telesecundaria vs Secundarias Generales y Técnicas. https://revista.universidadabierta.edu.mx/2016/11/30/la-telesecundaria-vs-secundarias-generales-y-tecnicas/. Accessed 2021-09-24.

Ainsworth, Robert, Rajeev Dehejia, Cristian Pop-Eleches, and Miguel Urquiola. 2020. "Information, Preferences, and Household Demand for School Value Added." Technical Report, National Bureau of Economic Research.

Almagor, Lelac. 2021, June. I Taught Online School This Year. It Was a Disgrace. https://www.nytimes.com/2021/06/16/opinion/remote-learning-failure.html. Accessed: 2012-01-14.

Avitabile, Ciro, and Rafael De Hoyos. 2018. "The Heterogeneous Effect of Information on Student Performance: Evidence from a Randomized Control Trial in Mexico." *Journal of Development Economics* 135:318–348.

Bau, Natalie, and Jishnu Das. 2020. "Teacher Value Added in a Low-income Country." *American Economic Journal: Economic Policy* 12 (1): 62–96.

Becker, Gary S. 2009. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago press.

Beg, Sabrin A, Adrienne M Lucas, Waqas Halim, and Umar Saif. 2019. "Beyond the Basics: Improving Post-primary Content Delivery through Classroom Technology." *National Bureau of Economic Research*.

Behrman, Jere R, Susan W Parker, and Petra Todd. 2020. "Impacts of Prospera on Enrollment, School Trajectories and Learning." *Working Paper*.

Behrman, Jere R, Susan W Parker, Petra E Todd, and Weilong Zhang. 2021. "Prospering through Prospera: CCT Impacts on Educational Attainment and Achievement in Mexico." Technical Report, Penn Institute for Economic Research.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (4, Part 1): 352–365.

Bianchi, Nicola, Yi Lu, and Hong Song. 2020. "The Effect of Computer-Assisted Learning on Students' Long-Term Development." *National Bureau of Economic Research*.

Björklund, Anders, and Robert Moffitt. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *The Review of Economics and Statistics*, pp. 42–49.

Boardman, Anthony E, and Richard J Murnane. 1979. "Using Panel Data to Improve Estimates of the Determinants of Educational Achievement." *Sociology of Education*, pp. 113–121.

Brinch, Christian N, Magne Mogstad, and Matthew Wiswall. 2017. "Beyond LATE with a Discrete Instrument." *Journal of Political Economy* 125 (4): 985–1039.

Bulman, George, and Robert W Fairlie. 2016. "Technology and Education: Computers, Software, and the Internet." In *Handbook of the Economics of Education*, Volume 5, 239–280. Elsevier.

Calderoni, Jose. 1998. *Telesecundaria: Using TV to Bring Education to Rural Mexico*. Education and Technology Team, Human Development Network, World Bank.

Cameron, Stephen V, and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints using Returns to Schooling." *Journal of Political Economy* 112 (1): 132–182.

Card, David. 1995. "Using Geographical Variation in College Proximity to Estimate the Return to Schooling." *Aspects of Labor Market Behavior: Essays in Honor of John Vanderkamp*, pp. 201–222.

Carneiro, Pedro, and James J Heckman. 2002. "The Evidence on Credit Constraints in Post-Secondary Schooling." *The Economic Journal* 112 (482): 705–734.

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101 (6): 2754–2781.

Carneiro, Pedro, Michael Lokshin, and Nithin Umapathi. 2017. "Average and Marginal Returns to Upper Secondary Schooling in Indonesia." *Journal of Applied Econometrics* 32 (1): 16–36.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic perspectives* 20 (1): 91–116.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg. 2018. "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance." *Journal of Political Economy* 126 (6): 2356–2409.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47.

Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Vapital: Evidence from College Openings." *The Quarterly Journal of Economics* 118 (4): 1495–1532.

de Cossio, Roger Diaz. 2007. "Repensar la Telesecundaria." *El Universal*, August 21.

de Hoyos, Rafael, Ricardo Estrada, and María José Vargas. 2021. "What do Test Scores Really Capture? Evidence from a Large-scale Student Assessment in Mexico." *World Development* 146:105524.

De Hoyos, Rafael, Vicente A Garcia-Moreno, and Harry Anthony Patrinos. 2017. "The Impact of an Accountability Intervention with Diagnostic Feedback: Evidence from Mexico." *Economics of Education Review* 58:123–140.

Duflo, Esther, Rema Hanna, and Stephen P Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102 (4): 1241–78.

Dustan, Andrew, Alain De Janvry, and Elisabeth Sadoulet. 2017. "Flourish or Fail? The Risky Reward of Elite High School Admission in Mexico City." *Journal of Human Resources* 52 (3): 756–799.

Engzell, Per, Arun Frey, and Mark D Verhagen. 2021. "Learning Loss Due to School Closures during the COVID-19 Pandemic." *Proceedings of the National Academy of Sciences* 118, no. 17.

Escueta, Maya, Vincent Quan, Andre Joshua Nickow, and Philip Oreopoulos. 2017. "Education Technology: An Evidence-based Review." *National Bureau of Economic Research*.

Estrada, Rafael Quiroz. 2003. "Telesecundaria: Los Estudiantes y los Sentidos Que Atribuyen a Algunos Elementos del Modelo Pedagógico." *Revista Mexicana de Investigación Educativa* 8, no. 17.

Estrada, Ricardo. 2019. "Rules Versus Discretion in Public Service: Teacher Hiring in Mexico." *Journal of Labor Economics* 37 (2): 545–579.

Estrada, Ricardo, and Jérémie Gignoux. 2017. "Benefits to Elite Schools and the Expected Returns to Education: Evidence from Mexico City." *European Economic Review* 95:168–194.

Fabregas, Raissa. 2020. "Broadcasting Education: The Long-term Effects of Mexico's Telesecundarias." *Working Paper*.

Fan, Jianqing, and Irene Gijbels. 1996. *Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability*. Volume 66. CRC Press.

Heckman, James J, Daniel Schmierer, and Sergio Urzua. 2010. "Testing the Correlated Random Coefficient Model." *Journal of Econometrics* 158 (2): 177–203.

Heckman, James J, and Edward J Vytlacil. 1999. "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences* 96 (8): 4730–4734.

———. 2001a. "Local Instrumental Variables." In *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya*, edited by K. Morimune C. Hsiao and J. L. Powell, 1–46. New York: Cambridge University Press.

———. 2001b. "Policy-relevant Treatment Effects." *American Economic Review* 91 (2): 107–111.

———. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73 (3): 669–738.

———. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to

Evaluate Social Programs, and to Forecast Their Effects in New Environments." *Handbook of Econometrics* 6:4875–5143.

Imbens, Guido W, and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, pp. 467–475.

Johnston, Jamie, and Christopher Ksoll. 2017. "Effectiveness of Interactive Satellite-Transmitted Instruction: Experimental Evidence from Ghanaian Primary Schools. CEPA Working Paper No. 17-08." *Stanford Center for Education Policy Analysis*.

Kane, Thomas J, and Cecilia Elena Rouse. 1995. "Labor-Market Returns to Two- and Four-year College." *The American Economic Review* 85 (3): 600–614.

Kling, Jeffrey R. 2001. "Interpreting Instrumental Variables Estimates of the Returns to Schooling." *Journal of Business & Economic Statistics* 19 (3): 358–364.

Maldonado, Joana Elisa, and Kristof De Witte. 2021. "The Effect of School Closures on Standardised Student Test Outcomes." *British Educational Research Journal*.

Michaelsen, Maren M, and Paola Salardi. 2020. "Violence, Psychological Stress and Educational Performance During the "War on Drugs" in Mexico." *Journal of Development Economics* 143:102387.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J Ganimian. 2019. "Disrupting Education? Experimental Evidence on Technology-aided Instruction in India." *American Economic Review* 109 (4): 1426–60.

Navarro-Sola, Laia. 2019. "Secondary School Expansion Through Televised Lessons: The Labor Market Returns of the Mexican Telesecundaria." *Working Paper*.

OECD. 2019. *Education at a Glance 2019: OECD Indicators*. OECD Publishing.

Parker, Susan W, and Petra E Todd. 2017. "Conditional Cash Transfers: The Case of Progresa/Oportunidades." *Journal of Economic Literature* 55 (3): 866–915.

Robinson, Peter M. 1988. "Root-N-consistent Semiparametric Regression." *Econometrica*, pp. 931–954.

Rodriguez-Segura, Daniel. 2020. "Educational Technology in Developing Countries: A Systematic Review." *Working Paper*.

Roy, Andrew Donald. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2): 135–146.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688.

Santos, Annette. 2001. "Oportunidades Educativas en Telesecundaria y Factores que las Condicionan." *Revista Latinoamericana de Estudios Educativos (México)* 31 (3): 11–52.

Sapelli, Claudio, and Bernardita Vial. 2002. "The Performance of Private and Public Schools in the Chilean Voucher System." *Cuadernos de Economía* 39 (118): 423–454.

Todd, Petra E, and Kenneth I Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *The Economic Journal*, pp. F3–F33.

UNESCO, Institute for Statistics. 2016. "The World Needs Almost 69 Million New Teachers to Reach the 2030 Education Goals." *(Fact Sheet No. 39)*.

# Appendices

### A. Analysis of Cheating and Differential Test-Taking

In this section, we show that the pattern of test scores shown in Table 1 is unlikely to be affected by either differential rates of cheating or differential rates of test-taking across school types.

The ENLACE exams are low-stakes, but the Secretariat of Education (SEP) nevertheless attempted to reduce cheating in a number of ways. External administrators oversaw the administration of the exam at each school. Teachers were forbidden from supervising students taking the exams, and exams were graded at a central location. Whenever possible, test-takers were arranged so that they sat adjacent to students in different grades who were taking different exams (de Hoyos, Estrada, and Vargas 2021).

In addition, the SEP developed a statistical tool to determine whether students were likely to have cheated on the exam. This method, the k-index method, identifies students in the same testing location with the same number of correct answers. For all these students, the k-index is calculated as the probability that a group of students have $n_1$ identical incorrect answers given that they have $n_2$ total incorrect answers. Students with a k-index value below a particular threshold are marked as likely having copied.

We find that cheating is slightly more prevalent at telesecundarias than traditional schools, although rates of copying are low at both schools: 1.41% (1.11%) of students at telesecundarias (traditional schools) are deemed to have cheated on their 7th grade exams. When we replicate the means in Table 1 among students who have not cheated, we find nearly identical results.

We next consider whether there are differential rates of test-taking across school types. A pattern whereby students at telesecundarias improve their scores relative to students at traditional schools could occur even in the absence of increased learning at telesecundarias if low-performing students at these schools were less likely to sit the exams. In Table A-2, we investigate this possibility by running several probit regressions. The dependent variable is a binary variable that takes the value of one if a student who is enrolled in the seventh grade sits the ENLACE exam and zero otherwise. In column (1), we include only a dummy for telesecundaria attendance. Column (2) adds controls for students' standardized sixth grade math and Spanish scores. Column (3) includes interactions between telescundaria attendance and sixth

Table A-1
Mean Scores on Exams, by School Type

| | Math | | | | Spanish | | | |
|---|---|---|---|---|---|---|---|---|
| | All | Traditional | Telesecundaria | Difference | All | Traditional | Telesecundaria | Difference |
| Grade 6 | 521.1 | 530.7 | 484.0 | -46.7 | 516.1 | 526.5 | 476.0 | -50.4 |
| Grade 7 | 499.7 | 501.2 | 494.0 | -7.20 | 499.5 | 503.4 | 484.2 | -19.2 |
| Grade 8 | 508.3 | 504.9 | 521.6 | 16.7 | 477.0 | 479.0 | 469.1 | -9.90 |
| Grade 9 | 524.3 | 518.5 | 546.6 | 28.1 | 499.5 | 501.1 | 493.3 | -7.70 |

The table shows mean scores on the math and Spanish ENLACE exams for public school students in Mexico who are presumed to not have cheated on the exam. Grade 6 is primary school. Starting in grade 7, public school students attend either traditional secondary schools or telesecundarias. The standard deviations on the math ENLACE exams are 118.3, 98.1, 101.2, and 116.4 in grades 6, 7, 8 and 9, respectively, for this sample. The standard deviations on the Spanish ENLACE exams are 103.9, 97.5, 105.8, and 98.1 in grades 6, 7, 8 and 9, respectively.

grade test scores, and column (4) includes additional controls for the student's age, number of siblings, sex, prospera status, family income, mother's education, number of books in the home, and whether the family has access to a computer.

The results in column (1) show that students who attend telesecundarias are no more or less likely to sit the seventh grade ENLACE exam. Adding covariates for prior exam scores in math and Spanish reveals that telesecundaria students are actually more likely to sit the exams than are students at traditional schools, conditional on past academic performance. Column (3) shows that, for students who attend telesecundarias, the effect of scoring an additional standard deviation higher on their sixth grade math exam is to increase the probability of sitting the seventh grade exam by 0.7 percentage points. Scoring one standard deviation higher on the Spanish sixth grade exam has an average effect of reducing the probability of sitting the exam by 1.4 percentage points. These effects are small given the large change in scores considered, and they are opposite in sign. As a result, we believe it is unlikely that differential exam-sitting at telesecundarias relative to traditional schools can explain the large gains in test scores for telescundaria students documented in Table 1 in the main text.

Table A-2
Probability of Sitting Seventh Grade ENLACE Exam

| | Average Derivative | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Tele | 0.0033 | 0.0130 | 0.0100 | 0.0244 |
| | (0.0019) | (0.0020) | (0.0021) | (0.0025) |
| Math Score (6th Grade) | | 0.0005 | -0.0012 | -0.0019 |
| | | (0.0013) | (0.0014) | (0.0014) |
| Spanish Score (6th Grade) | | 0.021 | 0.0235 | 0.0187 |
| | | (0.0014) | (0.0015) | (0.0015) |
| Tele $\times$ Math | | | 0.0076 | 0.0073 |
| | | | (0.0029) | (0.0030) |
| Tele $\times$ Spanish | | | -0.0141 | -0.0157 |
| | | | (0.0029) | (0.0029) |
| Controls | No | No | No | Yes |
| N | 145065 | 145065 | 145065 | 145065 |

The table shows the average marginal effect of telesecundaria attendance on the probability of sitting the seventh grade ENLACE exam. The sample is our main estimation sample augmented with individuals who are known to have attended secondary school in the seventh grade but for whom no score is observed. The dependent variable is a binary variable that equals one if the student sits the seventh grade ENLACE exam and zero otherwise. Tele equals one if the student is enrolled in a telesecundaria and zero otherwise. Sixth grade math and Spanish scores on the ENLACE exams are measured in standard deviations. Controls include age, number of siblings, sex, prospera status, family income, mother's education, number of books in the home, and an indicator for whether the family has access to a computer. Standard errors are calculated via 250 bootstrap replications.

## B. Additional Tables and Figures

Table B-1
Summary Statistics by Secondary School Type

| | General | Technical | Telesecundaria | Dropped |
|---|---|---|---|---|
| Proportion of Cohort | 0.49 | 0.27 | 0.17 | 0.06 |
| Mean Math Score (7th Grade) | 512.3 | 503.8 | 494.9 | - |
| Mean Spanish Score (7th Grade) | 515.2 | 505.6 | 484.9 | - |
| Mean Math Score (6th Grade) | 542.6 | 533.5 | 484.4 | 453.8 |
| Mean Spanish Score (6th Grade) | 540.8 | 527.5 | 476.3 | 451.4 |
| Fraction Female | 0.51 | 0.50 | 0.50 | 0.50 |
| Mean Age (2008) | 11.9 | 11.9 | 12.2 | 12.8 |
| Fraction Prospera | 0.13 | 0.19 | 0.66 | 0.40 |

The table displays characteristics of students who attend each of three secondary school types – General, Technical, and Telesecundaria – as well as students who drop out. The sample size for this table is $N = 186,483$. Based on information presented in this table, we consider General and Technical schools as a single alternative for the purposes of estimating value added between the sixth and seventh grades.

Table B-2
Summary Statistics

| | Telesecundaria | | Traditional | |
|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. |
| Math Score (7th grade) | 496.6 | (98.2) | 503.5 | (98.2) |
| Spanish Score (7th grade) | 487.4 | (97.1) | 505.5 | (96.9) |
| Math Score (6th grade) | 488.5 | (110.8) | 536.0 | (117.8) |
| Spanish Score (6th grade) | 480.4 | (95.6) | 531.8 | (102.6) |
| Relative Distance | -5.08 | ( 4.06) | 2.99 | ( 3.58) |
| Age | 12.17 | ( 0.81) | 11.91 | ( 0.60) |
| Siblings | 3.68 | ( 2.40) | 2.40 | ( 1.67) |
| Prospera | 0.65 | ( 0.48) | 0.16 | ( 0.37) |
| Female | 0.51 | ( 0.50) | 0.51 | ( 0.50) |
| Computer at Home | 0.13 | ( 0.33) | 0.42 | ( 0.49) |
| Rural Residence | 0.67 | ( 0.47) | 0.09 | ( 0.29) |
| Northern State | 0.05 | ( 0.22) | 0.29 | ( 0.46) |
| Speaks Language other than Spanish at Home | 0.16 | ( 0.36) | 0.04 | ( 0.20) |
| Books in the Home : $\leq 10$ | 0.71 | ( 0.46) | 0.47 | ( 0.50) |
| Books in the Home : 20 | 0.18 | ( 0.38) | 0.26 | ( 0.44) |
| Books in the Home : 50 | 0.07 | ( 0.25) | 0.16 | ( 0.36) |
| Books in the Home : $\geq 100$ | 0.05 | ( 0.21) | 0.12 | ( 0.32) |
| Mother's Education : Primary | 0.76 | ( 0.43) | 0.41 | ( 0.49) |
| Mother's Education : Middle | 0.19 | ( 0.39) | 0.28 | ( 0.45) |
| Mother's Education : Secondary | 0.04 | ( 0.20) | 0.23 | ( 0.42) |
| Mother's Education : Postsecondary | 0.01 | ( 0.11) | 0.07 | ( 0.26) |
| Income (Pesos/mo) : $\leq 2500$ | 0.56 | ( 0.49) | 0.23 | ( 0.42) |
| Income (Pesos/mo) : 2500-2999 | 0.27 | ( 0.44) | 0.30 | ( 0.46) |
| Income (Pesos/mo) : 3000-7499 | 0.12 | ( 0.32) | 0.32 | ( 0.47) |
| Income (Pesos/mo) : $\geq 7500$ | 0.04 | ( 0.20) | 0.15 | ( 0.36) |

The table displays summary statistics for outcome variables, covariates, and the instrument – relative distance – for students in our estimation sample. The statistics are broken down by the type of secondary school attended.

## C. Estimation of MTE and Treatment Effects under the Assumption of Joint Normality

The fully parametric approach to estimating the MTE specifies the unobservables in the selection and outcome equations as jointly normally distributed:

$$\begin{pmatrix} U_1 \\ U_0 \\ V \end{pmatrix} \sim N \left[ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{10} & \sigma_{1V} \\ & \sigma_0^2 & \sigma_{0V} \\ & & 1 \end{pmatrix} \right] .$$

Under these assumptions, the marginal treatment effect has the following simple functional form:

$$MTE(X, u_D) = X(\beta_1 - \beta_0) + (\sigma_{1V} - \sigma_{0V})\Phi^{-1}(U_D) .$$

We estimate the parameters $\beta_1, \beta_0, \sigma_{1V}, \sigma_{0V}$ via a two-step method that first estimates the propensity score via probit and then includes control functions in the outcome equations as follows:

$$\begin{aligned} \mathbb{E}(Y_1 \mid D = 1, X, Z) &= X\beta_1 + \mathbb{E}(U_1 \mid D = 1) , \\ &= X\beta_1 + \sigma_{1V} \left( -\frac{\phi(\Phi^{-1}(P(Z)))}{P(Z)} \right) , \\ \mathbb{E}(Y_0 \mid D = 0, X, Z) &= X\beta_0 + \mathbb{E}(U_0 \mid D = 0) , \\ &= X\beta_0 + \sigma_{0V} \left( \frac{\phi(\Phi^{-1}(P(Z)))}{1 - P(Z)} \right) . \end{aligned}$$

After estimating $(\beta_0, \beta_1, \sigma_{0V}, \sigma_{1V})$, we construct parametric estimates of common treatment parameters as follows:

$$\begin{aligned} ATE &= \bar{X}(\beta_1 - \beta_0) , \\ TT &= \frac{1}{N_T} \sum_{i=1}^{N_T} D_i \left[ X_i(\beta_1 - \beta_0) + (\sigma_{1V} - \sigma_{0V}) \left( -\frac{\phi(\Phi^{-1}(P(Z)))}{P(Z)} \right) \right] , \\ TUT &= \frac{1}{N_C} \sum_{i=1}^{N_C} (1 - D_i) \left[ X_i(\beta_1 - \beta_0) + (\sigma_{1V} - \sigma_{0V}) \left( \frac{\phi(\Phi^{-1}(P(Z)))}{1 - P(Z)} \right) \right] , \end{aligned}$$

where $D_i = 1$ if individual $i$ attends a telesecundaria and $0$ otherwise, and $N_T$ ($N_C$)

denotes the number of students attending telesecundarias (traditional schools). Standard errors for these treatment parameters are estimated via the nonparametric bootstrap.

## D. Evidence of Sorting on Gains

A marginal treatment effect that is nonconstant in $U_D$ is evidence of a pattern of sorting on gains. We test formally for evidence of selection using methods developed in Heckman, Schmierer, and Urzua (2010). As explained in Heckman and Vytlacil (2005), the Local Average Treatment Effect (LATE) introduced by Imbens and Angrist (1994) is simply the integral of the MTE over a specific region of the domain of $U_D$. One way of testing for evidence of selection is to test whether LATEs defined by integrating the MTE over different intervals of $supp(U_D)$ are equivalent. Tables D-1 and D-2 display the results of these tests for math and Spanish, respectively. To perform these tests, we partition the support of $U_D$ into 25 intervals of width 0.04 and test whether the integrated MTEs on adjacent (but not overlapping) intervals differ. We then test that all LATEs are jointly equal to each other. The tests are conducted by using 250 bootstrapped data sets and computing simulated p-values under the null hypothesis of equivalent LATEs.

Table D-1
Tests for Selection on Unobservables: Math

| | | | | | | |
|---|---|---|---|---|---|---|
| Range of $U_D$ for $LATE^j$ | (0, 0.04) | (0.08, 0.12) | (0.16, 0.2) | (0.24, 0.28) | (0.32, 0.36) | (0.4, 0.44) |
| Range of $U_D$ for $LATE^{j+1}$ | (0.08, 0.12) | (0.16, 0.2) | (0.24, 0.28) | (0.32, 0.36) | (0.4, 0.44) | (0.48, 0.52) |
| Difference in LATEs | 22.2 | 9.50 | 6.28 | -0.71 | -3.54 | 1.55 |
| $p$-value | 0.00 | 0.01 | 0.00 | 0.74 | 0.18 | 0.48 |
| Range of $U_D$ for $LATE^j$ | (0.48, 0.52) | (0.56, 0.6) | (0.64, 0.68) | (0.72, 0.76) | (0.8, 0.84) | (0.88, 0.92) |
| Range of $U_D$ for $LATE^{j+1}$ | (0.56, 0.6) | (0.64, 0.68) | (0.72, 0.76) | (0.8, 0.84) | (0.88, 0.92) | (0.96, 1) |
| Difference in LATEs | 2.84 | -6.77 | 1.33 | 4.11 | 9.43 | 63.3 |
| $p$-value | 0.30 | 0.00 | 0.64 | 0.24 | 0.20 | 0.00 |
| Joint $p$-value | | | 0.000 | | | |

The table shows the results of tests for equality of LATEs for math value added defined by adjacent and non-overlapping regions of the domain of $U_D$. Given an interval $[L_j, H_j]$, the LATE for that interval is given by $LATE^j = \mathbb{E}[Y_1 - Y_0 | X = \bar{X}, L_j \leq U_D < H_j]$, which is simply the average of the MTE between $L_j$ and $H_j$ evaluated at $X = \bar{x}$. $p$-values test the hypothesis that the difference between adjacent LATEs is equal to zero. $p$-values are obtained by 250 bootstrap replications.

The tables reveal that many adjacent LATEs differ from one another and that, jointly, the LATEs are not equal at conventional significance thresholds for either math or Spanish. For example, the entry in column 1 of Table D-1 indicates that

$$\mathbb{E}(Y_1 - Y_0 | X = \overline{x}, 0.08 < U_D \leq 0.12) - \mathbb{E}(Y_1 - Y_0 | X = \overline{x}, 0 \leq U_D \leq 0.04) = 22.2 ,$$

Table D-2
Tests for Selection on Unobservables: Spanish

| | | | | | | |
|---|---|---|---|---|---|---|
| Range of $U_D$ for $LATE^j$ | (0, 0.04) | (0.08, 0.12) | (0.16, 0.2) | (0.24, 0.28) | (0.32, 0.36) | (0.4, 0.44) |
| Range of $U_D$ for $LATE^{j+1}$ | (0.08,0.12) | (0.16, 0.2) | (0.24, 0.28) | (0.32, 0.36) | (0.4, 0.44) | (0.48, 0.52) |
| Difference in LATEs | 5.55 | 6.46 | 4.88 | 1.16 | -0.626 | -3.61 |
| $p$-value | 0.39 | 0.04 | 0.03 | 0.62 | 0.81 | 0.16 |
| Range of $U_D$ for $LATE^j$ | (0.48, 0.52) | (0.56, 0.6) | (0.64, 0.68) | (0.72, 0.76) | (0.8, 0.84) | (0.88, 0.92) |
| Range of $U_D$ for $LATE^{j+1}$ | (0.56, 0.6) | (0.64, 0.68) | (0.72, 0.76) | (0.8, 0.84) | (0.88, 0.92) | (0.96, 1) |
| Difference in LATEs | -5.03 | -8.70 | -0.18 | 3.62 | 17.80 | 50.90 |
| $p$-value | 0.04 | 0.00 | 0.94 | 0.37 | 0.02 | 0.00 |
| Joint $p$-value | | | 0.000 | | | |

The table shows the results of tests for equality of LATEs for Spanish value added defined by adjacent and non-overlapping regions of the domain of $U_D$. Given an interval $[L_j, H_j]$, the LATE for that interval is given by $LATE^j = \mathbb{E}[Y_1 - Y_0|X = \bar{X}, L_j \leq U_D < H_j]$, which is simply the average of the MTE between $L_j$ and $H_j$ evaluated at $X = \bar{x}$. $p$-values test the hytpothesis that the difference between adjacent LATEs is equal to zero. $p$-values are obtained through 250 bootstrap replications.

and that the p-value that this difference differs from 0 is $p = 0.00$. The joint $p$-value for the test of the hypothesis that all LATEs are equivalent is $p = 0.00$ for both math and Spanish. As a result of these tests, we reject the hypothesis of no selection on unobservables. Attendance in telesecundarias is correlated with unobserved determinants of student achievement.

The pattern of reverse sorting that we find occurs when students who select into the treatment benefit less than untreated students. Such a pattern cannot be explained by a choice model in which individuals attend telesecundarias if $Y_1 > Y_0$, as in the original Roy (1951) model. In this section we argue that the pattern of sorting we uncover is consistent with both the existence of biased expectations and with families basing their decisions on factors other than test score gains.

Recent papers that investigate the consequences of educational choices made early in the child's life have found evidence that parents often choose academically inferior educational options for their children. Both Cornelissen et al. (2018) and Ainsworth et al. (2020) find that a lack of information and inaccurate beliefs about the relative quality of options available to the child can explain these findings. In Mexico, the predominant view expressed in the media and bolstered by academic research is that telesecundarias are of low quality (de Cossio 2007). Santos (2001) argues that they are worse than traditional schooling options on the grounds that telesecundaria students perform worse academically in the cross section, a finding we confirm for the seventh grade. While this analysis fails to control for the relative deprivation of students who attend telesecundarias, the perception of telesecundarias as lower quality alternatives seems ingrained (Acta Educativa 2016). It is trivial to show that if decision-makers have flipped expectations, for example if they based their decisions on $-(Y_1 - Y_0)$ rather than $(Y_1 - Y_0)$, a pattern of reverse sorting will result.

We investigate this hypothesis by applying methods developed by Brinch, Mogstad, and Wiswall (2017) to delve further into the source of reverse sorting. As shown in this paper, we can rewrite the marginal treatment effect as

$$MTE(X, U_D) = X(\beta_1 - \beta_0) + k_1(U_D) - k_0(U_D) \ , \tag{D-1}$$

where $k_j(U_D) = \mathbb{E}[U_j|U_D]$ for $j = 1, 2$. Here $k_1(U_D)$ can be thought of as the average unobserved match quality between students and telesecundarias for students with a particular resistance to attending telesecundarias (given by $U_D$). Similarly, $k_0(U_D)$ represents the unobserved match quality between students and traditional schools as a function of the same resistance variable.

We estimate $k_1(U_D)$ and $k_0(U_D)$ separately to determine whether the shape of the MTE curve is determined primarily by variability in match qualities between students and telesecundarias or between students and traditional schools. As shown in Heckman and Vytlacil (2007) and Brinch, Mogstad, and Wiswall (2017), $k_1(U_D)$ and $k_0(U_D)$ can be estimated using a control function approach on each of the $D = 1$ and $D = 0$ subsamples. Under Assumption (A-2)',

$$\mathbb{E}[Y_j|X = x, P(Z) = p, D = j] = X\beta_j + K_j(p) \ ,$$

for $j = 0, 1$, where

$$K_1(p) = E\left(U_1 \mid U_D \leq p\right) \ ,$$

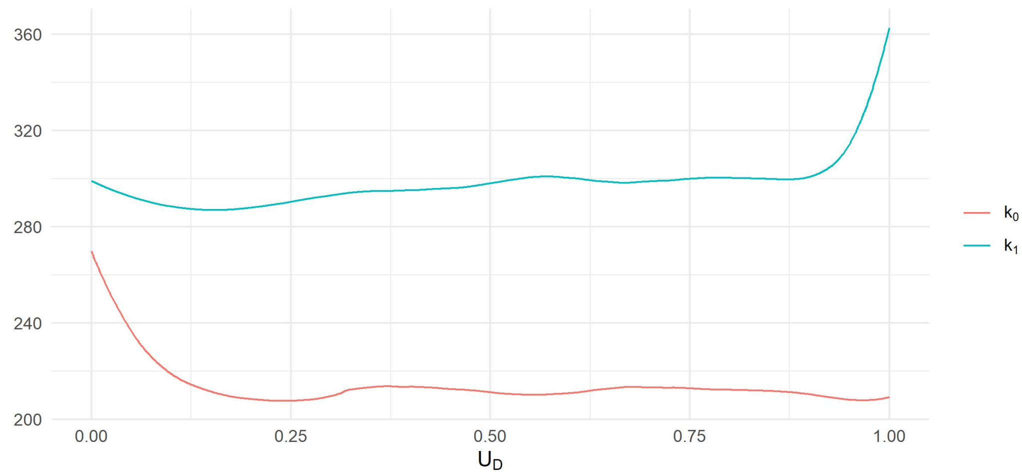
and

$$K_0(p) = E\left(U_0 \mid U_D > p\right) \ .$$

We can obtain $k_1$ and $k_0$ from $K_1$ and $K_0$ using the following identities in Brinch, Mogstad, and Wiswall (2017):

$$k_1(p) = p\frac{\partial K_1(p)}{\partial p} + K_1(p)$$

Figure D-1
The Source of Reverse Selection




The figure plots $k_1(U_D) = \mathbb{E}[U_1 \mid U_D]$ and $k_0(U_D) = \mathbb{E}[U_0 \mid U_D]$ evaluated at $X = \overline{x}$ on the vertical axis against $U_D$ on the horizontal axis. $U_1$ is the child's unobserved outcome in the equation for math value added in telesecundarias. $U_0$ is the child's unobserved outcome in the equation for math value added in traditional schools. Details on the estimation of $k_1(\cdot)$ and $k_0(\cdot)$ are provided in Appendix B..

$$k_0(p) = -(1-p)\frac{\partial K_0(p)}{\partial p} + K_0(p) \; . \tag{D-2}$$

Figure D-1 shows estimates of $k_1$ and $k_0$ for math as an outcome variable. Each estimated curve, $k_j$ for $j = 0, 1$, is obtained via two semiparametric regressions on the subsample with $D = j$. We compute the conditional expectation of $Y - X\beta_j$ given the estimated propensity score, $p$, using Local Linear Regression to obtain $K_j(p)$. We then compute the derivative of the conditional expectation of $Y - X\beta_j$ given $p$ via Local Quadratic Regression to obtain $\frac{\partial K_j(p)}{\partial p}$. $k_j(p)$ is then obtained from the identities in D-2. The bandwidth, 0.317, is the same as in the estimation of the MTE function in Section VI.

Figure D-1 demonstrates that variation in mean unobserved outcomes at telesecundarias, $k_1(U_D)$, and traditional schools, $k_0(U_D)$, occur at opposite ends of the domain of $U_D$. Students with low values of $U_D$ have a similar match quality at both telesecundarias and traditional schools. However, as $U_D$ increases and students become less likely to attend telesecundarias, the quality of the match at traditional schools declines rapidly, while the match quality at telesecundarias stays roughly constant. As $U_D$ increases still further, from $0.25$ to $0.85$, there is a steady increase in $k_1(U_D)$, while $k_0(U_D)$ remains constant. Then, for high values of $U_D$, the quality of the match at telesecundarias improves significantly as $U_D$ approaches values near one.

Figure D-1 reveals that the pattern of reverse sorting is driven by variability in outcomes at *both* telesecundarias and traditional schools. This pattern, which is not unprecedented in the literature, is consistent with decision-makers having flipped expectations regarding the quality of telesecundarias, a scenario that is plausible based on perceptions of their quality in Mexico.

An alternative explanation for the pattern of sorting is that families base their decision on factors other than learning. The findings in section VI-B. of the main text reveal that sorting on gains for ninth grade enrollment is nonmonotonic but that estimates of treatment on the treated exceed those of treatment on the untreated (TT>TUT). When choosing a school, families may therefore place more weight on educational attainment than academic learning. A large literature, surveyed in Cunha and Heckman (2007), documents the multiplicity of skills in explaining important life outcomes. If a child's idiosyncratic fit at a particular school type has a strong effect on the noncognitive skills affecting grade progression, it could explain the large estimates of TT for educational attainment along with the reverse sorting on gains for test scores.

## E. Sensitivity Analysis

We now consider alternative specifications, including those designed to investigate the validity of Assumption (A-2)′ in the main text. Our main analysis omits from the sample all children who attend a secondary school more than fifteen kilometers from their primary school to eliminate the threat of endogeneity caused by families who may move to be closer to a specific school based on their child's realization of $(U_1, U_0)$.[19]

To further alleviate this concern, we replicate our analysis on subsamples composed of individuals who are unlikely to change residence for the purpose of their children attending a different school. These subsamples are the set of children living in rural areas and the set of children whose parents earn under $2500$ pesos per month (approximately $220$ USD in 2008). Table E-1 presents estimates of the ATE, TT, and TUT for math ENLACE scores on these subsamples. In both subsamples, there is evidence, consistent with our main results, of reverse selection given by the fact that TT<ATE<TUT. Estimates of the treatment parameters are lower for the rural subsample than for the main sample (Table 4), which are in turn lower than the estimates for the low income subsample. However, these differences are not statistically significant, as 90% confidence intervals generated by the bootstrapped standard errors overlap for all estimated parameters.

We find the same qualitative results for the subsample analysis with Spanish ENLACE scores as a dependent variable (Table E-2). A pattern of reverse sorting on gains is found, and none of the estimated treatment parameters differ statistically from those estimated in the main analysis. These results suggest that our main findings are unlikely to be affected by families moving to take advantage of schools where their children would have an unexpectedly good fit, as given by $(U_0, U_1)$.

If students must walk to school, as is common in many parts of Mexico, an additional concern may be that traveling long distances to school causes academic performance to suffer. To alleviate concerns that the instrument is directly correlated with academic outcomes, we augment our outcome equations, (1) and (2), with a measure of the distance actually traveled to secondary school as an explanatory variable.

We find that including distance traveled to secondary school makes little differ-

[19] Evidence of behavior of this sort has been used to criticize the validity of distance to college as an instrument in the United States (Cameron and Taber 2004, Carneiro and Heckman 2002), although it is less of a concern in Mexico. The source of funding for secondary schools in Mexico – 79% from the federal government with the remainder coming from state governments – reduces the incentive for hyper-local sorting driven by differences in school budgets (OECD 2019).

Table E-1
Subsample Analysis: Math

| | Rural | | Low Income | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| Average Treatment Effect | 28.5 | (3.30) | 38.6 | (2.98) |
| Treatment on the Treated | 22.6 | (3.46) | 27.2 | (2.22) |
| Treatment on the Untreated | 38.5 | (4.37) | 42.7 | (3.55) |

The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on scores on the math ENLACE exam in the seventh grade for two subsamples. The subsamples are the set of children living in rural areas and the set of children growing up in households earnings under $2,500$ pesos per month. We first estimate semiparametric MTEs that are specific to each subsample. The bandwidth, 0.317, is the same as in the main sample. Then we calculate treatment parameters by integrating the subsample-specific MTEs with respect to the corresponding densities for each subsample: $f(X, U_D)$, $f(X, U_D \mid D = 1)$, and $f(X, U_D \mid D = 0)$. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the semiparametric MTE. Standard errors are obtained by 250 bootstrap replications.

ence in the estimated MTEs or treatment effects for math and Spanish. Figures E-1 and E-2 show that the estimated MTE curves for these specifications are very similar to those in our main analysis. The similar point estimates of treatment parameters (Tables E-3 and E-4) also reassure that our estimates of the causal effects of telesecundarias on learning are little affected by any correlation between the distance traveled to secondary school and unobserved determinants of test scores.

Throughout the paper we have measured distance "as the crow flies." We therefore conduct a final robustness check with a new instrument that is the relative distance between the nearest traditional and telesecundaria schools where distance is measured along roads and paths by Google maps. In most cases, the use of this new school measure does not cause a change in the identity of the nearest school of each type, but, for a small percentage of students, the nearest school as measured by "traveled distance" differs from the nearest school measured "as the crow flies." Nevertheless, the estimates of the MTE curve and treatment parameters are little changed when we use this alternative measure of "traveled distance." We display the MTE curves for this specification in Figures E-3 and E-4, and the treatment parameters in Tables E-5 and E-6. The MTE curves and the estimated parametric and semiparametric treatment effects for math and Spanish do not differ significantly from our main estimates.

Table E-2
Subsample Analysis: Spanish

| | Rural | | Low Income | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| Average Treatment Effect | 19.9 | (3.16) | 24.1 | (2.57) |
| Treatment on the Treated | 18.3 | (3.48) | 16.6 | (2.30) |
| Treatment on the Untreated | 22.5 | (4.00) | 26.8 | (3.01) |

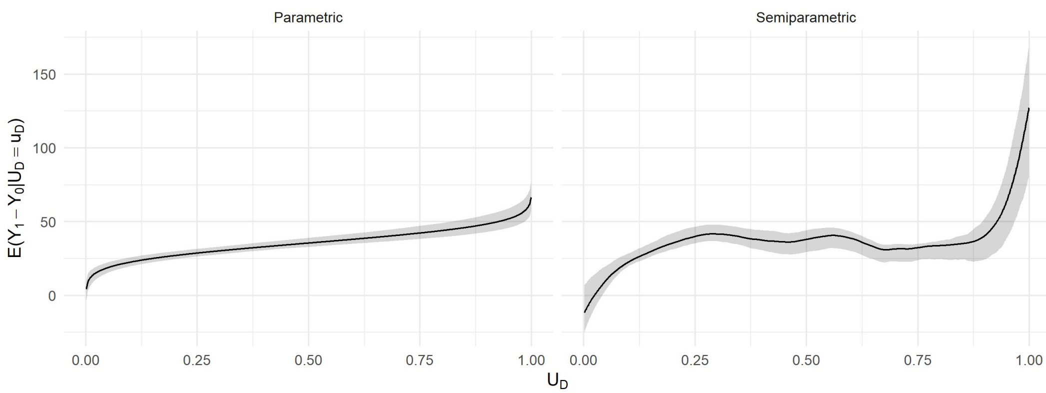
The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on scores on the Spanish ENLACE exam in the seventh grade for two subsamples. The subsamples are the set of children living in rural areas and the set of children growing up in households earnings under $2,500$ pesos per month. We first estimate semiparametric MTEs that are specific to each subsample. The bandwidth, 0.317, is the same as in the main sample. Then we calculate treatment parameters by integrating the subsample-specific MTEs with respect to the corresponding densities for each subsample: $f(X, U_D)$, $f(X, U_D \mid D = 1)$, and $f(X, U_D \mid D = 0)$. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the semiparametric MTE. Standard errors are obtained by 250 bootstrap replications.

Table E-3
Estimated Treatment Effects Controlling for Distance Traveled: Math

| | Parametric | | Semiparametric | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| Average Treatment Effect | 35.5 | (1.85) | 36.8 | (2.48) |
| Treatment on the Treated | 29.9 | (1.55) | 26.1 | (1.87) |
| Treatment on the Untreated | 36.8 | (1.97) | 39.2 | (2.83) |

The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on raw scores on the seventh grade math ENLACE exam. It differs from Table 4 in the main text in that this specification conditions on the distance actually traveled to secondary school in the outcome equations. The three treatment parameters are obtained by integrating the MTE with respect to the densities displayed in Figure 1. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the semiparametric MTE. Standard errors are obtained by 250 bootstrap replications.

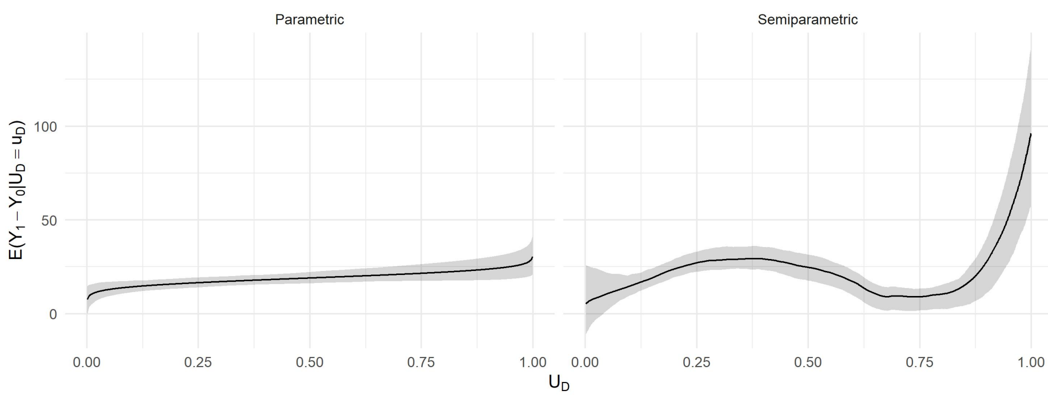
Figure E-1
Marginal Treatment Effect Controlling for Distance Traveled: Math




The dependent variable in the outcome equation is the raw score on the seventh grade ENLACE math exam. The outcome equations include controls for sixth grade math and Spanish scores, the age, sex, and number of siblings of the child, family income, mother's education, the number of books in the home, whether the family has access to a computer at home, the distance actually traveled to secondary school, and dummies for rural residence, residence in a Northern state, speaking a language other than Spanish at home, and participation in the conditional cash transfer program Prospera. The school choice model includes the same controls and also includes the relative distance between the nearest telesecundaria and nearest traditional secondary school as an exclusion restriction. The school choice model is estimated via probit. The parametric MTE is estimated using a two-step Least Squares method that controls for selection on unobservables. The semiparametric MTE is estimated using Local Quadratic Regression and an Epanechnikov kernel with a bandwidth of 0.317. Both MTEs are evaluated at the mean value of the covariates, $X = \overline{x}$. Ninety percent confidence intervals are computed via nonparametric bootstrap with 250 draws.

Figure E-2

Marginal Treatment Effect Controlling for Distance Traveled: Spanish




The dependent variable in the outcome equation is the raw score on the seventh grade ENLACE Spanish exam. The outcome equations include controls for sixth grade math and Spanish scores, the age, sex, and number of siblings of the child, family income, mother's education, the number of books in the home, whether the family has access to a computer at home, the distance actually traveled to secondary school, and dummies for rural residence, residence in a Northern state, speaking a language other than Spanish at home, and participation in the conditional cash transfer program Prospera. The school choice model includes the same controls and also includes the relative distance between the nearest telesecundaria and nearest traditional secondary school as an exclusion restriction. The school choice model is estimated via probit. The parametric MTE is estimated using a two-step Least Squares method that controls for selection on unobservables. The semiparametric MTE is estimated using Local Quadratic Regression and an Epanechnikov kernel with a bandwidth of 0.317. Both MTEs are evaluated at the mean value of the covariates, $X = \overline{x}$. Ninety percent confidence intervals are computed via nonparametric bootstrap with 250 draws.

Table E-4
Estimated Treatment Effects Controlling for Distance Traveled: Spanish

| | Parametric | | Semiparametric | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| Average Treatment Effect | 19.0 | (1.84) | 22.8 | (2.67) |
| Treatment on the Treated | 17.0 | (1.54) | 16.5 | (2.06) |
| Treatment on the Untreated | 19.5 | (1.97) | 24.2 | (3.01) |

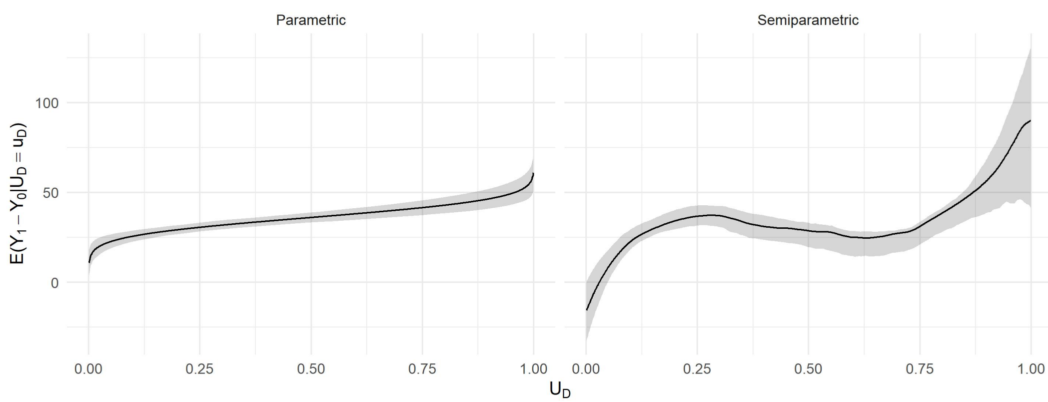
The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on raw scores on the seventh grade Spanish ENLACE exam. It differs from Table 5 in the main text in that this specification conditions on the distance actually traveled to secondary school in the outcome equations. The three treatment parameters are obtained by integrating the MTE with respect to the densities displayed in Figure 1. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the semiparametric MTE. Standard errors are obtained by 250 bootstrap replications.

Table E-5
Estimated Treatment Effects with "Traveled Distance" IV: Math

| | Parametric | | Semiparametric | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| Average Treatment Effect | 36.1 | (1.80) | 34.1 | (2.70) |
| Treatment on the Treated | 31.2 | (1.48) | 26.6 | (2.10) |
| Treatment on the Untreated | 37.1 | (1.93) | 35.7 | (3.02) |

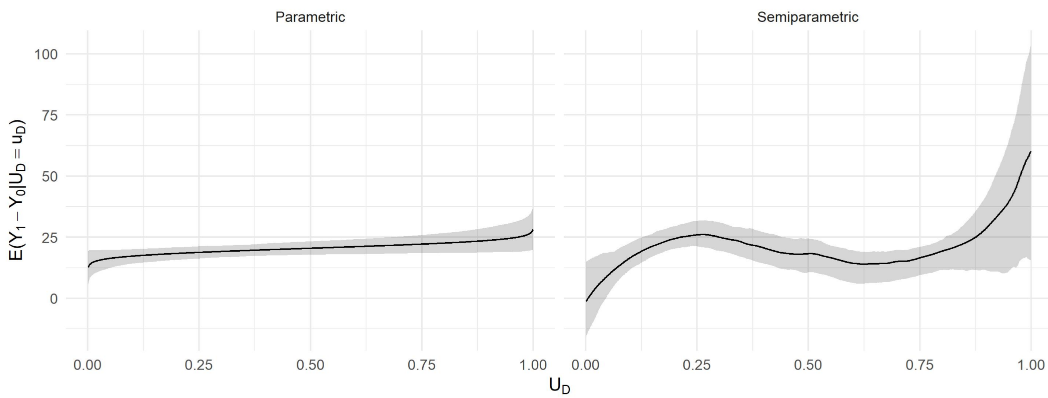
The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on raw scores on the seventh grade math ENLACE exam. It differs from Table 4 in the main text in that this relative distance measure is computed by differencing the distances *measured along roads and paths* to the nearest secondary school of each type. The three treatment parameters are obtained by integrating the MTE with respect to the densities displayed in Figure 1. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the semiparametric MTE. Standard errors are obtained by 250 bootstrap replications.

Figure E-3
Marginal Treatment Effect with "Traveled Distance" IV: Math




The dependent variable in the outcome equation is the raw score on the seventh grade ENLACE math exam. The outcome equations include controls for sixth grade math and Spanish scores, the age, sex, and number of siblings of the child, family income, mother's education, the number of books in the home, whether the family has access to a computer at home, the distance actually traveled to secondary school, and dummies for rural residence, residence in a Northern state, speaking a language other than Spanish at home, and participation in the conditional cash transfer program Prospera. The school choice model includes the same controls and also includes the relative distance *as measured along roads and paths* between the nearest telesecundaria and nearest traditional secondary school as an exclusion restriction. The school choice model is estimated via probit. The parametric MTE is estimated using a two-step Least Squares method that controls for selection on unobservables. The semiparametric MTE is estimated using Local Quadratic Regression and an Epanechnikov kernel with a bandwidth of 0.317. Both MTEs are evaluated at the mean value of the covariates, $X = \overline{x}$. Ninety percent confidence intervals are computed via nonparametric bootstrap with 250 draws.

Figure E-4
Marginal Treatment Effect with "Traveled Distance" IV: Spanish




The dependent variable in the outcome equation is the raw score on the seventh grade ENLACE Spanish exam. The outcome equations include controls for sixth grade math and Spanish scores, the age, sex, and number of siblings of the child, family income, mother's education, the number of books in the home, whether the family has access to a computer at home, the distance actually traveled to secondary school, and dummies for rural residence, residence in a Northern state, speaking a language other than Spanish at home, and participation in the conditional cash transfer program Prospera. The school choice model includes the same controls and also includes the relative distance *as measured along roads and paths* between the nearest telesecundaria and nearest traditional secondary school as an exclusion restriction. The school choice model is estimated via probit. The parametric MTE is estimated using a two-step Least Squares method that controls for selection on unobservables. The semiparametric MTE is estimated using Local Quadratic Regression and an Epanechnikov kernel with a bandwidth of 0.317. Both MTEs are evaluated at the mean value of the covariates, $X = \overline{x}$. Ninety percent confidence intervals are computed via nonparametric bootstrap with 250 draws.

Table E-6

Estimated Treatment Effects with “Traveled Distance” IV: Spanish

| | Parametric | | Semiparametric | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| Average Treatment Effect | 20.5 | (1.64) | 21.0 | (2.72) |
| Treatment on the Treated | 18.9 | (1.45) | 16.6 | (2.02) |
| Treatment on the Untreated | 20.8 | (1.74) | 21.9 | (3.06) |

The table displays three treatment parameters corresponding to the effect of telesecundaria attendance on raw scores on the seventh grade Spanish ENLACE exam. It differs from Table 5 in the main text in that this relative distance measure is computed by differencing the distances *measured along roads and paths* to the nearest secondary school of each type. The three treatment parameters are obtained by integrating the MTE with respect to the densities displayed in Figure 1. The simulation method of Carneiro, Lokshin, and Umapathi (2017) is used to integrate the semiparametric MTE. Standard errors are obtained by 250 bootstrap replications.