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

October 28, 2020

PRELIMINARY AND INCOMPLETE
DO NOT DISTRIBUTE OR CITE WITHOUT PERMISSION

Abstract

In 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 state capacity is increasing but low levels of past public provision mean many private schools already exist. 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 the characteristics of the contractor randomly assigned to build the project. 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 level of public provision affects the overall level of quality in the market, we specify and estimate an empirical model of demand (students choosing schools) and supply (schools choosing whether to enter, stay open and what price to charge). We use the model estimates to calculate the level of public provision that maximizes learning. Due to equilibrium competitive

*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. Yiren Ding, María Elena Guerrero, Astrid Pineda, and Nestor Viñals provided excellent research assistance. We would like to thank Liran Einav and Matthew Gentzkow, and seminar participants at Stanford University, Harvard University, University of Chicago Booth, University of Michigan, the Urban Economics Association, and the Virtual BREAD/CEPR/STICERD/TCD Conference on Development Economics 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.

effects, we find that the optimal level is non-monotonic in the quality of the increased public schooling.

Contents

1 Introduction

Economists have long debated how much of essential services (education, healthcare, utilities, etc.) the government should provide directly versus outsource to the private sector. While profit incentives may lead private providers to make cost-reducing investments, the public sector may gain a cost advantage by exerting input market power. Even if the sectors have identical cost functions, public or private provision may be optimal. Public providers may forgo exercising output market power by selling at cost, or by imposing a price cap to ensure universal access to social services. However, a subsidized public option, may lead to inefficiently low levels of quality by divorcing the payers from the consumers.

Co-existing public and private sectors can provide further benefits. Subsidized public provision, with higher-quality private options, allows the government to redistribute income to lower-income families (Besley and Coate, 1991). The presence of multiple options may also lead to fiercer competition. Whether competition from a subsidized sector leads to better outcomes is unclear, however. Increased demand elasticities may cause private providers to lower prices and improve quality (Neilson, 2013). But if private providers cannot remain profitable when competing with a subsidized option, then high-quality options may be crowded out (Dinerstein and Smith, 2018). Thus, the competitive impacts of increased public provision are complicated and may be non-monotonic in the level of provision.[1]

These issues are particularly relevant in low-income countries where state capacity to deliver services may be limited. Across many 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 policy. Indeed, a higher quality private sector might encourage less public provision (and perhaps 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 channels 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 (DR) that increased government spending on education from 2.5% to 4% of GDP in a single year.[2] This 78% expansion of public capacity allows us to examine

[1] We discuss later whether the competitive impacts necessarily arise from public provision or whether subsidized private competition has similar impacts.

[2] As a fraction of GDP, this is nearly identical to the Indonesian school construction program (Duflo, 2001).

the effects of public provision on a student's 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 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, information not typically available in administrative data. 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 were not affiliated with a construction firm. We thus use the randomized assignment of projects to contractor types – specifically, whether the randomized builder is a firm and the builder's pre-lottery log employment and log employee-months – as instruments 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 shifts, 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. For a neighborhood that received a new public school, we estimate that the number of private options decreased by 0.7. Among private schools that remain open, we see a decrease in the prices they charge, and an increase in test score value-added.

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 likely 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, private schools maximize profits by choosing whether to enter the market, 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, especially given the discreteness of the entry and exit choices. 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). Schools keep track of the expected (exponentiated) representative value students receive from attending a competing private school, where schools are uncertain about which competitors will remain open. This behavioral assumption, with consistent beliefs, yields an equilibrium that simplifies computation and does not depend on school leaders solving high-dimensional complex problems.

We estimate the demand model via simulated method of moments. We use aggregate enrollment share moments supplemented by several sources of exogenous variation. We use the same builder characteristics from the event study analysis. 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 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 most schools have significant market power in pricing, capturing nearly all revenues as variable. But our estimates of the fixed cost distribution indicate that total profits are quite low, not surprising given the large exit response to the new public schools.

With our estimated model, we revisit our question of the optimal level of public provision. We consider the government's problem as choosing the amount of public provision (schools) to maximize student learning, subject to a technological constraint that determines the quality of the new schools.[3] To explore this problem, we conduct counterfactual exercises that vary the level of public provision, and at different levels of quality. Our benchmark estimates vary the level of public provision holding fixed the quality at the low levels observed in the schools that opened under the construction initiative.[4]

We explore the consequences of varying the level of public provision from a low level – matching the pre-construction level in the DR – all the way to an expansion three times as large as the actual construction program.[5] We find that the private school exit rate is fairly unresponsive

[3]Future drafts will incorporate the financial cost of provision.

[4]The new schools have mean value-added of -0.06 student standard deviations (σ), the incumbent public schools have mean value-added of -0.02σ, and the private schools have mean value-added of 0.002σ.

[5]The implemented program was not quite half as large as the planned expansion. Thus, larger programs were

to increases in public provision at low levels, but then around the size of the actual policy, exit rates quickly double and increase steadily for higher levels of public provision. Prices follow an opposite path, starting high but then dropping once public provision expands. We evaluate the effect on student learning – which is ex-ante ambiguous – and find that learning (as measured by enrollment-weighted market mean value-added) falls for small expansions in public provision before increasing once the expansion is a bit larger than the actual policy. It peaks just before a policy three times as large as the observed one, before falling slightly. This non-monotonicity reflects several forces – students switching to the new public schools that tend to be lower quality, high-quality private schools lowering prices and attracting more students, and crowd-out of private schools, starting with low-quality but then affecting some high-quality schools.

We end by characterizing how the optimal level of public provision varies with the technology (quality) of the new public schools. While higher quality new schools tend to lead to equilibria with more student learning, this is not always the case. At very high levels of provision, the high quality public schools crowd out enough high quality private schools that learning actually falls (relative to lower quality provision). Further, the relationship between learning and the amount of public provision varies depending on the quality level. While lower quality provision produces more learning at higher levels of provision (near three times the observed policy), the optimal level of higher quality schools is closer to two times the observed policy. These non-monotonicities highlight the importance of considering equilibrium responses when choosing levels of public provision.

1.1 Related Literature

This paper contributes to four different strands of literature. A first and growing strand of literature has delved into the interplay between public and private providers in the delivery of public services. Recent research has shown increased competition from the state to have mixed effects on the prices of private providers.[6] Atal et al. (2020) find the entry of public pharmacies in Chile to lead to an increase in prices and market segmentation, while Jiménez-Hernández and Seira (2020) find the entry of new government milk suppliers in Mexico to lead to price decreases in the private market. These results highlight the fact that price effects are theoretically ambiguous and will depend on consumers' underlying preferences and market structure. Price effects in the private market have been documented not only in response to new public entrants, but also to changes in the pricing schemes of existing public providers (Clemens and Gottlieb, 2017; Jiménez-Hernández

under consideration and are of direct policy interest.

[6]On the flip side, other work has highlighted the quality and cost benefits of outsourcing public services to private providers. Romero et al. (2020) documents increases in test-scores from outsourcing school management in Liberia and Banerjee et al. (2019) document cost reductions as a result of outsourcing food service delivery in India. The positive impacts of increased competition among the outsourced firms have also been documented by Busso and Galiani (2019) in the context of the delivery of basic consumption goods in the Dominican Republic.

and Seira, 2020).

A second, and related, strand of literature examines interactions between public and private schools. Existing research has focused on whether school quality responds to competition (Hoxby, 1994; McMillan, 2005; Card et al., 2010; Neilson, 2013), on how students sort themselves between public and private sectors (Hoxby, 2003; Epple et al., 2004), and on whether school funding reform affects private school enrollment (Downes and Schoeman, 1998; Hoxby, 2001; Estevan, 2015) and entry and exit decisions (Dinerstein and Smith, 2018).

A third strand of relevant literature studies the impacts of public school construction in developing countries. Duflo (2001) shows that a massive school construction program in Indonesia increases educational attainment, and leverages time and cohort exposure variation to estimate the returns to schooling. Kazianga et al. (2013) and Kazianga et al. (2019) examine the impacts of increasing the number of classrooms in existing schools in Burkina Faso and find large increases in enrollment and test-scores. Burde and Linden (2013) find similar effects for a school construction program in Afghanistan.

A fourth strand of literature examines the role and incentives of private schools in developing countries. Existing research has documented large test-score gains of attending private schools in rural Pakistan (Andrabi et al., 2020), and private voucher schools in Chile (Hsieh and Urquiola, 2006) and Colombia (Angrist et al., 2002). Our paper contributes to the nascent literature of equilibrium models of private school competition. Neilson (2013) provides a framework to estimate demand for schools in Chile, Bau (2019) develops and estimate a model of horizontal competition of private schools in Pakistan, and Allende (2020) provides an model of demand and supply of private schooling in Peru that accommodates preferences for peers in equilibrium.

We study the tradeoffs between public and private schools and examine the optimal level of public education provision. We move beyond the existing literature by accounting not only for responses to public school construction by private schools on the intensive margin (e.g. price responses), but also on the extensive margin (i.e. the decision of private firms to enter or exit the market). This paper builds upon Dinerstein and Smith (2018) by adding private school price 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). School attendance 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).[7] 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, which divided students into shifts, or a private school. Perhaps related, education outcomes lagged behind other countries, as reflected by the Dominican Republic ranking last in student skills among all participating countries in the international TERCE in 2013 and PISA test in 2015 (UNESCO, 2015; OECD, 2016).[8]

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).[9] JEE is a program to transition the country from a half-day schooling model to one of full-time schooling, intended to broaden educational 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.[10] The expansion required to move away from the multiple shift model 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 scheduled to be built.

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

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

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

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

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 select firms 20 percent of the value of awarded contracts in advance and, subsequently, the construction of many schools was severely delayed or never begun.[11] 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 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.[12] Accordingly, over 90% 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 for parents. The average private school charges $650 yearly for secondary grades, and must attract enough students to cover costs. Private schools prices are transparent, as they are reported to the government; there is minimal use of price discrimination.

[11] The Dominican Office of Procurement Services provides more details about the prior procurement process here.

[12] Nationally, private schools account for a 23% and 25% of primary and secondary schooling enrollment, respectively.

3 Data

We combine several datasets to characterize the schools and student populations affected by the expansion policy and to determine its impact along several margins. 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 exams, which we use to create measures of school enrollment and quality. We combine these data with administrative records of private school prices and investment strategies. We use detailed data from the school construction program to track the construction progress and inauguration status of the new public schools. Finally, we complement our data with in-depth surveys to parents, students, and school principals.

3.1 Census and Conditional Cash Transfer Program:

Census: We use the 2010 Census micro-data to create schooling markets and estimate the distribution of household types by geography. We classify the population into two groups, according to educational attainment: less than high school or at least some high school. 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 their children's school attendance. 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 across geography. Second, this data contains detailed information on the school attended of every individual in the program, which we use to construct micro-moments for demand model estimation.

3.2 Schools and Students

Enrollment: We use administrative school enrollment data for all students from the 2010 to 2019 school years.[13] 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

[13] Unless indicated otherwise, we refer to school years by their spring year – e.g., 2012 for the 2011-2012 school year.

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 enrollment by school and grade level over time.

National Examinations: We have data on 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 elicited the education of students' parents and other individual characteristics.

Private Schools: Every year, the Ministry of Education surveys private schools to keep track of their pricing and expenditures. 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.

Student and Parent Surveys: We use in-depth surveys of rising 8th graders and their parents that collected information on household demographics, students' assessment of the returns to education, and high school choice (Berry et al., 2020). These survey questions form several micro moments we use in estimating our demand model. In particular, the questions about school choice include hypothetical choices if the student's first option were not available or if students could attend any school for free. These hypothetical choices reveal information about how students would substitute across schools if market structure were to change.

Principal Surveys: We conducted three waves of surveys of public and private school principals. These waves spanned different points in the construction process such that some principals already had a new public competitor that had opened while others anticipated one. The survey responses cover principals' assessment of their competitors, estimates of demand elasticities and fixed operating costs, and responses to over-capacity enrollment. These surveys motivate many of our modeling choices and allow us to validate some of our estimated model parameters with principals' own assessments.

3.3 Lotteries and School Construction

Construction: We collected administrative information across several government agencies from all stages of the school construction program. At the project assignment stage, we observe the project's original specifications in terms of the number of classrooms and expected budget. We further have lottery records that detail every contractor assigned to a project, the backup contractors set to take over in case the original assignee cannot execute the project, and all contractors who entered the lottery but did not win a project. We then match these contractors to the full panel of individual and firm tax records. These records span the period before, during, and after construction and include revenues, employment, government contracts, and supplier networks.

We then supplement the assignment data with data on construction progress. We observe when the government has purchased the land such that construction may begin, annual construction progress, and any subsequent changes to the project budget. Finally, we observe when the new school was finally opened for enrollment.

4 Event Study Analysis

We start the analysis by estimating the causal effect of a new school opening on several 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 a 1 kilometer radius around the (eventual) location of a new public school. As students travel an average of 1.6 kilometers from home to school, this definition means that neighborhoods contain very close competing schools.[14] 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. Within the building process, there are three different events: when the project was assigned by lottery; when the government identified and purchased land such that building could begin; and when the school opened to students. Because the school opening is likely the relevant event for students' enrollment choices, we define the event as a new public school opening and accepting students. But for school outcomes that may be forward-looking, we will also include results where the event is either when the project was assigned or when construction started. In our neighborhood-level analysis, we occasionally have multiple new public schools open near each other. 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

[14] At the end of this section, we extend the results to consider larger, non-overlapping neighborhoods according to administrative boundaries.

new schools are opening in the same neighborhood.

Given this neighborhood and event definition, we specify the following event study model for neighborhood n in year t:

$$y_{nt} = \sum_{\tau=-3}^{3} \beta_\tau 1\{YearInaugurated_n = t + \tau\} + \theta_n + \theta_t + \epsilon_{nt} \tag{1}$$

where y_{nt} is a neighborhood-level outcome and $YearInaugurated_n$ is the year that the neighborhood's first new public school opens.

We could estimate this model via OLS and identify the parameters by comparing how outcomes change in neighborhoods that receive public schools sooner rather than later. But we have a couple of reasons to worry that the timing of treatment may be related to neighborhood trends. First, as the Dominican Republic has grown richer, some urban neighborhoods have become much denser. These areas, with particularly acute overcrowding, were the focus of the earliest waves of new school assignments. Second, within these growing neighborhoods, inexpensive land was difficult to find and the government was concerned about holdup problems with current landholders. Thus, the government often placed the new school in an area of declining density, within a broader growing neighborhood. Because these patterns point in opposite directions for most outcomes, they may not violate parallel trends assumptions. But for count outcomes (e.g., enrollments, number of schools) we find that OLS event study results are sensitive to how narrowly we define the neighborhood in a way that may be consistent with complicated underlying trends.[15]

We thus use the details of the public procurement process to construct instruments for when a neighborhood receives a new school. Because the project assignments were via lottery, the characteristics of the assigned builder are independent of the project characteristics. And because the scale of the construction required so much building, many projects were assigned to contractors with minimal prior experience. We characterize contractor heterogeneity by whether the contractor was part of a firm or registered as an individual, whether the contractor had any employees in 2012, the contractor's log employment in 2012 (if employment was positive), and the contractor's 2012 log employee-months.[16] As a balance test that the assignment of contractors to projects was indeed random, we regress project budgets upon assignment and updated budgets post-assignment, as well as the number of classrooms, on the contractor characteristics. We control for the wave-

[15]Estimating the event study model with OLS also likely leads to treatment effect heterogeneity related to when a school opens, as smaller projects likely finish faster but also have smaller effects on students and competing schools. When estimating with OLS, we follow Sun and Abraham (2020) in interacting the model with event cohort dummies. We find very similar results to the naive event study OLS estimates. While we do not make similar adjustments to the instrument variables event study analysis, the instruments potentially balance treatment effect heterogeneity more evenly across event cohorts.

[16]Employees are typically hired at the monthly level and because the construction industry often offers irregular work, employee-months captures the intensity of labor usage, conditional on annual employment. Employment and employee-months are recorded for firms and individuals who employ others.

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, as expected, the characteristics of the assigned contractor are not statistically related to the project characteristics.

Firms, especially those with higher levels of pre-lottery employment, may be more efficient at building schools due to past experience, the selection of the best builders, or well-developed supplier networks. But if project heterogeneity or government regulations overwhelm the cross-builder variation, our instruments may not shift when a neighborhood's new school opens. We assess this 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. The builder characteristics thus appear to matter for how soon a neighborhood's new school opens.

To incorporate the instruments, we return to our discussion of events. Our endogenous regressors are the vector of event dummies, $1\{YearInaugurated_n = t + \tau\}$, using the school's opening as the event and indexing event time by τ. The instruments matter in terms of how quickly the school is built, once assigned by the lottery. Builders assigned projects in later lotteries may be efficient but are unlikely to finish faster than projects assigned several lotteries earlier. In specifying the first stages, we therefore rely on a different event – when the project was assigned – and index this event time by κ. We specify the set of first stage equations as:

$$1\{YearInaugurated_n = t + \tau\} = \sum_{\kappa=-5}^{4} \beta_\kappa 1\{YearAssigned_n = t + \kappa\} \sum_l z_{nl} + \pi_n + \pi_t + \upsilon_{nt} \quad (2)$$

for κ from -3 to 3 where z_{nl} is the l-th contractor characteristic for neighborhood n's new public school project. For cases where there are multiple new schools per neighborhood, the first stage is based on the first assigned project.

We present first stage regressions (showing just firm status as the builder characteristic) 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 necessary for demonstrating that public provision can have direct effects on students and equilibrium effects in terms of competition. Because the quality level of a new school will only affect student outcomes once the school is open, the event study specification still properly summarizes the timing of the treatment, even if it is a composite treatment. That said, quality differences would be relevant for considering how treatment effects might be heterogeneous across different opening cohorts. Here, the Dominican government's quality inspection process is important. If a new school building's quality was too low, the inspectors required changes prior to the school opening. Indeed, this is a key factor explaining why some inexperienced builders took longer to finish their projects. Thus, the quality inspection process likely compresses much of the variation in building quality once the school opens.

4.1 Effects on Enrollment

We start our neighborhood-level analysis by examining whether the new schools changed where students enroll. In Table 3 we show the estimates for the effect of a new school on enrollment outcomes, and we present corresponding event study plots in Figure 4. The first column is the enrollment of the new school. Unsurprisingly, the new school's enrollment increases once it is opened.[17] Once open, the new schools attract an additional 402 students, with the enrollments increasing by nearly 700 students in the new school's third year.

This large 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. Neighborhood private school enrollments fall by 81 students when the new public school opens, with the decrease almost twice as big two years later (column 3). The average neighborhood has 501 private school students, such that this enrollment decrease is a substantial shift across sectors. In column 4, we show the effect on the private sector's enrollment share in the neighborhood and we see an immediate 4 percentage points drop. This decrease is concentrated in areas with initially high private school shares; but even in these neighborhoods, the decrease is over 10% of initial enrollment. The drop in private school enrollment closely tracks the treatment effects on the number of private schools (column 2). We see high levels of crowd-out, with a new public school leading to 0.27 fewer private schools in the neighborhood after the first year and 0.78 fewer private schools after three years. The average neighborhood has 3 private schools, so this is a remarkable loss in private school supply.

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. Interestingly, we don't find increases, which could indicate that school choice is determined

[17] If all schools were entirely new entities we would have no data prior to opening. Many of the new schools are expansions, so we have pre-period outcomes.

primarily by geography (i.e., no substitution across neighborhoods) or that the small increase in overall supply limit the impact of a new public school on making the neighborhood more attractive in aggregate for potential enrollees. The lack of an effect on drop out may also reflect that (a) before the public school expansion a variety of options still existed and (b) the pre-expansion rationing of hours meant that marginal students considering whether to enroll in school or enter the work force could do both.

4.2 Effects on Private School Supply

Given such a large demand response to the presence of a new school, private schools' competitive positions are shifted. This could lead to the public sector replacing private provision, private provision becoming more efficient, or some combination. We have already seen neighborhood-level evidence that the supply of private schools falls. In this subsection, we explore how the schools respond. We alter our event study specifications to be operate at the school level, where an observation is a private school j in year t, the school's "neighborhood", $n(j)$, is a 1 km radius around the private school, and the second stage event is the first new public school in the neighborhood:

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

$$1\{YearInaugurated_{n(j)} = t + \tau\} = \sum_{\kappa=-5}^{4} \beta_\kappa 1\{YearAssigned_{n(j)} = t + \kappa\} \sum_{l} z_{n(j)l} + \pi_j + \pi_t + \upsilon_{jt}. \tag{4}$$

We start by unpacking the extensive margin response. The effect on the number of private schools could come from reduced entry or increased exit. To test for increased exit rates, we set the private school sample as all schools active at the beginning of our data (2010) and define y_{jt} as the probability school j has exited by year t. We present the estimated event study coefficients in the top panel of Figure 5. Once the new public school opens, we see that the private school cumulative exit rate increases by almost 5 percentage points per year. By the fourth year since the new school has opened, neighboring private schools are over 15 percentage points more likely to have exited, off of a mean sample exit rate of 10.3%. We also see some evidence of a small pre-trend, with exit rates increasing in neighborhoods that *will* receive their new public schools sooner. To the extent that exit decisions are forward-looking and made in anticipation of schools to open in coming years, the extensive margin supply response may precede the new public schools opening. The instruments still lead to exogenous variation in when a new public school is *about* to open, as the builder's progress is observable to the community. Thus, we interpret the pre-trend as an estimate of the forward-looking nature of the exit response.

We further explore the exit response by assessing how its timing relates to the assignment of the new school. For this analysis, we re-define the event to be when the new public school project was assigned to a builder. Because our instruments shift construction completion time, conditional on when the assignment lottery was held, we do not have an instrument for time since lottery and thus run OLS regressions that exploit variation in timing across lotteries. We present the results in the bottom panel of Figure 5. We see that prior to the lottery, there is little evidence of differential pre-trends across private schools soon to receive new or expanded public competitors. Then in the years following the lottery, we see an increase in exit rates. By the second year after the lottery, when some projects are finishing and others are nearing completion, the cumulative exit rate is 5 percentage points higher and increases linearly with each year post-assignment of the new public school. We therefore find consistent evidence that the new public schools led to spikes in the private school exit rate.[18]

Exiting the market is potentially a last-resort response, as schools would likely prefer to remain in operation, if profitable. That such exit occurs implies that many schools did not have a way to adjust prices or quality to stay viable. But if the surviving private schools were able to avoid exit by making their schools more attractive to students, then this improvement is an important consequence of providing more public options.[19] We thus test for whether surviving private schools' characteristics changed in response to a new public school. Because our event study specification includes school fixed effects, we can identify the treatment effects provided selection on the extensive margin is based on time-invariant factors (e.g., a school with persistently high fixed costs). If schools exit in response to the new public school based on time-varying school-specific shocks, then our estimates would be picking up the effect of a new public school on the composition of the private sector.

We start by examining private school prices in Figure 6. Dominican private schools typically charge both an annual enrollment fee and a monthly fee. We see large price drops in enrollment fees (top panel) and monthly fees (bottom panel) following the opening of a new public school, with both types of prices falling by about a third of the average level. The figures show no evidence of a pre-trend. This price drop could reflect a large reduction in market power from the introduction of a new (public) competitor. If so, the new public school's impact would extend even to students not in the public system by inducing a large transfer from the private school to its enrollees. On the other hand, the price drop could come from cost decreases. If private schools lower spending on

[18]While the project's assignment to a builder is an important event, neighborhoods still often were not sure if they would receive a new public school because the government did not commit to specific locations until it had purchased the land. The more relevant event for forward-looking private schools may therefore be when the government purchases the land and construction begins. In a future draft, we plan to analyze how exit rates and other outcomes shift in response to this event, using an OLS estimation strategy.

[19]Even with a large increase in public school competition, surviving private schools may not necessarily improve their product. The exit of other private competitors might increase the survivors' market power. Further, a loss in enrollment could leave the school with students who have strong reasons to attend the private school and are less price or quality elastic (McMillan, 2005).

productive inputs, their costs may fall, at the expense of school quality. Private school students' learning outcomes may then decrease.

We test whether private schools quality decreases using students' test score. For each high school and year, we construct the school's test score value-added (μ_{jt}) as the school's mean change in student test scores unexplained by students' lagged scores and demographics:[20]

$$y_{it}^{12} = h(y_{i,t-4,M}^{8}, y_{i,t-4,S}^{8}; \beta^{VA}) + \beta_X^{VA} X_{ijt} + \mu_{jt} + \eta_{it} \tag{5}$$

where y_{it}^{12} is the student's mean 12th grade score across math and Spanish, $y_{i,t-4,M}^{8}$ and $y_{i,t-4,S}^{8}$ are the student's 8th grade scores in math and Spanish, respectively, $h()$ is a cubic polynomial in each argument, and X_{ijt} includes the student's gender, a polynomial in the student's age, and school characteristics including the type of session the student attended.

As students only take national exams in grades 8 and 12, we can only construct value-added measures for high schools. This lowers our sample size to the point where our instrumental variables estimates are very noisy. For testing for effects on school value-added, where therefore estimate our event study model with OLS and present the estimated coefficients in the dashed lines in Figure 7. Immediately upon inauguration of a new public school, we see an increase for grade 12 value-added of over 0.07σ for private schools. This increase is slightly larger than the difference in mean value-added across the private and public sectors and potentially represents a meaningful change to school quality.

Private schools may achieve test score increases by providing better quality instruction (even with reduced revenue per student) 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 4 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 student 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.

While the private sector may be most responsive to changes in competition, the public schools may still be affected. We therefore run a similar analysis for how a new public school affects an incumbent public school's value-added and plot the coefficients in the solid gray line in Figure 7. We find a fairly precise zero effect on local public school quality.[21] Without the profit incentive,

[20] 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.

[21] That local public school value-added does not fall also suggests the private school effects are not merely changes

public schools may not be responsive to changes in competition. But even without strategic considerations, incumbent public schools may be affected by having to share inputs, especially access to productive teachers, with the new schools. Several aspects of the pre-expansion teacher labor market made it possible to fill the classrooms with certified teachers. First, when public schools had multiple sessions, teachers sometimes only taught a single session; these teachers could therefore increase their hours to a full day. Second, the public system had an oversupply of teachers such that there was a waiting list for positions.[22] The government could therefore pull teachers from the waitlist to staff the new schools.[23] Despite these available actions, the government still faced an undersupply of teachers and took short-term measures to increase the supply, such as advancing student-teachers to immediate certification.[24] The lack of change in local public school value-added indicates that either these new teachers were of similar quality to the more experienced teachers or that the effects of teacher quality were spread across larger geographic areas than our event study analysis captures.

Finally, we examine a margin of competition that is particular to this context – school hours. Prior to the expansion, private schools offered slightly longer sessions than public schools. We run an event study analysis where we test whether incumbent schools increased the length of the school day to compete with full-day pubic schooling. We plot the estimated coefficients in Figure 8, with the dashed lined for private schools and the solid gray line for incumbent public schools. As expected, the incumbent public schools increased their hours once the new public schools relaxed capacity constraints in the public system. The coefficients decline over time as (control) areas without new local public schools open are able to offer full-day instruction because enough schools have opened elsewhere in the market. On the private side, we see no evidence of a change in hours offered. Thus, private schools go from offering more hours than the public schools on average before the school construction policy to offering fewer after. The treatment effects on hours appear misaligned with the value-added results: only private value-added increases while only public hours increase. Families may value the additional public hours for several reasons, including childcare, but the hours do not seem to translate into better test score outcomes. This is consistent with USAID's pre-expansion assessment that even with the short 4-5 hour shifts, students only received active instruction for 2 hours, 40 minutes per school-day (USAID, 2018). The low baseline hours therefore may not have been constraints on the amount of academic instruction in the short-run.

in sorting of students across sector.

[22] Dinerstein et al. (2020) estimates large negative impacts of waiting on teacher productivity.

[23] The public sector could also potentially hire teachers previously in the private system. The private schools, however, rely in part on foreign teachers that are ineligible for a public school teaching position.

[24] We are in the process of analyzing teacher data that tracks teachers across schools.

4.3 Effects on Student Outcomes

We find that the new public schools had large impacts on the education market. Student learning may have improved, as the new schools may have attracted students from lower-quality schools, private schools improved their quality and lowered price, and some low-quality private schools closed. But the new schools also drew students from higher-quality options, and some of the crowd out hit the high-quality part of the private sector. To test how the effects add up, we estimate the effects of a new school on students who live nearby, without conditioning on where they go to school. For this analysis, we need data on precise student locations. We thus use the Prosoli conditional cash transfer data that links low-income students and their residential locations to their school choices. As we only have one test score per student, we omit student fixed effects from our event study specification and estimate with OLS.[25] We present the estimated coefficients in Figure 9.[26] We see a noisy zero effect on students' test scores.[27] But because this result hides substantial changes to the number and type of options available to students, a different level of public provision might result in a higher or lower effect on student outcomes. We analyze this by specifying and estimating a structural model.

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, entry and exit (supply), and value-added (technology).

In each school year t, the schooling economy exists of a set of (rising) 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 non-overlapping education markets is difficult in urban areas when physical distance is a relevant characteristic. In principle, we could forgo separating students and schools into markets

[25] Students take two national exams, in 8th and 12th grades. But our Prosoli data is a single cross-section of students enrolled as of 2018, such that we lack locations for all students who took 12th grade tests before 2018. We measure students' exposure to the new public schools based on which schools were open when the student was in 8th grade.

[26] Because we have a single 2018 cross-section, we have almost no students still in school in 2018 who were in 8th grade several years prior to a new public school opening. The event study thus has a pooled pre-period.

[27] We are in the process of expanding these results to include more data so that our estimates are more precise.

and estimate a single equilibrium. But dividing the country into markets yields considerable savings in computational time. We thus follow an iterative procedure to separate schooling markets into cities where students are extremely unlikely to attend school in a different market from which they live (Neilson, 2013). For each self-contained market, we then specify a set of geographic nodes on a 400m by 400m grid and estimate the distribution of students by age, mother's education, and Prosoli eligibility. We describe this procedure in Appendix B.

5.2 Demand: 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 we segment mother's education level to be binary as well (0 if mother completed at most primary school, and 1 if mother attended secondary school or more). 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.[28] 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:

$$u_{ijt} = -\alpha_i p_{jt} + \beta_i^{VA}\mu_{jt} + \beta^{hr} hr_{jt} + \beta^{priv} priv_{jt} - \gamma dist_{ijt} + \xi_{jt} + \epsilon_{ijt} \tag{6}$$

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

We specify the preference coefficients as:

$$\beta_i^{VA} = \bar{\beta}^{VA} + \beta_1^{VA} x_{it}^p + \beta_2^{VA} x_i^e + \beta_3^{VA} x_{it}^p x_i^e$$
$$\alpha_i = \bar{\alpha} + \alpha_1 x_{it}^p + \alpha_2 x_i^e + \alpha_3 x_{it}^p x_i^e + \nu_i$$

where $\nu_i \sim^{iid} logN(\mu_\alpha, \sigma_\alpha^2)$. We let $\epsilon_{ijt} \sim^{iid} T1EV$.

This demand model incorporates sufficient preference heterogeneity to accommodate flexible substitution patterns within and across sectors.

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

5.3 Supply: Private School Entry, Exit, and Pricing

Motivated by incumbent public schools' lack of value-added change in response to a new public school (Figure 7), we assume that public schools are non-strategic; i.e., their supply (and characteristics) are determined exogenously by government policy. In a first stage, we model location-specific potential entrant private schools as deciding whether to pay an entry cost and enter the market. In a second stage, we model new entrants and incumbent private schools as simultaneously choosing whether to remain open and what price to charge to maximize profits.[29] We treat this choice as a static problem such that schools make choices to maximize this year's profits only.

We describe the schools' actions in stages and refer to the timeline in Figure 12 for help in exposition. At the beginning of the year, schools' exogenous characteristics (VA_{jt}, hr_{jt}), demand unobservable (ξ_{jt}), marginal cost (mc_{jt}), and fixed operating cost (FC_{nt}) are drawn.[30] Other than the fixed operating cost, these characteristics are public information to all schools. FC_{nt} is drawn from a known distribution G and varies by location (n). Thus, two schools in the same location have the same fixed operating cost. Each school knows its own location's fixed cost but only knows that other locations' fixed costs are drawn from G. Thus, the information set at time 0, Ω_{0j}, consists of $(\vec{VA}, \vec{hr}, \vec{\xi}, \vec{mc}, FC_{n(j)t})$.[31]

Armed with information Ω_{0j}, a set of potential entrants decides whether to pay the entry cost, χ, to enter the market. This entry cost includes registration fees and any costly learning of how to run a school. We allow for free entry at all locations in the market, such that each location has a potential entrant every year. The assumption of free entry reflects the nature of the Dominican urban private schools. Many of these schools are run out of the school leader's home, and often the school leader gave up a for-profit business unrelated to education to start the school. These features lead to a large number of locations that could accommodate a school plus the large pool of potential school leaders. The potential entrant's decision is:

$$\max_{Enter_{jt}} Enter_{jt} \cdot \left(E_{\Omega_0} \left[\hat{\tilde{\Pi}}_{jt} \right] - FC_{nt} - \chi \right) \tag{7}$$

Note that we place modifiers over the variable profit, Π. These reflect beliefs and behavioral

[29] 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.

[30] In future drafts, we hope to relax this assumption and, in particular, allow private schools to make endogenous investments that shift their value-added. We also plan to specify public school value-added to be a function of the aggregate level of public provision to capture input scarcity, though our event study results show little evidence that public school quality varies with the local level of provision.

[31] This choice to have marginal costs be public information but fixed costs be private information 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.

assumptions, which we will return to once we describe the other stages. We draw the potential entrants' characteristics from the incumbents' distribution of characteristics; in future drafts, we plan to estimate differences in these distributions.

After the first stage, entry decisions are realized and observed by all market participants. When observing which locations have added an entrant and which locations have not, schools update their beliefs on the distribution of others' location-specific fixed costs. Concretely, schools calculate $E_{\Omega_0}\left[\hat{\tilde{\Pi}}_{jt}\right] - \chi$ for each potential entrant and use this to bound FC_{nt} from above if entry occurred and from below if entry did not. Thus, the information set at $\tau = 1$, Ω_1, includes everything in Ω_0, plus knowledge of the set of competitors and bounds on other locations' fixed costs. This belief updating captures the idea that market participants likely have some information about other locations' costs, and it will be important when we take the model to data in summarizing geographic heterogeneity.

In the second stage, the entrants and incumbents choose whether to exit and, if not, what price to set. For market size N_{mt}, schools choose price and whether to exit:

$$\max_{Exit_{jt}, p_{jt}} (1 - Exit_{jt}) \cdot [\underbrace{(p_{jt} - mc_{jt}) \cdot N_{mt} \cdot E_{\Omega_1}[\tilde{s}_{jt}(\vec{p}, \vec{VA}, \vec{hr}, \vec{\xi})]}_{\equiv \tilde{\Pi}_{jt}} - FC_{nt}] \tag{8}$$

where their perceived enrollment, as a function of price and other characteristics, is $N_{mt} \cdot E_{\Omega_1}[\tilde{s}_{jt}(\vec{p}, \vec{VA}, \vec{hr}, \vec{\xi})]$.

Because schools do not know each location's fixed cost exactly 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, $N_{mt} \cdot E_{\Omega_1} s_{jt}$, 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 changes in market structure 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 adapt Sánchez (2018) by assuming that schools keep track of their own type (fixed cost draw), bounds on the other locations' fixed costs, and the expected market equilibrium.[32] Specifically, let $V_{ikt} = u_{ikt} - \epsilon_{ikt}$ be student i's representative utility to attending school k if the

[32] 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."

school is open and $V_{ikt} = -\infty$ if closed. Then let

$$\lambda_{ijt} = E_{FC_{-n,t}} \sum_{k \neq j, \text{k private}} exp(V_{ikt}) \tag{9}$$

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 i chooses private school j in school year t is:

$$\begin{aligned}\tilde{P_{ijt}} &= \frac{exp(V_{ijt})}{1 + exp(V_{ijt}) + \sum_{\text{k public}} exp(V_{ikt}) + E_{FC_{-j,t}} \sum_{k \neq j, \text{k private}} exp(V_{ikt})} \\ &= \frac{exp(V_{ijt})}{1 + exp(V_{ijt}) + \sum_{\text{k public}} exp(V_{ikt}) + \lambda_{ijt}}\end{aligned} \tag{10}$$

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 $N_{mt} \cdot \tilde{s_{jt}} = \sum_i \tilde{P_{ijt}}$. This is the share function that the schools' use when deciding whether to enter and then the optimal pricing and exit decisions.

For schools that decide to remain in the market, they incur the fixed cost, FC_{nt}. Then, given the set of schools that have decided to be open and their characteristics (including price) students choose their most preferred option. At this point, school enrollments are realized and private school j earns its variable profits, Π_{jt}.

We return to the entry stage and clarify why the "hat" is above perceived profits. Potential entrants, when deciding whether to enter, consider the competitive environment they will enter and how their own entry choice may affect the equilibrium. In particular, they predict the second stage pricing and exit equilibrium, were they to enter. They also account for the fact that if they enter an already-occupied location, both schools will know each other's fixed cost with certainty. But to avoid assuming the potential entrants can solve an incredibly complex problem, we impose a restriction on the degree to which potential entrants are forward-looking. We assume that potential entrants do not consider how other potential entrants would shift the second stage equilibrium if they were to enter. This means that potential entrants do not integrate over the possible sets of other entrants – and the associated changes they would cause to other schools' beliefs – when deciding whether to enter. This assumption allows for the possibility of regret in entry, where a potential entrant is surprised that so many other potential entrants choose to enter and realized profits are below predicted profits.

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 (as we did in Section 4:

$$y_{it}^{12} = h(y_{i,t-4,M}^{8}, y_{i,t-4,S}^{8}; \beta^{VA}) + \beta_X^{VA} X_{ijt} + \mu_{jt} + \eta_{it} \tag{11}$$

where y_{it}^{12} is the student's mean 12th grade score across math and Spanish, $y_{i,t-4,M}^{8}$ and $y_{i,t-4,S}^{8}$ are the student's 8th grade scores in math and Spanish, respectively, $h()$ is a cubic polynomial in each argument, and X_{ijt} includes the student's gender, a polynomial in the student's age, and school characteristics including the type of session the student attended. We assume that $\eta_{it} \perp\!\!\!\perp \nu_i, \epsilon_{ijt}$, such that unobservable student heterogeneity in demand and production are independent.

5.5 Equilibrium: Static "Oblivious" Equilibrium

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

1. $Entry_{jt}$ solve 7 for each potential entrant j
2. $p_{jt}, Exit_{jt}$ solve 8 for each private school j
3. $\lambda_{ijt} = \sum_{k \neq j, \text{k private}} (1 - Pr(Exit_k(\lambda))) exp(V_{ikt})$ 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.

6.1.1 Technology

We start by estimating each school's value-added for each school year by running OLS regressions of Equation 11 and recovering the estimated fixed effects, $\hat{\mu_{jt}}$.[33]

[33] Estimating value-added in the presence of dropout introduces complications. As we do not find changes in dropout or composition (based on lagged scores) across the two sectors, we defer these challenges to the next draft.

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.

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 graduated 8th grade but 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} ($= \bar{\beta^{\mu}}\mu_{jt} + \beta^{hr} hr_{jt} + \bar{\beta^{priv}} priv_{jt} + \xi_{jt}$) in the inner loop.[34]

As school pricing decisions likely depend on ξ_{jt}, which might also be correlated with μ_{jt}, we specify instruments for price and quality. The first set of instruments comes from the policy and match the variation we used in the event study analysis. For each private school, we find the first opening public school in its neighborhood (if there is one) and use the builder's characteristics as instruments. We use the same characteristics as in the event study analysis – whether the builder is a firm, log employment, whether employment is positive, and log employee-months. 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}).[35]

[34] Typically $-\bar{\alpha}p_{jt}$ would be included in the mean utility. Because we use micro moments to target identification of the price coefficient, we keep it in the outer loop.

[35] Future versions will incorporate a third price instrument. Public school teacher salaries increased during our

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 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 the Prosoli data, which is matched to student residential location.[36] The in-depth student survey 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.[37] 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).

6.1.4 Supply

We specify G, the CDF of fixed costs FC_{nt}, as:

$$FC_{nt} \sim \begin{cases} \infty \text{ w.p. } p^{\infty} \\ logN(\mu_{FC}, \sigma^2_{FC}) \text{ w.p. } 1 - p^{\infty} \end{cases} \tag{12}$$

We include the possibility for an infinite fixed cost to allow the model, if necessary, to rationalize exits among very high variable profit schools without needing to stretch the fixed cost distribution out so far that it loses its ability to predict exit for smaller profit schools.

We estimate the model with method of simulated moments, where we choose the moments as the event study coefficients (plus the mean exit rate) from Figure 5 and neighborhood-level event study estimates for private school entry. With this specification, we estimate the supply model using an adaptation of the iterative procedure described in Sánchez (2018). We start by guessing the distribution of incumbents' marginal costs and drawing a vector of marginal costs for potential entrants. We then solve for the first stage (potential entry) equilibrium. We use these entry decisions to update beliefs on fixed costs and solve for the second stage exit equilibrium taking prices as given in the data. Once we find the equilibrium, we invert the pricing first-order condition to recover marginal costs for each school in each year.[38] We then compare the recovered

sample, and particularly for teachers measured to be high quality. As public and private schools may compete for teachers, private school costs likely increased.

[36]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.

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

[38]Even for many exiting schools we are able to invert the first-order conditions because we observe prices. Our

marginal costs to the assumed marginal cost distribution and iterate back and forth between the first and second stages until the marginal cost distribution hits a fixed point. Once we have estimated marginal costs, we estimate the counterfactual equilibrium where no new public schools opened. In this step, we solve for prices and exit decisions, and then we use the model predictions to calculate the moments. In each step where we solve for an equilibrium, we fix $\vec{\lambda}$, solve for pricing and exit decisions as a function of $\vec{\lambda}$, and then update $\vec{\lambda}$ to be consistent with these choices. We 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 and value-added instruments exploit policy variation in procurement lottery process 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. The randomized lottery used to assign school construction projects to builders provides variation unrelated to local conditions. 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.[39]

The other micro moments map 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 model, 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

data provides this unique opportunity because many private schools report planned prices for the following school year to the government. In cases where this is not possible, we give schools that exited (without revealing a price) its marginal cost from the prior year. For non-price characteristics – value-added, hours, and ξ_{jt}, we do not observe these for schools that exit. When predicting choice shares were the school to be open counterfactually, we use the prior year's value.

[39]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. In future drafts we will assess robustness to using only the variation from the hypothetical survey question for identifying the price coefficient.

by assuming that the observed exit decisions reflect a trade-off between static variable profits and fixed costs. The policy variation serves as a shifter of variable profits, as some neighborhoods receive the shock of a new public school while others do not. By having our model match the event study moments, we trace out how shifts in variable profits translate into exit probabilities, and thus we pin down the location and scale of the fixed cost distribution. In the event studies, we saw large increases in exit rates, even for demand shocks that didn't necessarily siphon off all of a school's students (according to our demand estimates). Thus, the variation implies that a large mass of schools have fixed costs fairly close to pre-expansion variable profits. The probability of an infinite fixed cost is pinned down by the percentage of exits among very high variable profit schools. While the log-normal fixed cost distribution could rationalize such exits, it would stretch out the distribution sufficiently to miss the event study moments.

Finally, the entry cost (χ) is the value that rationalizes the differential responses of entry and exit to the policy variation. Further, our specification of neighborhood-specific fixed operating costs implies that if $\chi = 0$ then the bounds provided by observed entry decisions should be consistent with the observed exit decisions, up to variation in value-added, hours, and ξ_{jt}. The extent to which these entry-implied bounds deviate from realized exit identifies the size of χ.

In estimation, we always find an equilibrium, though we do not have a general existence proof. In terms of equilibrium uniqueness, we have tried multiple starting values and arrived at the same equilibrium but this requires further exploration and we believe it likely that multiple equilibria exist.

6.3 Estimates

We present preliminary estimates from estimation of our structural model. We start with the demand model (Equation 6) and present the estimates in Table 5. We estimate mean disutilities to distance and price, with a kilometer of distance valued the same as 6 USD in annual tuition. This lack of price sensitivity in preferences is important when considering schools' incentives to lower prices.

In terms of preference heterogeneity, we estimate a similar price coefficient for all demographic groups, except children with more educated mothers and not Prosoli eligible, for whom the price coefficient is less than half as large. For value-added, we estimate a large and precise positive coefficient. We find that the preference is higher for children not Prosoli eligible but don't find large differences based on mother's education. We also find that students value hours when choosing a school, which is important because the new schools allow the public system to shift to full-day instruction.

On the supply side, we estimate considerable variation in school-year markups of price relative to marginal cost (Figure 13). 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 40%. 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 estimate a mean location fixed cost of 92.5. But because most locations don't get an entrant (or incumbent), the mean location fixed cost, conditional on having an operating school, is much higher, at 273.0. This could reflect that the best locations, for demand, are also the most expensive. It also means that despite very large markups over marginal cost, total profits are much lower (Figure 14). In fact, over 90% of variable profits are dissipated by large fixed costs.

We estimate no probability of an infinite fixed cost, and the entry cost is just 2.82. This very small level comes because the wedge between potential entrant and incumbent location-specific variable costs is roughly the right size to explain the increased exit rate from the policy.

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 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 main 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 no expansion (0) to the observed expansion (1) to three times as big (3), we present the counterfactual private school mean prices and annual exit rates in Figure 15.[40] We find that the private school exit rate is fairly unresponsive to increases in public provision at low levels, but then around the size of the actual policy, exit rates quickly double and increase steadily for bigger levels of public provision. This largely reflects a mass of small schools with relatively high fixed costs whose supply is particularly sensitive to moderate changes in market conditions. Once these schools have exited, the remaining schools are slightly less subject to exit risk. Prices follow an opposite path, starting high but then dropping once public provision expands.

We evaluate the effect on student learning – which is ex-ante ambiguous – and find that learning (as measured by enrollment-weighted market mean value-added) falls for small expansions in public provision before raising once the expansion is a bit larger than the actual policy. It peaks just before a three times policy before falling slightly. This non-monotonicity reflects several forces – students switching to the new public schools that tend to be lower quality, high-quality private schools lowering prices and attracting more students, and crowd-out of private schools, starting with low-quality but then affecting some high-quality schools.

7.2 Varying the Public Technology

We end by characterizing how the optimal level of public provision varies with the technology (quality) of the new public schools. In Figure 16 we plot the learning outcomes for different levels of public provision, by quality of the new public schools. The "Low" technology reflects the observed value-added of the new public schools in the data. The "High" technology corresponds to new public schools that are the same quality as the mean private school, while the "Mid-high" technology is equivalent to the quality of the old public schools that operated prior to the expansion.

While higher quality new schools tend to lead to equilibria with more student learning, this is not always the case. At very high levels of provision, the high quality public schools crowd out enough high quality private schools that learning actually falls (relative to lower quality provision). Further, the relationship between learning and the amount of public provision varies depending on the quality level. While lower quality provision produces more learning at higher levels of provision (near three times the observed policy), the optimal level of higher quality schools is closer to two times the observed policy. These non-monotonicities highlight the importance of considering equilibrium responses when choosing levels of public provision.

[40] We place new public schools in the same locations as the observed expansion and scale the size of the public provision in utility terms.

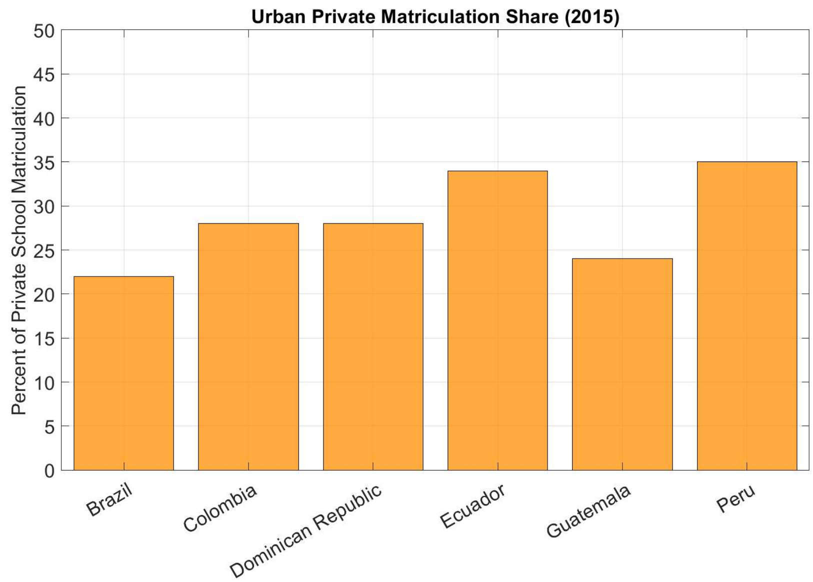
8 Conclusion

In this paper, we examined the how public and private provision of education interact to reach equilibrium. 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 how this relationship varies with the quality of the new public schools. In particular, we found that the learning-maximum level of (costless) public provision is potentially an intermediate amount, but that the optimal point varies inversely with the quality of the new public schools.

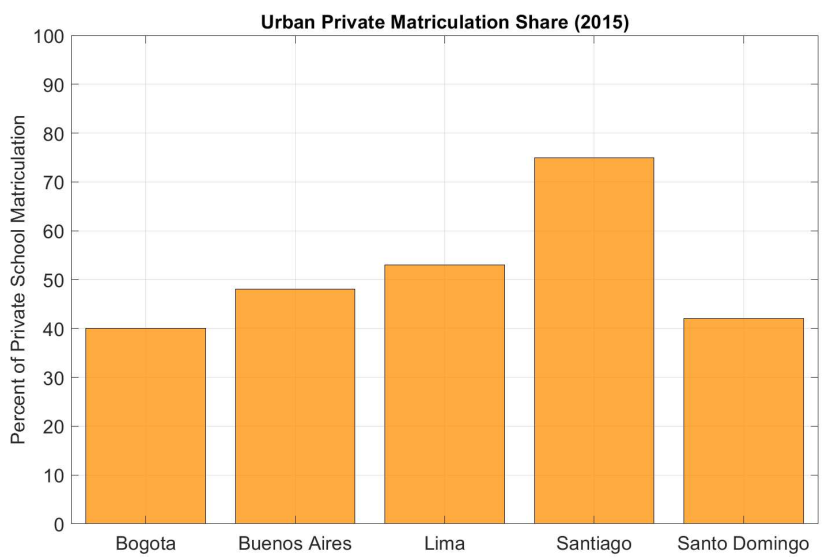
Future drafts will incorporate heterogeneous treatment effects in the event study analysis, extend the amount of preference heterogeneity in the demand model, and expand the counterfactual analysis. We also hope to add investment in quality as an important strategic choice that could explain some of the observed equilibrium effects.

Figures and Tables

Figure 1: Private Share of Primary Enrollment in LAC



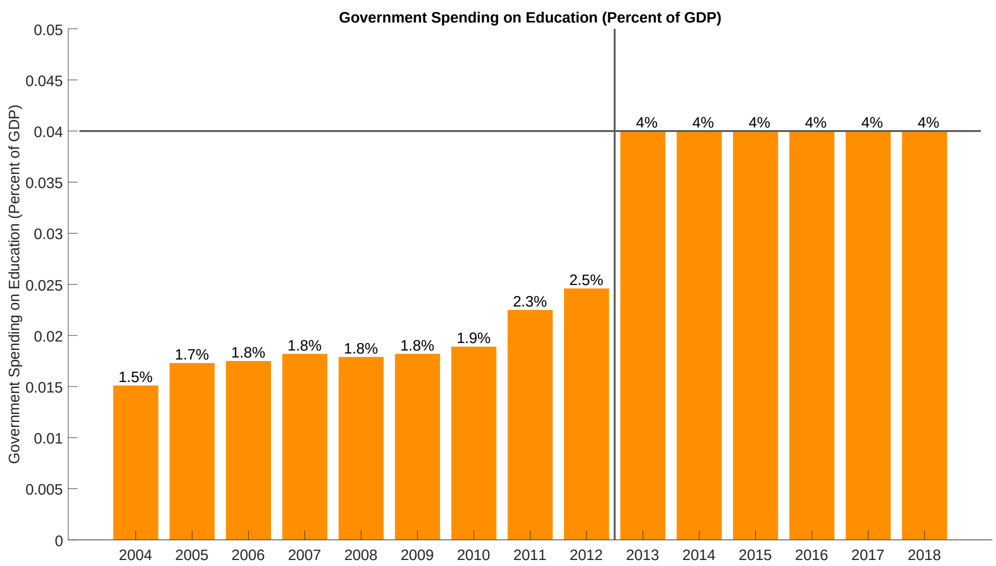
(a) Countries



(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

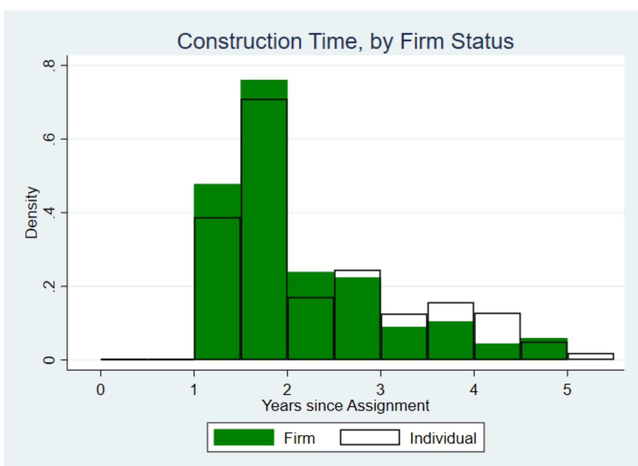


Figure 4: IV Event Study Estimates: Enrollment Outcomes

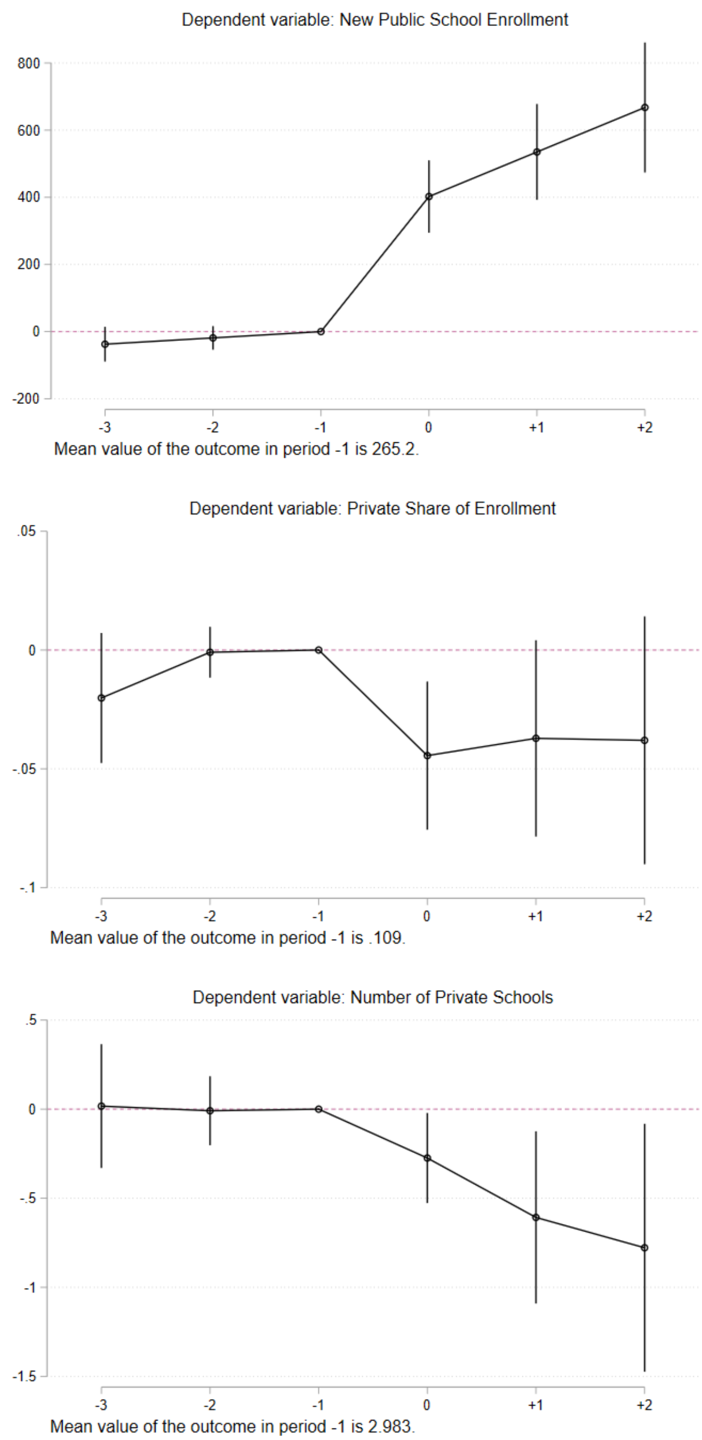


Figure 5: IV and OLS Event Study Estimates: Exit

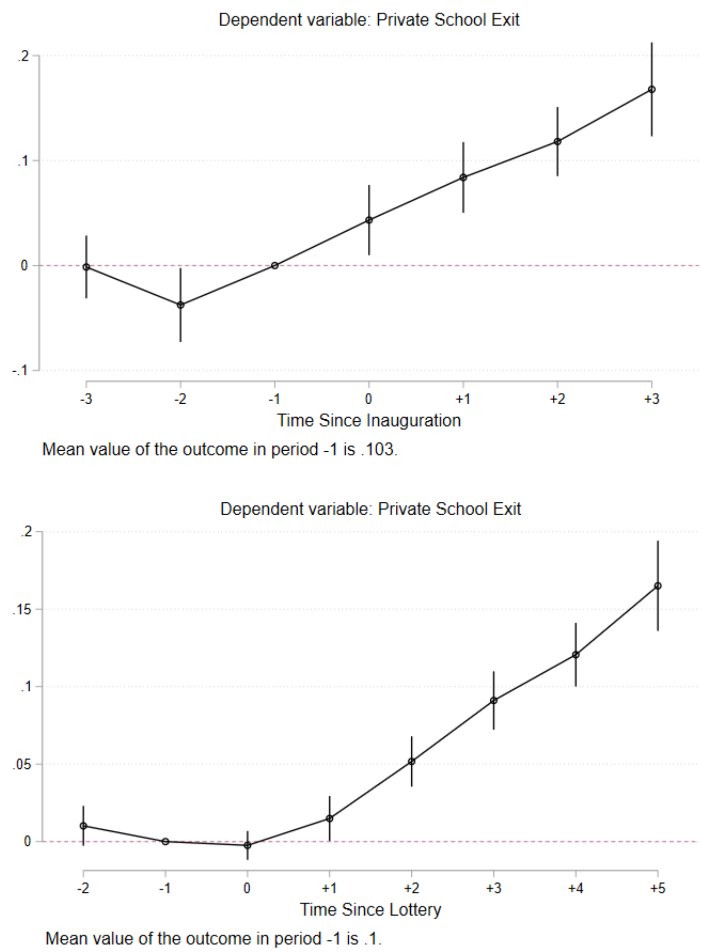


Figure 6: Event Study Estimates: Prices

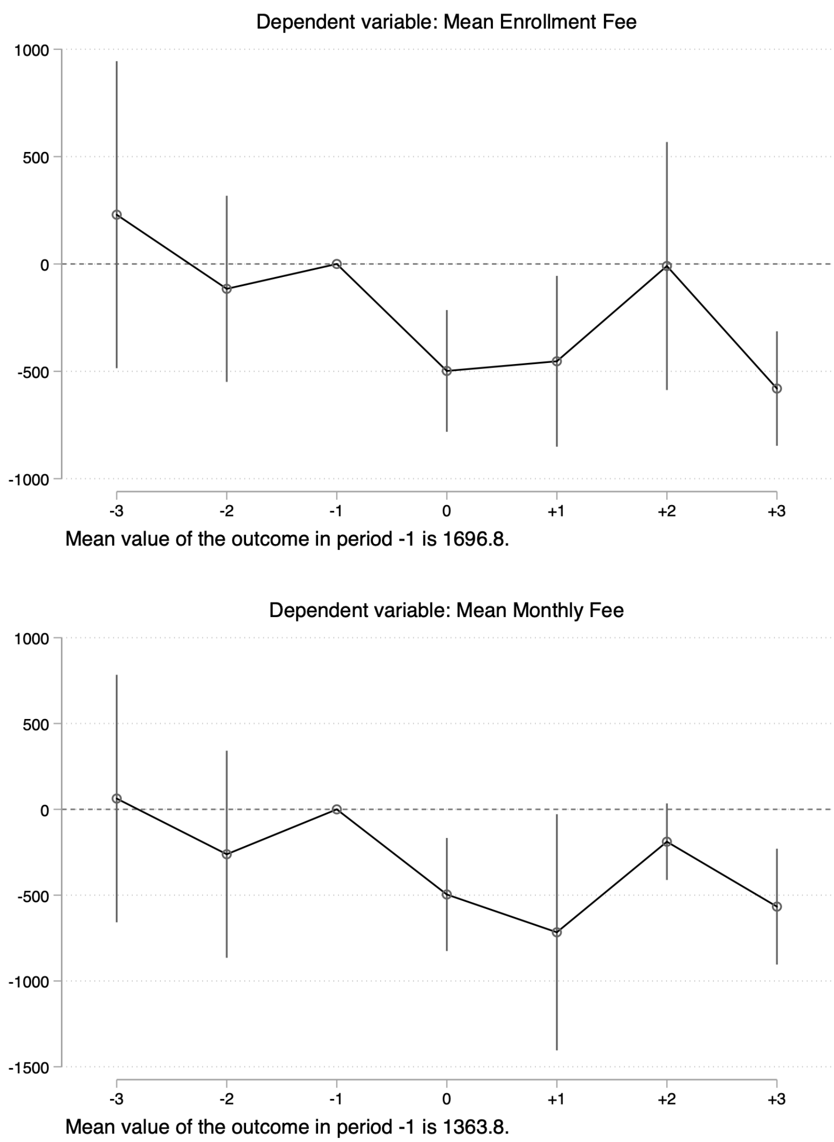


Figure 7: Event Study Estimates: Value-Added

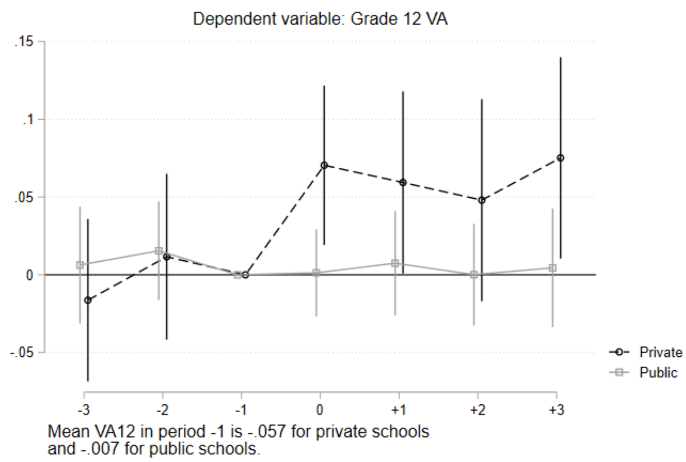


Figure 8: Event Study Estimates: Hours

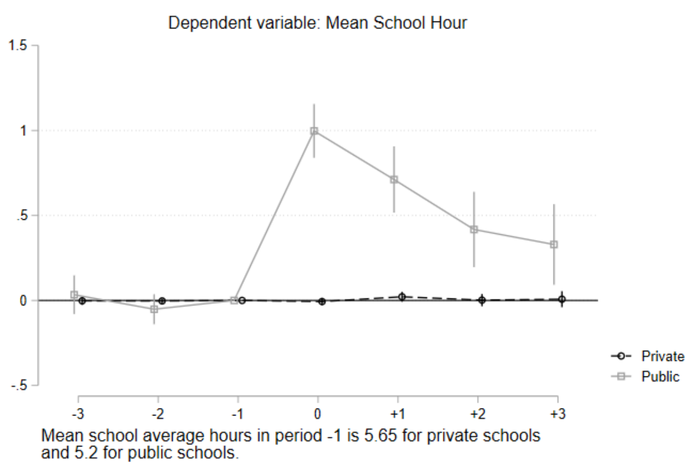


Figure 9: Student Event Study Estimates: 8th Grade Test Scores

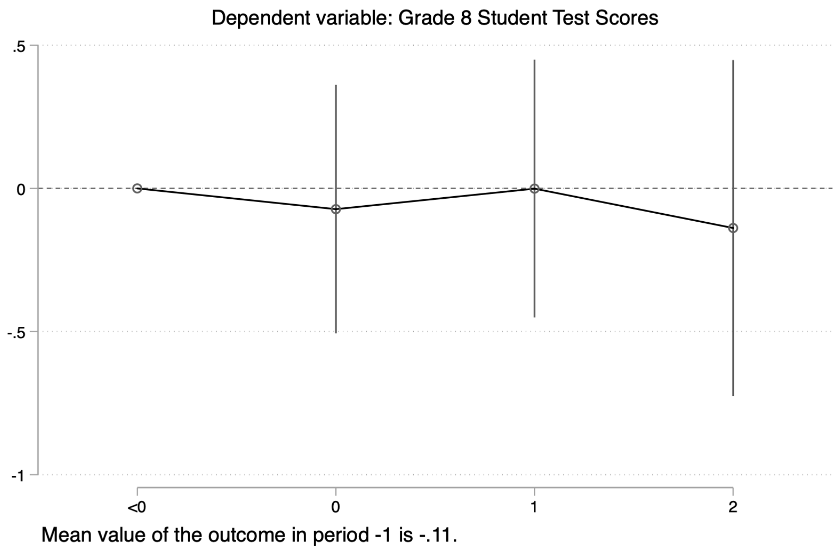


Figure 10: Urban Barrios/Parajes

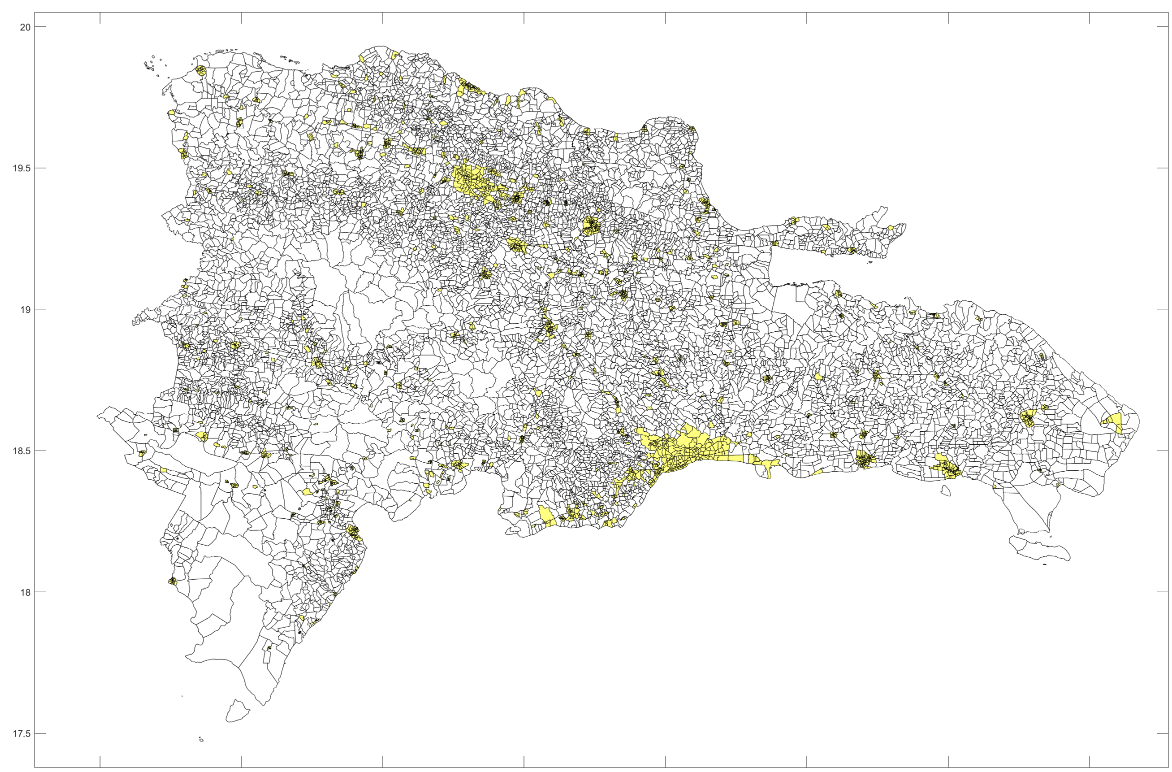


Figure 11: Market definition: Santo Domingo and Boca Chica

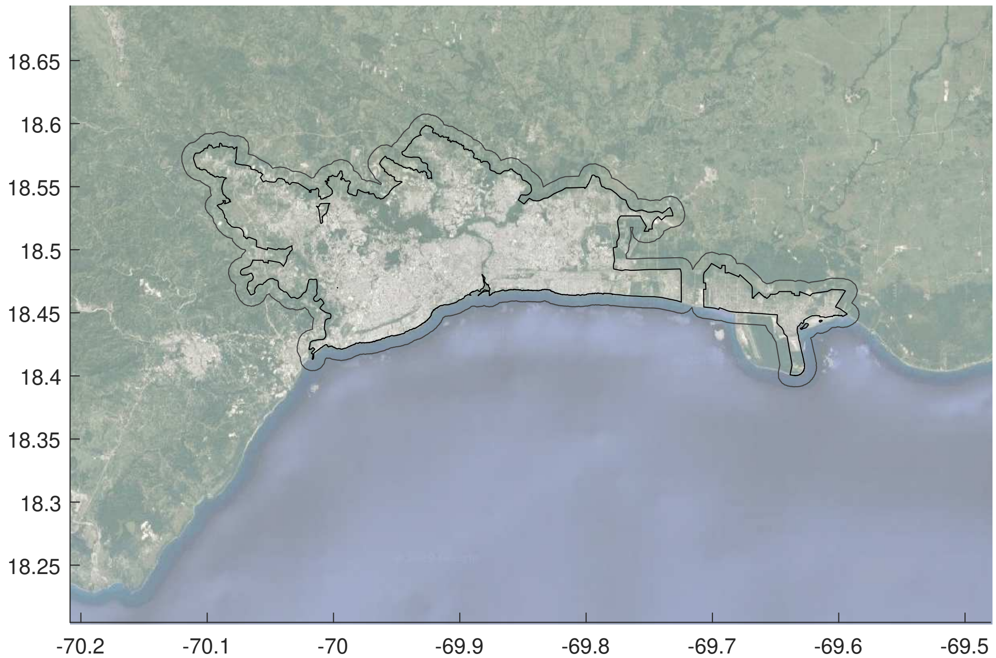


Figure 12: Model Timeline

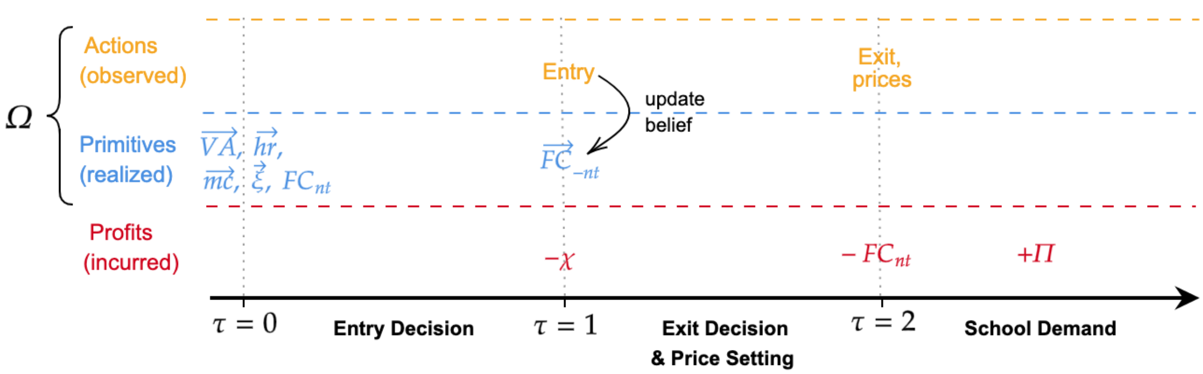


Figure 13: Distribution of Estimated Markups

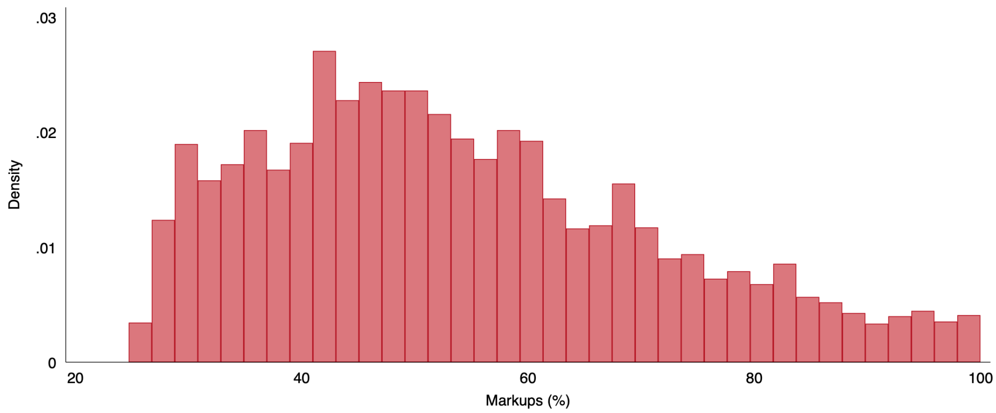


Figure 14: Total Profits Estimates

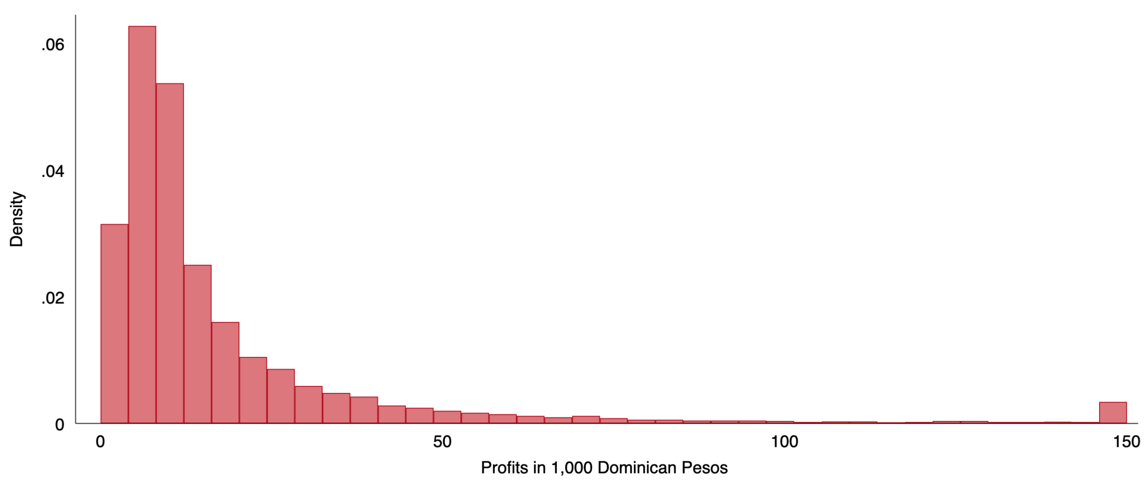


Figure 15: Counterfactual 1: Increase Level of Public Provision

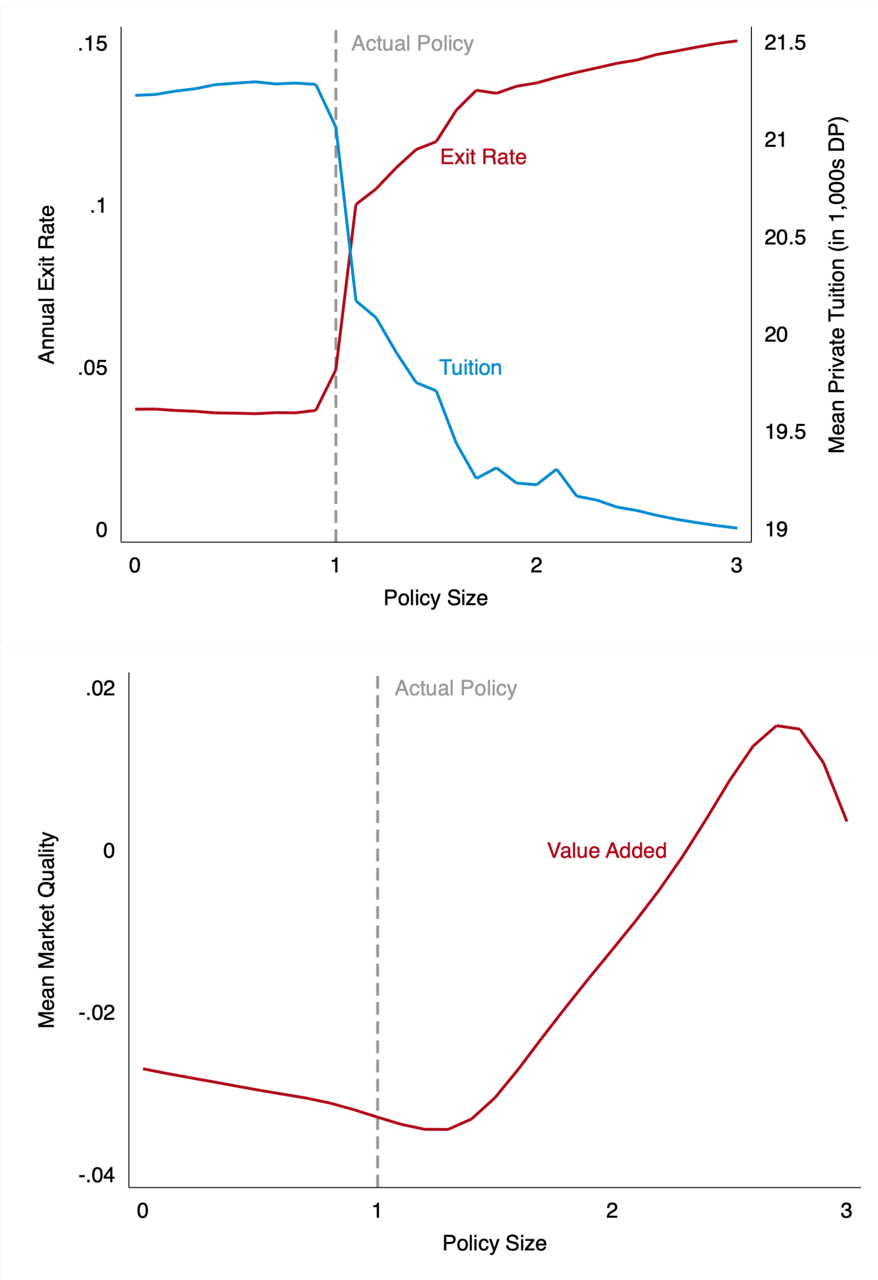


Figure 16: Counterfactual 2: Vary Technology

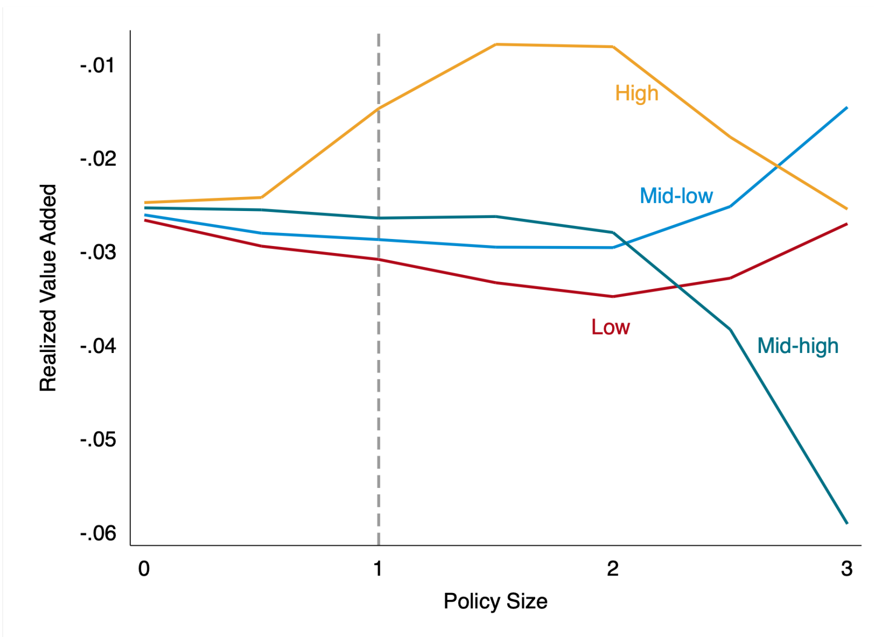


Table 1: Balance Test

	(1) Budget 2015	(2) Budget 2016	(3) Budget 2017	(4) Budget 2018	(5) #Classrooms
Firm	-0.154	4.019	-3.836	1.535	0.160
	(1.151)	(2.520)	(2.679)	(1.106)	(0.539)
log(Employment)	0.0604	-0.00891	-0.0164	0.0675	0.558
	(2.060)	(4.764)	(5.488)	(2.080)	(1.015)
(Employment>0)	-3.726	0.411	-12.85	-3.223	-6.506
	(15.83)	(36.63)	(45.02)	(16.05)	(7.819)
log(Employee Months)	-1.098	-1.458	-2.768	-1.202	-0.994
	(1.386)	(3.206)	(4.351)	(1.413)	(0.686)
Mean y	41.72	45.20	47.82	43.49	15.48
R-squared	0.569	0.235	0.528	0.593	0.232
N	1669	1878	414	1901	1937

Province-sorteo fixed effects included.

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

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 Enrollment	(4) Private Share of Enrollment	(5) Neighborhood Total Enrollment
Inauguration t-3	-37.46	0.0170	7.983	-0.0202	155.3
	(26.41)	(0.177)	(40.76)	(0.0139)	(95.49)
Inauguration t-2	-18.72	-0.00888	22.50	-0.000934	85.83*
	(17.89)	(0.0986)	(22.19)	(0.00545)	(47.54)
Inauguration t-1	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
Inauguration t	402.3***	-0.274**	-81.17***	-0.0444***	50.81
	(54.89)	(0.129)	(29.63)	(0.0159)	(74.92)
Inauguration t+1	535.1***	-0.608**	-133.3**	-0.0372*	-37.64
	(72.72)	(0.246)	(65.65)	(0.0211)	(123.9)
Inauguration t+2	667.7***	-0.778**	-158.6*	-0.0380	-238.8
	(98.68)	(0.355)	(87.83)	(0.0266)	(181.9)
Mean y	293.8	3.041	500.6	0.110	2167.1
#Neighborhoods	890	890	890	890	890
N	7568	7568	7568	7568	7568

Standard errors clustered at the neighborhood level. Neighborhood and year fixed effects included.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: 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 5: Demand Estimates

Parameter	Estimate	SE
Distance (per km)	0.378	(0.009)
Price (per 1000 Dom pesos)	0.093	(0.019)
Price x HighIncome	0.001	(0.003)
Price x HighEduc	-0.009	(0.003)
Price x HighIncome x HighEduc	-0.051	(0.005)
Price logN RC μ_α	-3.216	(6.457)
Price logN RC σ_α	0.813	(5.200)
VA (test s.d.)	5.186	(0.975)
VA x HighIncome	0.904	(0.119)
VA x HighEduc	0.225	(0.379)
VA x HighIncome x HighEduc	-0.534	(0.520)
Hours	0.118	(0.019)
Private	0.168	(0.044)
Constant	-0.613	(0.114)

Notes: Demand model estimated via simulated method of moments.

References

Allende, C. (2020). Competition under social interactions and the design of education policies. *Working Paper*.

Andrabi, T., N. Bau, J. Das, and A. I. Khwaja (2020). Upping the ante: The equilibrium effects of unconditional grants to private schools. *American Economic Review Forthcoming*.

Andrabi, T., J. Das, A. I. Khwaja, S. Ozyurt, and N. Singh (2020). Private schooling, learning, and civic values in a low-income country. *Working paper*.

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

Atal, J. P., J. I. Cuesta, F. González, and C. Otero (2020). The economics of the public option: Evidence from local pharmaceutical markets. *Stanford Working Paper*.

Banerjee, A., R. Hanna, J. Kyle, B. A. Olken, and S. Sumarto (2019). Private Outsourcing and Competition: Subsidized Food Distribution in Indonesia. *Journal of Political Economy 127*(1), 101–137.

Bau, N. (2019). Estimating an equilibrium model of horizontal competition in education. *Working Paper*.

Berry, J., L. Coffman, D. Morales, and C. Neilson (2020). Marketing schooling: An at-scale experiment in the dominican republic.

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.

Besley, T. and S. Coate (1991). Public provision of private goods and the redistribution of income. *The American Economic Review*, 979–984.

Burde, D. and L. L. Linden (2013, July). Bringing education to afghan girls: A randomized controlled trial of village-based schools. *American Economic Journal: Applied Economics 5*(3), 27–40.

Busso, M. and S. Galiani (2019). The causal effect of competition on prices and quality: Evidence from a field experiment. *American Economic Journal: Applied Economics 11*(1), 33–56.

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.

Clemens, J. and J. D. Gottlieb (2017). In the Shadow of a Giant: Medicare?s Influence on Private Physician Payments. *Journal of Political Economy 125*(1), 1–39.

Dinerstein, M., R. Megalokonomou, and C. Yannelis (2020). Human capital depreciation. *NBER Working Paper*.

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.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review 91*(4), 795–813.

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.

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

Jiménez-Hernández, D. and E. Seira (2020). Should the government sell you goods? evidence from the milk market in mexico. *Stanford Working Paper*.

Kazianga, H., D. Levy, L. L. Linden, and M. Sloan (2013, July). The effects of "girl-friendly" schools: Evidence from the bright school construction program in burkina faso. *American Economic Journal: Applied Economics 5*(3), 41–62.

Kazianga, H., L. L. Linden, A. Protik, and M. Sloan (2019). The medium-tem impacts of "girl-friendly" schools: Seven-year evidence from school construction in burkina faso. *NBER Working Paper* (26006).

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*.

Romero, M., J. Sandefur, and W. A. Sandholtz (2020, February). Outsourcing education: Experimental evidence from liberia. *American Economic Review 110*(2), 364–400.

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.

Sun, L. and S. Abraham (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects.

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

USAID (2018). Dominican republic country development cooperatin strategy, fy 2014-2021. Technical report.

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 6 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 6: Lotteries Description

Lottery Process	Date	Lots	Budget (in 1,000s RD$ $\approx$ 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

Table 7 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 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 7: 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

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 4 groups (based on two levels of educational level attained interacted with Prosoli eligibility) that may predict heterogeneous responses to the policy. In Appendix B we explain in greater detail the markets definition and the linkage of the different datasets.

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.

Figure B.1: Urban Barrios/Parajes

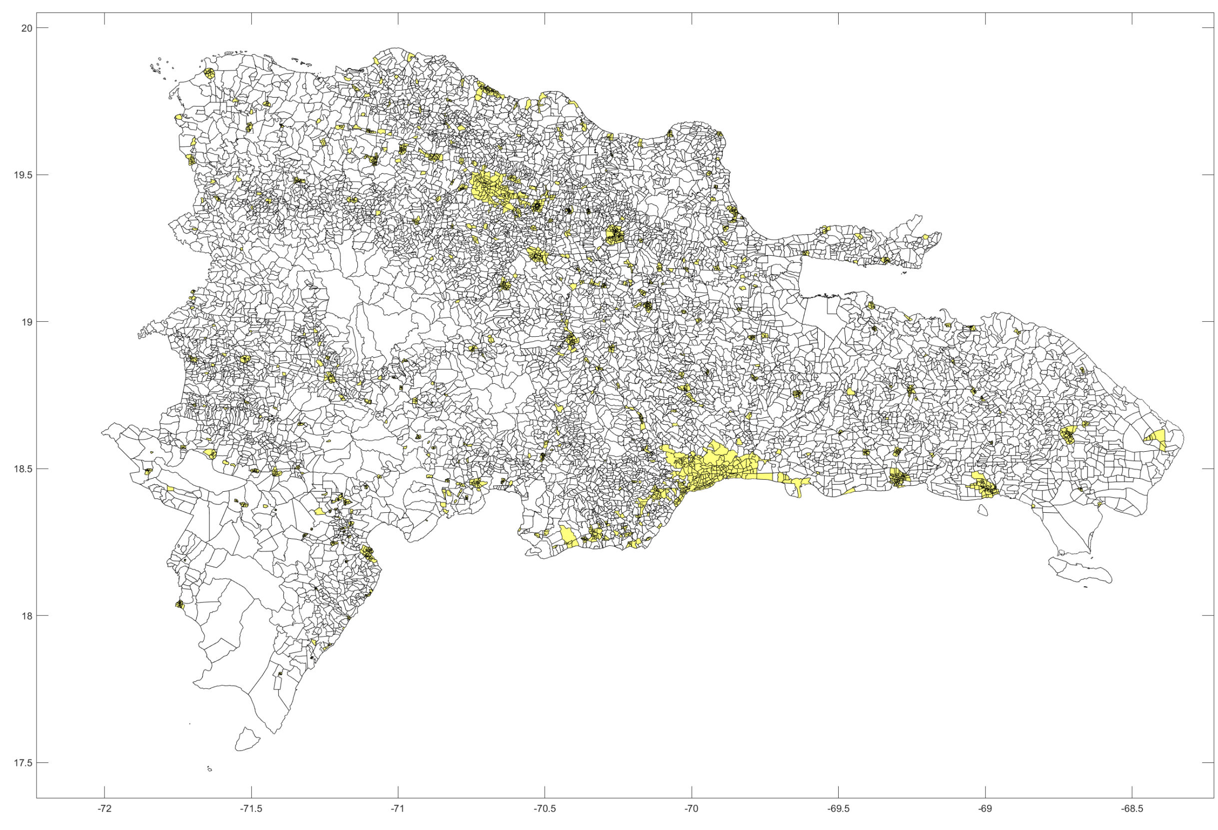
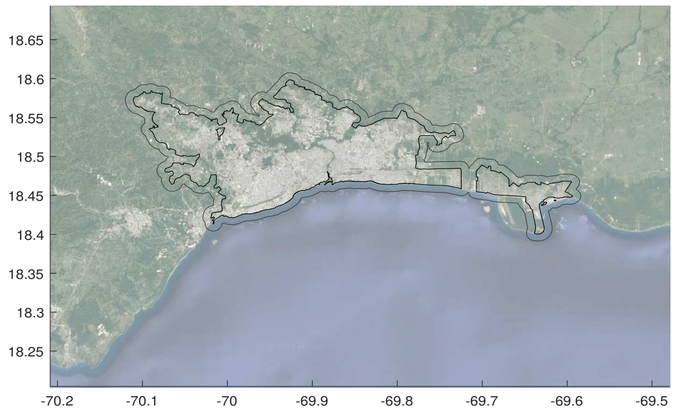


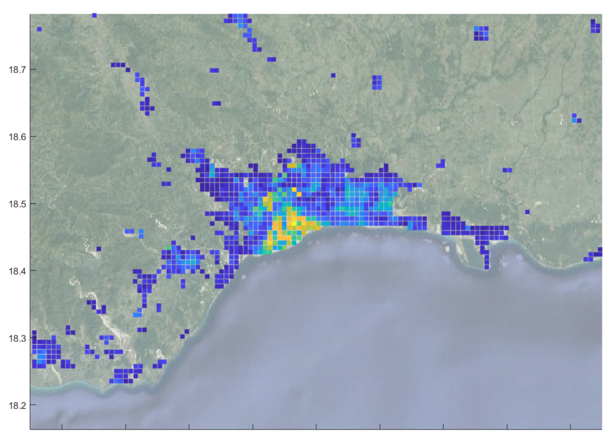
Figure B.2: Market definition: Santo Domingo and Boca Chica



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.

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



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

The population of the markets is binned into 4 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 secondary education and poor
- Type 2: Less than secondary education and not poor
- Type 3: Secondary or above and poor
- Type 4: 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 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[41]. 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.

[41] 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.