The Equilibrium Effects of Public Provision in Education Markets: Evidence from a Public School Expansion Policy*

Michael Dinerstein⋆ University of Chicago

Christopher Neilson‡ Princeton University

Sebastián Otero§ Stanford University

March 31, 2020

PRELIMINARY AND INCOMPLETE
DO NOT DISTRIBUTE OR CITE WITHOUT PERMISSION

Abstract

In a variety of markets with private options, the optimal level of public provision may require balancing a tradeoff between reducing private options' market power with the possibility of crowding out potentially high-quality products. These considerations are particularly relevant in many developing countries' education systems where private schools capture high market shares while public schools are overcrowded. We study the equilibrium effects of public provision in the context of a large expansion of public schools in the Dominican Republic. Over a five-year period, the government aimed to increase the number of public school classrooms by 78%. Using an event study framework, we estimate the effect of a new public school on neighborhood outcomes and competing private schools, where we instrument for how quickly the public school construction project finished with whether the procurement lottery randomly

*First version: February, 2020. This project would have not been possible without the support and assistance of Ryan Cooper, Juan Ariel Jiménez and Ancell Schecker. María Elena Guerrero and Astrid Pineda provided excellent research assistance. We would like to thank Liran Einav and Matthew Gentzkow, and seminar participants at Stanford University, Harvard University and University of Michigan for valuable comments and suggestions. We also wish to thank Ministerio de Educación (MINERD) of the Dominican Republic and Agencia de Evaluación de la Calidad Educativa for facilitating joint work between government agencies that produced the data from the Dominican Republic used in this study. All remaining errors are our own.

⋆University of Chicago. Email: mdinerstein@uchicago.edu. ‡Princeton University. Email: cneilson@princeton.edu. §Stanford University. Email: sotero@stanford.edu.

assigned the project to a firm or an unaffiliated individual. We find that a new public school increased public sector enrollment significantly. As public enrollment increased, a large number of private schools closed while the surviving schools lowered prices and increased school quality. To study how the provision of high quality schools varies with the level of public provision, and to compare the effects to the alternative policy of public financing, we specify and estimate an empirical model of demand (students choosing schools) and supply (schools choosing whether to stay open and what price to charge).

Contents

1 Introduction

Many industries – healthcare, energy, education, etc. – include a combination of public and private entities, and even within-industry, governments have chosen different ways to organize these markets, ranging from full privatization to public-private partnerships to direct public provision. Public provision may ensure equal access to important social services while also investing any extra revenue back into the system, while private provision may prove more efficient if profit incentives lead to more innovation or higher quality. In this paper we study the optimal level of public provision in primary and secondary education and compare the effectiveness of public provision relative to public financing.

These issues are particularly relevant in the developing world where state capacity to deliver services may be limited. Across many countries and large cities in Latin America, private schools capture large market shares (Figure 1). And because in many cases private options are not subsidized by the government, this could indicate that they are providing higher quality instruction. Yet, many countries' public school systems are overcrowded such that seats or hours in the public sector are rationed. Thus, the private sector may be only a temporary substitute for too little public provision. Whether private schools out-compete public schools on quality or space is important for designing optimal policy. Indeed, a higher quality private sector might encourage more public financing of private options while a private sector that gains market share by having excess capacity might lead governments to expand public provision.

Separating these stories and more generally understanding how the public and private sectors interact is difficult because private provision is determined in equilibrium and may respond to public policies. In this paper, we therefore leverage a large investment in public school infrastructure in the Dominican Republic that increased government spending on education from 2.5% to 4% of GDP in a single year. This 78% expansion of public capacity allows us to examine the effects of public provision on students' choice of sector and learn whether initially high private sector shares reflected insufficient public provision. Further, because the initial equilibrium involved a large private school enrollment share reaching close to 40% in urban areas, the policy effects of the public school expansion likely depend in part on the impacts on and responses by the private sector. We thus evaluate how the public expansion changes the market equilibrium in terms of which private schools are open and what characteristics they operate with.

To answer these questions, we collect data from a variety of sources. We combine administrative enrollment and test score data that is linked to individual students over time with administrative data on eligibility for public benefits and household demographics. We also use administrative data on private schools that includes their prices and investments, variables that are typically not available in administrative data in many countries. To further understand students' choices and how they depend on school characteristics, we bring in detailed data from surveys of students and principals.

We identify the policy effects by exploiting details of the procurement process. The projects were assigned in four waves, and in each wave the government held a lottery to determine which contractor would execute which project. These lotteries induced substantial entry among contractors such that over 80% of potential builders did not yet have a construction firm. We thus use the randomized assignment of projects to contractor types – specifically, whether the randomized builder is a firm or not – as an instrument that led some projects to finish faster than others. This exogenous variation lets us compare local areas all designated to receive a new public school but for which the year the school opened varied for reasons orthogonal to area characteristics.

We embed such variation in an event study framework and estimate the impact of a new public school on a variety of outcomes. We start by showing that the new public schools enrolled large numbers of students such that incumbent schools saw a substantial decrease relative to pre-policy levels. To further understand the sources of these enrollment increases, we assess whether the private schools were affected by – and how they responded to – the new schools. We see that private schools near new public schools saw a large decrease in enrollment, much of which came in the form of private school closures. Among private schools that remain open, we see a decrease in the prices they charge, and an increase in test scores (both unadjusted and adjusted with student characteristics).

These results imply that much of the policy impact may have been mediated through effects on the private schooling sector. How these changes affect overall quality is potentially non-monotonic. The increased competitive pressure raises school quality, while the crowd out of (unsubsidized) private schools has an ambiguous effect on available quality. To understand the optimal level of public investment, given this potential tradeoff driven by supply side responses, we extend the analysis by estimating a structural model of demand and supply. In the demand model, students choose a school for 9th grade based on the school's characteristics and the student's heterogeneous preferences. On the supply side, schools maximize profits by choosing price and whether to remain open. As schools operate in a large but highly differentiated market with many competitors, predicting the strategic responses of all relevant competitors is a high-dimensional problem. We thus make a behavioral assumption on how schools assess their competition. We use a notion of a static "oblivious" equilibrium as introduced in Sánchez (2018). Essentially, schools keep track of the expected (exponentiated) value students receive from attending a competing private school. This behavioral assumption, with consistent beliefs, yields an equilibrium that simplifies computation.

We estimate the demand model via simulated method of moments. We use aggregate enrollment share moments supplemented by two price instruments. Specifically, we use policy variation from the expansion of the conditional cash transfer program that increased the transfer families receive if their child attends secondary school. We also leverage the increased hours offered by the public school system as a shock to competition. We further construct micro moments from survey data that ask students what school they would choose if (1) their preferred school were not available or (2) if all schools had zero prices. These responses to hypothetical choice scenarios reveal significant information about preference heterogeneity and how students trade off price versus other characteristics.

On the supply side, we assume optimal price-setting to maximize (perceived) profits. This allows us to invert price first-order conditions at the observed price levels and back out marginal costs. We find that some schools have significant market power, capturing nearly all revenue as profit, while others have relatively low markups. We estimate the fixed cost distribution by relating variable profits to the closure rates observed in the data.

With our estimated model, we conduct counterfactual exercises that vary the level of public provision or the policy instrument between public provision and public financing (vouchers). We find that increased public provision initially yields large crowd out of the private sector but this eventually stabilizes such that the main effect is a very small pro-competitive decrease in private school markups. Public financing, on the other hand, leads to intense price competition that significantly reduces private schools' markups. Despite the public subsidy, many private schools can no longer stay open under this intense competition so we predict fairly large exit rates even under public financing.

1.1 Related Literature

This paper relates to several large literatures.[1] The first strand examines interactions between public and private schools and has focused on whether school quality responds to

[1] The current literature review is highly incomplete and will be extended in the next draft.

competition (e.g., Hoxby (1994), McMillan (2005), Card et al. (2010), Neilson (2013)) and how students sort between the public and private sectors (e.g., Hoxby (2003), Epple et al. (2004)). The second strand evaluates school funding reforms and whether spending affects student outcomes. Work on school funding reforms and effects on private school enrollments includes Downes and Schoeman (1998), Hoxby (2001), and Estevan (2015). This paper also builds upon Dinerstein and Smith (2018) by adding intensive margin private school responses.

2 Background and the School Construction Program

The Dominican Republic has seen the fastest economic growth of any Latin American country over the last two decades, at an average rate of 5.3 percent per year (World Bank, 2018). Education access has improved in line with this growth, with gross enrollment rates of 102 percent at the primary level and 77 percent at the secondary level (World Bank, 2018).[2] But while more students have enrolled in school, the system's public capacity had been limited, such that most students either attended an oversubscribed public school that divided students into shifts or a private school. Perhaps related, education outcomes lagged behind other countries, as reflected by the Dominican Republic claiming last place in student skills among all participating countries in the international TERCE in 2013 and PISA test in 2015 (UNESCO, 2015; OECD, 2016).[3]

2.1 The School Construction Program

Improving educational outcomes became an important issue in the 2012 presidential election, and Danilo Medina won with a promise to allocate 4% of GDP to education. Figure 2 shows a dramatic increase in the share of GDP allocated to education from 2.5% in 2012 to 4% in 2013. This allocation has been used primarily to finance school construction and renovation for two flagship education programs: *Jornada Escolar Extendida* (JEE) and *Quizqueya Empieza Contigo* (QEC).[4] JEE is a program to transition the country from a half-day schooling model to one of full-time schooling, intended to broaden educational

[2]The primary gross enrollment rate above 100 percent reflects the high prevalence of over-age enrollment.

[3]TERCE is an international standardized examination that compares student learning among Latin American countries.

[4]The initial budget allocated more than 2.5 billion dollars to infrastructure.

offerings and improve performance on pedagogical management indicators. QEC is an early childhood development (ECD) program tasked with increasing the coverage and quality of services provided to children between 0 and 5 years of age, based on the notion that the nature of learning is cumulative (Shonkoff and Phillips, 2000) and that environments that do not stimulate young children place them at an early disadvantage (Heckman, 2006).

In November 2012, the government issued Decree Number 625-12, which created the National School Construction Program (*Programa Nacional de Edificaciones Escolares*, PNEE hereafter). The PNEE mandated the construction of 28,000 classrooms across primary and secondary schools over a four-year period, which would be needed to meet the demands of full-time schooling under the JEE. Prior to the reform, school buildings functioned in two or three shifts to accommodate high student demand. In 2012, about half of the public enrollment attended a morning shift, while just 2% of public enrollment was in full-time instruction.[5]

The expansion required to move away from the multiple shift model–without overcrowding–corresponds to a 78 percent increase in the number of classrooms available in 2013. To achieve this goal, 425 schools were refurbished or expanded, and over 1,300 new schools were built. An additional 100 new ECD centers built to satisfy the needs of QEC. This represents an increase of 87% relative to the 114 ECD centers that existed in 2013.

2.1.1 Construction Lotteries

The years leading up to the PNEE saw several corruption claims related to the procurement of school construction contracts. Cases were brought to court in which the government paid selected firms 20 percent of the value of awarded contracts in advance and, subsequently, the construction of many schools was severely delayed or never begun. To promote procurement transparency in the aftermath of these corruption cases, the government decided to allocate new school construction contracts for the PNEE and QEC through open lotteries carried out by the Office of Procurement Services and the Ministry of Education. Civil engineers, architects, and construction companies fulfilling certain minimum requirements were invited to participate in the lotteries, either as firms or as individuals.

School construction, expansion, and renovation projects were grouped into lots to be drawn together and allocated to a single firm or individual through the lottery process. Lots

[5]The matriculation shares by shift are: 49% to morning shifts, 44% to evening, 5% to night, and 2% to full-time school days.

had different budgets determined by the number of classrooms they included. Each lot could only involve construction work related to one school. Overall, 1,833 lots of construction contracts for the PNEE and QEC were allocated through lotteries between 2012 and 2014. These lots were then drawn through four different lottery waves for the PNEE and one lottery wave for QEC, for a total of five lottery waves. See Appendix A for more details on the lottery assignment.

2.2 Urban markets in the Dominican Republic

The private sector is a key actor in the primary and secondary school sector in the Dominican Republic, especially in urban markets where it accounts for more than 30% of total enrollment.[6] Accordingly, nearly 70% of private schools are located in these areas. Private schools are usually independently run and tend to be smaller, with a median per-grade enrollment of 40 students compared to 147 in public schools. In contrast to some other countries in the region, private schools are not subsidized via vouchers and their tuition fees translate one-to-one to out-of-pocket expenses from parents. The average private school charges $650 yearly for secondary grades, and must attract enough students to cover costs.

3 Data

We combine several datasets to characterize the schools and student population affected by the expansion policy and to determine its impact along several margins and outcomes of interest. We use census data together with records from a nation-wide conditional cash transfer program to define schooling markets and to characterize the population served by primary and secondary schools. We also have access to detailed administrative education data including enrollment records of the universe of primary and secondary schools linked with students' performance on the national high-school exams (PN), which we use to create measures of enrollment and quality of schools. We combine these data with administrative records from private school prices and investment strategies. We use detailed data from the school construction program to track the construction progress and inauguration status of every new public school in every educational markets in the country. Finally, we complement our data with in-depth surveys to parents, students and school principals.

[6] Private schools account for a 23 and 25% of the primary and secondary schooling enrollment throughout the country.

3.1 Census and Conditional Cash Transfer Program:

Census: We have access to the Census micro-data from 2010. We leverage the granularity and geographic nature of these data to create geographic schooling markets. We also classify the population into three groups, according to educational attainment: less than primary, completed primary, and high school or more. Combining this measure with population counts, we are able to characterize every educational market in terms of size and educational demographics.

Conditional Cash Transfer Program: Prosoli is a conditional-cash transfer program that provides families with income conditional on the school attendance of the children in the family. We have access to the universe of beneficiaries of the program. We use this dataset with two goals. First, we merge household locations to Census data to quantify the share of prosoli beneficiaries in each of the markets, and use it as an income proxy. Second, this data contains detailed information on school-choice of every individual in the program, that we use to construct micro-moments that will be used for the estimation of the model.

3.2 Schools and Students

Enrollment: We use administrative school enrollment data for all students from the 2010 to 2019 school years. This dataset is provided by the Ministry of Education and contains information on every student enrolled at any primary or secondary education institution in the Dominican Republic. We have access to student identifiers that allow us to track individuals over time and across schools, and to link them to other data sources. In addition to the student level data, we also observe school characteristics such as its location, district, sector (private or public), and the shifts they serve (morning, afternoon, or evening). We use these data to construct a panel of aggregate matriculation by school and grade level over time.

National Examinations: We have access to the universe of students taking national exams (*Pruebas Nacionales* - PN hereafter) since 2010. These examinations were mandatory for promotion for 8th graders until 2016, and are still required for 12th graders. The data includes students' course GPA and standardized exam scores. We link these data to the enrollment records to create value-added measures for every high school in our sample. In 2016 the research team included a short questionnaire to the universe of students taking PN that allows us to recover the education of their parents as well as other individual characteristics.

Private Schools: Every year, the Ministry of Education surveys private schools to keep track of their pricing and expenditure strategies. The principal of the school is requested to report the posted prices for each of the grades offered by the school, as well as the expected tuition fee for next year. In addition, private schools need to report the investment they are expecting to make to justify the price increases. The investment form is broken down by infrastructure, equipment, teacher training, labor conditions of teachers and administrative staff.

In-depth Surveys: We performed in-depth surveys to student and principals from both private and public schools. We use this information to construct model moments and to validate our demand and supply model estimates.

3.3 Lotteries and School construction

Construction: We collected information from several stages of the school construction program. In particular, for each construction lot we observe the contractor assigned by the lottery, as well as every other lottery applicant. From the lottery records we observe whether the contractor is a firm or an individual, together with a tax identifier that allows us to match to several characteristics from tax records. We also observe the expected and realized budget for each project, as well as the number of classrooms specified in the original project. Finally, we merge our data to records from the Ministry of Public Works to track the progress of every construction over time. Specifically we observe the initial and final dates of construction, annual construction progress and the inauguration date of the school.

4 Event Study Analysis

We start the analysis by estimating the causal effect of a new school opening on a variety of equilibrium outcomes. We employ an event study empirical design that identifies effects based on the differential timing of when local areas had a new public school open.

The complexity of the policy – several waves of assignments of new or expanded schools – plus heterogeneous student substitution patterns across schools require some choices in summarizing the treatment and who might be affected. We first define a "neighborhood" as the Census's administrative unit of a "barrio/paraje." These neighborhoods vary somewhat in size with a mean area of 3.8 square kilometers throughout the country but 0.7 square kilometers in urban areas. As students travel an average of 1.6 kilometers from home to school, these neighborhoods are small and schools in the same neighborhood would reasonably be competing for students that live in a similar area.[7] The analysis will test for effects of new public schools on students or other schools in the same neighborhood. As distance is just one factor determining students' choices, we may not perfectly capture exposure to a new public school. But we opt for the transparency of using a simple exposure definition and later explore robustness.

The second choice is how to define the timing of the event. For each neighborhood, we will use the opening of the first new school as the event. Thus, post-event periods may have larger treatment effects both because the impact of the initial new school grows over time and because additional new schools are opening in the same neighborhood.

4.1 Effects on Enrollment

We start by examining whether the new schools changed where students enroll. For school j in year t, we find the first assigned new public school in the same neighborhood n as j. Let $YearInaugurated_n$ indicate when the public school opened and first enrolled students. For some outcome y_{jt}, we specify the following event study model:[8]

$$y_{jt} = \sum_{\tau=-3}^{3} \beta_\tau 1\{YearInaugurated_{n(j)} = t + \tau\} + \theta_j + \theta_t + \epsilon_{jt}$$

If the government assigned new schools according to local (time-varying) educational needs, the year of inauguration may not be orthogonal to neighborhood shocks. Therefore, we use the details of the public procurement process to construct an instrument for when a neighborhood receives a new school. First, because the project assignments were via lottery, the characteristics of the assigned builder are independent of the project characteristics. Second, because so many projects were assigned, many assigned builders were not part of a firm but rather entered the procurement lottery as an individual. We compare project characteristics – budgets upon assignment and updated budgets post-assignment as well as

[7] Results are qualitatively similar if instead of using administrative neighborhoods, we define overlapping treatment units based on how far students or schools are from a new public school.

[8] We also use the same specification for neighborhood outcomes y_{nt} where we include neighborhood instead of school fixed effects.

the number of classrooms – across projects assigned to firms or individuals. We control for the wave-province in which the project was assigned and use only within-lottery variation. We present the results of these balance tests in Table 1. We see that whether a project was assigned to a firm is not statistically related to the project characteristics.

Post-assignment, if firms, based on experience, selection of the best builders, or well-developed supplier networks, are more efficient at building schools, then neighborhoods whose new schools were assigned to firms may receive their finished schools sooner than neighborhoods whose new schools were assigned to individuals. We assess this possibility in Figure 3 where we plot histograms of the number of years between when a new school project was assigned and when it was inaugurated, separately by whether the assigned builder was a firm or an individual. We see overlap – plenty of individuals finish projects quickly – but on average firms complete their projects slightly faster. We present the first stage regressions in Table 2. The endogenous regressors are time until or since inauguration while the instruments are whether the project was assigned to a firm, interacted with how many years it has been since *project assignment*. We see that projects that are assigned to firms are more likely to open two years post-assignment than projects assigned to individuals. This increase of 9 percentage points borrows largely from opening probabilities in the subsequent two years.

Before proceeding to instrumental variable estimates, we consider the exclusion restriction. Building time may not be the only project outcome affected by the assigned builder. In particular, firms may build faster and produce higher quality buildings. If this were the case, we would find it difficult to separately identify the effect of the opening of a homogeneous new school from the effect of a higher quality school. Separately identifying these effects is not necessarily crucial for demonstrating that public provision can have direct effects on students and equilibrium effects in terms of competition. But for interpreting the treatment within our model, we need to take a stand on what variation the builder generates. Here, we rely on the quality inspections that the Dominican government carried out. If school building quality was too low, the inspectors required changes prior to the school opening. This inspection process potentially limits variation in building quality upon opening.

We now present our event study results, using whether the project was assigned to a firm to generate variation in when a new school opened. In Table 3 we show the estimates for the effect of a new school on enrollment outcomes. The first column is the enrollment of the new school. Unsurprisingly, the new school's enrollment increases once it is opened.[9] This increase in new public school enrollment appears to come at the expense of private schools, as the neighborhood's number of private schools and their within-neighborhood enrollment shares both decrease after the new public school opens. The effect on the number of private schools grows stronger over time.

Additionally, because our neighborhoods are small geographic units, and because some students are on the margin of dropping out, we test whether the new public school increases total neighborhood enrollment and we indeed find a large effect. We show the results visually with event study coefficient plots in Figure 4.

The effect on the number of private schools could come from reduced entry or increased exit. We therefore run exit regressions where the outcome is whether a school exited in year t and present the results in Table 4 (and the last panel of Figure 4). We see very strong increases in the private school exit rate, increasing linearly with each year post-inauguration of the new public school.

These event study estimates indicate that the new public schools induced a large reallocation of enrollment toward the new schools and away from the private schools.[10] This reallocation has the potential to change private schools' profits in a way that led to extensive margin responses (exit). But if there are also intensive margin responses to increased competition, these may lead to better student outcomes.

4.2 Effects on School Characteristics and Student Outcomes

To assess the effect on school characteristics beyond enrollment, we must condition on a school being open. We therefore have variation only within neighborhoods that have private schools (if the characteristic is specific to private schools). This sample restriction costs us considerable statistical power. Thus, we drop our instrumental variables strategy and instead refine the time fixed effects to vary with the new public school's assignment lottery wave, l. Thus, we estimate the event studies assuming within-lottery variation in when a school opens is orthogonal to time-varying local schooling outcomes (except through the presence of the new school). Specifically, we estimate:

[9] If all schools were truly "new" we would have no data prior to opening. Many of the new schools are in fact expansions, so we have pre-period outcomes.

[10] We find little effect on student dropout.

$$y_{jt} = \sum_{\tau=-3}^{3} \beta_\tau 1\{YearInaugurated_j = t + \tau\} + \theta_j + \theta_{lt} + \epsilon_{jt}$$

We start by examining private school prices in Table 5. Dominican private schools typically charge both an annual enrollment fee and then a monthly fee. We include prices from pre-school ("inicial"), primary school ("basica"), and secondary school ("general"). We see precipitous price drops in monthly fees following the opening of a new public school, with monthly secondary school prices falling by more than half of the average monthly price. These price drops could reflect reduction in market power but could also entail less spending on productive inputs that generate learning outcomes. In Table 6 we present event study results for the effect of the new public school on private school test scores.[11] We see a moderate increase for grade 12 test scores.[12]

Private schools may achieve test score increases by providing better quality instruction or by selecting higher achieving students. This could be particularly likely in a context, like this, where there is a large enrollment reallocation across sectors. In Table 7 we test for changes in sorting of students by sector. For each high school, we calculate the mean 8th grade test score of its 9th grade entering cohort. These tests were taken when the attended a different school for primary education and thus are not subject to mechanical changes from improvements in private high school instruction. Our estimates are somewhat noisy, but we do not find any strong evidence that students are changing how they sort to sector, at least according to past test scores. Thus, it appears that the reduced revenue may reflect reduction in market power rather than lowering quality.

Another margin of competition could be in school hours. We thus test whether private schools increased the length of the school day to compete with full-day public schooling. In Table 8 we see that while the number of public school hours increased with the new schools, private schools did not adjust. Private schools go from offering more hours on average before the school construction policy to offering fewer after.

In summary, we find that the new public schools had large impacts on the market. They increased public sector enrollment, which led to private schools either closing or lowering their prices and increasing their quality.

[11]As is standard in the education literature, we standardize test scores into z-scores based on grade-year means and standard deviations. Thus, a 0.1 coefficient reflects a change by one-tenth of a student standard deviation.

[12]We see similar results if we use grade 12 scores, residualized by the student's grade 8 score.

5 Empirical Model

The event study analysis demonstrates that the new public schools affected student outcomes by changing the schooling market equilibrium. To isolate the impacts of each mechanism and to conduct counterfactual policy analysis, we specify an empirical model of student choice in secondary education (demand), school pricing and exit (supply), and value-added (technology).

In each school year t, the schooling economy exists of a set of 9th grade students indexed by $i = 1, ..., I_t$ and sets of potential schools indexed by $j = 1, ..., J_t$. We let $j = 0$ represent the option of not attending any school. Each student and school belongs to a market m. Let I_{mt} and J_{mt} denote the set of students and schools, respectively, in market m in school year t.

5.1 Educational Markets

Defining the market is a difficult task in many settings when physical distance is a relevant characteristic. It is generally not easy to find a boundary where one market ends and one begins in broad urban areas. Papers that study retail markets typically have used political or administrative boundaries to define markets such as cities or counties (Davis, 2006). In some cases, such as small isolated communities, this works well. However, in large urban areas consumers close to the border of a county might also be close to firms in the next county. In these cases, it is possible for consumers to choose to cross market lines to buy from firms in neighboring markets. In this application, we follow Neilson (2013) and take advantage of the relatively sparse distribution of the population in the Dominican Republic where communities tend to be far from each other. This creates a natural definition of a market based on the idea that consumers in one city will not travel very far across rural areas to go to school in another city.[13]

We define a schooling market m as follows:

1. Geographic boundaries B^m (a polygon).
2. A set of schools J^m that operate within at any point in time.
3. A set of I^m students of K observable types that live inside the market.

[13]In principle, we could forgo separating students and schools into markets and estimate a single equilibrium. But dividing the country into markets yields considerable savings in computational time.

4. A distribution of student types across markets. The distribution is described by Π^m which is a vector of length K containing the shares of each type of student in the market m. We have that $\sum_k^K \Pi_k^m = 1$ for each market m and $\sum_k^K S_k^m = S^m$.

5. A set of N^m nodes spread evenly within the boundaries of the market that describe where students are located.

6. A distribution of student types across nodes within each market. This distribution is described by w_k^m which is a vector of length N^m containing the share of students of type k of the market m that are located at each node n. We have that $\sum_n^{N^m} w_{nk} = 1$ and $\sum_k^K \sum_n^{N^m} w_{nk} \Pi_k S_k^m = S^m$

We use neighborhoods from the 2010 Census (BPs hereafter, for *Barrio/Paraje*) as the building blocks of our markets since they are small enough to be entirely classified as either urban or rural. The average BP has an area of 3.8 km^2 and is populated by 250 households. For our analysis, we define a BP as urban if the neighborhood meets any of the two restrictions: i) an urban BP as classified by the census, or ii) population density of over 1000 inhabitants per km^2. Figure 5 highlights all the selected urban BPs. The mean area of the selected neighborhoods is 0.7 km^2.

Next, we proceeded to join all urban neighborhoods that were separated by 2 km or less at their closest distance. This resulted in 302 non-overlapping markets that could comprise a single neighborhood (isolated urban areas) or any number up to 272 neighborhoods. Around each market, we added a 1 km buffer. Figure 6 shows an example of the market of Santo Domingo. After defining the market boundaries, we overlay a grid of squared nodes on top to have a standardized geographic unit that is consistent over time. The nodes are $400m \times 400m$ and thus have an area of $0.16 km^2$.

For each market we recover the number of households and total population from Census 2010 data. We combine it with Prosoli data to recover the number of poor individuals in the area. The population of the markets is binned into 6 groups (based on 3 levels of educational level attained interacted with poverty status) that may predict heterogeneous responses to the policy in place. In Appendix B we explain in greater detail the markets definition and the linkage of the different datasets.

5.2 Demand Side: Student Choice of School

Ninth grade students are heterogeneous in their observable characteristics and preferences. Students' observable characteristics are whether their family qualifies for the Prosoli program x_{it}^{p}, the education level of their mother x_i^e, and their residential location l_{it}. Prosoli eligibility is a binary status while mother's education level takes on three values (1 if mother did not complete primary school, 2 if mother completed primary school only, and 3 if mother completed secondary school). Students' unobservables preferences are the vector ν_i.

Students choose a single school to attend (or dropout) under an open enrollment system where we assume capacity constraints do not bind.[14] Schools are differentiated in terms of their price p_{jt}, value-added μ_{jt}, hours of instruction hr_{jt}, whether they are private $priv_{jt}$, a location l_{jt}, and an unobservable (to the econometrician) ξ_{jt}. Let $dist_{ijt} = d(l_{it}, l_{jt})$ where $d()$ is the geodetic distance function.

Student i's utility from attending school j in school year t is:

$$u_{ijt} = -\alpha_i p_{jt} + \beta_i^{\mu}\mu_{jt} + \beta^{hr} hr_{jt} + \beta_i^{priv} priv_{jt} - \gamma dist_{ijt} + \xi_{jt} + \epsilon_{ijt} \tag{1}$$

with $u_{i0t} = \epsilon_{i0t}$ representing the utility from not attending school (dropping out).

We specify the preference coefficients as:

$$\begin{aligned}
\beta_i^{\mu} &= \bar{\beta}^{\mu} + \beta_p^{\mu} x_{it}^{p} + \beta_{e1}^{\mu} 1\{x_i^e = 1\} + \beta_{e2}^{\mu} 1\{x_i^e = 2\} + \sigma^{\mu}\nu_i^{\mu} \\
\beta_i^{priv} &= \bar{\beta^{priv}} + \beta_p^{priv} x_{it}^{p} + \beta_{e1}^{priv} 1\{x_i^e = 1\} + \beta_{e2}^{priv} 1\{x_i^e = 2\} + \sigma^{priv}\nu_i^{priv} \\
\alpha_i &= \bar{\alpha} + \alpha_p x_{it}^{p} + \alpha_{e1} 1\{x_i^e = 1\} + \alpha_{e2} 1\{x_i^e = 2\} + \sigma^{p}\nu_i^{p}
\end{aligned}$$

with $(\nu_i^{\mu}, \nu_i^{priv}, \ln(\nu_i^{p})) \sim^{iid} N(0,1)$. We let $\epsilon_{ijt} \sim^{iid} T1EV$.

5.3 Supply Side: Private School Supply

We assume that public schools are non-strategic; i.e., their supply (and characteristics) are determined exogenously by government policy. We model private schools as simultaneously choosing whether to remain open and what price to charge to maximize profits.[15]

[14] In our principal surveys, most schools report finding space for a student even if already at capacity.

[15] Unlike private schools in the US and other countries, Dominican private schools are usually for-profit entities. Most schools do not have religious or other affiliation that might lead to alternate objectives.

We treat this choice as a static problem such that schools make choices to maximize this year's profits only.

Let private school j's marginal and fixed costs in school year t be mc_{jt} and FC_{jt}, respectively. Define $Exit_{jt}$ to be an indicator for whether the school shuts down and let $Q_{jt}(p_{jt}, p_{-j,t}, Exit_{-j,t})$ be the total number of students enrolling as a function of own price (p_{jt}), other schools' prices ($p_{-j,t}$), and other schools' exit decisions ($Exit_{-j,t}$).

We assume that schools' marginal costs (mc_{jt}) are public information while schools' fixed costs (FC_{jt}) are private information, drawn from a known CDF, F.[16] Because schools do not know each other's fixed costs and choices are made simultaneously, schools must predict the enrollment they would get as a function of their own decisions.

A natural choice would be for schools to maximize expected enrollment, EQ_{jt}, where the expectation is taken over other schools' fixed cost distributions. But because markets can include many schools, this expectation is over such a high-dimensional space that we consider it unlikely schools are sophisticated enough to make such a calculation. An alternative would be to have schools only consider their close neighbors' fixed cost distributions. Schools, however, have overlapping sets of neighbors such that even change in schools far away could shift the demand for a local school.

We instead make a behavioral assumption that simplifies the school's problem while preserving much of the strategic considerations of competing against products with varying degrees of differentiation. We follow Sánchez (2018) by assuming that schools keep track of their own type (fixed cost draw) and the expected market equilibrium.[17] Specifically, let $V_{ikt} = u_{ikt} - \epsilon_{ikt}$ be student i's utility, without the ϵ_{ikt}, to attending school k if the school is open and $V_{ikt} = -\infty$ if closed. Then let

$$\lambda_{ijt} = E_{FC_{-j,t}} \sum_{k \neq j, \text{k private}} exp(V_{ikt}) \tag{2}$$

be the expected total $exp(V_{ikt})$ over all other private schools, where the expectation is over the other schools' fixed cost distributions (or, the probability other schools will be open). With our Type I Extreme Value assumption on ϵ_{ijt}, the perceived probability that student

[16]This choice reflects the government's regulation of private schools which requires schools to report their prices, as well as any changes to costs that affect prices. The gathering of such data into a public database potentially allows schools to observe each others' marginal costs.

[17]In future versions, we will pursue a hybrid approach that has schools keep track of their own type, their 5 closest neighbors' types, and the expected equilibrium in a "fringe."

i chooses private school j in school year t is:

$$\begin{aligned}\tilde{P_{ijt}} &= \frac{exp(V_{ijt})}{1 + exp(V_{ijt}) + \sum_{\text{k public}} exp(V_{ikt}) + E_{FC_{-j,t}} \sum_{k \neq j, \text{k private}} exp(V_{ikt})} \\ &= \frac{exp(V_{ijt})}{1 + exp(V_{ijt}) + \sum_{\text{k public}} exp(V_{ikt}) + \lambda_{ijt}}\end{aligned} \tag{3}$$

This vector λ_{ijt} has a separate element for each student, which allows schools to consider carefully the heterogeneity in its potential students. With this perceived probability for each student, the school's perceived quantity of students is $\tilde{Q_{jt}} = \sum_i \tilde{P_{ijt}}$.

Given the behavioral assumption, private school j simultaneously chooses whether to exit and its price to maximize perceived profits:

$$\max_{Exit_{jt}, p_{jt}} (1 - Exit_{jt}) \left[(p_{jt} - mc_{jt}) \tilde{Q_{jt}}(p_{jt}, \lambda) - FC_{jt} \right]. \tag{4}$$

where we have replaced $Q_{jt}(p_{jt}, p_{-j,t}, Exit_{-j,t})$ with $\tilde{Q_{jt}}(p_{jt}, \lambda)$.

Finally, we assume that the other time-varying school characteristics – hours of instruction and value-added – evolve exogenously. In future drafts, we hope to relax this assumption and, in particular, allow private schools to make endogenous investments that shift their value-added.

5.4 Technology: Value-Added

We model the school's technology in producing learning outcomes as evolving exogenously. Following the literature, we define a school's value-added (μ_{jt}) as the school's causal effect on test scores in school year t controlling for a flexible function of past test scores:

$$y_{it}^{12} = \beta_1^{VA} y_{i,t-4}^8 + \beta_2^{VA} (y_{i,t-4}^8)^2 + \beta_3^{VA} (y_{i,t-4}^8)^3 + \beta_4^{VA} + \mu_{jt} + \nu_{it} \tag{5}$$

where y_{it}^{12} is student i's 12th grade test score in year t and $y_{i,t-4}^8$ is student i's 8th grade test score in year $t-4$. Because students in the Dominican Republic do not take annual tests, we control for scores from 8th grade.

5.5 Equilibrium: Static "Oblivious" Equilibrium

A static "oblivious" equilibrium is a set of prices ($p_{jt} \forall$ j private), exit decisions ($Exit_{jt} \forall$ j private), and beliefs ($\lambda_{ijt} \forall i,$ j private) such that:

1. $p_{jt}, Exit_{jt}$ solve 4 for each private school j
2. $\lambda_{ijt} = \sum_{k \neq j}(1 - Pr(Exit_k(\lambda)))exp(V_{ikt})$ for each student i and private school j (consistent beliefs).

6 Estimation and Identification

6.1 Estimation

We separately estimate the technology, demand, and supply.[18]

6.1.1 Technology

We start by estimating each school's value-added for each school year by running OLS regressions of Equation 5 and recovering the estimated fixed effects, $\hat{\mu_{jt}}$. We further estimate shrunken estimates, $\tilde{\mu}_{jt}$, to use in our demand model that account for estimation error.

6.1.2 Household Locations

Students vary according to three observables: whether their household qualifies for Prosoli, mother's education, and household location. We simulate student observables using the Prosoli and Census data. Specifically, we define a grid with nodes 400m apart and use the Census to estimate the number of households closest to each node and the distribution of mother's education. We then merge the distribution of Prosoli status onto each node. For more details, see Appendix B.

[18]Future versions will jointly estimate demand and supply.

6.1.3 Demand

We estimate demand using simulated method of moments, where we simulate from the distribution of ν_i. We combine aggregate share moments with instrumental variable moments and micro moments and use a nested fixed point estimation routine as in Berry et al. (2004).

For the aggregate shares, we calculate each school's 9th grade enrollment share for each market-year using the administrative enrollment data. For the dropout, or outside option, share, we count the number of enrolled 8th graders from the prior year who did not enroll in 9th grade. For each year, we thus have J_t moments. We use these market share moments to recover mean utilities δ_{jt} $(= -\bar{\alpha}p_{jt} + \bar{\beta}^{\mu}\tilde{\mu}_{jt} + \beta^{hr}hr_{jt} + \beta^{\bar{priv}}priv_{jt} + \xi_{jt})$ in the inner loop.

As school pricing decisions likely depend on ξ_{jt}, we specify instruments for price.[19] The first instrument takes advantage of 2013 changes to Prosoli, the conditional cash transfer program, which expanded the program to families with high school students. Thus, the transfer families received for students attending high school suddenly increased in 2013. For each private school j we calculate the share of families in neighborhood $n(j)$ who are eligible for Prosoli. We form $z^p_{jt} = ShareProsoli_{n(j)}1\{t > 2013\}$.[20] The second instrument leverages the school expansion policy's effect on hours offered at public schools, even those that were not new. Once the new schools started to open, many of the other public schools converted to full-day instruction, on a staggered basis. Thus, even if there were no new public school (yet) in a neighborhood, the incumbent public school may offer more hours of instruction and thus exert competitive pressure on the private school. While a natural response might be for private schools to adjust their own instructional hours, we do not see much evidence of this. Instead, we allow increased competition through hours to affect private schools' pricing decisions. We thus construct z^h_{jt} as the mean number of instructional hours per student offered by public schools in private school j's neighborhood in year t. For these instruments, we impose that they are orthogonal to a private school's unobserved demand shock (ξ_{jt}).[21]

Finally, we specify a set of micro moments. Using our Prosoli administrative data and the 2016 survey of test-takers, we can match school choices to individual students

[19] In future versions we will also instrument for $\tilde{\mu}_{jt}$.

[20] We also include $ShareProsoli_{n(j)}$ directly in the utility model so that the instrument captures differential effects of being in high Prosoli-eligible neighborhoods after the policy change.

[21] Future versions will incorporate a third price instrument. Public school teacher salaries increased during our sample, and particularly for teachers measured to be high quality. As public and private schools may compete for teachers, private school costs likely increased.

and their demographics. We construct micro moments for the mean school characteristics (price, quality, private, number of hours) for each demographic group (Prosoli eligibility and mother's education). We supplement these mean choice characteristics with the mean distance traveled to school from our in-depth student survey.[22] This survey also asks students what school would be their second choice, after the one they are actually choosing. We use this second choice data to construct covariances in school characteristics (price, quality, private, number of hours) between first and second choices. Lastly, the survey asked students if school prices made them choose a school that is different from the one that they would choose if prices were not an issue. We calculate the change in the probability of wanting to attend a private school if prices were removed.[23]

For the survey-based moments, we calculate them within the model using the same populations (e.g., 2016 students for the moments based on the 2016 survey of test-takers). We estimate the model with two-step simulated method of moments where we use an optimal weighting matrix.

6.1.4 Supply

We specify F, the CDF of fixed costs FC_{jt}, as a distribution with 9 mass points, and we estimate the probabilities of each. We specify the mass points (in thousands of Dominican pesos) as $(0,100,300,500,700,1000,2000,4000,\infty)$. We opt for a distribution with mass points instead of a continuous but parametric distribution to avoid extrapolating far from the underlying variation.

With this specification, we estimate the supply model using the iterative procedure described in Sánchez (2018). For fixed $\vec{\lambda}$, we estimate marginal costs by inverting the pricing first-order condition. We note that our data provides a unique opportunity to invert first-order conditions even for schools that exit. Many private schools report planned prices for the following school year to the government.

With marginal costs, we estimate variable profits for each school and year conditional on remaining open. The fixed cost mass point probabilities are then chosen to match the empirical exit probabilities by bin of estimated variable profits (e.g., variable profits between 100,000 and 300,000 Dominican pesos). We then update $\vec{\lambda}$ to be consistent with

[22] Because we place household locations at discrete nodes, we calculate the distance traveled between a student's assigned node and school attended. The distance is very similar to the household to school distance.

[23] In the model, we match this moment by calculating choice probabilities when all schools have 0 prices.

the marginal cost and fixed cost distribution estimates and iterate until we find a fixed point in $\vec{\lambda}$.

6.2 Identification

In this subsection, we highlight the sources of variation in the data that prove useful for identifying model parameters. The price instruments exploit policy variation in the conditional cash transfer program and the move to full-day public schooling. The identification assumption is that the neighborhoods most affected by the policy changes did not differ – in terms of the private schools – from the neighborhoods less affected. For the expansion of the conditional cash transfer program, areas with more eligible households are likely different in terms of their schooling options. But once we control for eligibility, we assume that the timing of the policy variation is unrelated to other shocks to the private school market. For the change in public school hours, we assume that the staggered roll-out was unrelated to the local supply of private schools. Additionally, our survey question that asks for choices if prices did not matter yields considerable information about how students trade off price and other characteristics. The hypothetical nature of the survey question also provides clean variation in price that holds everything else, including potential equilibrium responses from more standard instruments, fixed.[24]

The other micro moments map fairly clearly into standard arguments for identifying demand models with heterogeneous preferences. The extent to which students from different demographic groups choose different types of schools pins down preference heterogeneity based on demographics. In terms of the random coefficients, the second choice survey responses are crucial. If unobservable preference heterogeneity for school characteristics is large in magnitude, then we would expect that conditional on mean choice probabilities, covariances in choice characteristics across first and second choices will be high.

For the supply estimates, our assumption of optimal pricing and the oblivious equilibrium allow us to invert first-order conditions to recover marginal costs. We identify the fixed cost distribution by assuming that the observed exit decisions reflect a trade-off between static variable profits and fixed costs.

[24]Because the other price instruments could plausibly lead to changes on the extensive margin of which schools are open, it is possible that ξ_{jt} becomes correlated with the instruments through selection. We therefore will assess robustness to using only the variation from the hypothetical survey question for identifying the price coefficient.

6.3 Parameter Estimates

We present preliminary estimates from estimation of our structural model. We start with a demand model that maintains some of the preference heterogeneity from Equation 1. Specifically, we let preferences over prices, quality (value-added), and whether a school is private vary with whether the student's mother has completed primary school.[25] We present the estimates in Table 9. We estimate significant mean disutilities to distance and price, with a kilometer of distance valued the same as 1,800 USD in annual tuition. In terms of preference heterogeneity, we estimate that children with more educated mothers are less price sensitive (with half the coefficient as the children with less educated mothers) and have higher valuations on school value-added and whether the school is private. These patterns are unsurprising given students with higher educated mothers are more likely to enroll in private schools, especially those with higher prices and higher value-added.

On the supply side, we estimate considerable variation in school-year markups of price relative to marginal cost (Figure 7). The distribution of price minus marginal cost (as a fraction of price) is somewhat evenly distributed between 0 and 1, with a median private school having a markup above 50%. Schools are sufficiently differentiated, mainly in space, that they are able to exert quite a bit of market power. Thus, adding another (public) option to students' choice sets has the potential to induce large increases in competition and lower prices, as seen in Section 4. For the estimated fixed cost distribution, we plot estimated exit rates as a function of variable profits in Figure 8. While we did not impose it, we estimate a monotonic relationship between exit rates and variable profits. We find that schools with variable profits below 2,000 USD have an exit probability of 7% each year. This decreases to a probability of 2% for schools with variable profits above 6,000 USD, reflecting a fairly large concentration of schools with fixed costs between 2,000 and 6,000 USD.

7 Counterfactuals

With our estimated model, we can conduct several policy counterfactuals that assess how the supply of private schools is affected by the level and type of policy intervention. For each counterfactual, we solve for a counterfactual equilibrium for the year 2018. We

[25] Future versions will expand the preference heterogeneity to all 3 categories of mother's education, to eligibility for Prosoli, and to the unobservable random coefficients, ν_i.

take the set of schools open in the data and use the estimated equilibrium $\vec{\lambda}$ as a starting value in searching for the counterfactual equilibrium. When we vary starting values across a broad range of values for $\vec{\lambda}$ we always return to the same equilibrium.

7.1 Increased Level of Public Provision

The first counterfactual assesses how private school prices and exit rates vary with the size of the public sector expansion. In addition to providing more options for students, which may increase welfare through any match effects (or, e.g., less distanced traveled) the size of the public sector has direct consequences for the equilibrium supply of private schools. The pro-competitive impact may lead to reduced private school market power and prices, which transfers savings to households and possibly reallocates students toward the (on average) higher quality private sector. On the other hand, expansion of the subsidized sector may crowd out high quality private schools, which might lower student outcomes.

For different levels of public sector expansion, ranging from the observed expansion (100%) to three times as big (300%), we present the counterfactual private school mean prices and three-year exit rates in Figure 9.[26] We estimate that the pro-competitive effects on lower market power increase nearly linearly with the size of the public school expansion, though the scale of the y-axis shows that these effects are fairly small. The private school exit rates though are large and mostly concentrated in the initial public school expansion. This largely reflects a mass of small schools with relatively high fixed costs whose supply is particularly sensitive to market conditions. Once these schools have exited, the remaining schools are less subject to exit risk. Thus, expanding the public sector even further has the potential to be pro-competitive, albeit with fairly large drops in market power.

7.2 Vouchers

Whether direct public provision or a different policy lever is most effective at reducing private school markups while ensuring adequate supply of high quality schools is an empirical question. Direct provision offers an additional differentiated product, not just on price but also location, which may cut severely into private school market shares unless they lower prices. But by lowering private school profits, some high quality schools may

[26]We place new public schools randomly in the geographic markets and draw their characteristics from the empirical distribution of the actual new public schools.

no longer be able to compete. Private school vouchers, instead, may have smaller effects on closure rates by subsidizing the private sector, while still possibly lowering markups if the competition within the private sector (at low prices) is strong enough. We assess this empirically by implementing voucher counterfactuals that vary the size of the voucher as a fraction of the median price in the data.

In Figure 10 we plot the average private school price and three-year exit rates from different voucher sizes. We see that the prices fall dramatically, and linearly, with larger vouchers. The subsidy eliminates a key source of differentiation, and there are enough private schools remaining that competition becomes fierce to the point that markups fall substantially. Exit rates are lower than we predict under public provision, but they are still increasing in the voucher size (at least at low levels). This might appear counter-intuitive, as the subsidy raises profits holding prices fixed, but the price competition becomes so intense that many small school can no longer compete.

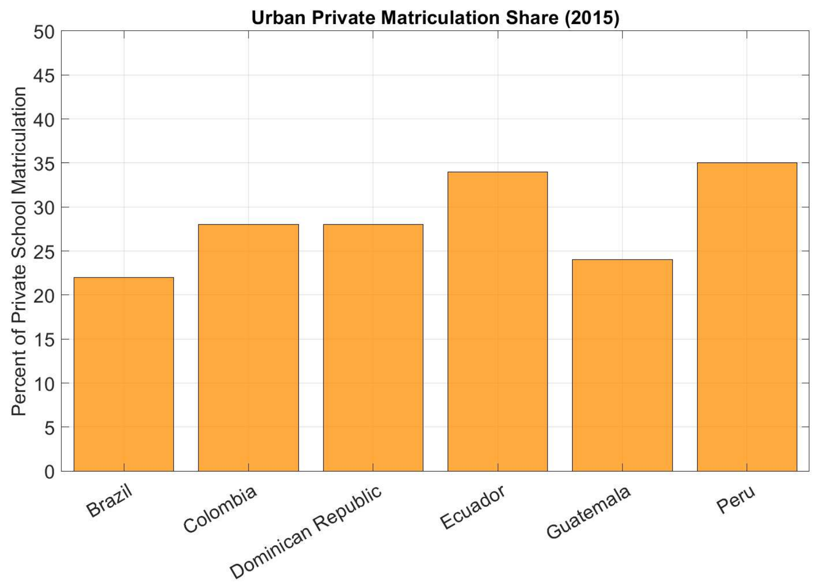
8 Conclusion

In this paper, we examined the interplay between public and private provision of education. Using a large public school construction initiative in the Dominican Republic, we found that increased public provision crowded out part of the private sector while also exerting competitive pressure that reduced private market power and increased quality. We specified and estimated a supply and demand model of education that enabled us to study how equilibrium effects scale with the amount of public provision and the specific funding instruments.

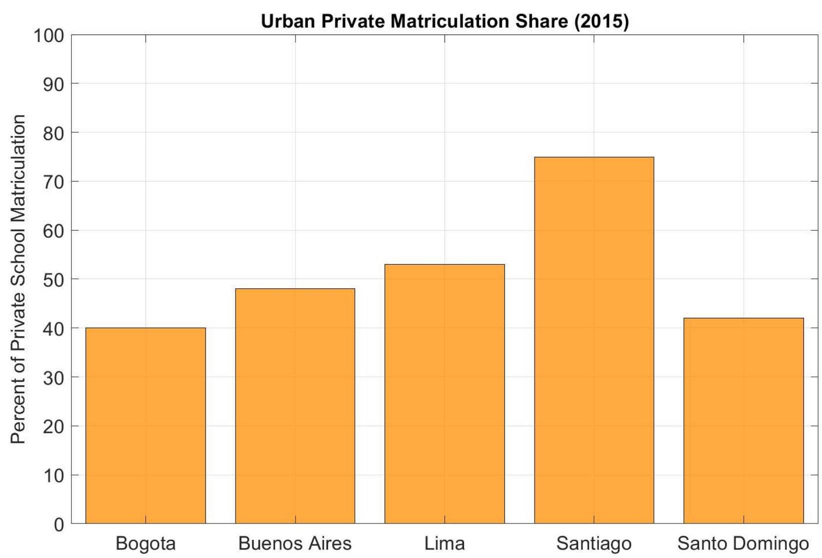
Future drafts will incorporate heterogeneous treatment effects in the event study analysis, extend the amount of preference heterogeneity in the demand model, and expand the counterfactual analysis.

Figures and Tables

Figure 1: Private Share of Primary Enrollment in LAC



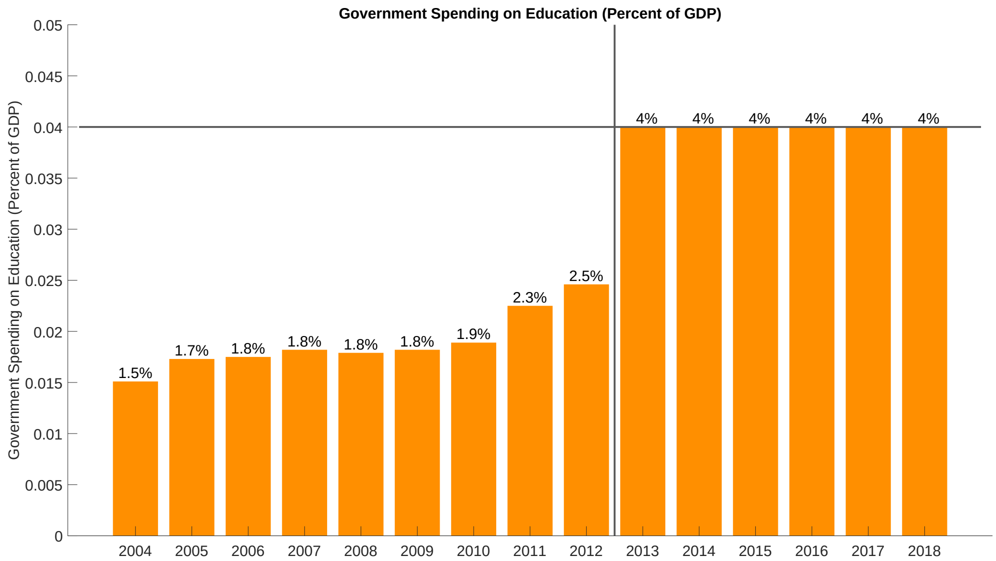
(a) Countries



(b) Large Cities

Notes: This figure shows the private school share in Latin America. Panel a) shows the share across countries, and Panel b) across large cities.

Figure 2: Government Spending on Education



Notes: This figure presents the evolution of the GDP share allocated to education over time for the Dominican Republic.

Figure 3: Histogram of Years from Assignment to Inauguration, by Builder Type

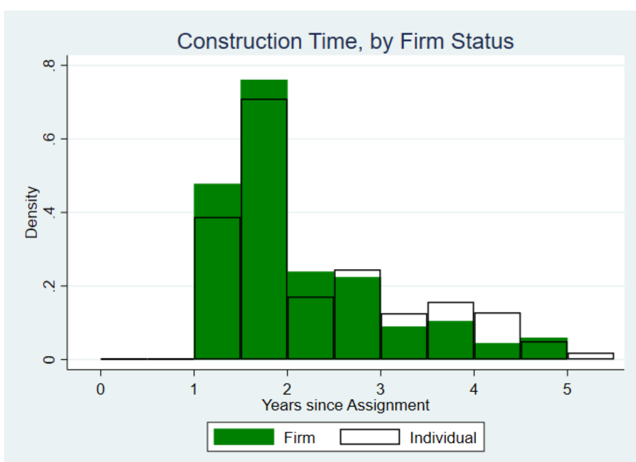


Figure 4: IV Event Study Estimates: Enrollment Outcomes

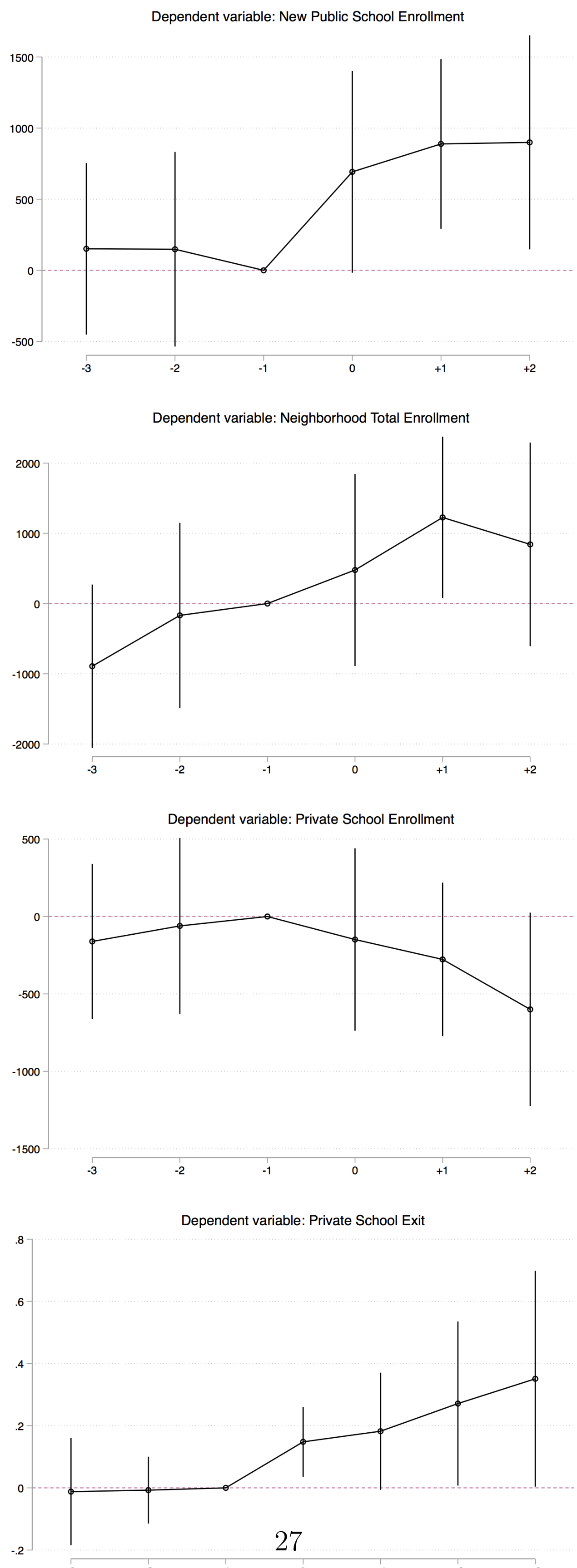


Figure 5: Urban Barrios/Parajes

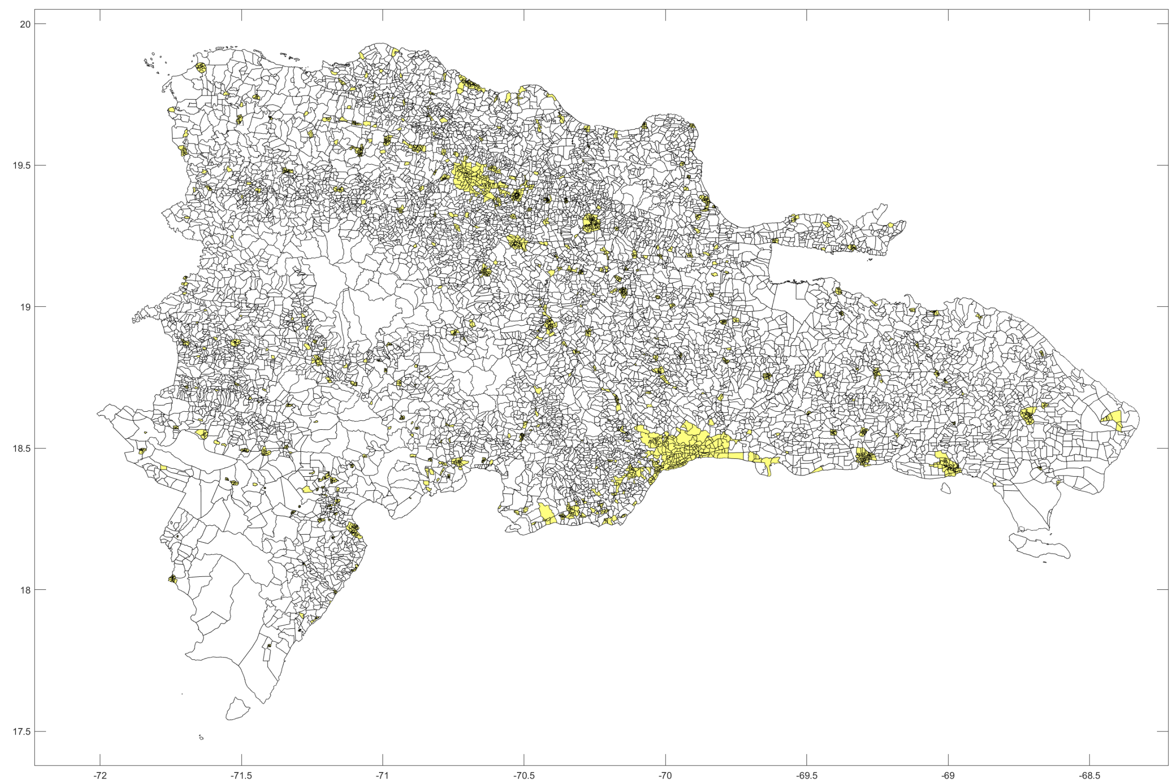


Figure 6: Market definition: Santo Domingo and Boca Chica

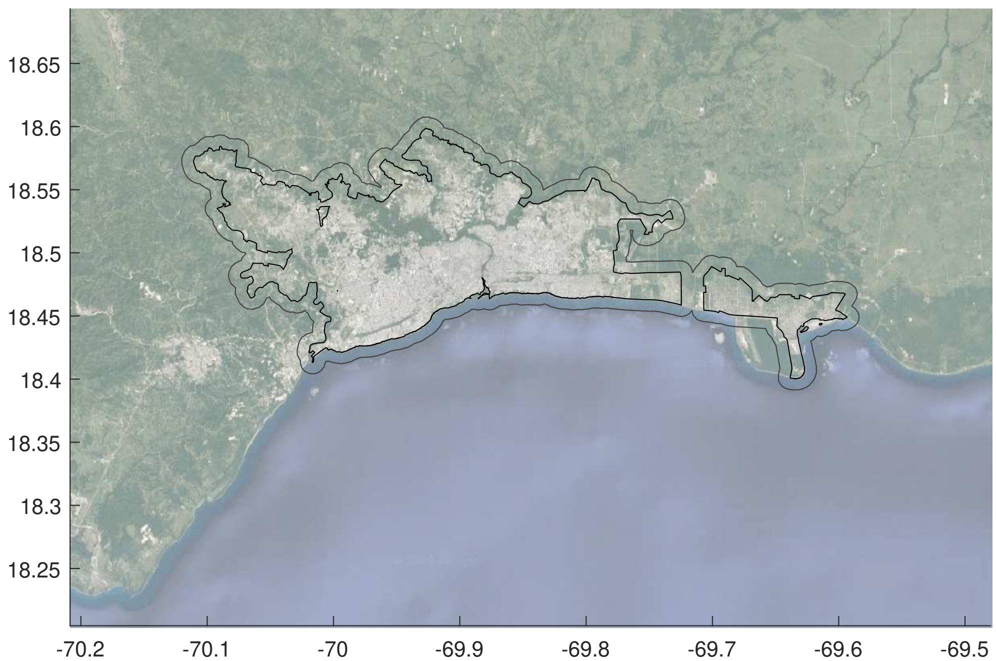


Figure 7: Distribution of Estimated Markups

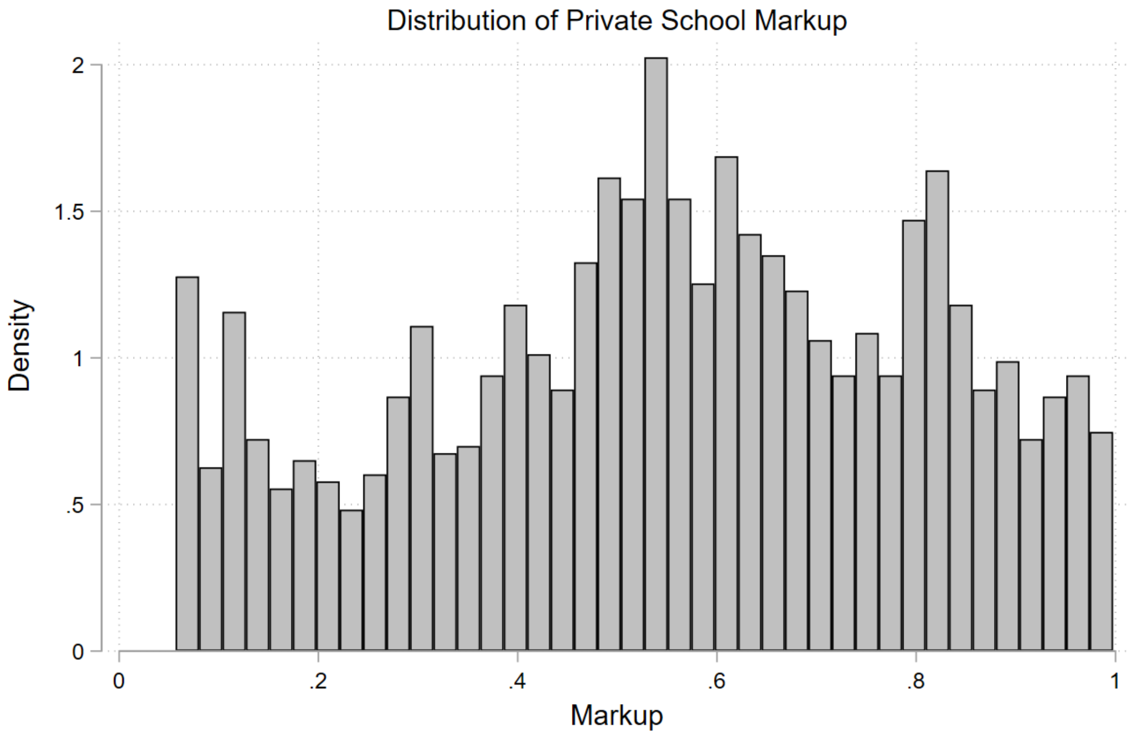


Figure 8: Supply Fixed Cost Estimates

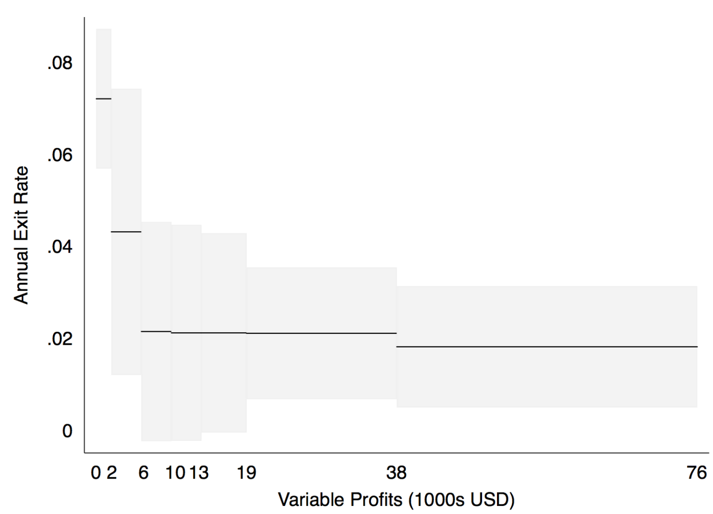


Figure 9: Counterfactual 1: Increase Level of Public Provision

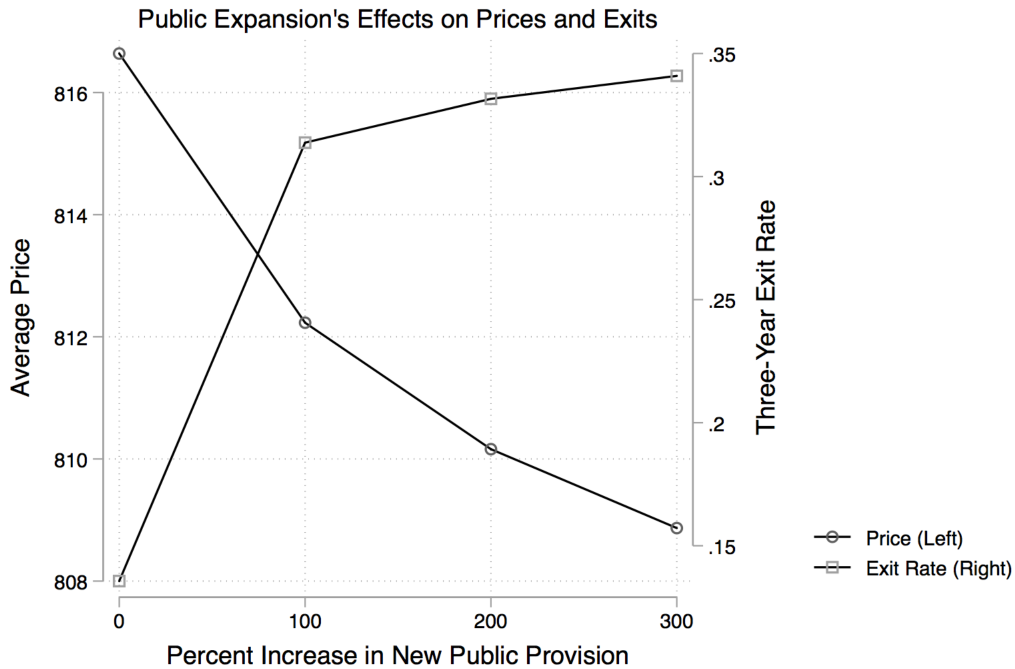


Figure 10: Counterfactual 2: Vouchers

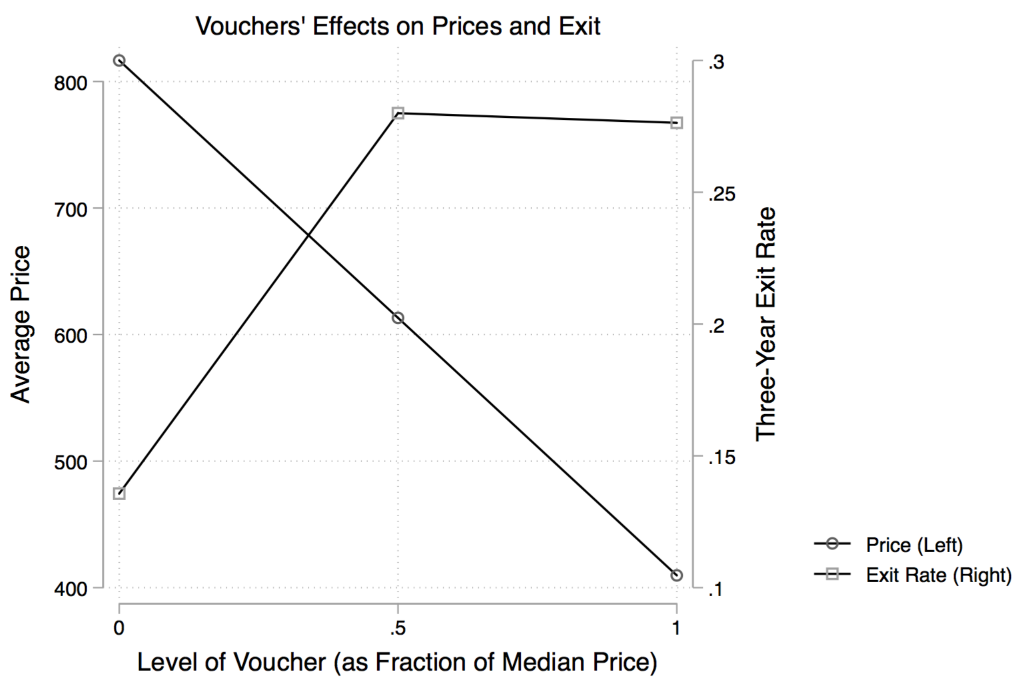


Table 1: Balance test

	(1) Budget_2015	(2) Budget_2016	(3) Budget_2017	(4) Budget_2018	(5) # Classrooms
Firm	-2.025	0.468	0.365	-1.650	-1.126
	(1.410)	(3.122)	(2.697)	(1.421)	(0.751)
ymean	37.05	38.82	40.15	36.90	16.63
r2	0.577	0.240	0.589	0.570	0.292
N	845	845	366	850	850

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: IV Event Study Analysis: First Stage

	(1) Inauguration t-3	(2) Inauguration t-2	(3) Inauguration t	(4) Inauguration t+1	(5) Inauguration t+2
FirmX5 years before lottery	-0.0129	0.00639	0.0212	-0.0105	-0.00335
	(0.145)	(0.172)	(0.147)	(0.130)	(0.101)
FirmX4 years before lottery	-0.00229	-0.00439	-0.00655	0.0139	-0.00583
	(0.136)	(0.161)	(0.139)	(0.122)	(0.0952)
FirmX3 years before lottery	0.0281	-0.00276	-0.0103	-0.00481	-0.00373
	(0.0356)	(0.0422)	(0.0362)	(0.0319)	(0.0249)
FirmX2 years before lottery	0.00402	-0.00296	-0.000389	-0.000152	-0.000143
	(0.0311)	(0.0369)	(0.0317)	(0.0279)	(0.0218)
FirmX1 year before lottery	0.0298	-0.0267	-0.000175	-0.0000226	-0.0000116
	(0.0311)	(0.0369)	(0.0317)	(0.0279)	(0.0218)
FirmXLottery year	-0.0627**	0.0925**	-0.00286	-0.000144	-0.0000108
	(0.0311)	(0.0369)	(0.0317)	(0.0279)	(0.0218)
FirmX1 year after lottery	-0.0280	-0.0347	-0.0268	-0.00284	-0.000153
	(0.0311)	(0.0369)	(0.0317)	(0.0279)	(0.0218)
FirmX2 years after lottery	0.000834	-0.0288	0.0922***	-0.0264	-0.00294
	(0.0311)	(0.0369)	(0.0317)	(0.0279)	(0.0218)
FirmX3 years after lottery	0.0285	-0.0281	-0.0403	0.0996***	-0.0303
	(0.0314)	(0.0372)	(0.0320)	(0.0282)	(0.0220)
FirmX4 years after lottery	0.0666	-0.0249	-0.0830*	0.00793	0.124***
	(0.0443)	(0.0524)	(0.0450)	(0.0396)	(0.0309)
R_Squared	0.780	0.288	0.387	0.463	0.590
N	7632	7632	7632	7632	7632
#Neighborhoods	899	899	899	899	899

Standard errors in parentheses. Lottery-province, year, and school fixed effects included.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: IV Event Study Analysis: Enrollment Outcomes

	(1) New Public School Enrollment	(2) Number of Private Schools	(3) Private Share of Enrollment	(4) Neighborhood Total Enrollment
Inauguration t-3	151.3	-0.518	0.115	-892.2
	(307.5)	(1.103)	(0.110)	(592.9)
Inauguration t-2	148.2	-0.185	0.0715	-168.4
	(348.9)	(1.251)	(0.125)	(672.6)
Inauguration t-1	0	0	0	0
	(.)	(.)	(.)	(.)
Inauguration t	692.4*	-0.333	-0.120	477.5
	(361.8)	(1.297)	(0.130)	(697.4)
Inauguration t+1	888.7***	-0.541	-0.191*	1225.9**
	(304.3)	(1.091)	(0.109)	(586.7)
Inauguration t+2	899.6**	-2.377*	-0.296**	842.5
	(383.8)	(1.376)	(0.137)	(739.8)
Mean y	293.4	3.011	0.110	2147.7
#Neighborhoods	899	899	899	899
N	7632	7632	7632	7632

Standard errors in parentheses. Lottery-province, year, and school fixed effects included.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: IV Event Study Analysis: Exit

	(1)
	Exit
Inauguration t-3	-0.0122
	(0.0879)
Inauguration t-2	-0.00727
	(0.0549)
Inauguration t-1	0
	(.)
Inauguration t	0.148***
	(0.0574)
Inauguration t+1	0.182*
	(0.0961)
Inauguration t+2	0.271**
	(0.135)
Inauguration t+3	0.351**
	(0.177)
Mean y	0.0694
#Schools	706
N	6354

Lottery-year, year, and school fixed effects included.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Event Study Analysis: Private School Prices

	(1) Inicial Enrollment Fee	(2) Basica Enrollment Fee	(3) General Media Enrollment Fee	(4) Inicial Monthly Fee	(5) Basica Monthly Fee	(6) General Media Monthly Fee
Inauguration t-3	-198.4	-236.9	-159.7	-74.80	-86.40	305.6
	(180.8)	(188.5)	(296.3)	(107.1)	(120.7)	(315.9)
Inauguration t-2	-52.55	-68.49	68.74	-6.013	1.718	428.2*
	(110.7)	(115.3)	(125.0)	(58.59)	(64.21)	(225.0)
Inauguration t-1	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)
Inauguration t	-80.09	-114.3	-133.1	-93.01**	-110.4**	-278.9**
	(86.02)	(89.68)	(151.0)	(40.50)	(48.41)	(130.7)
Inauguration t+1	-36.62	-45.30	-225.6	-152.4**	-132.3	-613.3**
	(146.5)	(161.8)	(263.0)	(72.75)	(102.2)	(283.4)
Inauguration t+2	-3.030	6.922	-280.0	-211.8*	-186.1	-858.2**
	(216.1)	(235.7)	(389.9)	(112.2)	(141.1)	(404.8)
Inauguration t+3	116.9	155.3	-130.8	-222.2	-232.0	-1177.0**
	(304.9)	(333.6)	(522.3)	(155.1)	(197.9)	(559.8)
Mean y	1689.4	1689.5	2361.9	1315.0	1350.9	1996.7
#Schools	235	234	111	234	233	110
N	803	800	402	803	801	400

Table 6: Event Study Analysis: Standardized Private Schools Test Score

	(1) Math 8 Score	(2) Math 12 Score	(3) Spanish 8 Score	(4) Spanish 12 Score
Inauguration t-3	0.0230	-0.0228	0.0108	-0.0217
	(0.0623)	(0.0576)	(0.0460)	(0.0407)
Inauguration t-2	-0.0128	-0.0322	-0.000966	-0.0217
	(0.0441)	(0.0412)	(0.0326)	(0.0291)
Inauguration t-1	0	0	0	0
	(.)	(.)	(.)	(.)
Inauguration t	0.0231	0.0549	-0.0285	0.0411
	(0.0465)	(0.0429)	(0.0343)	(0.0303)
Inauguration t+1	0.0500	0.100*	0.0117	0.103**
	(0.0664)	(0.0599)	(0.0490)	(0.0423)
Inauguration t+2	0.0633	0.0795	0.0522	0.105*
	(0.0954)	(0.0851)	(0.0704)	(0.0601)
Mean y	0.0811	0.148	0.173	0.164
#Schools	219	122	219	122
N	1331	751	1331	751

Table 7: Event Study Analysis: Selection of Students

	(1) 9th Graders' Mean 8th Grade Math Score	(2) 9th Graders' Mean 8th Grade Spanish Score	(3) 9th Graders' Mean 8th Grade Math Score	(4) 9th Graders' Mean 8th Grade Spanish Score
Inauguration t-3	-0.0335	0.0546	0.0259	0.0280
	(0.211)	(0.114)	(0.191)	(0.0915)
Inauguration t-2	-0.0352	0.0250	-0.0291	0.0221
	(0.114)	(0.0613)	(0.0857)	(0.0476)
Inauguration t-1	0	0	0	0
	(.)	(.)	(.)	(.)
Inauguration t	0.0681	0.00580	-0.0555	-0.0283
	(0.111)	(0.0594)	(0.0813)	(0.0436)
Inauguration t+1	0.0337	-0.0921	-0.102	-0.0327
	(0.211)	(0.114)	(0.167)	(0.0824)
Inauguration t+2	0.121	-0.0314	-0.0677	0.0177
	(0.316)	(0.163)	(0.251)	(0.124)
Inauguration t+3	0.114	-0.0626	-0.144	-0.0131
	(0.427)	(0.218)	(0.339)	(0.164)
Sample	Private	Private	Public	Public
Mean y	0.122	0.240	-0.0367	-0.142
#Schools	153	153	410	410
N	709	709	1821	1821

Table 8: Event Study Analysis: School Hours

	(1) Private	(2) Public	(3) Overall
Inauguration t-3	0.00789	-0.0807	-0.0270
	(0.0247)	(0.0873)	(0.0653)
Inauguration t-2	0.000929	-0.121*	-0.0806
	(0.0183)	(0.0693)	(0.0504)
Inauguration t-1	0	0	0
	(.)	(.)	(.)
Inauguration t	-0.0219	1.184***	0.743***
	(0.0189)	(0.0688)	(0.0508)
Inauguration t+1	-0.0204	1.274***	0.879***
	(0.0251)	(0.0850)	(0.0645)
Inauguration t+2	0.00142	1.387***	1.104***
	(0.0327)	(0.107)	(0.0823)
Inauguration t+3	0.0331	1.560***	1.453***
	(0.0433)	(0.138)	(0.107)
Mean y	5.699	6.207	6.026
#Schools	702	1300	1992
N	5495	9964	15461

Table 9: Demand Estimates

Parameter	Estimate	SE
Distance Coeff (per km)	0.71	(0.08)
Mean Price Coeff (per 1000 USD)	0.40	(0.12)
Price Coeff (High)	-0.20	(0.10)
VA Coeff (High)	1.72	(0.60)
Private Coeff (High)	1.32	(0.34)

Notes: Demand model estimated via simulated method of moments.

References

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy 112*(1), 68–105.

Card, D., M. D. Dooley, and A. A. Payne (2010). School competition and efficiency with publicly funded catholic schools. *American Economic Journal: Applied Economics 2*(4), 150–176.

Davis, P. (2006). Spatial competition in retail markets: Movie theaters. *The RAND Journal of Economics 4*(37), 964?982.

Dinerstein, M. and T. Smith (2018). Quantifying the supply response of private schools to public policies.

Downes, T. A. and D. Schoeman (1998). School finance reform and private school enrollment: Evidence from california. *Journal of Urban Economics 43*(3), 418–443.

Epple, D., D. Figlio, and R. Romano (2004). Competition between private and public schools: Testing stratification and pricing predictions. *Journal of Public Economics 88*(7), 1215–1245.

Estevan, F. (2015). Public education expenditures and private school enrollment. *Canadian Journal of Economics/Revue canadienne d'économique 48*(2), 561–584.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science 312*(5782), 1900–1902.

Hoxby, C. M. (1994). Do private schools provide competition for public schools? *NBER Working Paper 4978*.

Hoxby, C. M. (2001). All school finance equalizations are not created equal. *The Quarterly Journal of Economics*, 1189–1231.

Hoxby, C. M. (2003). School choice and school productivity (or could school choice be a tide that lifts all boats?). In C. M. Hoxby (Ed.), *The Economics of School Choice*. University of Chicago and NBER Press.

McMillan, R. (2005). Erratum to 'competition, incentives, and public school productivity'. *Journal of Public Economics 89*(5), 1133–1154.

Neilson, C. (2013). Targeted vouchers, competition among schools, and the academic achievement of poor students. *Unpublished Working Paper*.

OECD (2016). Pisa 2015 assessment and analytical framework. *PISA*.

Sánchez, C. (2018). Understanding school competition under voucher regimes.

Shonkoff, J. P. and D. A. Phillips (2000). *From neurons to neighborhoods: the science of early child development*. Washington, D.C.: National Academy Press.

UNESCO (2015). Terce 2013: Tercer estudio regional comparativo y explicativo. *TERCE*.

World Bank (2018). World development indicators. *The World Bank Group*.

Appendices

A School Lotteries

Each lottery was divided into the 32 provinces that make up the country. Provinces had different numbers of construction lots depending on their size and existing school infrastructure. For example, in the first lottery round, Santo Domingo (the province home to the country's capital city of the same name) included 43 lots while Dajabón included just three. In any given lottery round, applicants were only allowed to participate in a single province of their choosing.

For every lottery round, each of the 32 provinces held simultaneous lotteries. Applicants were required to attend the lottery in the province in which they participated. The lottery process worked as follows. Each applicant who fulfilled the minimum requirements received a lottery number that was posted online the day before the draw. The day of the lottery, all numbers were then placed in an urn and, for every lot, three applicants were randomly drawn. The applicant holding the first drawn number was assigned as the winner of that lot and their number was removed from the urn. In case the winner proved unable to complete the contract, the applicants holding the numbers drawn second and third were assigned as possible replacements. The backups' numbers were then put back inside the urn. As a result, lottery winners could obtain a contract for at most one school, while those in second and third places could still compete for another contract.

Table 10 provides a summary of each lottery round, including its budget, the number of lots it included, and the date on which it was held. The first lottery took place in November 2012, and the last one in December 2014. Each lottery round included between 100 lots – for the QEC round – and 548 lots and had average contract values ranging between 26 and 58 million Dominican pesos (or 0.602 and 1.347 million USD). The number of lots included in each lottery increased in later rounds as did the size of the contracts, reflecting the fact that that later lotteries included more contracts for the construction – as opposed to renovation – of classrooms.

Table 11 presents the number of participants and winners per lottery round. Naturally, the number of winners in each lottery round is the same as the number of lots, while the number of second and third places is smaller as any given participant could be drawn in second or third place for multiple lots. Across all lottery rounds, the majority of applicants

Table 10: Lotteries Description

Lottery Process	Date	Lots	Budget (in 1,000s RD$ ≈ 23 USD in 2013) total	mean	min	p25	p75	max
PNEE #1	11-30-2012	372	15,166,190	40,770	3,689	30,704	56,634	70,335
PNEE #2	01-31-2013	548	14,349,634	26,185	359	22,521	30,748	73,883
QEC #1	09-13-2013	100	32,826,944	32,826	32,826	32,826	32,826	32,826
PNEE #3	11-19-2013	401	23,494,580	58,590	23,349	44,769	68,364	74,177
PNEE #4	12-17-2014	462	26,324,082	56,979	20,100	32,826	79,910	141,946

Notes: Add Notes

were individuals, although the share of applicants that were firms increased from 13 percent in the first round to 20 percent in the last round.

Table 11: Lotteries Participants and Winners

	All				Individuals				Firms			
Lottery	All	1st	2nd	3rd	All	1st	2nd	3rd	All	1st	2nd	3rd
PNEE #1	3427	371	353	268	3029	310	305	233	398	61	48	35
PNEE #2	8423	548	521	527	7130	474	442	443	1293	74	79	84
QEC #1	6053	100	100	100	-	79	89	86	5241	21	11	14
PNEE #3	9737	401	394	390	8111	329	315	324	1626	72	79	66
PNEE #4	13354	462	453	453	11157	368	362	384	2197	94	91	69

Notes: Add Notes

Given the random nature of the assignment, we observe a similar distribution of firms and individuals among winners. Although winners were excluded from draws for subsequent lots, they were still able to participate in future lottery rounds as long as they had delivered the contracted classrooms and terminated their previous contract beforehand. The probability of participating in another lottery round conditional on having participated at all varies between 70 and 75 percent.

B Data preparation for demand estimations

This appendix contains:

- An explanation of the steps followed in the delimitation of the schooling markets' boundaries
- A description of the data build that combines census and Prosoli data to characterize households in markets

B.1 Market construction - Standard methodology overview

The standard set of steps to define a schooling market m is as follows:

1. Geographic boundaries B^m (a polygon).
2. A set of schools F^m that operate within at any point in time.
3. A set of S^m students of K observable types that live inside the market.
4. A distribution of student types across markets. The distribution is described by Π^m which is a vector of length K containing the shares of each type of student in the market m. We have that $\sum_k^K \Pi_k^m = 1$ for each market m and $\sum_k^K S_k^m = S^m$.
5. A set of N^m nodes spread evenly within the boundaries of the market that describe where students are located.
6. A distribution of student types across nodes within each market. This distribution is described by w_k^m which is a vector of length N^m containing the share of students of type k of the market m that are located at each node n. We have that $\sum_n^{N^m} w_{nk} = 1$ and $\sum_k^K \sum_n^{N^m} w_{nk} \Pi_k S_k^m = S^m$

B.2 Geographic Boundaries (B^m)

The census map data in the Dominican Republic is divided into Regions (10), Provinces (32), Municipalities (155), Districts (386), Sections (1,565), Neighborhoods (12,565), Polygons, Supervision Areas, and Segments. We use neighborhoods (BPs hereafter, for *Barrio/Paraje*) as the building blocks of our markets since they are small enough to be entirely classified as either urban or rural. The average BP has an area of 3.8 km^2 and is populated by 250 households.

The first step is to select the urban BPs. Using the census' classification of urban areas yielded 2,620 of them as a starting point. However, overlaying the roads map on top suggested that this definition was too restrictive as it was ignoring a significant portion of urban sprawl. Therefore, we extended the definition so that the starting point to build the markets would be the set of BPs classified as urban in the census plus any BP with a population density of over 1000 inhabitants per km^2. Figure B.1 highlights all the selected urban BPs (before joining them to generate markets). The mean area of the selected neighborhoods is 0.7 km^2.

Then, we proceeded to join all urban neighborhoods that were separated by 2 km or less at their closest distance. This resulted in 302 non-overlapping markets that could comprise a single neighborhood (isolated urban areas) or any number up to 272 neighborhoods. Around each market, we added a 1 km buffer. Figure **??** shows as an example the market of Santo Domingo.

B.3 Nodes within markets (N^m)

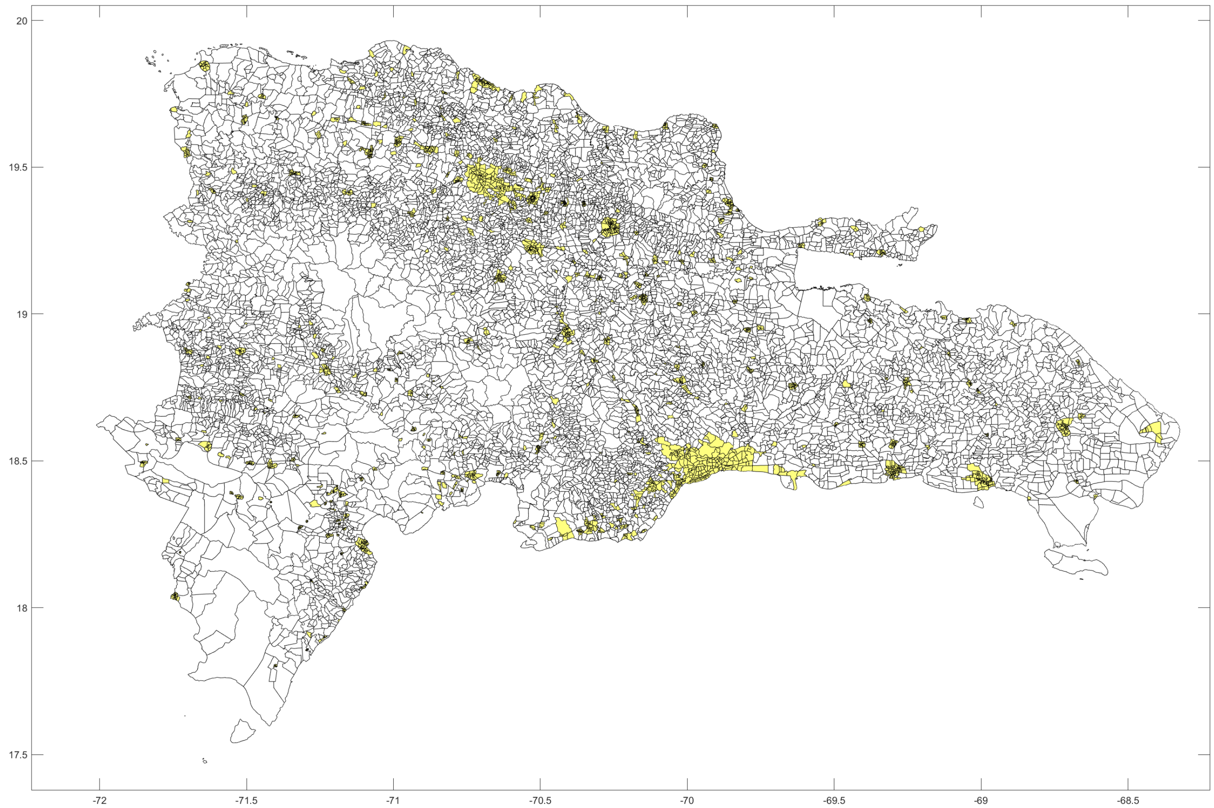
Once the market boundaries have been defined, we overlay a grid of squared nodes on top to have a standardized geographic unit that is consistent over time. The nodes are $400m \times 400m$ and thus have an area of $0.16km^2$. Figure B.3 shows the same market from Figure **??**, divided into homogeneous nodes. It also shows that these nodes allow for a detailed characterization of demographic heterogeneity.

B.4 Construction of Π^m (Market-level data) and w_k^m (Node-level Data)

The population of the markets is binned into 6 groups based on two characteristics (highest educational level attained, poverty status) that may predict heterogeneous responses to the policy in place. The types are defined as:

- Type 1: Less than primary education and poor
- Type 2: Less than primary education and not poor

Figure B.1: Urban Barrios/Parajes



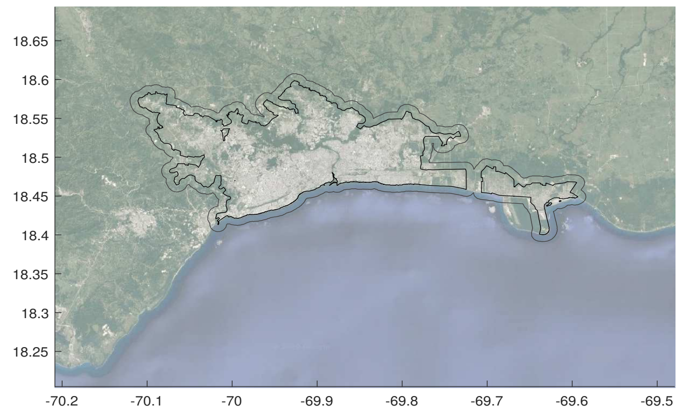
- Type 3: Primary education and poor
- Type 4: Primary education and not poor
- Type 5: Secondary or above and poor
- Type 6: Secondary or above and not poor

For each market, we can recover Π and w_k^m by combining two sources of microdata, and making some assumptions. Here we describe these data and explain the specific steps we followed.

Step 1: Clean census data and get the number of adults over 25 **by Segmento**. Break this number down into N of adults that finished any of the **3 educational levels specified in the types of "Option 1"**. *Output: Segment-level dataset with ONE code, ONE name of barrio, number of adults in each of the 3 education categories (*).*

Step 2: Clean Prosoli data to get the N of poor adults (over 25) by Barrio/Paraje. I get

Figure B.2: Market definition: Santo Domingo and Boca Chica



10,268 BPs. *Output: Barrio-level poor counts dataset with prosoli name*

Step 3: Fuzzy merge Barrio-level prosoli data to MapParajes (ONE registry of Barrios). So now I have the barrio-level dataset from step 2, with ONE code. *Output: Barrio-level poor counts dataset with ONE code*

Step 5: For each *barrio*, spread the number of poor people uniformly among the *segmentos* that intersect with it, using areas of intersection as weights. *Output: Segment-level poor counts dataset with ONE code*

Step 6: Merge output from step 5 to output from step 1[27]. Now we have poor counts and educational level counts by segmento. We assume independent marginal distributions and simulate the joint distribution.

Step 7: Aggregate into homogeneous nodes, again using areas of intersection as weights and assuming the population is uniformly distributed.

[27]One detail to consider is that the Census and Prosoli are 8 years apart. Therefore, all N from the Census data were adjusted to account for population growth using the average rate for the past 8 years in the following way: $x = x \times (1.015)^8$. A pending task is to look for growth rates for specific areas, as population growth most likely is not uniform. However, overcoming this limitation should be straightforward once we access the 2020 census.

Figure B.3: Percentage of moms with a college degree by Node