Competition Under Social Interactions and the Design of Education Policies*

Claudia Allende†

Job Market Paper

This version: November 15, 2019

Link to the latest version

Abstract

This paper studies families' preferences for peers and the implications for the distribution of academic outcomes. I develop an equilibrium model of school competition and student sorting under social interactions. In the model, families differ by human capital and income. Academic achievement depends on own characteristics, school productivity, and the characteristics of the peers. Geographic differentiation gives schools local market power to increase prices and decrease quality in the absence of close substitutes. On top of that, social interactions generate interdependencies in demand that add a new dimension for school differentiation. This modifies school incentives through two channels: it strengthens market power for some schools (direct channel) and it incentivizes a screening strategy that exploits heterogeneous responses to prices and quality to intensify differentiation (strategic channel). To study the empirical importance of these mechanisms, I estimate the model using administrative microdata from Peru. I address endogeneity of prices, quality, and peers by combining a regression discontinuity in the assignment of a scholarship with instruments that exploit the timing and local variation of a generous teacher payment reform and shocks to student sorting generated by a teachers' strike. I find that social interactions have sizable effects, increasing the income gap in academic achievement by 30 percentage points I use the predictions of the model to analyze the effects of counterfactual education policies in equilibrium. I then decompose the mechanisms to provide guidance on how to design education policies that improve the distribution of outcomes.

*I am very grateful for the guidance and helpful comments from Kate Ho, Adam Kapor, and Chris Neilson, and Miguel Urquiola. I thank Patrick Agte, Milena Almagro, David Arnold, John Asker, Peter Bergman, Leah Boustan, Nick Buchholz, Yeon-Koo Che, Sarah Cohodes, Chris Conlon, Natalie Cox Bachas, Ignacio Cuesta, Rebecca Diamond, Mateus Dias, Mike Dinerstein, Bó Honoré, Anders Humlum, Gastón Illanes, Jakub Kastl, Ilyana Kuziemko, Bentley MacLeod, Alex Mas, Suresh Naidu, Sebastián Otero, Stephen Redding, Jonah Rockoff, Mike Riordan, Bernard Salanié, Tobias Salz, Mica Sviatschi, Roman Andres Zárate, Owen Zidar and seminar participants at the joint IO/Labor Seminar and the IO, Labor, and Education Research Lunch at Princeton University, the IO, Applied Micro Theory, Applied Micro, and Development Colloquium at Columbia University for their valuable feedback. This paper would not have been possible without the support from the Ministry of Finance of Peru; I especially thank María Antonieta Alva, José Carlos Chavez, Jonathan Enriquez, Zoila Llempen, and Felipe Sarmiento for their collaboration and support. I thank Fabiola Alba for introducing me to the education system in Perú, David Abel Barraza and María Elena Guerrero for excellent research assistance, to Eduardo Vargas at J-PAL for his help during the fieldwork, and to Franco Calle for the help with the web scraping. I gratefully acknowledge support from the Education Policy Dissertation Research Fellowship and the CDEP Student Research Grant Program at Columbia University, Microsoft Research through Azure for Research Award: Data Science, and the Overdeck Education Research Innovation Fund through the Woodrow Wilson School at Princeton University. All errors are mine.

†Columbia Univeristy (Ph.D. Candidate) and Princeton Univeristy (Trailing Student). Email: cacruz@princeton.edu

1 Introduction

Under school choice, families should be able to avoid the options that fail to satisfy their demands. The proponents of these policies argue that choice will generate system-wide increases in school productivity and boost educational outcomes (Friedman, 1962; Chubb and Moe, 1990; Hoxby, 2003). Within a framework for a market for quality, the idea behind this principle is intuitive: schools face competitive pressure to improve their provision of quality or risk losing students to more preferred (higher quality) schools. These intuitive arguments inspired policymakers around the world to implement programs that expanded the degree of families' choice. Yet, the results from the empirical work evaluating these policies are scattered: from negative or ineffective (Abdulkadiroğlu et al., 2018; Aguirre, 2017) to highly promising (Bettinger et al., 2017; Neilson, 2013), with some showing moderate effects (Muralidharan and Sundararaman, 2015)[1].

The theoretical literature suggests that this intuition can fail in education markets. Student sorting, the distribution in the access to high achieving peers, and the incentives for the provision of quality can change when schools compete in an environment in which parents have peer preferences (Epple and Romano, 1998, 2008; Barseghyan et al., 2019). At the same time, the empirical literature suggests that that the peer composition of the school can be relevant for parents' choices and in the education production function (Abdulkadiroglu et al., 2019; Rothstein, 2006). However, there seems to be a gap between these strands of the literature: few papers take an empirical approach to study school incentives to provide quality, and, to the best of my knowledge, none of these have embedded social interactions in the model to study the implications over how schools compete and the distribution of outcomes in equilibrium.

In this paper, I connect the theoretical and empirical school choice literature to make progress on understanding the industrial organization of education markets under social interactions. I use the concept of social interactions to encompass the whole set of reasons why parents may value the peers in the school, emphasizing that this considers more than just classroom peer effects.[2] I do this by including frictions that are specific to education in a model of demand and supply for schools. I study how these frictions affect the market equilibrium, including prices, school quality, student sorting, and the distribution of academic outcomes in ways that the standard intuition from consumer product markets fails to capture.

In particular, I emphasize three features of education that do not generally apply together to other industries. First, education shares features with retail markets, in the sense that schools are both vertically and horizontally differentiated. Some are more effective at improving outcomes,

[1]See Epple et al. (2017); MacLeod and Urquiola (2019) for a review of the empirical evidence

[2] There are several reasons that have been identified in the literature (Altonji et al., 2015): classroom peer effects, indirect effects through teachers' effort and school accountability, and just preferences for some groups, not related to outcomes.

which makes families agree on which schools they like most (the vertical component). Yet, schools are geographically differentiated, and there are potential match effects. These two characteristics are specific to the family-school pair and make families disagree in their assessment of the schools (the horizontal component¡). Second, education is not provided in isolation: schools are characterized by social interactions, in the sense that the students that choose the same school will share the school infrastructure, teachers, and school days together. Then, the utility that the family gets when choosing a school depends —in part— on who else chooses that option with them. Third, public and private schools usually compete directly in the provision of educational services. This implies that some families can opt-out from the public provider and attend a private school from which the other families —who do not pay the fees— are excluded.

I exploit the Peruvian elementary education system as a setting to study school competition under frictions that are specific to education. This setting has several attractive features for this purpose. It combines a convenient regulatory environment, several sources of quasi-experimental and policy variation, and extensive administrative data. I start by providing evidence for the role of prices on families' school choice decisions. I use a regression discontinuity design that exploits the assignment of a generous private voucher program. The program targets middle and low-income families and covers 75% of the tuition for a network of schools. The eligibility rules generate large discontinuities in the relationship between academic test scores and the probability of being eligible for the voucher. I find that being offered a 75% discount for a set of schools increases the probability of going to a private school by 40 p.p. (percentage points), increases the quality of the schools chosen by 0.642 standard deviations, increases the sticker price by $90 a month, decreases the out of pocket price by $40 a month, and increases the percentage of high income and high human capital peers in the schools chosen by 30 and 35 p.p.

Overall, these results suggest that prices can be a relevant barrier for middle and low-income families to choose schools that have higher value-added and a higher share of high SES families. Under the intuition of a market for quality already discussed, a naive interpretation of these results could suggest that scaling up a similar policy should effectively reduce the income gap in the access to value-added. However, it is not clear how these individual-level effects translate into incentives for the schools and what the distribution of outcomes will look like in equilibrium.

To answer these questions, I provide a framework to evaluate the effects of education policies taking into account the complex relationship between the families' preferences and the trade-offs that schools face when making their supply decision. To do this, I specify and estimate a model of school competition under social interactions. In the model, families differ by human capital, income, and location. A combination of the first two characteristics generates observable family types. The model consists of three main components: a simple model for student outcomes, a model of demand for schools, and a model of school competition, all of which consider social interactions. In the first component, the model for student outcomes, academic achievement depends on the student's own characteristics and on the school effects, referred to as value-added in the education literature. The value-added, in turn, has two components: the productivity to increase outcomes, which is a function of the school inputs and independent of who else goes to the school, and the contribution of the peers' characteristics to other students' outcomes. Following the literature, I refer to the first one as school effectiveness or quality and the second one as the peer effects.

The second main component is a model of demand for schools. I follow Neilson (2013)'s approach to estimate families' preferences, allowing for rich observed and unobserved heterogeneity, and for schools and families to differ in their location in the market. I extend the specification of the student outcomes to include peer effects in the production function and the specification of demand to include preferences for the peers in the school. The model for students' outcomes allows me to recover a measure for school quality independent of the peer effects, so that quality and preferences for peers can be included as separate attributes in the utility function. I model families' beliefs as consistent with the peer composition that is realized in equilibrium. Other work has shown this to be a relevant factor (Black, 1999; Hastings and Weinstein, 2008; Abdulkadiroglu et al., 2019), but, to the best of my knowledge, no previous paper has embedded peers in an empirical IO model of demand and supply.

The final component is a model of supply. I start from the specification of the schools' first-order conditions in Neilson (2013). Under this framework, geographic differentiation gives schools local market power to increase prices and decrease quality in the absence of close substitutes. I expand the first-order conditions to account for social interactions in families' utility functions. Social interactions generate interdependencies in demand that modify school incentives through two channels. On the one side, social interactions add a new dimension for school differentiation, strengthening the market power for firms that are naturally better at attracting more desirable peers. I call this the *direct social effect* through which social interactions affect supply decisions. On the other side, the inclusion of preferences for social interactions in demand incentivizes a screening strategy that exploits heterogeneous responses to price and quality by family type. To the extent to which there is a correlation between how desirable an SES group is and the demand sensitivity of that group, schools will find it profitable to change their price and quality to intensify differentiation on the student body margin. I call this one the *strategic social effect* through which social interactions affect supply decisions.

To study the empirical importance of these mechanisms, I estimate the model using administrative microdata on students' choices and schools' characteristics from urban education markets in Peru. I deal with the endogeneity of prices, quality, and peers, by leveraging three sources of credible exogenous variation. The first set of instruments focus on the endogeneity of the peers' composition. I exploit shocks to student sorting using lagged variation in the local exposure to a teachers' strike, combined with variation in the local market structure (exogenous characteristics of the other schools) and on the local distribution of student characteristics. The second set of instruments consists of cost and quality shifters that leverage a generous teacher payment reform in the public schools, which, by design, introduced substantial variation in teachers' wages across time and within markets. Finally, I exploit the individual level variation on prices generated by the private vouchers' assignment rule. I construct moments based on indirect inference to connect the reduced form regression discontinuity estimates from the data with the predictions from the model for the effects of the voucher. This variation in prices helps to discipline the demand model, as it captures the trade-offs that parents make with respect to each of the school characteristics when they face different prices.

I use the model to study how private schools' incentives change under social interactions. The schools' first-order conditions can be arranged to understand the deviation of schools' supply decisions with respect to the ones under perfect competition. I look at price markups (upward deviations from competitive prices) and quality markdowns (downward deviations from competitive quality) to capture schools' market power. These elements are decomposed into the contribution of baseline market power due to geographic differentiation, direct changes in market power that results from stronger differentiation due to families' preferences for peers, and changes due to the schools' strategic responses to those preferences. To summarize the impact on equilibrium outcomes, I use the demand and supply estimated models to look at the same decomposition for the income gap in the value-added chosen by families.

I find that schools have significant market power in the baseline scenario: on average, prices have a 16.3% markup and qualities have a 11.5% markdown with respect to the quantities schools would have chosen under perfect competition, in the absence of horizontal differentiation. The direct social effect makes price markups to increase by 19.7 p.p. and quality markdowns to increase by 18.4 p.p., expanding the gap in value-added by 20.3 p.p. The strategic social effect contributes to markup increases of 3.4 p.p. and decrease in markdowns of 1.9 p.p. However, there is substantial heterogeneity in the responses of schools that serve low and high SES families, generating a relevant increase in the income gap in access to value-added of 9.0 p.p.

Adopting an equity and efficiency framework, I use the predictions of the model to analyze the effects of counterfactual education policies in equilibrium. I consider four policies that target the demand for schools (vouchers with different designs) and two policies that target the supply side (investments in public schools and entry subsidies for private schools). For each policy, I hold the ex-ante budget of the policy constant, simulate a new equilibrium, recover the distribution of the value-added of the schools to which low and high-income families go and compute the ex-post budget for comparison. I present two measures that summarize the effects of the policies: I look at the change in the value-added income gap as a measure of their impact on equity, and I look at the overall mean change in the quality provided as a measure of efficiency. The policy that performs best in terms of equity is targeted vouchers, but it is quite inefficient. The policy that performs best in terms of efficiency is an entry subsidy to high-quality private schools, but it has only a moderate performance in terms of equity. The rest of the policies range between these two extreme cases.

The discussion on introducing market incentives in education usually focuses on the impact evaluation of isolated policies. However, different policies may provide incentives in different dimensions that complement each other. I explore this idea by looking in more detail at targeted vouchers and entry subsidies, computing the correlation between the local gains from the policy in terms of equity and efficiency and the local market conditions. Targeted vouchers are particularly effective at increasing the value-added of the schools to which low-income families go for families located in neighborhoods with certain characteristics. These are locations close to a private school that has high quality in the baseline or close to a private school that has a low cost of increasing quality while being surrounded by a significant number of voucher eligible families. On the other side, private schools choose to enter into neighborhoods in which they face higher demand for their services. These tend to be neighborhoods in which there is a significant concentration of high-income families or neighborhoods in which there is substantial SES heterogeneity, so that private schools can focus on serving high-income families that prefer to pay a higher price to avoid the schools with low-income families.

These correlations suggest that vouchers cannot increase efficiency if there are no potential high-quality schools in the neighborhood that will respond to more competition. Entry subsidies will not provide incentives for firms that have low costs of providing quality to enter in low-income neighborhoods if there is no willingness to pay for their services. Motivated by these ideas, I simulate counterfactuals in which I assign a share of the budget to vouchers and the rest to entry subsidies. I find that the mix of policies outperforms the linear combination of them. The voucher increases the willingness to pay in low-income neighborhoods and eliminates the strategic responses to social interactions by closing the gap in the willingness to pay between families of low and high SES. This promotes the entry of effective schools in low-income neighborhoods, which families can attend without paying a high price or traveling a long distance.

Related Literature. This paper contributes to three strands of the literature. First, it contributes to the literature on the industrial organization of education markets. Early studies such as Manski (1992); Epple and Romano (1998, 2008); Nechyba (2000, 2003) provide theoretical and computational models on the general equilibrium effects of voucher programs. A second and more empirical set of studies in this literature uses actual data from existing school systems to learn about the economic implications of increased school choice,[3] most of which are reviewed by Epple et al. (2017).

[3]See e.g. Rouse (1998), Hoxby (2003), Angrist et al. (2002, 2006), Hsieh and Urquiola (2006), Ferreyra (2007), Rouse and Barrow (2009), Bravo et al. (2010), Lara et al. (2011), Correa et al. (2014), Dinerstein and Smith (2014), Gazmuri (2015), Altonji (2015) Barrera-Osorio et al. (2017), Bau (2017), Ferreyra and Kosenok (2017), Neilson (2017), Singleton (2017), Walters (2017), Abdulkadiroglu et al. (2018).

Closely related are recent papers that use discrete choice and industrial organization tools to answer questions related to school demand and supply. Among the first ones, Hastings et al. (2009), Walters (2018), and Abdulkadiroglu et al. (2019) focus on understanding the demand for schools. The second set of papers looks at different margins of private schools' supply responses to a broad set of public policies. Dinerstein et al. (2017) look at the effects of school funding reform on the entry and exit of private schools, Neilson (2013) looks at the impact of targeted vouchers on schools' market power, Sánchez (2017) studies the effects of voucher design on schools' pricing and program participation decisions, Kapor (2016) studies the effects of affirmative action and associated scholarships on colleges admission rules, Allende et al. (2018) study the effects of information interventions provided at scale on the provision of quality, and Singleton (2019) looks at the effects of funding policies on charter schools' supply. The first three papers exploit the variation from the policy of interest to understand its effects, while the three last ones exploit quasi-random variation in related but different policies to look at the effects of hypothetical counterfactual policies in equilibrium. This paper belongs to the second group and, to the best of my knowledge, is the first one to model schools choosing both quality and prices, while also considering how their incentives change when they face preferences over social interactions in their demand.

Second, this paper contributes to an extensive literature that deals with endogenous social effects in a diverse set of economic problems. Several elements of the paper are related to this literature: the estimation of peers effects on the education production function; the identification of demand with strategic complementarities (preferences for the student body); the computation of equilibrium models of sorting when such complementarities are present. The last element is necessary for understanding the schools' behavior and recover their marginal costs. A methodological challenge that is common to the first two exercises is what is referred to as *The Reflection Problem* in Manski (1993), which "arises when a researcher observing the distribution of behavior in a population tries to infer whether the average behavior in some group influences the behavior of the individuals that comprise that group. He concludes that "Given that identification based on observed behavior alone is so tenuous, experimental and subjective data will have to play an important role in future efforts to learn about social effects.", where observed behavior can be either academic performance or school choice.

The literature on the effects of peers on academic outcomes has taken two approaches to address this problem, following Manski's advice. The results are mixed results. One set of papers exploit random formation of groups and estimate either small positive effects (Epple and Romano, 2011; Sacerdote, 2011) or sizeable significant effects (Duflo et al., 2011; Carrell et al., 2013; Garlick, 2014). Angrist (2014) calls for caution on the interpretation of those studies: when students are randomized into groups, all groups are very similar by design, generating bias from weak instruments. Zárate (2019) addresses this concern directly in an experiment in selective boarding schools in Peru by guaranteeing systematic variation in group composition in the experiment design. He finds that higher-achieving peers do not improve academic achievement, and can even reduce it for low performing kids. These results are more aligned with the second group of papers that uses quasi-experimental variation in peer characteristics (Duflo et al., 2011; Abdulkadiroğlu et al., 2014). My approach also uses a quasi-experimental variation, but my results are more aligned with the first group of papers. However, I show that my instruments have a relevant first stage, reducing concerns about weak instruments. Also, because of the focus on certain samples in the studies that use quasi-experimental variation in peers (students in highly selective boarding schools) or margins (students near to the admission cutoffs in heavily over-subscribed exam schools), it is reasonable to expect that those results need not coincide with what I find in this paper.

There are several papers that use experimental or quasi-experimental methods to address the reflection problem in the identification of interdependencies in demand for a broad set of questions. Some examples include technology adoption (Conley and Udry, 2010; Guiteras et al., 2019), purchase of insurance (Mobarak and Rosenzweig, 2013), networks and the value of service (Björkegren, 2018), effort at work (Mas and Moretti, 2009), investments in energy efficiency (Allcott, 2011), among others. I take a similar approach by using instruments that exploit the strike shifter and variation in local demographics at the census block level.

The third set of papers in this broad literature on social effects looks at equilibrium models of sorting to recover preferences under interdependencies in demand and understand their impact on some market outcomes. Bayer et al. (2007) and Diamond (2016) use empirical industrial organization methods to study households' preferences for neighborhoods and the determinants of workers' labor supply to cities to understand how they translate into housing prices and the provision of endogenous amenities. Conceptually, this is similar to what I do when I estimate supply using first-order conditions for which the demand sensitivity to prices and quality incorporates the social interactions (the direct effect). However, what is novel in my approach is the inclusion of a strategic social effect, which I find to have relevant implications for inequality in the access to value-added.

This paper is also related to a growing literature on the study of equilibrium market outcomes and competition when firms are choosing more than one characteristic, and when the first-order conditions are potentially nonstandard. This form of endogenous product differentiation appears in several dimensions. Ho (2009) looks at the determinants of insurers' hospital network choices and the effects on pricing and hospital investment incentives. Fan (2013) analyzes the effects of ownership consolidation in the newspaper industry and shows that ignoring the adjustments in newspaper characteristics results in substantial changes in the ex-post outcomes. Wollmann (2018) shows the importance of allowing for endogenous product offering in the US commercial vehicle market to accurately predict the equilibrium outcomes as a response to public policies and changes in market structure. Crawford et al. (2019) look at the welfare distortions from endogenous quality choice in imperfectly competitive markets, concluding that quality is overprovided at a high price in the US cable television industry. My contribution is to provide a framework for the equilibrium analysis of schools' supply decisions when price and quality are chosen endogenously by firms that consider the indirect impact that those choices have on the student body composition and, as a consequence, the implications for overall profits.

The remainder of the paper is organized as follows. In Section 2, I provide an overview of the model and show the decomposition of the schools' incentives under social interactions. In Section 3, I describe the empirical setting and provide evidence for the role that price plays in school choice. In Section 4, I develop a model for student outcomes and an equilibrium model of demand and supply for schools under social interactions. In Section 5, I describe the empirical strategy and research design, and I present the results for the estimated model in Section 6. In Section 7, I use the model to study the outcomes under counterfactual policy designs. Finally, Section 8 concludes.

2 Framework

2.1 Overview of the Model and Intuition

In the following sections, I will specify and estimate a model of school competition under social interactions. In the model, schools are characterized by three types of attributes: the exogenous attributes, the ones that the school chooses directly —price and quality—, and ones that are determined in equilibrium and are only indirectly affected by the school's choices —the peers' characteristics. Families and schools also differ in their location, generating geographic differentiation between schools that is specific to each family which I treat as exogenous[4]. This gives rise to local market power to increase prices and decrease quality in the absence of close substitutes. On top of that, social interactions generate interdependencies in demand that modify school incentives through two channels:

(i) *Direct social effect*: social interactions add a new dimension for school differentiation, changing demand sensitivity to prices and quality. The direction of this effect depends on the preferences. If the families are willing to pay for certain peers, the schools that are naturally better at attracting them (for example, due to location or an unobservable component) have now more market power to raise prices and decrease quality. In other words, schools exploit the fact that stronger differentiation makes families less likely to leave the schools that attract desirable peers when other attributes (like price and quality) change.

[4] A topic for future research is to extend the model to account for endogenous entry of schools.

(ii) *Strategic social effect:* Different types of families can have different preferences over the school attributes. Under certain relationships between the willingness to pay for a particular peer type and that particular type's preferences, schools have incentives to engage in indirect screening by adjusting other characteristics. For example, under a negative (positive) gradient between how desirable a family type is as a peer group and how sensitive that type is to price (quality), schools may choose to raise their prices (quality) to increase their product differentiation in terms of the peer group.

The main applications of this framework will use the model predictions to analyze the effects of counterfactual policies. To guide the discussion for the rest of the paper and understand the intuition of the model, I present a preview of the decomposition of the schools' incentives under social interactions, implied by the estimates in Section 6.The school incentives are described in terms of how they change the markup of price over marginal cost and the markdown of the quality over the one that would have been provided under perfect competition. I summarize the effects that the change on incentives has in equilibrium by looking at the gap in the value-added of the schools chosen by families of high and low income.

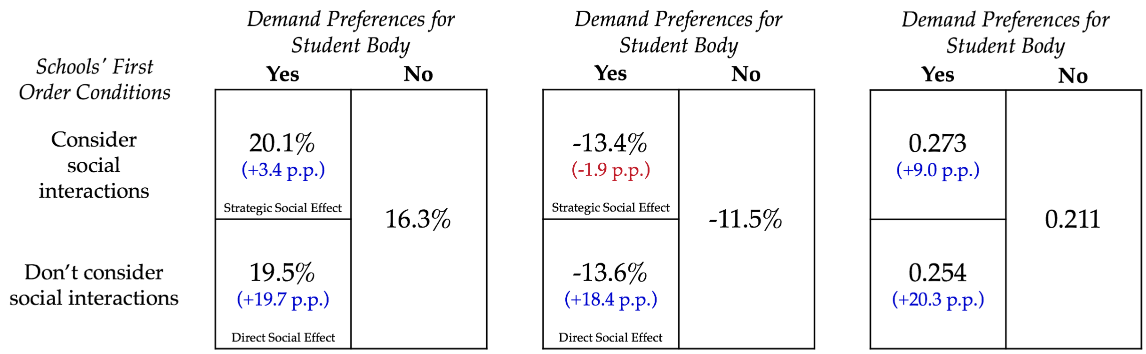
To describe the intuition behind the decomposition o the incentives, I distinguish between two types of school attributes. On the one hand, school attributes are *direct substitutes (complements) in demand* if the school can keep its market share when an indirect attribute increases by reducing (increasing) the one that is controlled directly. On the other hand, school attributes are *strategic substitutes (complements) in demand* when an increase in an attribute that is chosen directly results in a reduction (increase) of the indirect attribute.

2.2 Preview of the Results: Decomposition of Schools Incentives

Figure 1 shows the decomposition of the schools' incentives implied by my estimates. Panel (a) presents the predicted price markup, panel (b) presents the equilibrium quality markdown, and panel (c) presents the predicted equilibrium gap in the value-added between the schools chosen by families of high and low income. Each panel presents scenarios that differ in two dimensions: the families either consider (left-column) or ignore (right-column) the student body of the school, and the schools are either passive (bottom-row) or strategic (top-row) at responding to preferences for social interactions.

For the baseline scenario (right boxes), the simulations in Figure 1 show that, even in the absence of preferences for social interactions, the schools have market power to raise prices (and decrease quality) above (below) the competitive price. The price markup is 16.3% of the marginal costs, and the quality markdown is -11.5% of the competitive quality. Schools' prices and provisions of quality in equilibrium, together with families' preferences, generate a baseline gap in access to value-added equivalent to 0.21.

Figure 1: Decomposition of Schools' Incentives



(a) Price Markup **(b)** Quality Markdown **(c)** Gap

The second scenario (bottom-left boxes) allows for social interactions in demand, which adds a new dimension for school differentiation. In the estimated demand model, all family types value high SES peers and schools with high value-added, but dislike prices. This implies that peers and value-added are *direct substitutes in demand*: if the share of high SES peers of a school goes up and all else stays equal, the school can decrease the quality and keep its market share. The opposite occurs with the price. Prices are *direct complements in demand* with peers. If the concentration of high SES peers drops, the school may be able to keep the market share by decreasing the price. Under this type of demand, including preferences for the student body composition changes the sensitivity of demand to price and quality for schools that naturally attract more high SES families in equilibrium. This makes price markups go up by, on average, 20 p.p. and quality markdowns to go down by, on average, 14 p.p. As a result, the income gap in the access to value-added is increased by more than 20 p.p.

The third scenario (top-left boxes) includes strategic responses from the schools. In the estimated model, I find that demand satisfies strictly increasing differences in the preferences. All types prefer to go to a school with higher shares of high SES families. However, the parameters imply that high SES families value the social interactions in the school (i.e., other high SES families) and school quality more than low SES families. High SES families are also less sensitive to prices. This makes prices and peers *strategic complements in demand:* if the price of a school goes up, and all the rest remains equal, the quality of the peers should also go up. The same thing happens with quality: if the quality of a school goes up, the peer composition of the school should also improve. Then, schools should find it profitable to charge higher prices, as prices screen out low SES families. Consecutively, the SES composition of the school gets more desirable, and the families that remain in the school are willing to pay the higher price, without leaving to the next school.

The case of quality is similar to price. Some schools find it profitable to increase quality, as they attract high SES families and can charge a higher price for the school. The results show that markups go up by an additional 3.4 p.p of the baseline value. However, quality markdown goes down by 1.9 p.p. of the baseline value. Even though these magnitudes seem small compared to the direct effect, the impact on the income gap in the access to value-added is relevant, increasing by 9.0 p.p. As I will show in Section 7, there is a relevant heterogeneity in the magnitudes of the direct and strategic incentives, which leads to seizable redistributional effects in equilibrium.

Note that the direct and strategic effects for price go in the same direction while, for quality, they have opposite effects. Then, even under demand that has the same properties as the ones implied by the estimates, the total effect on quality is an empirical question.

3 Education Markets in Peru

3.1 Relevant Features of the Setting

The empirical application in this paper focuses on the urban education markets in Peru. This education system enrolls more than 600,000 first grade students every year under a mix of public and private providers. More than half of the students go to schools located in urban areas. Descriptive statistics for each education market are provided in Appendix O-1. The context can be explained by examining four events: three changes in regulation and a teachers' strike.

3.1.1 Law for the Promotion of Private Investments in Education (Law N-882)

Peru underwent a major education reform in the late nineties. Law N-882 was made effective in 1998.[5] It aimed at promoting the participation of private providers in education and modified several aspects of the existing regulation. This law established that any individual or organization —including for-profit entities— has the right to start and manage a school. The owner is free to decide how to operate and organize the school, within minimal legal restrictions. Along with this, the document states that teachers in private schools are subject to labor laws for private employers. Their contracts are not affected by the agreements between the Government and public school teacher unions, and there are no legal requirements in terms of teachers' certification for the private sector. Finally, the reform also introduced tax incentives for private schools. Firms that offer education services receive credit for 30% of the reinvested utilities, in addition to import and sales tax exemptions.

Law N-882 contributed to relevant changes to the market structure of the education system in Peru. The country experienced an extensive de-facto privatization, despite the fact that the Gov-

[5] See *Decreto Legislativo N 882 - Minedu, Ley de Promoción de la Inversión en la Educación*

ernment does not provide subsidies for families to attend schools in the private sector.[6] Figure 3 shows the share of students enrolled in private elementary schools in Peru since 1990. This number increased from a stable 13% before 1998 to 28% in 2015. The rates of privatization are larger in urban areas: in Lima, the capital and largest city of Peru, the share of private schools increased from 23% in 2000 to 51% in 2017. These increases coincide with years of steady economic growth,[7] but in which there was no consistent increase in the spending per children in the public sector as a percentage of the GDP, except after 2012 when the Government implemented a teacher payment reform.

3.1.2 Law for the Teachers' Payment Reform (Law N-29944)

Public spending per children in the public sector started to increase in 2012 when the Government initiated a comprehensive teacher payment reform.[8] The main objective was to establish a teaching career in the public sector based on merit and to promote the professional development of teachers. The reform had three components: (i) an immediate change that guaranteed better working conditions, more benefits, and higher base salaries to all teachers, (ii) a gradual implementation of tenure-track teacher contracts, with higher salaries based on tenure and performance, and (iii) the implementation of a centralized application system for teacher assignment.

These changes generated substantial variation in wages across time and across schools. Figure 4 shows the evolution of wages and contract types over time. Panel (a) and (b) show the mean wages for non-tenured and tenured teachers, which go up by 120% and 75% between 2012 and 2018 respectively. Both of these contracts start with a base hourly wage of $15 USD, which also represents a significant increase with respect to the pre-reform mean salary. Panel (c) shows the number of contracts of each type by year, and panel (d) shows the variation in wages within the tenure-track positions. Figure 6 shows an example of the time and geographic dispersion in local wages generated by the reform for the city of Cuzco. This shows that the variation in wages is not only across time but also across locations within a market.

Appendix A1 provides additional details on the design of the teachers' wage structure and centralized application system.

[6]The only exception is the non-for-profit network of Catholic schools "fe y alegría", which provide educational opportunities for poor kids in 19 countries. These schools receive direct subsidies from the Government, but they represent less than 0.01% of total enrollment.

[7]The average GDP growth for the period 2000-2017 was 4.9%

[8]See *Ley de Reforma Magisterial, N-29944.*

3.1.3 New Regulation for Primary Schools (Directiva N 014-2012-MINEDU)

Besides the Teachers' Payment Reform, the Ministry of Education introduced a new regulation for schools in 2012.[9] These rules include two elements that are relevant as modeling assumptions:

1. **Price discrimination**: schools are forced to to make their enrollment conditions public and uniform for all students. This includes application dates, fees, and other payments.

2. **Selection from the school side**: schools are not allowed to engage into active selection or discrimination of students.

Appendix A1 provides more details on the contents of the new regulation and provides supportive evidence for its enforcement.

3.1.4 2017 Teachers' Strike

A massive teachers' strike took place in the 2017 school year.[10] As a consequence, the affected public schools remained closed for up to 60 days. The strike had an important political component, and generated a crisis that ended with the impeachment of the Minister.[11] The decision of the teachers to join the strike was heavily influenced by the teachers' union membership, political affiliation, and political networks. As a result, there is a substantial geographical variation in the intensity of the strike in different schools, both in terms of duration and number of teachers involved.

At the time of the strike, both the Government and experts were worried that many public school families that could afford a private school would switch to the private sector, leaving the lower SES families behind. The effects of the strike can be appreciated in the event studies presented in Figure 5. Panel (a) shows that public schools that had a very high (high) baseline predicted probability of experiencing an intense strike (using an index that will be introduced in Section A6.1) lost 17 (9) students on average. Panels (c) and (d) show that private schools located close to public schools that had a very high (high) baseline predicted probability of experiencing an intense strike (Strike Exposure Index II) increased the share of high SES students enrolled in their schools. However, panel (b) in Figure 6 shows no evidence of negative effects on teachers scores in the next academic year for schools that went on strike.

Appendix A1 provides more details on the origins, timing, and local variation for the strike.

[9]See Directiva N 014-2012-MINEDU/VMGP *"Normas y Orientaciones para el Desarrollo de Año Escolar 2013 en la Educación Básica"*

[10]In Peru the school year starts in March and finishes in December each year.

[11]Ricardo Cuenca, General Director of IEP, an influential thinktanks in Peru claimed that *"This is a political strike, it is not a union strike. Then, the political operators of the Government should start working to fight this situation politically"* .

3.2 Quasiexperimental Evidence: the Impact of Prices on School Choice

In this section, I provide evidence on the restrictions that prices impose on the school choice decisions of middle and low-income families' in Peru. This motivates the analysis in the rest of the paper. I focus on the relationship between changes in the prices that families face and the changes in the characteristics of the schools that they choose.

I use a regression discontinuity design that exploits the assignment of a generous private voucher program. The program is run by an NGO and targets middle and low-income families. The private voucher covers 75% of the tuition for a predetermined network of schools in Lima and eight other big cities in the country. The tuitions covered range between 120 USD and 220 USD, and the voucher is paid directly to the school by the NGO. The application system for the program is simple. There is no pre-application. Interested students show up with their parents at a specific location and time. There is a calendar with application dates for the fall each year. Parents take a household survey and students take an academic test. These application events are highly publicized and held every Saturday from June to November in multiple geographic locations.

The assignment is done by using a two-stage discontinuous rule. Appendix A1 provides more details on the selection process. For this exercise, I exploit the fact that the assignment process generates large discontinuities in the relationship between the academic test scores in the first stage and the probability of being eligible for the voucher.

Panel (a) in Figure A5 presents estimates of the change in the probability of being eligible for the scholarship for students who are above the cutoff, where $Score \geq Cutoff$ is a dummy variable that equals one if the students' academic score is above the cutoff for scholarship eligibility. Here and in what follows, I use optimal bandwidths and robust confidence intervals proposed by Calonico et al. (2014). Being above the cutoff increases the probability of receiving the scholarship by 37% percentage points.

I present the results in Figure 7 and Table 5. The figures displays binned mean of observable characteristics of the schools chosen by students' academic score relative to the cutoff. Table 5 presents estimates of the probability of choosing different school characteristics. Panel A contains reduced form estimates and Panel B contains instrumental variable estimates, where the discontinuity is used as an instrumental variable for being eligible for the scholarship. Panel (a) in Figure 7 shows the probability of choosing a private school. There is a clear jump in the probability of choosing a private school for students above the cutoff. Column 1 in Table 5 shows that being above the cutoff increases the probability of attending a private school by 13.5% in panel A, almost doubling the mean probability. The IV estimate of the effect of getting the scholarship, shown in panel B, is 40.9%. This means that students that get the scholarship close to 100% likely to attend a private school. Panel (b) in Figure 7 shows the effect on the value-added of the schools chosen. Again, we see a jump in the mean value-added of the school chosen when crossing the threshold. Column 2 in Table 5 shows that the reduced form effect of being above the first cutoff is an increase of 0.21 in the value-added of the school chosen and 0.64 for the IV estimation of the effect of receiving the scholarship. Panel (c) in Figure 7 shows the price of the school chosen by parents, which also jumps when they cross the academic test cutoff. Panel A of Column 3 in Table 5 shows that being above the cutoff increases the price of the school chosen by parents by 30 USD and the IV estimate for getting the scholarship is 90 USD. Panel (d) in Figure 7 and Column 4 in Table 5 show the effect on the out-of-pocket price payed. The reduced form impact is a decrease of 13 USD and for the IV a decrease in 40 USD. This is explained by the fact that the mean parent has a 56% probability of choosing a private school, so getting a 75% percent discount decreases the out of pocket price paid even if they are choosing more expensive schools. Panel (e) and (f) in Figure 7 and Column 5 and 6 in Table 5 show the effect on the share of high SES peers in the schools chosen, in terms of family income and education. The reduced form impact is an increase of 0.11 and 0.12 and for the IV an increase in 0.33 and 0.38.

Appendix A1 provides more details on the scholarship program and presents robustness checks for the regression discontinuity design.

4 A Model of School Competition and Social Interactions

To move beyond the small scale policy analysis, I need a framework to understand how these individual-level effects translate into incentives for the schools when the program is expanded and implemented at scale. To do so, I develop a model of how families choose schools and how schools compete in a market for primary education under social interactions. The model will allow me to examine how potential education policies originally aimed at improving equity and efficiency of the education system interact with the direct and strategic incentives when firms compete under social interactions.

In each period, schools play a two-stage game in which they first choose what quality to offer and then what prices to charge. The agents take the following actions under specific information environments.

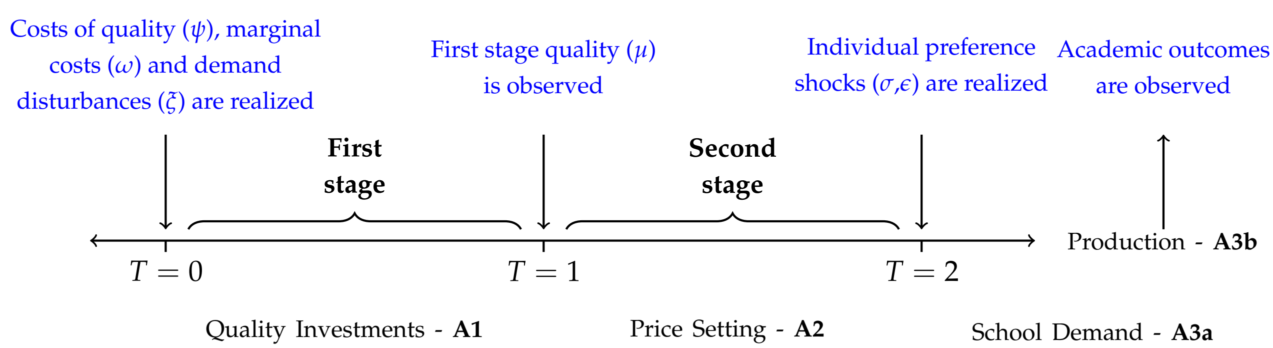
A1 : **Quality Investments**. In the first stage, shocks to fixed costs of quality, marginal cost shocks not related to quality, and demand disturbances are realized and observed by schools. Schools invest in quality simultaneously to maximize expected profits.

A2 : **Price Setting.** In the second stage, schools observe quality investments. Private schools choose a vector of school prices to maximize profits conditional on the quality chosen in the first stage.

A3a : **School demand**. I.i.d. individual preference shocks are realized and observed by families. Families sort into schools based on schools' prices, quality, and their expectation for the student body composition. School profits are realized.

A3b : **Education Production**. Schools provide education services to students, who interact with each other in the classroom. Short-run students' academic outcomes are realized and observed in the form of test scores.

Figure 2: Timing of the Model



4.1 Production: A Simple Model for Student Outcomes

I present a general model for student outcomes that takes student body effects into account. This model is parsimonious; its objective is to recover a measure of school quality and not to fully specify the education production function. I adopt a specific understanding of what is quality: it is the school's effectiveness to produce gains in observable outcomes (test scores, in this case), cleaned of the effect of individual and peer characteristics. In other words, it is the average treatment effect on outcomes of going to a particular school, unconditional on who else goes there. I will refer to this object as the *school observable quality*. This definition acknowledges the fact that there could be other dimensions of quality that are not observed by the econometrician. The model presented in this section will allow for such unobservables to affect parents' choices, with the limitation of not being able to separate the effectiveness to improve unobserved outcomes from other unobserved reasons why parents may value schools.

4.1.1 School Value-Added

Let Y_{ij} be the potential value of an outcome that student i would achieve if she attended school j at time t. The outcome is determined by a vector of characteristics of the student, X_{it}, and a school quality component common to all students in the school:

$$Y_{ijt}(\tau) = \mathbf{X}'_{it}\pi^x + \theta_{jt}(\tau) + \epsilon_{ijt} \tag{1}$$

Where $\theta_{jt}(\tau)$ is the value-added of the school, common for all students. Following Altonji et al. (2015), I allow this term to vary as a function of the policy in place, which is indexed by τ (for example, a voucher program that targets low SES students). As in their setting, I do not observe policy changes that vary the degree of choice in my empirical application. All the estimation of the model is done under the baseline scenario denoted by $\tau = 0$. However, in section 7, I will use the parameters of the model to explore how the distribution of student achievement changes under counterfactual school choice policies. I keep the term τ in the notation to highlight the mechanisms in the counterfactuals, which would require simulating how student outcomes change due to both changes in sorting and changes in the supply for quality. To do so, I need to decompose the value-added into the contribution of the school and the contribution of the student body.

4.1.2 Input Decomposition of the Value-Added

The estimated value-added, θ_{jt}, captures both the effects of the student body over outcomes and the effects of school inputs and technology over outcomes (teacher's quality, pupil ratio, infrastructure, principal's quality, curriculum, among others). The vector $\boldsymbol{x_i}$, which is a subset of $\mathbf{X}_{it}$, includes characteristics that are assumed to have a potential impact on the outcomes of other students. This vector includes a two-dimensional socioeconomic type $\boldsymbol{x_i} = (x_i^y, x_i^e)$, composed by two binary variables that indicate whether the family has high income and/or high human capital.[12]. To characterize the peers in the school, I define $\mathbf{z_{jt}}$, a vector that includes the mean for the characteristics in $\boldsymbol{x_i}$ for school j in time t. To decompose these terms, I project the value-added, $\theta_{jt}(\tau)$, into a linear function of two types of inputs (peers and school inputs) and a residual:

[12] $x_i^y = \mathbf{1}(y_i \geq \bar{y})$ and $x_i^e = \mathbf{1}(e_i \geq \bar{e})$. These variables could have included nonlinear functions of underlying variables and can be equal to or a subset of X_{it}. I include a restricted set of variables to keep the model tractable.

$$\theta_{jt}(\tau) = \mathbf{z}_{jt}(\tau)'\pi^z + \underbrace{\mathbf{I_{jt}}(\tau)'\pi^I(\tau) + \epsilon^s_{jt}(\tau)}_{\mu_{jt}(\tau)} \tag{2}$$

The component of school value-added that does not include peers is μ_{jt} and coincides with the concept of *school observable quality*. This includes a vector of school inputs, $\mathbf{I_{jt}}(\tau)$,[13] how these inputs translate into value-added (the "technology", π^I) and a residual ϵ^s_{jt} that includes unobservables that contribute to school effectiveness. However, for the analysis in this paper it is sufficient to have a measure of μ_{jt}, without decomposing it into other inputs.

This specification provides a simplified framework, ignoring aspects that have been proven to be important in the literature. One is indirect peer effects. As discussed by Rothstein (2006), differences in outcomes between schools can be attributed to the student body composition in three ways: (i) the effect of individual student characteristics on their own outcomes, (ii) direct effects of peer characteristics on outcomes, and (iii) indirect effects of the school composition on outcomes through inputs that respond to the student body (for example, better peers may lower the cost of attracting more effective teachers). It is hard to separate these channels, without better measures of the school inputs and instruments for them. Future versions of the paper could, for example, allow marginal costs to vary as a function of peer composition.[14] This specification also ignores potential student-school match effects. If students select into the schools based on those effects in a way that is not captured by $\mathbf{X}_{it}$, they will also be part of θ_{jt}.

Outcomes Under Counterfactuals In counterfactual exercises (represented by τ), I look at the distribution of the value-added of the schools chosen by families of different SES groups. Changes in θ_{jt} under different τ's come through two effects: changes in the characteristics of the peers and changes in the school observable quality. The former comes from responses in the equilibrium allocation of students under the counterfactual, which in turn shifts $\mathbf{z}_{jt}(\tau)$. I present a model of demand for schools in the next section. The latter is clear from Equation 2. The mix of inputs and the technology may vary as a function of the counterfactual in place. For example, a student busing program reduces demand sensitivity to distance, introducing more competition in quality. Schools may respond by either changing their observed inputs, changing how they use those inputs, or changing the unobserved inputs. I will not distinguish between this channels.

[13]This vector can include non-linear terms and interactions

[14]For example, Singleton (2019) estimates cost differentials across student populations and shows that charters schools enter in areas with low-cost populations).

4.2 Demand for Schools

4.2.1 Preferences

The proposed model of demand for schools follows Berry et al. (1995) and Berry et al. (2004) and allows for geographical differentiation as Neilson (2013). Each family i has one child who will enroll in school as a student[15] and is characterized by a two-dimensional socioeconomic type $\boldsymbol{x_i} = (x_i^y, x_i^e)$, which includes binary variables that indicate whether the family has high income and/or high human capital.[16] Families also differ in their location in the city, loc_i, which I assume to be exogenous. The location defines the distance between the family and each school D_{ij}. Finally, families also differ on unobservable characteristics, including the vector $\boldsymbol{v_i}$ and another characteristic v_i^{net}. To simplify notation, family observable type $\boldsymbol{x_i}$ is sumarized as a discrete type $k = \{1,2,3,4\}$.[17]

Each school j is characterized by a set of observable characteristics, including a measure of school quality in terms of the discussion in Section 5.3.1, μ_{jt}, price, p_{jt}, variables that characterize the student body in terms of the share of families with high income and high human capital (mothers' education) in the school, z_{jt}^y and z_{jt}^e, the location of the school, loc_j, an indicator for whether the school in the private voucher network, $net_{jt} = \mathbf{1}(j \in \Omega_{network})$, and other school characteristics, r'_{jt}, which include whether the school is private, has a religious orientation, is for-profit, coeducational, integrated with a high school, has morning shift. I also include dummies for the four main chains and for schools that belong to other chains. Finally, there is an unobservable quality component, ξ_{jt}. The utility that family i derives from chosing school j is a linear function of the school's characteristics:

$$U_{ijt} = \beta_i^\mu \mu_{jt} - \alpha_i p_{jt} + \beta_i^{z^y} \hat{z}_{jt}^y + \beta_i^{z^e} \hat{z}_{jt}^e + \beta_i^d D_{ij} + \beta_i^{net} net_{jt} + r'_{jt}\beta^r + \xi_{jt} + \epsilon_{ijt} \tag{3}$$

Where

$$\alpha_i = \bar{\alpha} + \sum_{l=y,e} \alpha_l x_i^l, \qquad \text{same for } \beta_i^d$$

$$\beta_i^\mu = \bar{\beta}^\mu + \sum_{l=y,e} \beta_l^\mu x_i^l + \beta_u^\mu v_i, \quad \text{same for } \beta_i^{z^y} \text{ and } \beta_i^{z^e}, \quad \text{with } \boldsymbol{v_i} \sim \ln\mathcal{N}\left(\boldsymbol{m}_v, \Sigma\right)$$

$$\beta_i^{net} = \bar{\beta}^{net} + \sigma^{net} v_i^{net}, \quad \text{with } v_i^{net} \sim \mathcal{N}\left(0,1\right)$$

[15] I assume a unitary actor for the household and refer interchangeably to families and students as decision makers.

[16] $x_i^y = \mathbf{1}(y_i \geq \bar{y})$ and $x_i^e = \mathbf{1}(e_i \geq \bar{e})$

[17] k represents the following combination of $\boldsymbol{x_i}$: $\{(LowY, LowE), (HighY, LowE), (LowY, HighE), (HighY, HighE)\}$

Note that utility is over the expectation for the student body composition. The preference for quality and the two dimensions of the student body are heterogeneous along an unobserved family characteristic distributed lognormal with parameters $\boldsymbol{m}_v$ and Σ, which allows for a correlation between the preferences for the student body. Preference for the schools in the private voucher network are heterogeneous along an unobserved family characteristic distributed normal. Finally, ϵ_{ijt} is a random shock to the individual preference of family i for each school j and time t, assumed to be independent of school attributes and from each other.

4.2.2 Market Shares

The family will choose school j to maximize U_{ijt} in their choice set Ω_{it}. Markets are defined by the urban areas of the 32 main cities in the country, which represent 87% of the urban population. I follow Neilson (2013) to characterize the distribution of types across markets. In practice, each market is characterized by a set of nodes (N_m) and schools (N_s), where w^m_{nk} is the distribution of students of type k across nodes and π^m_k is the proportion of the students in the market who are of type k. The model estimation exploits the detailed node-level geographic data by defining choice sets separately for every node: while choice sets include every school in the market, the distance is school-node pair specific, making the distance attribute vector vary among families in different nodes.

By assuming that ϵ_{ijt} has a standard extreme-value distribution we can calculate the probability that a family of observable type k, who lives at node n and has unobservable type $\boldsymbol{v_i}$ will select school j as follows:

$$\mathcal{S}^{nk}_{ijt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y, \mathbf{z}^e) = \left(\frac{\exp(U_{ijt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y, \mathbf{z}^e))}{\sum_{l \in \Omega_{it}} \exp(U_{ilt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y, \mathbf{z}^e))} \right) \tag{4}$$

By integrating these probabilities over i we will obtain the share of families of type k who live at node n at time t who will select school j. We can obtain the shares of each school at the market level, taking the sum of probability that a family of type k will select school j across nodes and both observable and unobservable types (N_v), which have weights ω^v.

$$s^k_{jt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y, \mathbf{z}^e) = \sum_n^{N_m} s^{nk}_{jt} \cdot w^m_{nk} \tag{5}$$

$$s_{jt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y, \mathbf{z}^e) = \sum_k^K \sum_n^{N_m} s^{nk}_{jt} \cdot w^m_{nk} \cdot \pi^m_k \tag{6}$$

Where Equation 5 is the market shares of the school j for each type k and Equation 6 is the share for the total market. Also,

$$s_{jt}^{nk} = \sum_{i=1}^{N_v} \mathcal{S}_{ijt}^{nk} \cdot w^v$$

4.2.3 Beliefs about the Student Body Composition

The student body characteristics of the school are equilibrium outcomes, determined jointly with demand. I assume that families observe the characteristics of the schools in their choice set and the individual preference shocks. Based on their information set, families form beliefs that are consistent with the expectation of the realized student body composition o the school. Then, $\mathbf{z}^y$ and $\mathbf{z}^e$ are vectors that contain the fixed points that satisfy the following system of equations:

$$\hat{z}_{jt}^x = \mathcal{Z}_{jt}^x(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y, \mathbf{z}^e) = \frac{N \times \pi_{k=x}^m \times \sum_n^{N_m} \sum_i^{N_v} \mathcal{S}_{ijt}^{n,k=x}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y, \mathbf{z}^e) \times w^v \times w_{n,k=x}^m}{N \times \sum_k^K \sum_n^{N_m} \sum_i^{N_v} \mathcal{S}_{ijt}^{nk}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y, \mathbf{z}^e) \times w^v \times w_{nk}^m \times \pi_k^m} \tag{7}$$

For $x = \{y, e\}$ and $j = \{1, ..., N_j^m\}$

4.3 Supply for Schools

4.3.1 Fixed and Marginal Costs

I follow Neilson (2013) and model private schools as profit-maximizing. I define the cost functions of school j as:

$$MC_j(\mu_{jt}) = \sum_l \gamma_l w_{jt}^l + \gamma_\mu \mu_{jt} + \omega_{jt} \tag{8}$$

$$FC_j(\mu_{jt}) = \sum_l \lambda_l w_{jt}^l + \psi_{jt}\mu_{jt},, \quad \text{where } \psi_{jt} = \bar{\psi}_j + \Delta\psi_{jt} \tag{9}$$

I assume that the marginal cost of providing education is both independent of the number of students and linear in a vector of cost characteristics. I decompose these characteristics into a subset observed by the econometrician, which includes the quality chosen —μ_{jt}— as well as other characteristics —the vector of marginal cost shifters w_{jt}^l—, and an unobserved component, ω_{jt}. I allow them to be correlated with ξ_{jt}, as schools with higher unobserved quality might engage in larger costs.

The fixed cost is also a linear function of school characteristics. For the analysis in this paper, the relevant characteristic in the fixed cost is quality. The fixed cost of providing quality μ_{jt} is given by an unobservable shock ψ_{jt}. This shock can be decomposed into a school-specific persistent component, $\bar{\psi}_j$, and a deviation, $\Delta\psi_{jt}$.

4.3.2 Second Stage: Pricing Decision

The profit function for a firm j in a market with N students is given by:

$$\pi_{jt}\Big(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p})\Big) = \Big(p_{jt} - MC_{jt}(\mu_{jt})\Big) \times N \times s_{jt}\Big(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p})\Big) - FC_{jt}(\mu_j) \tag{10}$$

The first order condition for price for school j is:

$$\frac{d\pi_{jt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{dp_{jt}} = s_{jt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p})) + \frac{ds_{jt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{dp_{jt}} \times (p_{jt} - MC_{jt}(\mu_j)) = 0 \tag{11}$$

The first-order condition can be rearranged as the following pricing equation:

$$p^*_{jt} = \underbrace{\sum_l \gamma_l w^l_{jt} + \gamma_\mu \mu_{jt} + \omega_{jt}}_{\text{Marginal Cost}} + \underbrace{s_{jt} \times \left(-\frac{ds_{jt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{dp_{jt}}\right)^{-1}}_{\text{Price Markup with Social Interactions}} \tag{12}$$

The second term in Equation 12 is the price markup and represents how much the optimal price chosen by a firm with market power deviates (upwards) from the competitive price (equal to the marginal cost). The markup considers the demand sensitivity to price, which is given by the total derivative of shares with respect to prices, $\frac{ds_{jt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{dp_{jt}}$. I call this second term the *markup with social interactions*, as it takes into account the effects that social interactions have on the school's pricing strategy. This derivative differs from the version without social interactions, as it considers both the direct and strategic effects of including the terms $\mathbf{z}^y(\mathbf{p})$ and $\mathbf{z}^e(\mathbf{p})$ in the demand. In practice, the total derivative is obtained by aggregating over observed and unobserved types and nodes:

$$\frac{ds_{jt}}{dp_{jt}} = \sum_k^K \sum_n^{N_m} \sum_{i=1}^{N_v i} w^m_{nk} \times \pi^m_k \times w_{v_i} \times \frac{ds_{ijt}}{dp_{jt}}$$

The nature of the incentives under social interactions is clear by looking at the individual level term $\frac{ds_{ijt}}{dp_{jt}}$:

$$\frac{ds_{ijt}}{dp_{jt}} = \underbrace{\overbrace{\frac{\partial s_{ijt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{\partial p_{jt}}}^{\text{Sensitivity to price - (a)}}}_{\text{I - Direct Effect (Baseline + Social)}} + \sum_{l \in \Omega_{it}} \overbrace{\frac{\partial s_{ijt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{\partial z^y_{lt}}}^{\text{Sensitivity to peers (Poor) - (b)}} \times \overbrace{\frac{dz^y_{lt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{dp_{jt}}}^{\text{Change in peers (Poor) - (c)}}$$
$$\underbrace{+ \sum_{l \in \Omega_{it}} \overbrace{\frac{\partial s_{ijt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{\partial z^e_{lt}}}^{\text{Sensitivity to peers (M.Educ) - (d)}} \times \overbrace{\frac{dz^e_{lt}(\boldsymbol{\mu}, \mathbf{p}, \mathbf{z}^y(\mathbf{p}), \mathbf{z}^e(\mathbf{p}))}{dp_{jt}}}^{\text{Change in peers (M.Educ) - (e)}}}_{\text{II - Strategic Social Effect}} \quad (13)$$

The first term is the direct effect of prices on demand (I). This term reflects a combination between the baseline market power and the direct social effect. The first one emerges because of product differentiation, which exists even in the absence of social interactions. The second one is the result of an expansion in product differentiation due to social interactions, which changes the distribution of demand responses to prices. Note that this term is negative. As this term gets closer to zero, the less sensitive the demand is to prices, and the larger the market power. If peers and prices are *direct complements in demand*, schools that are better at attracting high SES peers gain market power under social interactions.

The second term is the strategic social effect of prices on demand (II). This term depends on how sensitive demand is to peers —terms (b) and (d)—, and how much do peers respond to prices —terms (c) and (e). If prices and peers are *direct and strategic complements in demand*, both terms are positive, reducing the denominator of the markup in Equation 12 for schools that are better at attracting high SES peers. Then, the strategic social effect will make those schools gain market power.

In Appendix A3 I show how I solve for the terms (a), (b), (c), (d), and (e).

4.3.3 First Stage: Quality Decision

In the first stage, schools simultaneously choose quality, considering that their actions will impact profits in the second stage. The first stage profit function for a firm j in a market with N students is given by:

$$\pi_{jt}\Big(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu}), \mathbf{z}^y(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu})), \mathbf{z}^e(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu}))\Big) =$$
$$\Big(p_{jt} - MC_{jt}(\mu_{jt})\Big) \times N \times s_{jt}\Big(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu}), \mathbf{z}^y(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu})), \mathbf{z}^e(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu}))\Big) - F_{jt}(\mu_j) \quad (14)$$

Where $\mathbf{p}^*(\boldsymbol{\mu})$ is the vector with the equilibrium second stage prices given the vector of first stage qualities in Equation 12. Note that second stage prices are a direct and indirect function of the vector $\boldsymbol{\mu}$, as quality also affects the peer composition, which in turn affects prices.

The first order condition for quality for school j is:

$$\frac{d\pi_{jt}}{d\mu_{jt}} = -\underbrace{\psi_{jt}}_{\text{Change in Fixed Cost}} - \underbrace{\gamma_\mu \times N \times s_{jt}(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu}), \mathbf{z}^y(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu})), \mathbf{z}^e(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu})))}_{\text{Change in per student Margin} \times \text{N. of Students}}$$

$$+ \underbrace{\left(p^*_{jt}(\mu) - MC(\mu_{jt})\right) \times N \times \frac{ds_{jt}(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu}), \mathbf{z}^y(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu})), \mathbf{z}^e(\boldsymbol{\mu}, \mathbf{p}^*(\boldsymbol{\mu})))}{d\mu_{jt}}}_{\text{Margin per student} \times \text{Change in N. of Students}} = 0 \tag{15}$$

And the optimal quality is given by:

$$\mu^*_{jt} = \frac{p - \sum_l \gamma_l w^l_{jt} - \omega_{jt}}{\gamma_\mu} - \frac{1}{\gamma_\mu} \times \frac{ds_{jt}(\mu, \cdot)}{d\mu_{jt}}^{-1} \times \left(\frac{\psi_{jt}}{N} + s_{jt}(\mu, \cdot) \times \gamma_\mu\right) \tag{16}$$

As opposed to price, the quality of the school enters the fixed cost. Then, the expression for the quality markdown is obtained by comparing the optimal quality in Equation 16 with the one that would have been provided under perfect competition (the one that makes profits equal to zero):

$$\text{Quality Markdown} = \mu^*_{jt} - \mu^{\text{competitive}}_{jt}$$

where

$$\mu^{\text{competitive}}_{jt} = \frac{p - \sum_l \gamma_l w^l_{jt} - \omega_{jt}}{\gamma_\mu + \dfrac{\psi_{jt}}{N \times s_{jt}(\mu, \cdot)}} \tag{17}$$

The total derivative of the shares with respect to quality $\frac{ds_{jt}(\cdot)}{d\mu_{jt}}$ is, again, obtained by aggregating the individual effect over observed and unobserved types and nodes:

$$\frac{ds_{ijt}}{d\mu_{jt}} = \underbrace{\overbrace{\frac{\partial s_{ijt}}{\partial \mu_{jt}}}^{\substack{\text{Sensitivity to}\\ \text{Quality - (f)}}}}_{\substack{\text{I - Direct Effect}\\ \text{(Baseline + Social)}}} + \underbrace{\sum_{l\neq j} \overbrace{\frac{\partial s_{ijt}}{\partial p_{lt}}}^{\substack{\text{Sensitivity to}\\ \text{Prices - (a)}}} \times \overbrace{\frac{\partial p^*_{lt}}{\partial \mu_{jt}}}^{\substack{\text{Direct Change}\\ \text{in Prices - (g)}}}}_{\text{II - Strategic Pricing Effect}} + \underbrace{\sum_{l\in\Omega_{it}} \overbrace{\frac{\partial s_{ijt}}{\partial z^y_{lt}}}^{\substack{\text{Sensitivity to}\\ \text{Peers (Poor) - (b)}}} \times \overbrace{\left(\frac{\partial z^y_{lt}}{\partial \mu_{jt}} + \sum_{r\in\Omega_{it}} \frac{dz^y_{lt}}{dp_{rt}} \times \frac{dp^*_{rt}}{d\mu_{jt}}\right)}^{\text{Change in peers (Poor)}}}_{\text{III - Strategic Social Effect (...)}}$$

$$\underbrace{+ \sum_{l\in\Omega_{it}} \underbrace{\overbrace{\frac{\partial s_{ijt}}{\partial z^e_{lt}}}^{\substack{\text{Sensitivity to}\\ \text{Peers (Educ)}}}}_{\text{(d)}} \times \overbrace{\left(\underbrace{\frac{\partial z^e_{lt}}{\partial \mu_{jt}}}_{(j)} + \sum_{r\in\Omega_{it}} \underbrace{\frac{dz^e_{lt}}{dp_{rt}}}_{(e)} \times \underbrace{\frac{dp^*_{rt}}{d\mu_{jt}}}_{(i)}\right)}^{\text{Change in peers (Educ)}} + \sum_{l\neq j} \overbrace{\frac{\partial s_{ijt}}{\partial p_{lt}}}^{\substack{\text{Sensitivity to}\\ \text{Prices - (a)}}} \times \overbrace{\sum_{r\in\Omega_{it}} \left(\frac{\partial p^*_{lt}}{\partial z^y_{rt}} \times \frac{\partial z^y_{rt}}{\partial \mu_{jt}} + \frac{\partial p^*_{lt}}{\partial z^e_{rt}} \times \frac{\partial z^e_{rt}}{\partial \mu_{jt}}\right)}^{\substack{\text{Indirect Change}\\ \text{in Prices - (k)}}}}_{\text{(...) Strategic Social Effect}} \tag{18}$$

The individual effect is given by three components. The first one (I) follows a similar intuition as the one for price. It combines baseline market power from product differentiation with expanded differentiation from social interactions. As quality and peers are *direct substitutes in demand*, market power increases for firms that are naturally better at attracting high SES peers, as the demand sensitive to quality is reduced with a higher share of high SES peers. The second component (II) is not specific to social interactions and results from the two-stage nature of the game. The third term (III) is again similar to the strategic social effect on prices, with two differences. First, it goes in the opposite direction of the direct effects. While both quality and peers are *strategic complements in demand*, in this case, the partial derivative is positive. This means that movements of the demand sensitivity towards zero will go in the opposite direction of the positive strategic effect. Then, the strategic social effect on quality reduces strategic incentives to exert market power and reduce quality. Second, it also takes into account the screening incentives in the pricing game in the second stage, which is the last term (a) $\times$ (k).

An important term in this expression is $\frac{dp^*_{ij}(\mu)}{d\mu_{jt}}$. This is the derivative of the school's optimal price in Equation 12 with respect to quality. The expression and discussion for the intuition is given in Appendix A3. Terms (a), (b), (c), (d) and (e) come from the previous section. In Appendix A3 I show how I solve for the terms (f), (g), (h), (i), (j), and (k).[18] The following section characterizes

[18] The generic version of Equation 15 is given by:

$$\frac{d\pi_j}{d\mu_j} = \frac{\partial \pi_j}{\partial \mu_j} + \underbrace{\left(\frac{\partial \pi_j}{\partial p^*_j} + \frac{\partial \pi_j}{\partial \mathbf{z}^y_j}\frac{d\mathbf{z}^y_j}{dp^*_j} + \frac{\partial \pi_j}{\partial \mathbf{z}^e}\frac{d\mathbf{z}^e}{dp^*_j}\right)}_{=0 \text{ (Envelope Theorem)}} \times \frac{dp^*_j}{d\mu_j} + \left(\frac{\partial \pi_j}{\partial \mathbf{p}^*_{-j}} + \frac{\partial \pi_j}{\partial \mathbf{z}^y}\frac{d\mathbf{z}^y}{d\mathbf{p}^*_{-j}} + \frac{\partial \pi_j}{\partial \mathbf{z}^e}\frac{d\mathbf{z}^e}{d\mathbf{p}^*_{-j}}\right) \times \frac{d\mathbf{p}^*_{-j}}{d\mu_j} + \frac{\partial \pi_j}{\partial \mathbf{z}^y}\frac{d\mathbf{z}^y}{d\mu} + \frac{\partial \pi_j}{\partial \mathbf{z}^e}\frac{d\mathbf{z}^e}{d\mu_j}$$

Note that the second term is equal to zero, because of Equation 11 (the Envelope Theorem). Because of this, the terms the equilibrium.

4.4 Equilibrium

4.4.1 *Definition*

A rational-expectations equilibrium is a tuple

$$\left\{ \{\mathbf{z}^y, \mathbf{z}^e\}_{j\in J}, \{\mathbf{p}\}_{j\in J}, \{\boldsymbol{\mu}\}_{j\in J} \right\}$$

that satisfies the following properties:

1. $\{\mathbf{z}^y, \mathbf{z}^e\}_{j\in J}$ are the vectors of student body compositions, which correspond to the fixed points in Equation 7 given the vector of school qualities $\boldsymbol{\mu}$ and prices $\mathbf{p}$. Equation 7 is the result of aggregating the individual choice probabilities under consistent expectations for the student body so, in equilibrium,

$$\hat{\mathbf{z}}_i^y = \mathbf{z}^y \text{ and } \hat{\mathbf{z}}_i^e = \mathbf{z}^e, \forall i \in I \tag{19}$$

2. $\{\mathbf{p}\}_{j\in J}$ is the Nash equilibrium of the second stage pricing game, in which schools maximize profits given the vector of first stage qualities $\boldsymbol{\mu}$, subject to student body compositions $\mathbf{z}^y(\mathbf{p})$ and $\mathbf{z}^e(p)$, which are also a function of prices:

$$p_j{}^*(\mathbf{p}^*_{-j}) = \underset{p}{\operatorname{argmax}}\, \pi_j(\boldsymbol{\mu}, p, \mathbf{p}^*_{-j}, \mathbf{z}^y(p, \mathbf{p}^*_{-j}), \mathbf{z}^e(p, \mathbf{p}^*_{-j})), \forall j \in J \tag{20}$$

3. $\{\boldsymbol{\mu}\}_{j\in J}$ is the subgame perfect Nash equilibrium qualities, which maximize first stage profits for each school j, subject to the equilibrium of the second stage pricing game:

$$\mu_j{}^*(\boldsymbol{\mu}^*_{-j}) = \underset{\mu}{\operatorname{argmax}}\, \pi_j(\mu, \boldsymbol{\mu}^*_{-j}, \mathbf{p}^*(\mu, \boldsymbol{\mu}^*_{-j}), \mathbf{z}^y(\mathbf{p}^*(\mu, \boldsymbol{\mu}^*_{-j})), \mathbf{z}^e(\mathbf{p}^*(\mu, \boldsymbol{\mu}^*_{-j}))) \tag{21}$$

4.5 Discussion on the model assumptions

Production function of education. In the model for student outcomes, I present the decomposition of potential changes in μ_{jt} into peer effects, other inputs, and technology for expositional purposes, so I can highlight the fact that quality can change in the counterfactuals. However, I

(a) × (g) and (a) × (k) in Equation 18 only consider the sum over $l \neq j$.

do not estimate a functional form for the education production function[19]. Rather, based on the model in section 4.3 I will infer marginal costs from markups and use those to simulate the school quality provided in equilibrium, without specifying how that quality is provided. The limitation of this approach is that the counterfactual simulations will take the marginal cost parameters as invariant. This may not hold if there are equilibrium effects. For example, policy changes that induce schools to hire better teachers so they can provide higher quality will increase the demand for that input. Wages can go up, changing schools' marginal costs.

Interpretation of the quality and student body parameters. The school characteristics should be discussed in more detail. I adopt a specific understanding of quality that is related to school effectiveness with respect to gains in test scores, clean from the effect of individual and peer characteristics. In section 5.3 I will discuss how I calculate μ_{jt} and under which assumptions this element is identified. On the other hand, z_{jt} is defined as a vector with the mean of observable characteristics (income and mothers' education).

By including a measure of school quality that does not take the student body into account, $\tilde{\beta}_i^z$ can be interpreted as a parameter that captures the total weight that parents put on the school body composition when choosing a school. There are several reasons that have been identified in the literature for why parents care about this (Altonji et al. (2015)): classroom peer effects, signaling (MacLeod and Urquiola (2015)), indirect effects through teachers' effort and school accountability, and just preferences for some groups, not related to outcomes[20]. The model cannot distinguish between these reasons, but, to the extent to which the weight does not change under the counterfactuals, the model can be used for the set of exercises proposed in this paper. I cannot think of an evident reason why those weights would change under the counterfactuals.

Capacity constraints. I assume no capacity constraints in the baseline equilibrium. While this may not be reasonable in some school districts in the US, in which some schools are highly oversubscribed, it is a plausible assumption in Peru. The main difference between the two settings is that private schools in Peru can use price as a rationing mechanism when oversubscribed and have more flexibility in expanding capacity. Then, once a market is in equilibrium, schools can anticipate their demand and adjust prices or expand capacity to maximize profits, alleviating capacity constraints, as opposed to a setting in which there are no prices. However, schools may have less flexibility under some counterfactuals in section 7, and in those cases, I would need to make assumptions about how students are assigned to schools. The case of public schools is different, as there are no prices to work as a rationing mechanism. However, enrollment in public schools has been decreasing dramatically, so it is unlikely that they are oversubscribed, and, in any case, public schools are required by law to enroll every student who applies.

[19] Estimating production functions requires microdata on firm inputs and instruments. See Collard-Wexler and De Loecker (2016) for a discussion of this type of models.

[20] As mentioned in section 5.3.1, I ignore changes in inputs due to peer composition

Price discrimination within schools. I assume that schools have a single price scheme. I discuss this assumption in Section 3 and provide supporting evidence in Appendix A1.

Selection from the supply side. I assume that schools do ot select students based on their ability or SES. I discuss this assumption in Section 3 and provide supporting evidence in Appendix A1.

Information frictions. I assume that parents are aware of all the options available in their choice set and their characteristics. I acknowledge that this assumption may not be realistic in education markets and in a developing country setting.[21] This affects the interpretation of the model parameters and imposes some restrictions on counterfactual analysis, but they are not particularly relevant for the questions discussed in this paper.

The parameters in Equation 3 should be viewed as weights that families place on characteristics rather than deep structural preferences. However, the relevant question is whether these weights are subject to change under the counterfactuals. In Appendix A4 I present a simple framework to interpret the parameters under uncertainty about school characteristics. Under this model, the estimated parameters are the preference for attributes scaled by a measure of how precise the beliefs for the attribute are. If the estimated weight on quality for low SES families is lower than the one for high SES families, it can either be because of weaker preferences or less precise information (or a combination of both). Another interpretation that is complementary to this one is that parents put weight on peers relative to quality because they do not have information about the latter, and use peers as a proxy. While understanding the reasons why parents value peers is interesting and can have important policy implications, it does not change the fact that firms take those preferences into account and respond strategically to them. I cannot think of evident reasons why these weights should change under the counterfactuals, so the model can be used to predict behavior. This also complicates welfare analysis. That is why I do not look at welfare in the analysis of the counterfactual simulations, and instead focus on the distribution of value-added as the main outcome.

5 Estimation

5.1 Administrative and Survey Data

The data used in this paper come from several sources. Most of the data were obtained through data-sharing agreements with different Governmental Agencies within the Peruvian Government, in collaboration with the Ministry of Finance. Then, the data was centralized and merged using individual identifiers at the student level and links between family members. Using the merged and de-identified data, I generate the student, school, and market-level elements used in the esti-

[21] See Allende et al. (2018)

mation procedure. More details about the data are provided in Appendix A5

5.1.1 Student Level

Student's education data. From 2012 to 2018, I have detailed information on student records, which include enrollment, demographics on the kids and their parents, and academic performance by subject. The data are merged with individual-level results for Peru's national standardized test.

Student's birth and health records. These records are merged with the education data and include: health conditions of a child at birth (birth weight, height, gestation age, type of birth, type of professional serving, hospital), family's background information, health outcomes for ages 0 to 6, and social programs in which the student is a participates (conditional cash transfer, free lunch, among others).

National Household Survey. This detailed survey is merged with the education data and includes a large set of questions for the socioeconomic characterization of the household.

National Census. In collaboration with INEI (Census Bureau) I associated students in the education system with the census block in which their household is located for the 2017 Census.

Scholarship application. Scholarship application records, socioeconomic survey, and other administrative data with individual identifiers for the years 2014-2018.

5.1.2 School Level

I put together information from different sources to create a panel of schools that includes relevant characteristics. I also use these datasets to construct the instruments.

School characteristics. Detailed data on school characteristics (enrollment, focus, staff, facilities, among others) from the School Directory (*Padrón Escolar*) and the School Census (*Censo Escolar*).

Tuition. Using several sources, including self-collected survey data, I construct a unique panel of school prices from 2014 to 2018. More details in the Appendix .

Teachers. I use extensive administrative data on teachers: payroll data for teachers in the public sector, linked to contract details; data on the teachers' recruitment process, including vacancies, teachers' test scores, ranked order lists for applications, and final assignment; data on teacher unions', including a dataset that describes the union leaders, their political networks, and maps for the unions' internal geographical divisions within a market; data on teachers' political affiliations, web scraped from *Infogob*,[22] and matched using teachers' National ID numbers.

[22] https://infogob.jne.gob.pe/Politico

Surveys. I use three surveys. The first one (2016) is focused on school principals and asks questions related to the admission process and on how they perceive competition. The second one is a price survey for private schools (2016). This survey collects price data for 2,380 private schools in Lima. The third one is a new survey to private schools (2019) that I implemented for 3,150 private schools to verify current prices and recover missing price data from 2014.

5.1.3 Market Level

National Census of Peru I use the Census to obtain the joint distribution of family SES characteristics (income, mother's education, etc.) for kids about to enter primary school in each node for each market. The two recent Censuses were in 2007 and 2017. The Census is used to define the main school markets in the country. I restrict my attention to urban areas. Boundaries of the markets are defined as in Neilson (2013). Maps for each market are shown in Appendix O-1.

College graduates. Data on all registered college graduates registered, including the program and year of graduation. This data is matched with the 2017 Census to obtain the census block in which the graduates live.

5.2 Identification

The estimation of demand presents the usual econometric challenge of endogeneity: school characteristics may be correlated with the unobservable demand shocks. On influential work estimating models of demand with unobserved heterogeneity Berry et al. (2004, 1995); Nevo (2001); Petrin (2002) deal with this problem using instrumental variables. The endogeneity in price and quantities is implied by an oligopoly model in which firms chose prices knowing the realizations of own and other firms unobservables and their effects on relevant elasticities (Berry and Haile, 2016b). A similar logic applies to the endogeneity of quality in the equilibrium model proposed in section 4.

The student body characteristics are also endogenous because they are the result of a sorting problem: $\bar{z}_{jt}$ is the outcome of aggregate choices of certain types of households. This is clear from Equation 3 and Equation 4: if different family types put different weights on the observable characteristics relative to the unobservable one, then aggregate choices will be correlated with the unobservables. How the unobservables relate to the price is particularly sensitive: schools with high unobservable mean utility will tend to charge higher prices if enough high SES families are willing to pay for it, then low SES families (who are more price-sensitive) will be excluded from that school. This generates a correlation between the $\bar{z}_{jt}$ and the unobservable characteristic of the schools. The problem is exacerbated when firms may choose prices strategically, as described in section 4.3.

Moreover, the endogeneity problem is also present in the supply side estimation. Given assumptions on the model of competition and the identification of demand, the marginal costs and cost of quality can be easily recovered as a function of shares, market size, and matrix of partial derivatives for price and quality, which are functions of the demand parameters (Berry and Haile, 2014). To compute the counterfactuals of interest, I need to identify the marginal cost and cost of quality functions. These costs vary with qualities and quantities and depend on unobservable cost shocks that are potentially correlated with unobservable demand shocks. Demand shocks show up in the marginal costs and cost of qualities that are recovered directly from Equation 11 and Equation 15, complicating the identification of the primitives of supply.

To illustrate this idea, think of the following example. A school has a principal with outstanding leadership skills. He is capable of attracting effective teachers (equivalent to a low cost of increasing quality shock, ψ_{jt}, which all else equal implies higher μ_{jt}) and active at promoting a bullying-free environment (equivalent to a high non-observable that parents value, ξ_{jt}). Providing more of these two dimensions of quality (observable and unobservable) increases school differentiation, shifts demand, and increase market shares. The school will partially respond by charging higher tuitions, but the market shares will still be larger than the ones the model would predict based on the observable characteristics. The endogeneity problem results from ignoring the correlation between the unobservable in demand and cost shifters. As a consequence, the preference for observable quality will be overestimated, and the sensitivity to price will be underestimated. If the principals' leadership skills also make him a good manager (equivalent to a low marginal cost shock, ω_{jt}) this will affect the pricing and quality decisions as well, reinforcing the selection problem.

The choice of instruments is based on nonparametric identification results for demand and supply models in Berry and Haile (2014). Those results are extended for estimation with microlevel data in Berry and Haile (2016a), who emphasize that the microlevel data allows for identification under weaker instrumental variable requirements by providing within market variation of choices under fixed market-level unobservables.

I proceed to present the instruments. I describe each instrument with reference to the endogenous characteristics it should be particularly useful at identifying and for which the discussion of the exclusion restrictions is more intuitive. However, all the IV moments contribute to the identification of all the parameters.

5.2.1 Student body characteristics

Equilibrium sorting arises in school choice and residential location problems. Bayer and Timmins (2007) propose to use the fixed attributes of other locations as instruments for the characteristic that generates the spillover (e.g., race, income, or education in the case of individuals, and industry classification in the case of firms). This type of instrument is used in Ferreyra and Kosenok (2018). In practice, the instruments proposed are an application of the BLP instruments. If there are several endogenous characteristics, fewer are left to be used as instruments. This challenges the fulfillment of the relevance condition. Other good examples of papers that deal with equilibrium sorting are Bayer et al. (2007), who use these BLP style instruments together with boundary discontinuity design, and Diamond (2016), who exploits local demand shocks using Bartik instruments for skill sorting.

The instruments. In the context of spatially differentiated firms, it is reasonable to exploit exogenous geographic shocks that i) provide independent and differential exposure for different schools, and ii) shift the behavior of families differentially by type. This will provide correlation between the instrument and the student body composition in a school (relevance condition), independent from the unobservables (exclusion restriction). I use to the following instruments (and their interactions), all of which are calculated as a distance weighted average:

i. Lagged strike exposure index (2 indexes, distance weighted).

ii. Local market structure: share of private and number of schools (distance weighted).

iii. Local demographics (2): share high SES (income and education, distance weighted).

There are two lagged strike exposure index instruments based on a predicted strike intensity index, which is a prediction for the total number of hours lost per student for the schools in the sample. Explanatory variables include the political affiliations of the teachers in the school, the teacher union affiliation, unions political affiliation (which vary by school district), and several other demographic variables for teachers. The *Strike Exposure Index I* is calculated for all schools and considers a distance weighted average of the predicted strike intensity for the public schools in the neighborhood, excluding the own school's index. The idea of the instrument is that the more intense the exposure to the strike, the more likely it is that the school will receive (or not lose) high SES students. On the other side, the *Strike Exposure Index II* instrument is the own predicted probability of strike, calculated for public schools only. Schools with a higher predicted probability of strike are more likely to lose students. These instruments rely on the assumption that competing schools' decisions to participate are unrelated to their own and others' demand shock for the next year. To avoid correlations with enrollment through other channels, the instruments consider only the predicted (and not actual) intensity of strike, which is explained by unions' political factors, which are plausibly excluded from parents' utility and schools' supply decisions.

Local demographics and local market structure instruments are natural in a context in which families do not travel far to go to school. Taking residential sorting and market structure as given, a school located in a neighborhood with a high concentration of high SES families and with less competition from other private schools should enroll a higher share of that group than a school located in a neighborhood with a lower concentration of high SES families and more competition from other private schools.

Appendix A6 discusses how the instruments are constructed, the exclusion restriction, and relevance conditions in detail.

5.2.2 *Instruments for price and quality*

Berry and Haile (2014) show that competitors' product characteristics are not sufficient on their own, as they provide independent variation in the shares, but other instruments are required to provide the independent exogenous variation in the prices[23]. An additional challenge in a setting in which other product characteristics are also considered endogenous is that there could be little variation left in the characteristics that are assumed exogenous (Fan, 2013).[24]

The instruments. I use instruments that shift both marginal costs and the cost of increasing quality which should not be correlated with the school unobservables ξ_{jt}. I also use two set instruments that exploit variation in local demographics and in market structure, in terms of the number and exogenous characteristics of the schools around them. Consistent with the demand model, local demographics, and local market structure should shift markups and markdowns.

I exploit the Teacher Payment Reform that was gradually implemented since 2013 and introduced subsequent changes that span until 2018. I also use other variables that characterize local teacher labor markets. As discussed in section 3, the reform introduced relevant variation in teachers' wages within markets and across time (see Figure 4 and Figure 6). Through the SUP and Nexus datasets, I can associate each public teacher contract and the wage paid to a specific latitude and longitude. The instruments that I use are the following:

i. Public school teacher wage index (distance weighted).

ii. Public school teacher vacancies for post-reform contracts (distance weighted).

iii. Stock of graduates with education degrees (t-1, distance weighted).

iv. Tests scores for teachers (t-1, distance weighted).

v. The flow of new graduates with education degrees in the market.

An example of the within markets spatial variation for the teachers' salaries is shown in Figure 6. The exclusion restriction for these instruments relies on the assumption that the design of

[23](Reynaert and Verboven, 2014) show the importance of having both types of instruments, even in a parametric setting.

[24]This is also related to work by Gandhi and Houde (2016), who argue that these type of instruments can have weak instruments problem, generating practical problems.

the teacher payment reform and workers sorting decisions are unrelated to current local school shocks. This is discussed in detail in Appendix A6.

5.2.3 *First stage*

Table 3 presents the first stage for the five endogenous characteristics of the schools. Column (1) and (2) show the share of high SES peers in terms of income and education, column (3) shows the price, column (4) shows the quality, and column (5) shows the shares. Each regression includes five sets of variables. The first group includes the exogenous school characteristics. The other four are the groups of instruments already described. The instrument group I includes the peers' shifters (strike instruments), the instrument group II includes the cost and quality shifters, instrument group III includes local demographics, and instrument group IV includes market structure variables: the weighted average local number of schools and the weighted average of the exogenous characteristics of the local schools. Some interactions are omitted from the table for space. In most of the cases, the instruments indicated as particularly helpful to shift certain endogenous variables are significant and have the expected sign. For each regression, I present the F-test for the null hypothesis that the first-stage coefficients for the instruments are zero, conditional on the exogenous school characteristics. For each regression, the F-stat is large, rejecting the null hypothesis.

5.2.4 *Using RDs to identify price parameters*

I exploit the individual-level variation in price generated by the discontinuity in the assignment of the private voucher program. This variation is similar to a small scale voucher quasi-experiment.

The NGO that manages the program surveys all parents and keeps the information of all applicants. I match their data to administrative records (enrollment, test scores, household survey, geocoded location of the household), what allows me to map the families to the types in the model (in terms of income, education, and location) and get data on their school choice decisions.

The next section describes how I construct moments based on indirect inference (Gourieroux et al., 1996) to exploit the discontinuity in the design of the program. I include additional moments that are particularly useful at helping me to identify the price sensitivity parameters. These moments also discipline the implied substitution patterns by informing on the trade-offs that families make, in terms of school characteristics, when they face different prices. A detailed discussion of the assumptions required to use this variation is provided in Appendix A6.

5.2.5 Dealing with Selection in the Scholarship Application Decisions

The assumption that the scholarship assignment is "as-good-as-random" around the cutoff justifies its use as a source exogenous variation in prices, and the results should be internally valid. However, the families that apply to the scholarship may come from a selected sample. This can bias the estimated demand parameters if the distribution of unobservables that define the price elasticity of the families that apply are different from the ones in the population.

It is reasonable to assume that the sample of applicants is selected. However, even under selection, the RD results contain valuable information. The challenge is that differences in observed choices between families that face different prices (families at each side of the cutoff) are affected, in part, by unobservables that affect both their selection into the program and their price elasticity. To deal with this selection problem, I implement a selection correction. I have data on the student household and the location in which they take the tests, as shown in Figure A6. We can see that students that live closer to the test location are more likely to apply. Then, as the distance to the test location is excluded from the school choice decision, it can be used as a valid instrument for scholarship application. This allows me to recover the distribution of unobserved preferences for the schools in the sample among the scholarship applicants, and correct for selection on those unobservables that also affect the simulated price elasticities. The model for selection into the private voucher program is presented in Appendix A7.

5.3 Estimating the School Observable Quality and Peer Effects

5.3.1 Value-added

To take the model in to the data I start from a general model of value-added, using the regression:

$$Y_{ijt} = X_{it}^{VA}\beta + \theta_{jt} + \epsilon_{ijt} \tag{22}$$

The value-added of the school is θ_{jt}, which can be obtained either by random or fixed effects and using different methods for adjustment (see Kane and Staiger (2008) and Chetty et al. (2014)). There are two key identifying assumptions behind this specification.

- *Selection on observables*: potential outcomes are independent of the school chosen, conditional on observed characteristics.
- *Stable Unit Treatment Value Assumption (SUTVA)*: the treatment applied to one unit does not affect the outcome of another unit.

Regarding the first assumption, the most widely used measure of value-added is the one that includes lagged values of the outcome, Y_{ijt-1}, in X_{it}^{VA}. Chetty et al. (2014) claim that selection on observables is more plausible for test scores than for longer-run outcomes for which lagged measures of the dependent variable are not available, and Abdulkadiroglu et al. (2019) find similar results when using rank-ordered control functions to construct selection-corrected estimates of value-added (there is less bias in outcomes for which there are lagged observations). Nevertheless, this measure may still be biased due to sorting (e.g., Rothstein (2006)). Recent evidence suggests that selection on observables can be a reasonable assumption: in comparisons to lotteries, value-added measures accurately predict the random assignment but are not perfectly unbiased (Deming, 2014; Angrist et al., 2016, 2017)

Unfortunately, lagged test scores are only available for one cohort in Peru. I will take the following approach. Following Neilson (2013) I will estimate value-added without lagged test scores but using an exceptionally rich set of controls for academic results, which include detailed information on socioeconomic background (including detailed parents' education, households structure, family composition, and assets), health at birth (Behrman and Rosenzweig, 2004; Currie, 2011; Almond and Currie, 2011; Bharadwaj et al., 2013) and health and development records under 6 years old. In the results section I show an approach to validate the value added estimation comparing the estimated value-added with and without lagged test scores.

5.3.2 *Separating the School Observable Quality and Peer Effects*

Regarding the second assumption, SUTVA, it is well known that θ_{jt} will capture: i) School inputs (teacher-pupil ratio, teacher quality, motivation, infrastructure, principal quality), ii) the effects of the student body over test scores and iii) a student-school specific match not captured by X_{it}^{VA} (I ignore this for now). As we have evidence that peer effects may be important for student outcomes, it is likely that SUTVA assumption is not met. Moreover, I need a measure for school quality that is separate from the effects of the peers, so I can include the student body as separate characteristics in the utility function.

To solve this I separate the school observable quality and peer effects in a second stage[25]. In a first stage, $\hat{\theta}_{jt}$ is estimated using Equation 1. In a second stage, I project $\hat{\theta}_j$ into peers and inputs as in Equation 2, plus a school fixed effect.

$$\hat{\theta}_{jt} = \mathbf{z}'_{jt}\pi^z_j + \mathbf{I_{jt}}'\pi^I + \alpha_j + \epsilon_{jt} \tag{23}$$

I use the instrumental variables in Section A6.1 to identify the effects of the student body composition on the measure for value-added. Then, I recover μ_{jt} by subtracting the predicted

[25] This is suggested but not implemented by Altonji et al. (2015) because of lack of data

contribution of peers to value-added, $\mathbf{z}_{jt}(\tau)'\hat{\pi}^z_j$, from $\hat{\theta}_{jt}$[26].

5.4 Estimation of the Demand and Supply Models

I augment the school choice model in Neilson (2013) to include the student body composition and estimate the supply side taking the cream-skimming incentives into account. I estimate the parameters from the utility equation and the marginal costs by using a method of moments estimator. I have three sets of parameters: the linear parameters in the utility function ($\theta_1 = \beta^r$), the non-linear parameters in the utility function ($\theta_2 = \tilde{\alpha}, \tilde{\beta}^q, \tilde{\beta}^z, \tilde{\beta}^d$) and the marginal cost function parameters ($\theta_3 = \gamma$). This method combines the aggregate, micro, and IV (demand) moments in Berry et al. (2004) with moments that exploit the discontinuous variation in price generated by the scholarship assignment and IV (supply) moments in Berry et al. (1995), in which I exploit the orthogonality between unobserved costs and the instruments.

5.4.1 Moments

I use the following moments:

Aggregate Moments for Shares. I need to choose the parameters such that for each year and school the model matches the predicted school market shares to observed shares. These moments will help me identifying the ξ_{jt} parameters.

$$G^1(\theta_2) = s^k_{jt} - s^k_{jt}(\theta_2) \tag{24}$$

Where $s^k_{jt}(\theta_2)$ is the expression in Equation 5.

Micro Moments. With these moments the model will choose parameters to fit the expected characteristics of the chosen schools (in terms of quality, student body, price, and distance) in each market, in each period, for each type k, to the characteristics in the data. These moments are particularly useful for identifying the preferences of different observed family types and for observed school characteristics. From the microdata, I have N_m observations in market m of students identified as type k at time t and their choices. I use the empirical averages of the quality, price, and distance chosen by these families and compare them to the simulated characteristic given the parameters of the model. This comparison defines moments for price, peers, quality, and distance:

[26] As discussed by Altonji et al. (2015) I deal with the reflection problem in Manski (1993) with the instruments discussed in detal in Section A6.1. The second step, in which I recover μ_{jt} will capture the vector of reduced-form coefficients of a model with reflection.

$$G_q^2(\theta_2) = \sum_{i \in N_{kt}^m} q_{ik} - \sum_{n}^{N_m} \sum_{j}^{N_{m,t}^f} s_{jt}^{nk}(\theta_2) \cdot w_{nk}^m \cdot q_{jn} \tag{25}$$

Same for price, peers (by income and education), and distance. $N_{m,t}^f$ are schools in each year t and market m.

Selection Correction for Scholarship Application Moments: In these set of moments, I exploit distance to the location of the for the scholarship application as a special regressor that enters into the application decision, but is excluded from the school choice one. The moment conditions are given by the score of the likelihood with respect to the cost of application parameters, which is described in detail in Appendix A7:

$$G^3(\theta_2, \eta) = \frac{1}{N_s} \nabla_\eta \sum_i log \ell_i(\theta_2, \eta) \tag{26}$$

These moments are useful at identifying the cost of application parameters, and allow me to recover the distribution of unobservable preferences for the schools in the network for the scholarship applicants.

Regression Discontinuity Quasi-experimental Moments. In this set of moments, I exploit the scholarship program as a voucher quasi-experiment, under the assumption that the price change around the cutoffs is as good as random. This moment would require the parameters in the model to match the model's predictions for the difference in the choice probabilities above and below the threshold to the estimates from the Regression Discontinuities. These moments are useful for identifying the price sensitivity of choices.

The estimated effect of getting the scholarship on the characteristics of the schools chosen using the discontinuities as an instrument for price is given by:

$$\hat{\beta}_l^{RD} = (X'ZWZ'Z)^{-1}(X'ZWZ'Y_l) \tag{27}$$

Where l are the school characteristics: quality, price, peers (income and education), and distance. The moment condition requires that the coefficients obtained from the administrative datasets satisfy the IV moment conditions using the simulated outcome:

$$G^4(\theta_2) = \frac{1}{N_s} \sum_{s=1}^{N_s} (Z'(Y_l^{sim}(\theta_2) - X\hat{\beta}_l^{RD'}))' WZ'(Y_l^{sim}(\theta_2) - X\hat{\beta}_l^{RD'}) \tag{28}$$

Where $Y_l^{sim}(\theta_2)$ is the vector of simulated characteristics of the schools chosen by families in the scholarship sample. These simulated characteristics include the selection correction in terms of the distribution of unobservable preferences for the network of schools that accept the scholarship. The detailed simulation procedure is detailed in Appendix A7.

IV Moments. I use two sets of IV moments: the ones for the demand unobservable (ξ_{jt}) and the ones from the pricing and quality decisions (ω_{jt} and ψ_{jt}). These moments express the orthogonality between the demand and cost side unobservable and the chosen instruments. In terms of the demand parameters, these moments are particularly useful at identifying the random coefficients in preferences, and in terms of the supply, this is the key variation to identify the cost parameters.

$$G^5(\theta_2) = \begin{pmatrix} \xi_{jt} \\ \omega^c_{jt} \\ \Delta\psi_{jt} \end{pmatrix} \cdot IV' \tag{29}$$

To construct these moments, ξ_j can be isolated by inverting the share equation as in Berry et al. (1995), and ω_j can be isolated from the FOC.

Note that this implementation exploits first the panel nature of the data, recovering a fixed school level unobservable that impacts the cost of quality ($\bar{\psi}_j$). Then, I exploit the orthogonality between the residual error and the instruments.

5.4.2 *Implementation of the Estimation*

The estimation of $\theta = (\theta_1, \theta_2, \theta_3)$ is done by GMM, following Berry et al. (2004, 1995). I use the SQUAREM acceleration method by Varadhan and Roland (2008) for the fixed point algorithm. The routine consists of an outer-loop to minimize the GMM objective function —whose components are the moments in Equation 25 to Equation 29— and an inner loop to evaluate this objective function at a given θ_2 by inverting the market share equations in Equation 24. The implied $\delta(\theta_2)$ and $s(\theta_2)$ are used to calculate the demand and cost disturbances. Finally, a normalization is required as there is no outside good, normalizing $\xi_{jt} = 0$ for one school in each market.

This version is implemented using data from all the cities for demand, and excluding Lima for supply. Appendix A8 describes how I deal with demand estimation in large markets as Lima. I use quadrature rules for numerical integration as in Skrainka and Judd (2011).

6 Results from the Model

6.1 Student Outcomes Model Results

Table 2 shows the regressions to estimate the value-added of the schools. These are estimates of test scores on a vector of individual student-level characteristics. School value-added is estimated as the school and year fixed effect. Column (1) shows the main specification, which is estimated using data from 2014 to 2018 and is the one used in the model. In order to validate my estimates of value-added I also estimate value-added using lagged test scores for the cohort for which they are available (2018), and compare the two results. This is shown in Figure 9 and in columns 2 and 3 of Table 2. The results do not display relevant changes between one specification and the other, and the plot shows a close and almost unbiased correlation between the two measures.

Table 4 shows the instrumental variables estimate for the peer effects. The instruments used are the ones presented in Table 3, which should address some of the methodological challenges described in the literature. For each of the three specifications for value-added, the peer effects are positive and significant.

6.2 Demand Model Results

Table 6 presents results for the estimated parameters for the school choice demand model. Panel A shows the results for the linear parameters θ_1(or mean preferences) for the main school attributes. All the attributes present the expected signs: parents value the school value-added and the fact that the school has more educated peers, and do not like price, peers that come from poor families and schools that are far from their home. Panel B shows the θ_2 parameters, which capture the observed heterogeneity in preferences. More educated families put a higher weight on value-added and educated peers, and are less price and distance sensitive. However, they put a larger negative weight on poor peers. On the other side, poor families care less about value-added, are more price and distance sensitive, and care less about the peers (they put less positive weight on peers with more educated mothers and less negative weight on peers from poor families). Panel C shows the unobserved heterogeneity in preferences. We see that for all attributes, there is heterogeneity in preferences that is not explained by observable characteristics. Finally, panel D shows the parameters that describe the private voucher application cost.

Panel (a) in Figure 10 shows the distribution of elasticities. The mean absolute value lies slightly below 2. Almost 90% of the estimated elasticities are above 1 in absolute value. The highest value for the elasticities is close to 3.5. In most settings, there are no prices for schools, implying that demand estimates for education are usually expressed in terms of willingness to travel. Then, there are not many results from other settings to compare these estimates.

These results are consistent with the restrictions discussed in section 2: under the estimated parameters, peers and value-added are *direct substitutes in demand*, while prices are *direct complements in demand* with peers. Both prices and peers are *strategic complements in demand*.

6.3 Supply Model Results

Table 7 shows the estimates for the supply model. As expected, providing quality is costly for the schools, increasing marginal costs. There are also systematic differences in costs faced by schools in different markets. For-profit schools have higher marginal costs, religious schools have lower marginal costs, schools that are integrated with a high school and schools that operate in the morning and offer a complete school day schedule have higher marginal costs. All of the chain effects are negative, suggesting scale economies of managing a network of schools.

Figure 10 shows the distribution of the estimated fixed cost of increasing quality. The results display substantial heterogeneity in the estimated values, suggesting large differences in school productivity.

6.4 Model Fit

I examine model fit by using the estimated parameters to simulate equilibrium outcomes and compare simulated to observed outcomes. All the demand side simulations are run on the complete estimating dataset. The supply-side simulations are run on the data for one market, Cuzco. The next section describes the procedure to compute the new equilibriums. The next section provides a description of the procedure to compute the new equilibriums under counterfactual policies.

Figure 11a, Figure 11b and Figure 11c present the fit to the micro-moments by family type. Overall, the model replicates the distribution of the characteristics of the schools chosen in each market. Figure 11d presents the fit for the Regression Discontinuity Moments. A key element for the model to capture the jump in the probability of choosing the schools in the network is the inclusion of the selection correction. Finally, Figure 11e and Figure 11f presents the fit for the predicted price and quality. The model has a reasonable performance fitting the schools' choices of price and quality, except for more expensive and higher-quality schools, for which the model predicts that the prices should be lower and the qualities larger than what I observe in the data. This suggests that there could be some convexity in the cost of providing quality that I do not account for in the model. This can be added to future versions of the paper.

7 Counterfactual Analysis of Education Policies

In this section, I use the estimated model to simulate the effects of counterfactual education policies on equilibrium outcomes and to examine the implications for relevant policy objectives. I also decompose the school incentives into the direct and strategic effects of the social interactions from the theoretical model. Then I characterize how they correlate with local market conditions in order to better understand the circumstances under which they are likely to play an important role in the distribution of academic outcomes.

7.1 Policy Objectives

The estimated model can be used to analyze counterfactual policies in equilibrium and guide the design of public policies. A relevant question is what outcomes should be used to compare alternative policies. I will focus on two metrics that are relevant form the social planner's point of view, which are commonly used in the public finance literature. The first is a measure of *equity* of the policy, defined as the reduction in the gap in access to VA by families of low and high SES (in terms of their income). The second is a measure of *efficiency* of the policy, defined as the mean increase in quality (which in turn affects the value-added).[27] In all the simulations, the ex-ante budgets of policies are held fixed. However, the policies could differ in their ex-post budgets due to incomplete take-up of the policy. To compare the policies by the dollar spent, I also present a measure of the efficiency measure that considers the ex-post cost of the policy. Figure 8 shows these outcomes in the baseline, including the whole distribution of value added by SES.

A usual approach in the literature is to look at the welfare impact of policies based on revealed preference theory. I would be reluctant to do that in this context.[28] The parameters in Table 6 are the ones that rationalize agents' observed actions. These parameters should be invariant to the proposed counterfactuals and can be used for a positive analysis of demand and supply. However, it is harder to interpret them as parameters of indirect utility for welfare analysis. Behavioral biases and information frictions[29] —which, as discussed in Section 4.5, have been extensively described in education— complicate the use of welfare measures. Instead, I focus on the measures of equity and efficiency proposed above and interpret the parameters as the weights that families

[27] It is important to note that all the changes in the value-added as a response to the policies studied in this paper are through changes in quality. While peers can change as a response to the policies, the population of students and their characteristics are fixed in the market. As the peer effects are modeled as linear and I rule out heterogeneity, the policies can have an equity impact through the redistribution of the peer effects, but there can not be a system-wide increase in value-added due to peers.

[28] See Finkelstein et al. (2019) and Hendren (2018) for a discussion on the limitations of willingness to pay estimations inferred from reveled preferences as the welfare-relevant metric for a normative analysis

[29] This type of behavior includes myopic behavior, inattention, complexity, al of which can vary by SES.

put to school characteristics rather than preferences. [30]

7.2 Solving for New Equilibriums

In all of the exercises, a counterfactual equilibrium is defined as a set of individual choice probabilities $\{\mathbf{s}\}$, and schools' student body compositions $\{\mathbf{z}^y, \mathbf{z}^e\}$, prices $\{\mathbf{p}\}$, and qualities $\{\boldsymbol{\mu}\}$ that satisfy the properties listed in Section 4.4.1. I search for a local equilibrium allowing schools to iteratively adjust their first order conditions in small steps.

To compute a new equilibrium, I iterate on the following steps, where for each iteration ι I proceed as follows:

1. Simulate the schools' student body compositions $\{\mathbf{z}_\iota^y, \mathbf{z}_\iota^e\}$ from the individual choice probabilities $\mathbf{s}_\iota$ given the policy in place (τ) and the vectors of prices and qualities $\boldsymbol{p}_{\iota-1}$ and $\boldsymbol{\mu}_{\iota-1}$. This step requires solving the fixed point problem in Equation 7, because families choose schools based on their expectation for the student body, which in turn depends on the choices of families. I proceed by: (i) computing simulated individual choice probabilities under the new policy for the student body composition at the starting point, (ii) computing the implied student body composition for each school from individual choice probabilities, (iii) computing simulated individual choice probabilities under the updated student body, (iv) repeating (ii)-(iii) until convergence of the student body compositions.

2. Simulate the quality investment decision for each school $\boldsymbol{\mu}_\iota^*$, by computing the optimal quality in Equation 16 using the simulated student body compositions and individual choice probabilities from Step 1.ii-iii. This step requires solving for the change in second stage prices when quality changes and the total derivative of the shares with respect to quality, as described in Appendix A3.2.

3. Simulate the pricing decision for each school $\boldsymbol{p}_\iota^*$, by using the student body compositions and individual choice probabilities from Step 1.ii and 1.iii and the qualities from Step 2 to compute the pricing equation in Equation 12. This step requires solving for the total derivative of the shares with respect to prices, as described in Appendix A3.1.

4. Compute the price and quality vectors for the next iteration by taking a small step $\alpha \in [0, 1]$ towards the simulated prices and qualities:

$$\boldsymbol{p}_\iota = \alpha \boldsymbol{p}_\iota^* + (1-\alpha)\boldsymbol{p}_{\iota-1} \qquad \boldsymbol{\mu}_\iota = \alpha \boldsymbol{\mu}_\iota^* + (1-\alpha)\boldsymbol{\mu}_{\iota-1}$$

5. Repeat Steps 1-4 until convergence of the simulated prices and qualities.

[30] A detail discussion the interpretation of the parameters under incomplete information is presented in Appendix A4.

7.3 Specific Market for Counterfactuals: Cuzco

The procedure described in the previous section implies solving for a fixed point in each step of the simulation, independently for each market. This slows down the computation of counterfactuals. This version of the paper presents results for a specific market: the city of Cuzco. In future versions, I will expand the set of markets to obtain a representative sample of the education system of Peru. Figure 12, panel (a) shows the distribution of private and public schools in the city. Panel (b) shows the distribution of family income. More details are provided in the Appendix O-1.

7.4 Decomposition of Schools Incentives

I presented a summary of the decomposition of the school incentives in Figure 1 in Section 2. Here I specify the three scenarios that I simulated in order to compute those results:

Baseline scenario under market power. I start by computing a counterfactual in which the demand for schools does not respond to the student body and in which there are no strategic effects in the schools' first-order conditions. I eliminate the differences in the student body composition between schools by setting them equal to the share of students of each type in the market. I then simulate the individual choice probabilities and market shares for each school and compute the total derivatives of the shares using only the direct effects.[31] I then compute the price markups, quality markdowns, and find the equilibrium qualities and equilibrium prices. Finally, I recompute the equilibrium individual choice probabilities and student body compositions to simulate the distribution of value-added of the schools chosen, which combines the school quality and the contribution of the student body.

Direct social effects. I then compute a counterfactual in which the demand for schools includes the student body, but the schools do not respond strategically. Then, the indirect effects in the first-order conditions are ignored. I simulate the individual choice probabilities and market shares, including the student body composition terms, but again compute the total derivatives of the shares using only the direct effects (which have changed with respect to the basic scenario due to changes in demand). I then compute the price markups, quality markdowns, and find the equilibrium qualities and equilibrium prices. Finally, I simulate the distribution of value-added of the schools chosen.

Direct + strategic social effects. Finally, I look at a counterfactual in which both the demand for schools includes the student body, and the schools respond strategically. The first-order conditions include both the direct and indirect effects. I simulate the individual choice probabilities and market shares, including the student body composition terms, and compute the complete version of the total derivatives of the shares using both the direct and indirect effects. I compute the price

[31] This Includes term (a) for price and terms (f), (a), (g) for quality.

markups, quality markdowns, and find the equilibrium qualities and equilibrium prices. Finally, I simulate the distribution of value-added of the schools chosen.

7.5 Evaluation of Counterfactual Education Policies

7.5.1 Description of the Counterfactual Policies

Table 1: Education Policies for Counterfactual Analysis

N	Policy	Mechanism	Counterfactual in the model	Examples
I	Universal Vouchers	Demand	Subsidies for all students $v_i = \bar{v}$	US[i], Denmark, NZ, Holland, Sweeden[ii]
II	Targeted Vouchers	Demand	Subsidies for poor students $v_i = \bar{v} \times \mathbb{1}[x_i^y = 0]$	Chile[iii], Colombia[iv], US[v]
III	Scholarships	Demand	Subsidies for skilled students $v_i = \bar{v} \times \mathbb{1}[x_i^e = 1]$	US, Peru (Private)[vi], UK (Public)[vii]
IV	Tuition tax deductions	Demand	Subsidies for non-poor students $v_i = \bar{v} \times \mathbb{1}[x_i^y = 1]$	US[viii], UK, Dominican Rep.
V	Public school investment	Supply	Change in market structure $\Omega'_{it} = \Omega_{it} \cup j', \mu_j > \bar{\mu}_j$	New York, Dominican Rep.[ix]
VI	Land lease subsidies	Supply	Entry incentives Δ in $F_{jt} \implies \Omega''_{it} = \Omega_{it} \cup j''$	India[x]

Notes: [i]In the US, fifteen states and Washington, D.C., have school voucher programs, see Epple et al. (2017) for more details. [ii]All these countries have large-scale voucher programs. For a review, see Epple et al. (2017). [iii]See Neilson (2013), Aguirre (2017) and Sánchez (2017) for detailed description of the targeted voucher program in Chile. [iv]See (Angrist et al., 2002). [v]Florida, Georgia, Indiana, Iowa, Oklahoma, Pennsylvania, Rhode Island have vouchers targeted to poor students or students enrolled in low performing schools. [vi]The Walton Family Foundation in the US and Peru Champs in Peru. [vii] 179 selective Direct Grant Grammar Schools existed in the UK between 1945 and 1976. 25% of the spots were given in the form of scholarships, directly funded by the Central Government. The rest of the spots had fees, which were paid by the Local Education Authority or private pupils. [viii]The most generous k-12 private school expenses state tax deduction programs are Louisiana and Wisconsin. At the Federal level, the Tax Cuts and Jobs Act of 2017 allows parents to use up to $10,000 per year from a 529 account to cover private K12 expenses .[ix]See Dinerstein et al. (2017) for a description on funding rules reform in New York and Dinerstein et al. (2019) for a description of the construction program in the Dominican Republic and implied changes in market structure. [x]As described by Rao (2019), many private schools in Delhi operate on land leased from the state (decades ago) in perpetuity at highly subsidized rates.

I look at 6 counterfactual policies summarized in Table 1. The third column describes the general mechanism through which the policy operates: four policies change the incentives for families to choose schools (whose aggregate behavior affects supply incentives indirectly), and two policies that change supply through either direct or indirect changes in market structure. The fourth column provides details on the mechanism through which the policy changes agents' decisions. The last column provides examples of real-world applications of these policies. As mentioned, all of these policies are implemented under a fixed budget[32].

Policy I: Universal vouchers. Under this policy, every student receives a voucher to attend any private school in the market.[33] For simplicity, schools can set one price without restrictions, but cannot engage in price discrimination. The school always receives $max\{p_j, v_i\}$. The price that the student pays is $max\{0, p_j - v_i\}$. The voucher is equal for each student and fixed at $\bar{v}$, which corresponds to 75% of the mean spending per student in the public sector. Then,

$$v_i = \bar{v}$$

Policy II: Targeted vouchers. Under targeted vouchers, only low-income students receive a voucher to attend private schools. The assumptions as from policy I hold, but I change the structure of the voucher (only low-income students receive the voucher, and the amount is increased to maintain the level of the total budget):

$$v_i = \bar{v}/(1 - \pi^y) \times \mathbb{1}[x_i^y = 0]$$

Policy III: Merit scholarships (selective vouchers). Merit scholarships are equivalent to vouchers for talented students. To keep this conterfactual simple, I assume that talented students are the ones that come from families with high human capital. Again, I hold the asumptions from policy I, changing only the structure of the voucher:

$$v_i = \bar{v}/\pi^e \times \mathbb{1}[x_i^e = 1]$$

Policy IV: Tax deductions of tuition expenses. In a simplified setting, tax deductions of tuition expenses can be thought of as vouchers for higher-income families, which are more likely to pay higher taxes. In a real-world version of this policy, the amount of the subsidy the family receives will be a function of their taxable income. I will assume that all higher-income families get a fixed tax deduction when they send their kids to a private school, in the form of a voucher that can only be used to pay that school (so that the assumptions in policy I hold). The structure of the voucher will be:

$$v_i = \bar{v}/\pi^y \times \mathbb{1}[x_i^y = 1]$$

Policy V: Investment in public schools. In this policy, I simulate the opening of high-quality public schools. I make several assumptions to infer the costs of construction and operation for schools in the top 85 and 70 percent of the distribution of value-added.[34] I divide the budget in policy I and divide it by the calculated cost per school to get the potential number of new schools that could be constructed, depending on the targeted value-added percentile. Then, for each node

[32]The policies are implemented so that under certain assumptions, the budget falls in a [-0.1,0.1] orders of magnitude of the budget of the first counterfactual. Details are provided in Appendix.

[33]I ignore the program participation margin. This has proven to be relevant in other settings, see (Sánchez, 2017).

[34]As I do not have estimates on the fixed or marginal costs for these schools. I will use data on public school expenses and recent public school construction from the Ministry of Education (MINEDU, 2017). I also use the data on teachers' payments, as wages constitute one of the largest categories of spending and have been shown to be correlated with teacher quality in Peru (Bobba et al., 2019).

in the city, I calculate the demand that a hypothetical public school would have. Then, I rank the nodes and assign the new public schools to the ones with higher expected demand. Version A (B) of the policy targets public schools in the top 85 (70) of the distribution of value-added. In other words, this policy expands the choice set, by adding high quality public schools in several locations:

$$\Omega'_{it} = \Omega_{it} \cup j', \mu_j > \bar{\mu}_j$$

Policy VI: Land lease subsidies for private schools. In this policy, I simulate a subsidy for the entry of high-quality private schools in underserved neighborhoods. I have estimates for marginal costs and for the fixed costs of increasing quality for all private schools, but, as in the previous counterfactual, I do not have estimates for other fixed costs of operation or costs of entry. I get numbers from Innova (2019), the bond issuing framework document by one of the big chains of schools. This document provides approximate average land costs for schools and costs of construction[35]. I calculate the number of schools that can enter covering 75% of the land costs. I ignore the nodes that are above the 75th percentile of the distribution of income, to focus on low and middle-income families. For each of the eligible nodes in the city, I calculate the revenues that the chain gets if a new school is opened in each node. Then, I rank the nodes and assign the new private schools to the ones with higher expected revenues.[36] This policy expands the choice set by lowering fixed costs of high-quality private schools and promoting the entry in several locations.

$$\Delta \text{ in } F_{jt} \implies \Omega''_{it} = \Omega_{it} \cup j''$$

7.5.2 Simulation Results

The results for the simulations are shown in Figure 13 and Table 8. Table 8 reports the distribution of the changes in the value-added of the schools chosen by families of low and high income, the reduction in the value-added income gap, and the mean quality provided by the schools chosen by families in the market. The last two columns show the ex-ante cost of the policy (as a fraction of the original budget for policy I) and the ex-post cost of the policy, which can vary due to incomplete policy take-up.

Figure 15 displays the policies in an equity and efficiency framework. The x-axis shows the mean increase in quality, and the y-axis shows the reduction in the SES value-added gap. For the filled circles, the increase in the mean quality is divided by the share of the budget spent ex-post.

Overall, the results show that the policy that performs best in terms of equity is targeted vouchers, but it is quite inefficient. The policy that performs best in terms of efficiency is an entry subsidy

---

[35] See A9 for details

[36] Note that I do not take into account unobservable entry costs and cannibalization considerations from the point of view of the chain.

to high-quality private schools, but it has only a moderate performance in terms of equity. The rest of the policies range between these two extreme cases.

To understand the role of social interactions, Figure 14 shows the heterogeneity in the effects of targeted vouchers on the value-added chosen under a simulation that ignores social interactions (light line) and one that takes into account social interactions (dark line). The figures show that ignoring social interactions leads to an underestimate of the change in value-added for the low SES families and an overestimation of the change in the value-added for families of high SES. In addition, it ignores large distributional effects: when social interactions are relevant, price and value-added dispersion go up significantly. Two economic forces explain the increase in quality dispersion. The first is a *redistribution of good peers* that arises because of different responses to targeted vouchers by eligible (low-income) families with low and high human capital. Low-income families with high human capital are more sensitive to quality and less sensitive to distance than low human capital families. This makes them more likely to take-up the vouchers and leave to a higher quality school. This redistribution of peers hurts the low-income, low human capital families left behind, increasing the inequality within low-income families. As the number of good peers is fixed in the market, this is a zero-sum redistribution.

The second is a redistribution of *market power*. In the absence of the policy, schools have local market power to markdown quality because of geographic differentiation. Low-quality schools (which are likely to have high costs of increasing quality) can survive in the market because higher quality alternatives may be far and/or charge a higher price, making them less attractive to low-income families who are sensitive to distance and price. When targeted vouchers eliminate price differences for low-income families, the ones with high human capital are more likely to take-up the program and move to a higher quality private school, potentially located farther away. Low quality schools may choose to compete only for the low human capital families after the policy, lowering quality further. The logic is that pre voucher, the potential demand from high human capital low-income families provides incentives for the provision of a certain level of quality. When these families leave, and if it is too costly for the school to raise the quality to compete to keep them, the school may choose to serve the quality inelastic low human capital families providing a lower level of quality.[37] Then, the low-income, low-education are exposed to low value-added, both because they were left behind with worse peers and because the schools in which they stayed became worst.

On the other hand, families who take up the voucher and go to a better school (who are more

[37](Hastings et al., 2017) describe a similar equilibrium in which some firms choose to compete for price inelastic individuals and charge the highest possible fee. In their settings, fees have two parts and there are caps. Then, the existence of that equilibrium imposes a numerical problem to find the solution: these firms best respond by charging fees on the boundary, so solutions to the zeros of the gradient of the profit functions may not exist. In this setting, there are no restrictions on the values that quality can take, which would be the equivalent problem. Then, I do not have the same problem to find the equilibrium in this paper.

likely to be low-income families with high-education) directly benefit from their choice, and potentially benefit other high-education families indirectly by increasing the competitive pressure for the schools in the neighborhood, exacerbating the dispersion among voucher eligible families.

7.5.3 *The Equity Efficiency Trade-off and the Combined use of Policies*

The results in the previous section show that targeted vouchers perform well on average in terms of improving equity, but the effects are heterogeneous. The policy combines gains and losses, to which social interactions have a substantial contribution. In this section, I use the model to understand when are those losses more relevant. Then, I use that intuition to propose and evaluate an alternative allocation of the budget to improve the results in the previous section.

The discussion on introducing market incentives in education usually focuses on the impact evaluation of isolated policies. However, different policies may provide incentives that work in different dimensions, making them more effective in some contexts than in others. Moreover, individual policies may complement each other. I explore this idea by looking in more detail at the two extreme policies under the equity and efficiency framework: targeted vouchers and entry subsidies.

The mechanisms that shape the incentives in the model suggest that vouchers cannot increase efficiency if there are no potential high-quality schools to which the families can go, if there are not enough eligible families who get the voucher in order to put enough competitive pressure on the existing schools to increase their quality or, if it is too costly for the existing schools to increase quality. On the other hand, entry subsidies will not provide incentives for firms that have low costs of providing quality to enter in low-income neighborhoods if there is no willingness to pay for their higher quality services or if they cannot exploit the strategic social incentives. This logic suggests that the incentives in these policies could complement each other to improve the results.

Motivated by these ideas, I simulate counterfactuals in which I assign a share of the budget to vouchers and the rest to entry subsidies. In the first simulation I assign 2/3 of the budget to targeted vouchers (policy II) and 1/3 of the budget to land lease subsidies for private schools (policy VI). I call this counterfactual *Combined policies A*. In the second simulation, I reverse the previous assignment and assign 1/3 for policy II and 1/3 for policy VI. I call this counterfactual *Combined policies B*.

Figure 15 shows that the combined policies outperform the linear combination of them in an equity and efficiency framework. The voucher increases the willingness to pay in low-income neighborhoods and eliminates the strategic responses to social interactions by closing the gap in the willingness to pay between families of low and high SES. This promotes the entry of effective schools in low-income neighborhoods, which families can attend without paying a high price or traveling a long distance. This is one example of the potential that the model has to understand families' responses and school incentives, which can be used to explore how to design education policies to improve the distribution of outcomes.

8 Conclusion

In this paper, I study families' preferences for peers and the implications for the distribution of academic outcomes. Despite the fact that policies that promote school choice are widespread and have been implemented for a long time in education markets all around the world, there is disagreement about their effects. Moreover, the design of the policies often lacks sophistication and is not adapted to the local market conditions. Despite efforts the reduce inequality, inadequate policy design may lead to unintended consequences, like high levels of socioeconomic sorting, low levels of provided quality, and inequality in the access to schools with high value-added.

I develop an equilibrium model of school competition and student sorting under social interactions, which I exploit in a variety of ways. First, I use it to decompose the mechanisms that shape market power and the incentives of private schools under social interactions. Second, I use the model's predictions to look at the bias in the policy evaluation when social interactions are ignored. While the mean effects are not systematically different, the naive evaluation misses important distributional effects. Third, I use the model to analyze the effects of a set of counterfactual education policies in equilibrium. Public policies have heterogeneous effects, combining gains and losses. Based on the simulations from the model, I compare the performance of policies in terms of equity and efficiency. Finally, I show that the equity and efficiency frontier can be potentially expanded by combining policies that have complementary effects.

My analysis shows how a combination of a theoretical framework, a setting well fitted for the question, careful research design, and extensive microdata can inform the design of education policies in choice-oriented education markets, by identifying relevant economic forces at work, and by measuring its implications and their relationship to relevant features of education markets.

The final comment is on avenues for future research. I take the market structure of the education market as given. More work is needed to incorporate the dynamic aspect of the interaction between school incentives, social interactions, and public policies, all of which may have a fundamental role in shaping the structure of education markets. A promising future avenue for work is to extend the model to account for endogenous entry of schools. That framework can be used to answer other questions related to schools' incentives for entry and exit in certain neighborhoods under social interactions, the interaction between schools supply decisions and residential segregation, and the implication that these choices have for the inequality in the access to school quality and the academic achievement of low and high-income kids.

References

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak, "The elite illusion: Achievement effects at Boston and New York exam schools," *Econometrica*, 2014, *82* (1), 137–196.

_, Parag A Pathak, and Christopher R Walters, "Free to choose: Can school choice reduce student achievement?," *American Economic Journal: Applied Economics*, 2018, *10* (1), 175–206.

Abdulkadiroglu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters, "Do Parents Value School Effectiveness?," *American Economic Review*, 2019, *Forthcoming.*

Aguirre, Josefa, "Can Progressive Vouchers Help the Poor Benefit from School Choice? Evidence from the Chilean Voucher System," *Evidence from the Chilean Voucher System (July 10, 2017)*, 2017.

Allcott, Hunt, "Social norms and energy conservation," *Journal of public Economics*, 2011, *95* (9-10), 1082–1095.

Allende, Claudia, Francisco Gallego, and Christopher Neilson, "The Equilibrium Effects of Information Interventions," *Working Paper*, 2018.

Almond, Douglas and Janet Currie, "Killing me softly: The fetal origins hypothesis," *Journal of economic perspectives*, 2011, *25* (3), 153–72.

Altonji, Joseph G, Ching-I Huang, and Christopher R Taber, "Estimating the cream skimming effect of school choice," *Journal of Political Economy*, 2015, *123* (2), 266–324.

Angrist, Joshua D, "The perils of peer effects," *Labour Economics*, 2014, *30*, 98–108.

_, Peter D Hull, Parag A Pathak, and Christopher R Walters, "Leveraging lotteries for school value-added: Testing and estimation," *The Quarterly Journal of Economics*, 2017, *132* (2), 871–919.

Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer, "Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment," *American economic review*, 2002, *92* (5), 1535–1558.

_, Peter Hull, Parag Pathak, and Christopher Walters, "Interpreting tests of school VAM validity," *American Economic Review*, 2016, *106* (5), 388–92.

Barseghyan, Levon, Damon Clark, and Stephen Coate, "Peer Preferences, School Competition and the Effects of Public School Choice," *American Economic Journal: Economic Policy*, 2019.

Bayer, Patrick and Christopher Timmins, "Estimating equilibrium models of sorting across locations," *The Economic Journal*, 2007, *117* (518), 353–374.

— , **Fernando Ferreira, and Robert McMillan**, "A unified framework for measuring preferences for schools and neighborhoods," *Journal of political economy*, 2007, *115* (4), 588–638.

Behrman, Jere R and Mark R Rosenzweig, "Returns to birthweight," *Review of Economics and statistics*, 2004, *86* (2), 586–601.

Berry, Steven and Philip A Haile, "Identification in differentiated products markets using market level data," *Econometrica*, 2014, *82* (5), 1749–1797.

— **and** — , "Nonparametric identification of multinomial choice demand models with heterogeneous consumers," *Working Paper, Yale University*, 2016.

— **and Philip Haile**, "Identification in differentiated products markets," *Annual review of Economics*, 2016, *8*, 27–52.

— , **James Levinsohn, and Ariel Pakes**, "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 1995, pp. 841–890.

— , — , **and** — , "Differentiated products demand systems from a combination of micro and macro data: The new car market," *Journal of Political Economy*, 2004, *112* (1), 68–105.

Bettinger, Eric, Michael Kremer, Maurice Kugler, Carlos Medina, Christian Posso, and Juan E Saavedra, "Can Educational Voucher Programs Pay for Themselves?," *Unpublished manuscript, Harvard Univ., Cambridge, MA*, 2017.

Bharadwaj, Prashant, Katrine Vellesen Løken, and Christopher Neilson, "Early life health interventions and academic achievement," *American Economic Review*, 2013, *103* (5), 1862–91.

Björkegren, Daniel, "The adoption of network goods: evidence from the spread of mobile phones in Rwanda," *The Review of Economic Studies*, 2018, *86* (3), 1033–1060.

Black, Sandra E, "Do better schools matter? Parental valuation of elementary education," *The Quarterly Journal of Economics*, 1999, *114* (2), 577–599.

Bobba, Matteo, Gianmarco Leon, Christopher Neilson, Marco Nieddu, and Camila Alva, "Teacher Wages, Student Achievement, and the Recruitment of Talent in Rural Peru," *Working Paper*, 2019.

Bucarey, Alonso, "Who Pays for Free College? Crowding Out on Campus," *Job market paper*, 2018, *Massachusetts Institute of Technology, Cambridge, Massachusetts.*

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Carrell, Scott E, Bruce I Sacerdote, and James E West, "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica*, 2013, *81* (3), 855–882.

Chetty, Raj, John N Friedman, and Jonah E Rockoff, "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates," *American Economic Review*, 2014, *104* (9), 2593–2632.

Chubb, John E and Terry M Moe, "America's public schools: Choice is a panacea," *The Brookings Review*, 1990, *8* (3), 4–12.

Collard-Wexler, Allan and Jan De Loecker, "Production function estimation with measurement error in inputs," Technical Report, National Bureau of Economic Research 2016.

Conley, Timothy G and Christopher R Udry, "Learning about a new technology: Pineapple in Ghana," *American economic review*, 2010, *100* (1), 35–69.

Crawford, Gregory S., Oleksandr Shcherbakov, and Matthew Shum, "Quality Overprovision in Cable Television Markets," *American Economic Review*, March 2019, *109* (3), 956–95.

Currie, Janet, "Inequality at birth: Some causes and consequences," *American Economic Review*, 2011, *101* (3), 1–22.

Deming, David J, "Using school choice lotteries to test measures of school effectiveness," *American Economic Review*, 2014, *104* (5), 406–11.

Diamond, Rebecca, "The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000," *American Economic Review*, 2016, *106* (3), 479–524.

Dinerstein, Michael, Christopher Neilson, and Sebastián Otero, "School Construction and Competition: Evidence from the Dominican Republic," *Working Paper*, 2019.

_ , Troy Smith et al., "Quantifying the supply response of private schools to public policies," *Working Paper, Stanford University*, 2017.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya," *American Economic Review*, 2011, *101* (5), 1739–74.

Epple, Dennis and Richard E Romano, "Competition between private and public schools, vouchers, and peer-group effects," *American Economic Review*, 1998, pp. 33–62.

_ and _ , "Peer effects in education: A survey of the theory and evidence," in "Handbook of social economics," Vol. 1, Elsevier, 2011, pp. 1053–1163.

_ and Richard Romano, "Educational vouchers and cream skimming," *International Economic Review*, 2008, *49* (4), 1395–1435.

_ , Richard E Romano, and Miguel Urquiola, "School vouchers: A survey of the economics literature," *Journal of Economic Literature*, 2017, *55* (2), 441–92.

Fan, Ying, "Ownership consolidation and product characteristics: A study of the US daily newspaper market," *American Economic Review*, 2013, *103* (5), 1598–1628.

Ferreyra, Maria Marta and Grigory Kosenok, "Charter school entry and school choice: the case of Washington, DC," 2018.

Finkelstein, Amy, Nathaniel Hendren, and Mark Shepard, "Subsidizing health insurance for low-income adults: Evidence from Massachusetts," *American Economic Review*, 2019, *109* (4), 1530–67.

Friedman, Milton, "Capitalism and freedom," *with the assistance of Rose D. Friedman.*, 1962.

Gallego, Francisco, Oswaldo Molina, and Christopher Neilson, "Decidiendo para un futuro mejor: el efecto de la información sobre la deserción escolar y el uso del tiempo," 2016.

Gandhi, Amit and Jean-Francois Houde, "Measuring substitution patterns in differentiated products industries," *University of Wisconsin-Madison and Wharton School*, 2016.

Garlick, Robert, *Academic peer effects with different group assignment policies: Residential tracking versus random assignment*, The World Bank, 2014.

Gourieroux, Monfort, Christian Gourieroux, Alain Monfort, Director Alain Monfort et al., *Simulation-based econometric methods*, Oxford university press, 1996.

Guiteras, Raymond, James Levinsohn, and Ahmed Mushfiq Mobarak, "Demand estimation with strategic complementarities: Sanitation in bangladesh," 2019.

Hastings, Justine, Ali Hortaçsu, and Chad Syverson, "Sales force and competition in financial product markets: the case of Mexico's social security privatization," *Econometrica*, 2017, *85* (6), 1723–1761.

Hastings, Justine S and Jeffrey M Weinstein, "Information, school choice, and academic achievement: Evidence from two experiments," *The Quarterly journal of economics*, 2008, *123* (4), 1373–1414.

Hastings, Justine, Thomas J Kane, and Douglas O Staiger, "Heterogeneous preferences and the efficacy of public school choice," *NBER Working Paper*, 2009, *2145*, 1–46.

Hendren, Nathaniel, "Measuring ex-ante welfare in insurance markets," Technical Report, National Bureau of Economic Research 2018.

Ho, Katherine, "Insurer-provider networks in the medical care market," *American Economic Review*, 2009, *99* (1), 393–430.

Hoxby, Caroline Minter, "School choice and school productivity. Could school choice be a tide that lifts all boats?," in "The economics of school choice," University of Chicago Press, 2003, pp. 287–342.

Innova, "Social Bond Framework," *Colegios Peruanos S.A. – Innova Schools*, 2019.

Kane, Thomas J and Douglas O Staiger, "Estimating teacher impacts on student achievement: An experimental evaluation," Technical Report, National Bureau of Economic Research 2008.

Kapor, Adam, "Distributional effects of race-blind affirmative action," in "Yale University, working paper" 2016.

MacLeod, W Bentley and Miguel Urquiola, "Reputation and school competition," *American Economic Review*, 2015, *105* (11), 3471–88.

_ **and** _ , "Is Education Consumption or Investment? Implications for School Competition," *Annual Review of Economics*, 2019, *11.*

Manski, Charles F, "Educational choice (vouchers) and social mobility," *Economics of Education Review*, 1992, *11* (4), 351–369.

_ , "Identification of endogenous social effects: The reflection problem," *The review of economic studies*, 1993, *60* (3), 531–542.

Mas, Alexandre and Enrico Moretti, "Peers at work," *American Economic Review*, 2009, *99* (1), 112–45.

MINEDU, "Inversión en Infraestructura Educativa 2011-2016," *Ministerio de Educación*, 2017.

Mobarak, Ahmed Mushfiq and Mark R Rosenzweig, "Informal risk sharing, index insurance, and risk taking in developing countries," *American Economic Review*, 2013, *103* (3), 375–80.

Muralidharan, Karthik and Venkatesh Sundararaman, "Editor's Choice The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India," *The Quarterly Journal of Economics*, 2015, *130* (3), 1011–1066.

Nechyba, Thomas, "School finance, spatial income segregation, and the nature of communities," *Journal of Urban Economics*, 2003, *54* (1), 61–88.

Nechyba, Thomas J, "Mobility, targeting, and private-school vouchers," *American Economic Review*, 2000, *90* (1), 130–146.

Neilson, Christopher, "Targeted vouchers, competition among schools, and the academic achievement of poor students," *Job Market Paper*, 2013.

Nevo, Aviv, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 2001, *69* (2), 307–342.

Petrin, Amil, "Quantifying the benefits of new products: The case of the minivan," *Journal of political Economy*, 2002, *110* (4), 705–729.

Rao, Gautam, "Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools," *American Economic Review*, 2019, *109* (3), 774–809.

Reynaert, Mathias and Frank Verboven, "Improving the performance of random coefficients demand models: the role of optimal instruments," *Journal of Econometrics*, 2014, *179* (1), 83–98.

Rivkin, Steven G, Eric A Hanushek, and John F Kain, "Teachers, schools, and academic achievement," *Econometrica*, 2005, *73* (2), 417–458.

Rockoff, Jonah E, "The impact of individual teachers on student achievement: Evidence from panel data," *American Economic Review*, 2004, *94* (2), 247–252.

Rothstein, Jesse M, "Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions," *American Economic Review*, 2006, *96* (4), 1333–1350.

Sacerdote, Bruce, "Peer effects in education: How might they work, how big are they and how much do we know thus far?," in "Handbook of the Economics of Education," Vol. 3, Elsevier, 2011, pp. 249–277.

Sánchez, Cristián, "Targeted Vouchers and the Pricing Response of Schools: Evidence from Chile," 2017.

Scotiabank, Studies Department, "Weekly Financial Report," 2017.

Singleton, John D, "Incentives and the supply of effective charter schools," *American Economic Review*, 2019, *109* (7), 2568–2612.

Skrainka, Benjamin S and Kenneth L Judd, "High performance quadrature rules: How numerical integration affects a popular model of product differentiation," 2011.

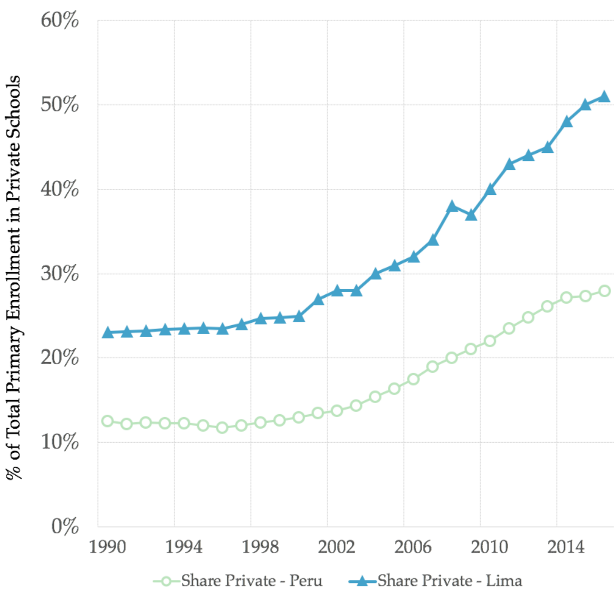
Varadhan, Ravi and Christophe Roland, "Simple and globally convergent methods for accelerating the convergence of any EM algorithm," *Scandinavian Journal of Statistics*, 2008, *35* (2), 335–353.

Walters, Christopher R, "The demand for effective charter schools," *Journal of Political Economy*, 2018, *126* (6), 2179–2223.

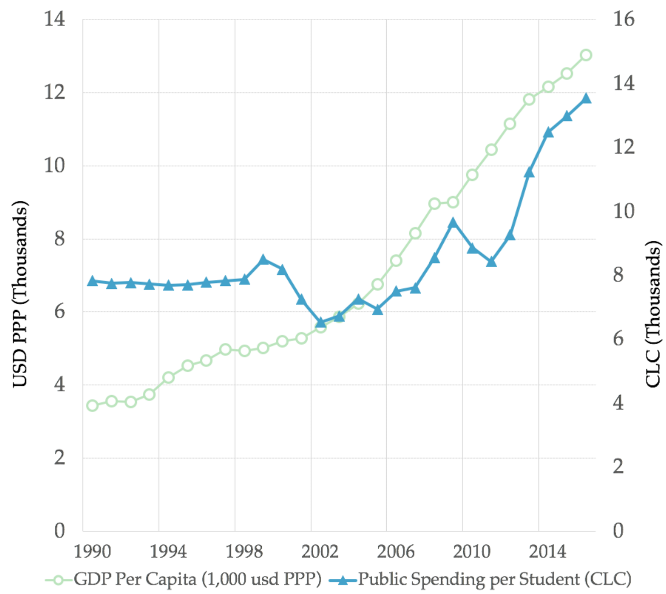
Wollmann, Thomas G, "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles," *American Economic Review*, 2018, *108* (6), 1364–1406.

Zárate, Román Andrés, “Social and Cognitive Peer Effects: Experimental Evidence from Selective High Schools in Peru,” 2019.

Figure 3: Private Education in Peru



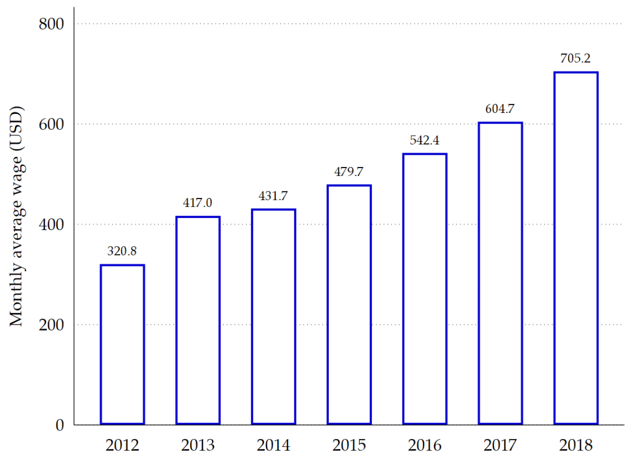
(a) Share of Schools that are Private



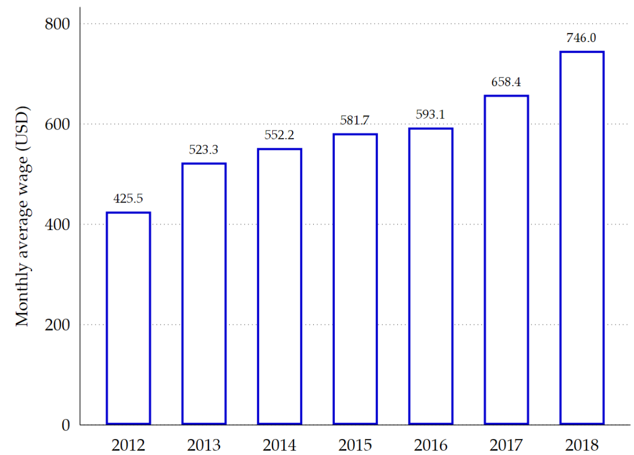
(b) GDP per Capita and Public Spending

Notes: The data in the left panel comes from the Ministry of Education of Peru and represents first-grade enrollment. The data in the second panel comes from the World Bank.

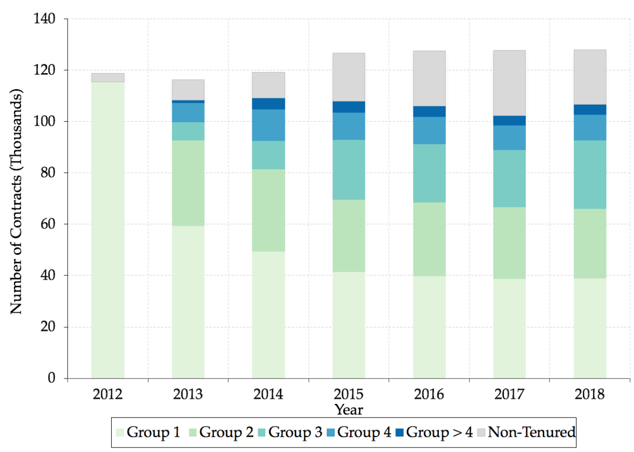
Figure 4: Evolution of Wages and Contract Types



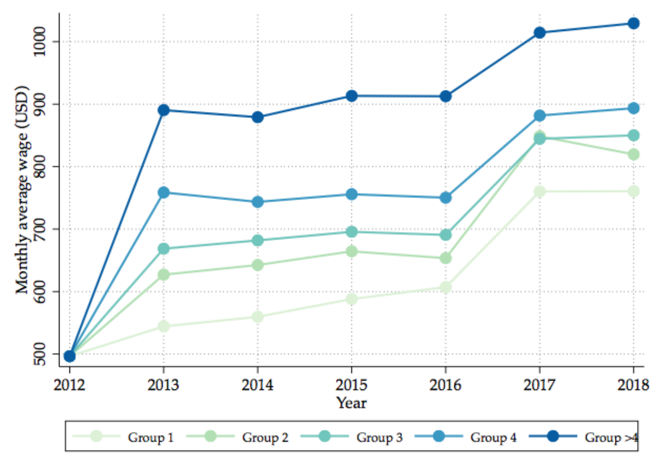
(a) Mean wages for non-tenured teachers



(b) Mean wages for tenured teachers



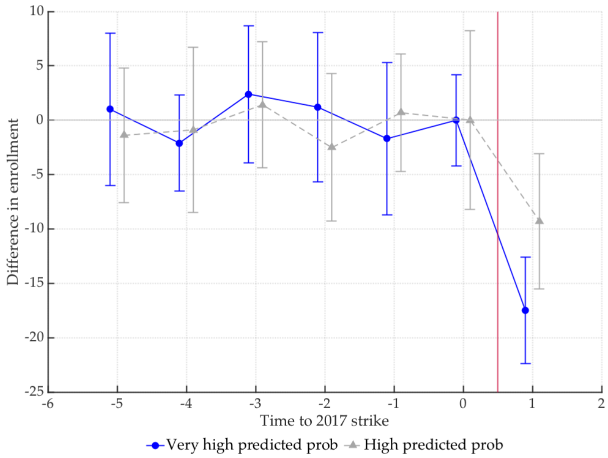
(c) Type of contracts by year



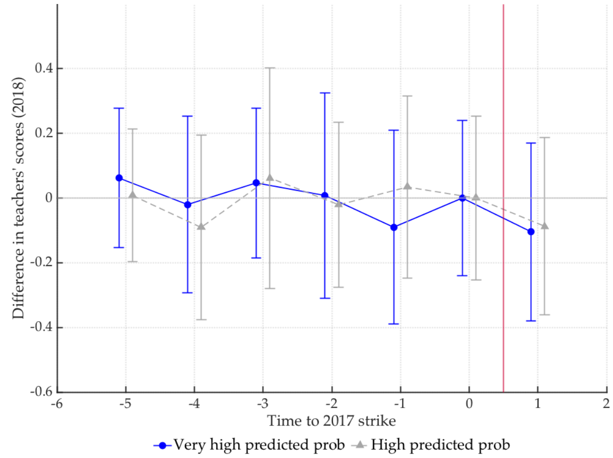
(d) Mean wages by performance group

Notes: Panels (a) and (b) in this figure display the evolution of the mean wages for teachers under tenured and non-tenured contracts. Panel (c) shows the evolution of the number of contracts by contract type. Groups 1 to 4 and greater than 4 represent tenure contracts. Non-tenured contracts are aggregated into one group. Panel (d) shows the evolution of mean wages by performance group type for tenured contracts.

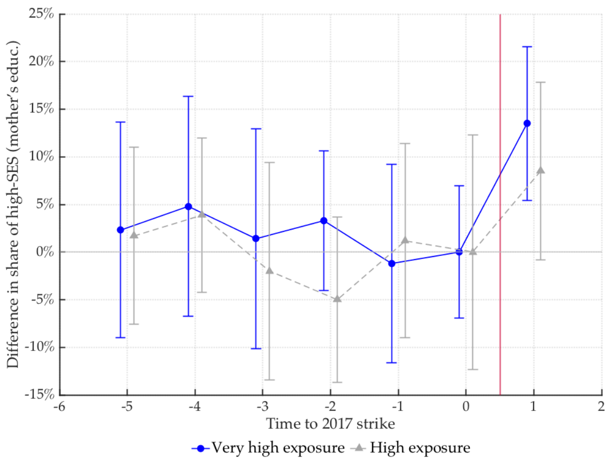
Figure 5: Difference for High Predicted/Exposed to Strike vs. Other Schools, by Intensity



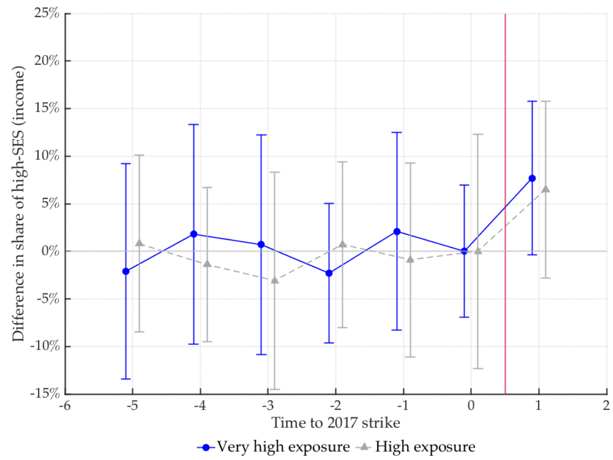
(a) Difference in Enrollment



(b) Difference in Teachers' Scores



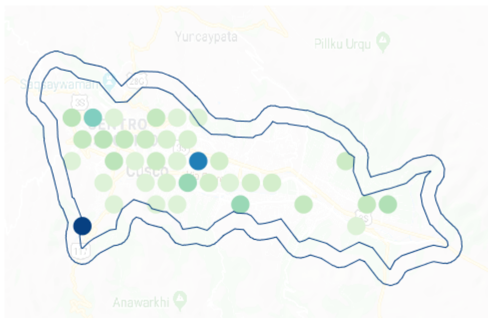
(c) Difference in z^e_{jt} (Share of High SES - Educ.)



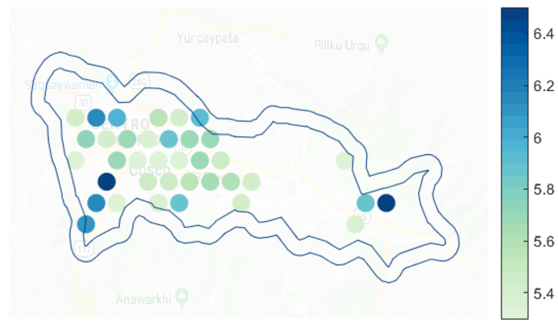
(d) Difference in z^y_{jt} (Share of High SES - Income)

Notes: These figures display the estimates and 95% confidence intervals of the coefficients β_s in the regression $Y_{jkt} = \sum_{k=-5}^{1} \beta_s D_i \tau_s + \gamma_1 X_{jt} + \gamma_2 Z_{jt} + \theta_k + \tau_s + \epsilon_{jkt}$, where Y_{ijt} is the outcome for school j in district k in year t, D_i equals one for schools that receive the treatment, τ_t are year fixed effects, X_{jt} it is a vector of school controls, Z_{jt} are district characteristics, including the total numer of kids enrolled and number of private schools, and θ_k are district fixed effects. The parameter β_0 is normalized to equal zero. Standard errors are clustered at the district level. For panel (a) the outcome Y_{ijt} is total enrollment in the school, for panel (b) the outcome Y_{ijt} is the mean of the standardized qualification for the teachers in the school (only public schools), for panel (c) the outcome Y_{ijt} is the share of mothers' in the school that have high education, for panel (c) the outcome Y_{ijt} is the share of families that have high income. In panels (a) and (b) the treatment variable D_i is an indicator for schools with either very high predicted probability of strike (top 15%, in solid blue) or high predicted probability of strike (top 25% - 15%, in dashed gray). In panels (a) and (b) the treatment variable D_i is an indicator for private schools that are one of the three closest schools to a school with either very high predicted probability of strike (schools with very high exposure, in solid blue) or high predicted probability of strike (schools with high exposure, in dashed gray).

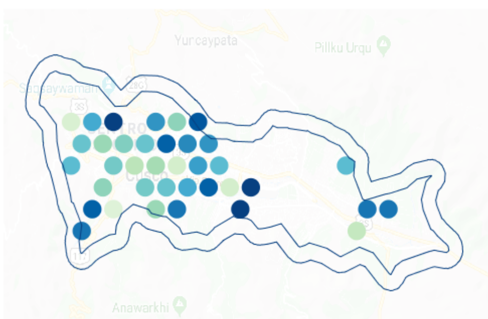
Figure 6: Wage Index by Node and Year (Cuzco)



(a) 2013

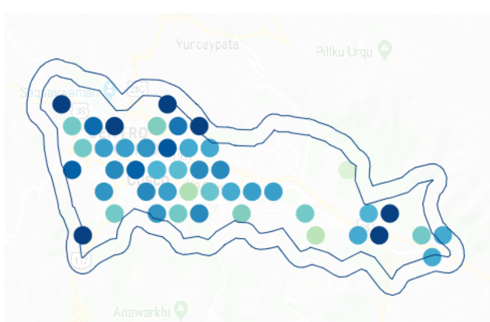


(b) 2014



(c) 2015

(d) 2016

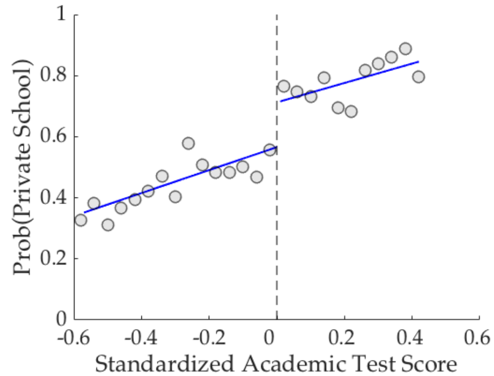


(e) 2017

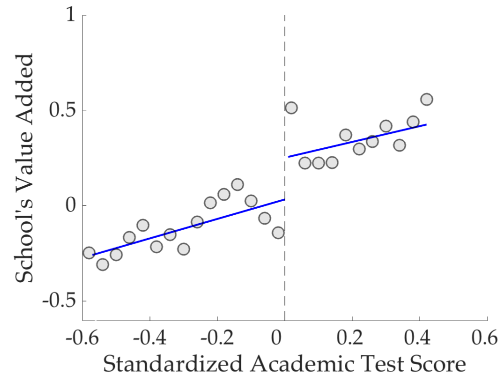
(f) 2018

Notes: This figure shows the distance weighted public school teacher wage index for the nodes of the city of Cuzco. The wage index is constructed by associating each public teacher contract and the wage paid to a specific latitude and longitude, and the, for each node, averaging the hourly payment in all the contracts in 1km, using linear distance weights.

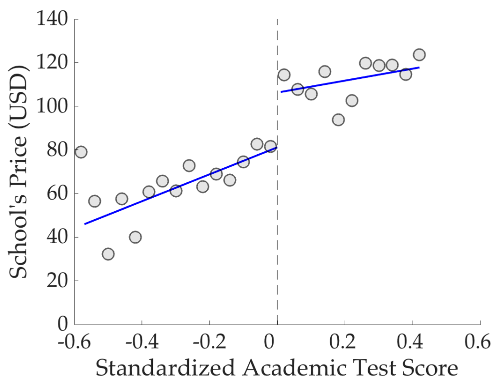
Figure 7: Visual Evaluation of The Threshold Crossing Effect (First Cutoff)



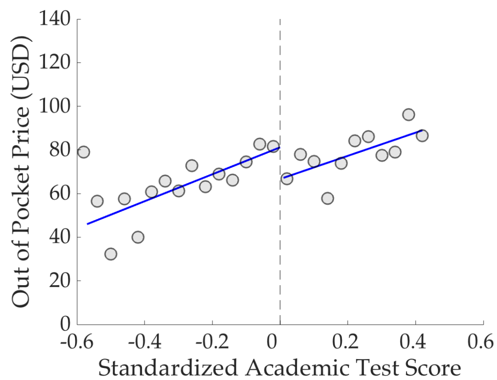
(a) Private School



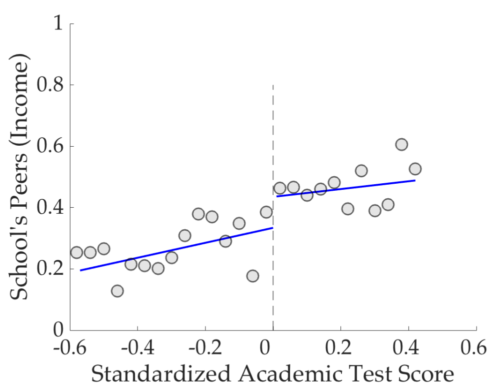
(b) Value-added



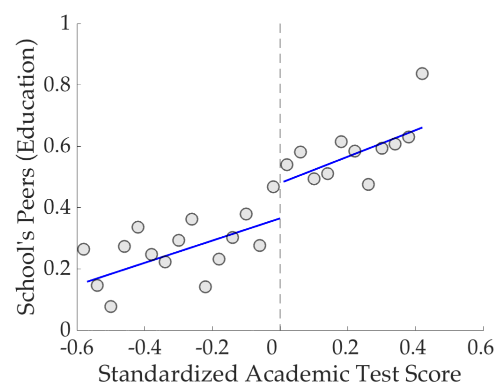
(c) School Price



(d) OOP Price



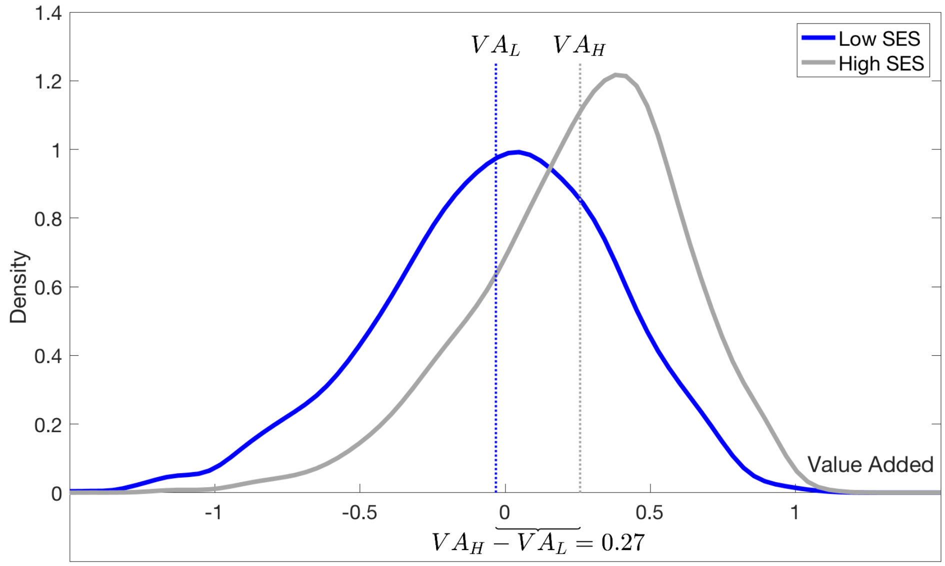
(e) Peers (Income)



(f) Peers (Education)

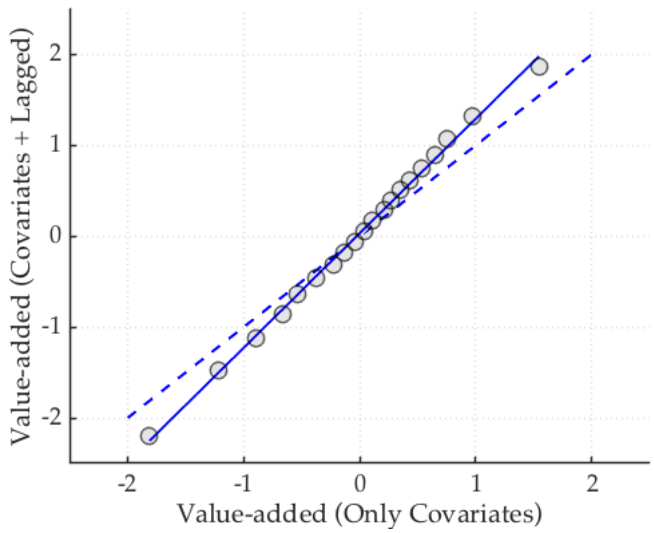
Notes: Panels (a) to (f) show the threshold crossing effect for different outcomes. White circles present the average in the outcome variable for individuals in equally spaced disjoint bins. Blue lines depict a first-order polynomial fit, separate for units at each side of the threshold.

Figure 8: Distribution of VA by SES



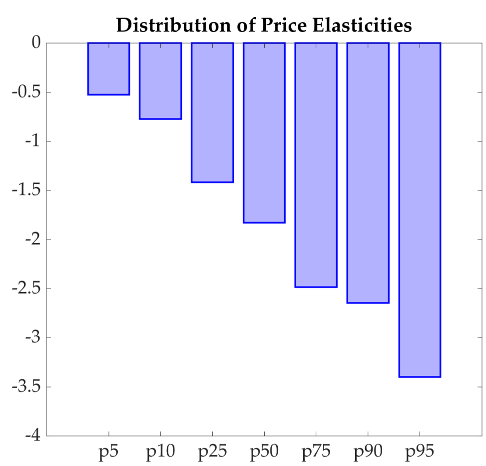
Notes: This figure shows the distribution of school value added in 2018 conditional on the students family income. The population of students is divided into two groups, those with low and high income. The term $VA_H - VA_L$ shows the difference of the mean value-added for each group.

Figure 9: Estimated Value-added

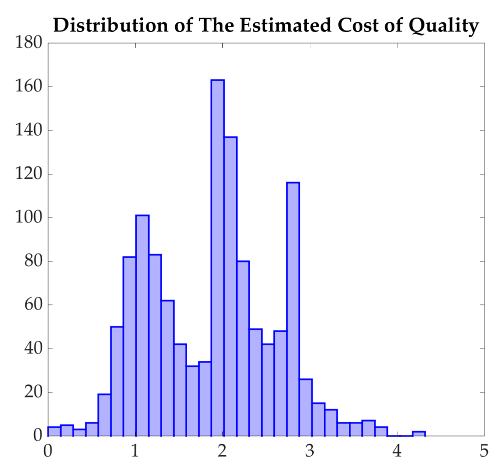


Notes: This figure shows a binned scatter plot for two measures of value-added using data for 2008. The measure in the x-axis is VA-I, which uses only non-academic student characteristics as explanatory variables. The measure in the x-axis is VA-II, which also includes lagged test scores. The regressions with which these measures are obtained are presented in Table 2

Figure 10: Model Results



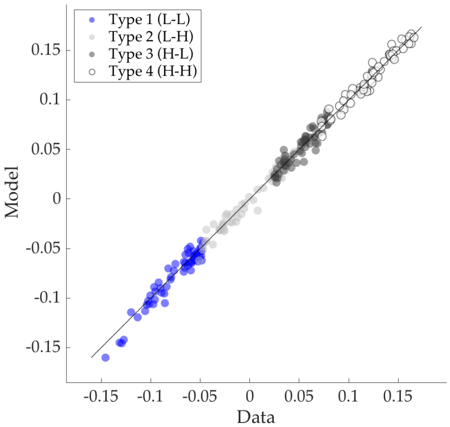
(a) Price Elasticities



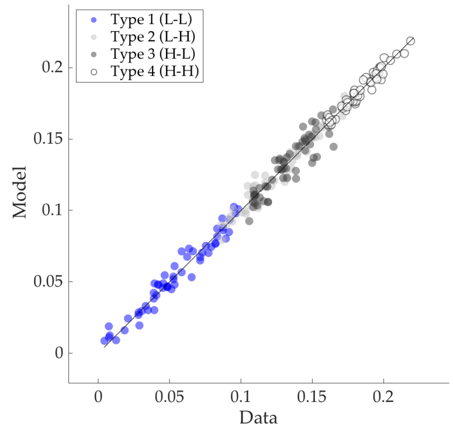
(b) Fixed Cost of Quality

Notes: This figure shows additional model results. Panel (a) shows the distribution of the price elasticities implied by the demand parameters. Panel (b) shows the distribution of the estimated fixed cost of increasing quality, $\bar{\psi}_j$, for private schools.

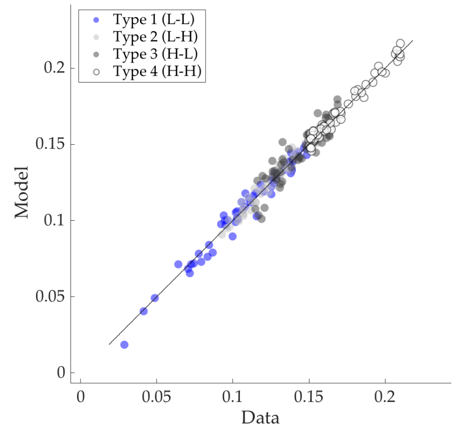
Figure 11: Model Fit



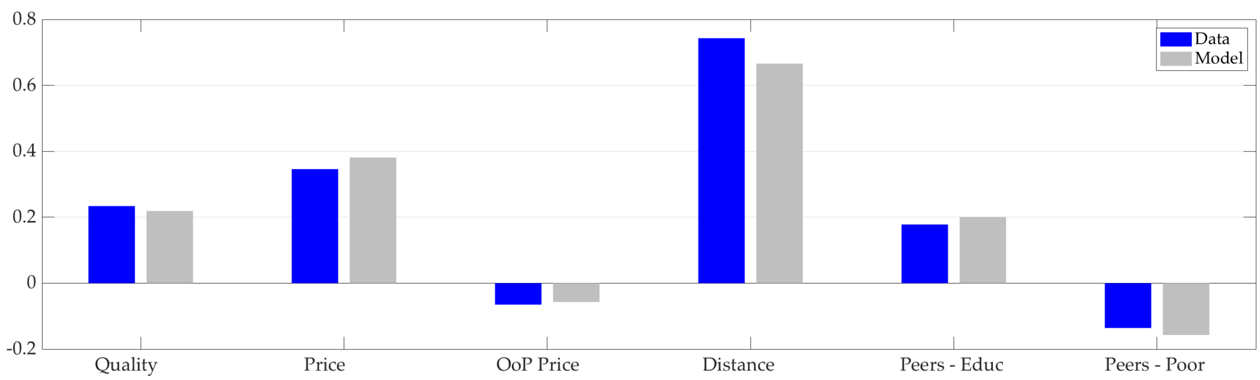
(a) Micro Moments - VA



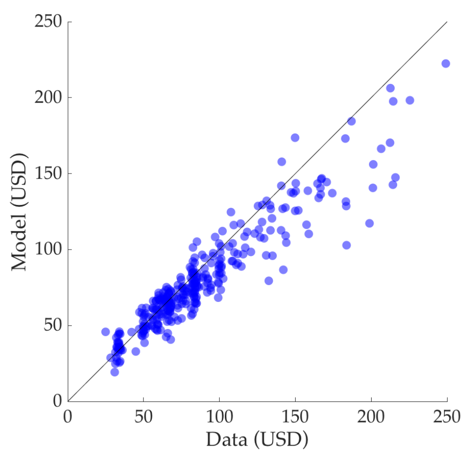
(b) Micro Moments - Price (No Zeros)



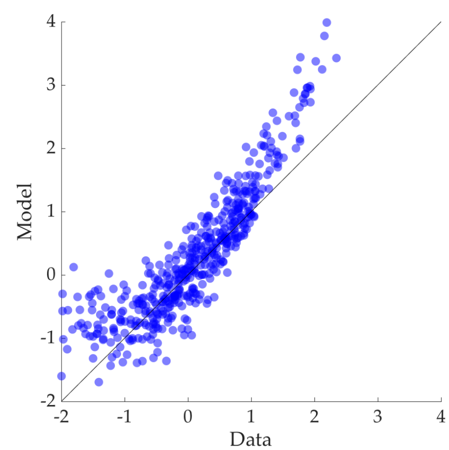
(c) Micro Moments - Distance



(d) Regression Discontinuity Moments



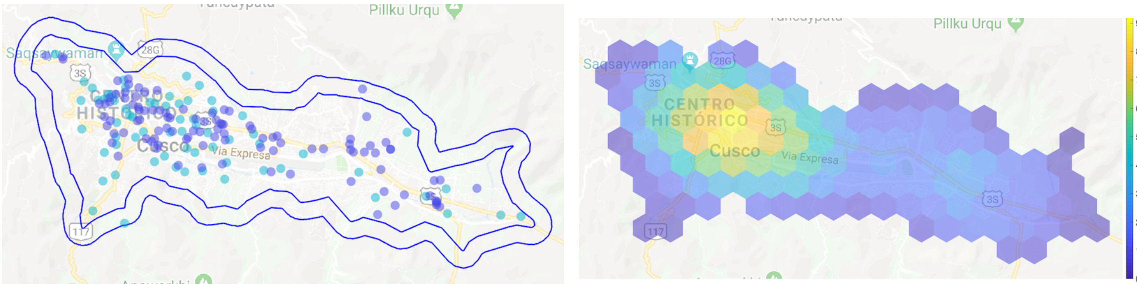
(e) Supply - Price (Market: Cuzco)



(f) Supply - Quality (Market: Cuzco)

Notes: This figure displays results for model fit. Predicted values are calculated as detailed in Section 6.4, using estimates from the model described in section 6.2. Panel (a)-(c) display observed and predicted values of the micro-moments for value-added, price, and distance for each market, averaged across years. Panel (d) displays observed and predicted regression discontinuity moments. Panel (e) and (f) display the observed and predicted values for the prices and qualities for one market (Cuzco)

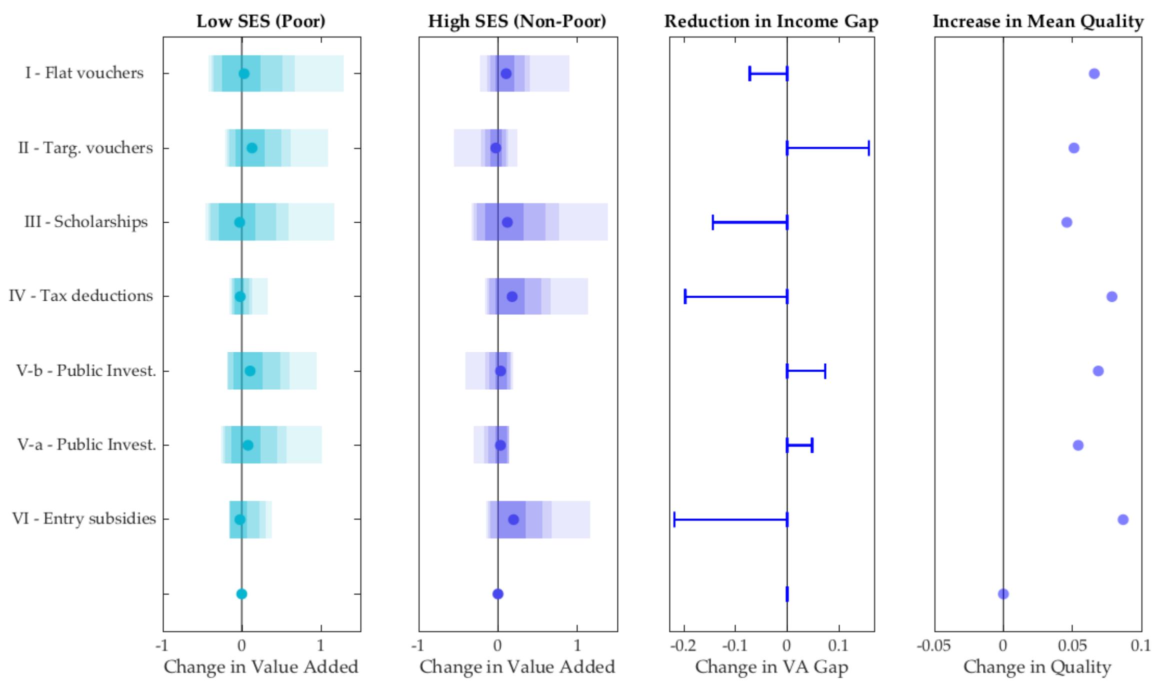
Figure 12: Market for Counterfactual Policy Simulation (Cuzco)



(a) Distribution of Private and Public Schools **(b)** Distribution of Family Income

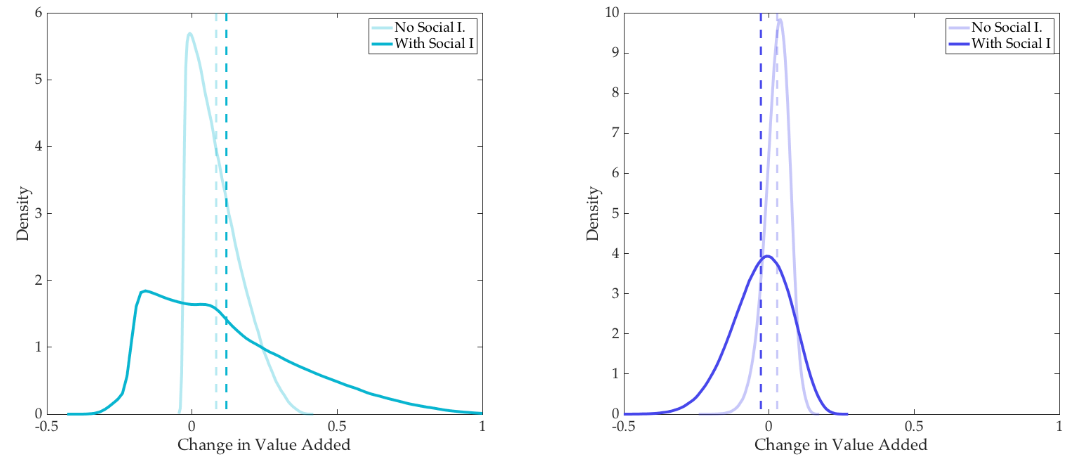
Notes: These figures describe the urban education market of Cuzco, which will be used to compute the counterfactual simulations. Panel (a) shows the distribution of public and private schools in the market. Panel (b) displays the geographic distribution of family income in the market.

Figure 13: Counterfactuals



Notes: This figure displays heterogeneity in the effects of the 6 counterfactual policies described in Table 1. The first two panels display the mean and distribution of individual-level changes in the value-added of the school attended for both low SES (first panel) and high SES kids (second panel). The mean is represented by a dark dot and the distribution by a bar with different shades. The dark center contains percentiles 25 to 75 (50% of the students). The slightly lighter area contains percentiles 10 to 25 and 75 to 90 (30% of the students). The next area contains percentiles 5 to 10 and 90 to 95 (10% of the students). The lightest shade represents changes in value-added bellow the percentile 5 and above percentile 95. The distribution displayed in this figure is trimmed bellow percentile 0.05 and above percentile 99.5. The third panel shows the reduction in the gap in the value to which families from low and high SES attend, calculated by subtracting the mean change in value-added for high SES families from the mean for low SES families. Finally, the fourth panel shows the average change in quality (including both low and high SES families) for each policy.

Figure 14: Heterogeneity in Effects of Targeted Vouchers on Value-added

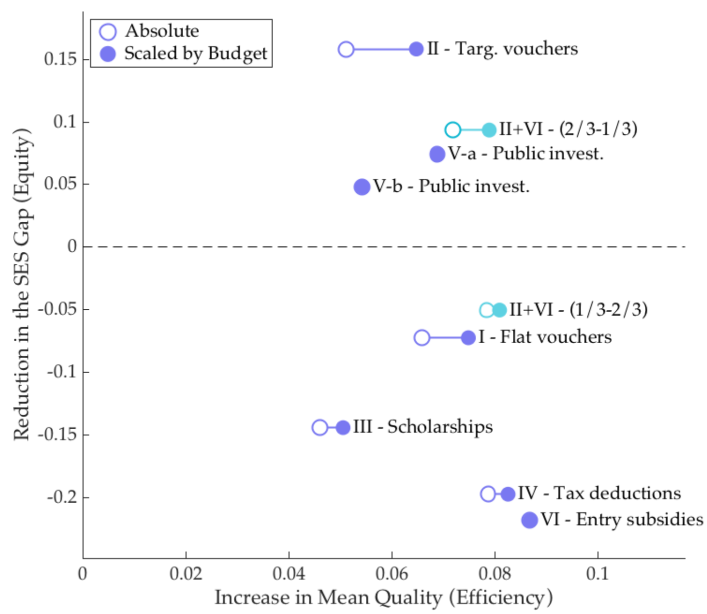


(a) Low SES

(b) High SES

Notes: These figures display the effects of targeted vouchers across students by SES. Panel (a) displays the results for high SES; panel B displays the results for low SES. Both figures compare outcomes under two simulations. In the first one (light line), I ignore social interactions to compute the simulation; in the second one (dark line), I allow for social interactions. The mean effects are shown in dashed lines.

Figure 15: Equity-Efficiency Trade-off for Counterfactual Policies



Notes: This figure displays the effects of each of the counterfactual policies in an equity and efficiency framework. The measure for equity is in the y-axis, and corresponds to the reduction in the gap in the value to which families from low and high SES attend presented in the third panel in Figure 13. The measure for efficiency is in the y-axis, and corresponds to the average change in quality presented in the third panel in Figure 13.

Table 2: Estimation of School Value Added

	VA-I (Full Period)		VA-I (2018 only)		VA-II (2018 only)	
	Coef.	StdErr	Coef.	StdErr	Coef.	StdErr
Lagged Score Language	-	-	-	-	0.344	(0.001)
Lagged Score Math	-	-	-	-	0.263	(0.001)
Native Language - Non Spanish	-0.156	(0.009)	-0.159	(0.007)	-0.050	(0.006)
Male	-0.031	(0.003)	-0.033	(0.003)	-0.012	(0.002)
Poor	-0.087	(0.003)	-0.086	(0.003)	-0.045	(0.003)
Mother's Educ - Prim. Comp.	0.006	(0.012)	0.004	(0.011)	0.004	(0.010)
Mother's Educ - Sec. Inc.	0.038	(0.009)	0.037	(0.012)	0.018	(0.010)
Mother's Educ - Sec. Comp.	0.138	(0.010)	0.133	(0.011)	0.075	(0.010)
Mother's Educ - HE 2Y. Inc.	0.214	(0.013)	0.210	(0.013)	0.121	(0.011)
Mother's Educ - HE 2Y Comp.	0.261	(0.012)	0.261	(0.012)	0.143	(0.010)
Mother's Educ - HE 4Y Inc.	0.228	(0.012)	0.225	(0.014)	0.142	(0.012)
Mother's Educ - HE 4Y Comp.	0.295	(0.010)	0.292	(0.012)	0.160	(0.010)
Mother's Educ - Grad	0.353	(0.021)	0.350	(0.025)	0.171	(0.021)
Father's Educ - Prim. Comp.	-0.020	(0.007)	-0.024	(0.006)	-0.021	(0.005)
Father's Educ - Sec. Inc.	0.007	(0.008)	0.006	(0.006)	-0.012	(0.005)
Father's Educ - Sec. Comp.	0.076	(0.006)	0.075	(0.005)	0.030	(0.004)
Father's Educ - HE 2Y. Inc.	0.170	(0.009)	0.166	(0.008)	0.082	(0.006)
Father's Educ - HE 2Y Comp.	0.185	(0.008)	0.185	(0.006)	0.090	(0.005)
Father's Educ - HE 4Y Inc.	0.195	(0.008)	0.191	(0.010)	0.100	(0.008)
Father's Educ - HE 4Y Comp.	0.226	(0.008)	0.221	(0.007)	0.113	(0.005)
Father's Educ - Grad	0.295	(0.016)	0.294	(0.019)	0.157	(0.016)
Married	0.004	(0.004)	0.005	(0.006)	0.005	(0.005)
Birth Weight	0.002	(0.000)	0.003	(0.002)	0.004	(0.005)
$(\text{Birth Weight})^2$	-0.000	(0.000)	-0.001	(0.001)	-0.001	(0.002)
Birth Gestation	0.008	(0.001)	0.008	(0.001)	0.006	(0.004)
$(\text{Birth Gestation})^2$	-0.000	(0.000)	-0.001	(0.002)	-0.002	(0.008)
Single Birth	0.031	(0.016)	0.033	(0.017)	0.022	(0.016)
Number of siblings	-0.041	(0.011)	-0.041	(0.013)	-0.022	(0.013)
Twin	-0.099	(0.021)	-0.097	(0.022)	-0.064	(0.018)
Birth Father's Age	0.012	(0.001)	0.012	(0.001)	0.009	(0.005)
Birth Mother's Age	0.007	(0.001)	0.009	(0.002)	0.009	(0.004)
R^2	0.394		0.386		0.607	
N obs	2,445,152		516,152		479,576	

Note: This table presents regression results for estimates of test scores on a vector of individual student-level characteristics. School quality is estimated as the school and year fixed effect (not presented in this table). VA-I uses only non-academic student characteristics as explanatory variables, VA-II also includes lagged test scores.

Table 3: First Stage for Endogenous School Characteristics

	Endogenous School Characteristics									
	(1)		(2)		(3)		(4)		(5)	
	Peers (Share of High SES)				Price		Quality		Shares	
	Income		Education							
Exogenous School Characteristics	Coef.	StdErr	Coef.	StdErr	Coef.	StdErr	Coef.	StdErr	Coef.	StdErr
Private	0.210	(0.022)	0.311	(0.055)	0.277	(0.029)	0.106	(0.047)	-0.106	(0.014)
For-profit	0.058	(0.016)	0.094	(0.023)	0.050	(0.009)	-0.011	(0.004)	0.018	(0.005)
Religious	0.021	(0.007)	0.109	(0.030)	-0.027	(0.013)	0.039	(0.016)	0.008	(0.003)
HS Integrated	0.008	(0.002)	0.011	(0.007)	0.002	(0.001)	0.024	(0.006)	0.017	(0.004)
Morning Shift	0.094	(0.045)	0.011	(0.001)	0.019	(0.005)	0.017	(0.004)	0.033	(0.014)
Instruments I: Peers' Shifters										
Strike Exp Idx I	0.101	(0.032)	0.075	(0.029)	0.105	(0.045)	-0.006	(0.003)	0.028	(0.011)
Strike Exp Idx II	-0.109	(0.038)	-0.093	(0.037)	0.000	(0.000)	0.012	(0.007)	0.012	(0.005)
Strike Exp Idx I $\times$ Strike Exp Idx II	0.021	(0.009)	-0.023	(0.012)	0.000	(0.000)	0.010	(0.005)	0.000	(0.000)
Instruments II: Cost and Quality Shifters										
Public School Teacher Wage Index	0.019	(0.005)	0.041	(0.009)	0.112	(0.016)	-0.084	(0.022)	0.039	(0.019)
Public School Tenure Vacancies	0.050	(0.014)	0.036	(0.009)	0.081	(0.010)	0.033	(0.010)	0.025	(0.012)
Wage Index $\times$ Vacancies	0.003	(0.002)	0.016	(0.002)	0.061	(0.018)	-0.009	(0.003)	0.010	(0.007)
Stock of Graduates in Education	0.012	(0.003)	0.071	(0.055)	0.000	(0.000)	0.009	(0.008)	-0.078	(0.048)
Lagged Local Teachers' Test Scores	0.026	(0.006)	0.020	(0.008)	-0.019	(0.010)	0.052	(0.039)	0.032	(0.015)
New Graduates in Education (Market)	-0.001	(0.004)	0.020	(0.012)	0.005	(0.002)	0.004	(0.002)	0.016	(0.010)
Instruments III: Local Demographics										
Share High SES (Income)	0.282	(0.093)	0.017	(0.003)	0.092	(0.023)	0.134	(0.030)	0.013	(0.007)
Share High SES (Educ)	0.123	(0.042)	0.218	(0.055)	0.030	(0.010)	0.102	(0.032)	0.028	(0.021)
Instruments IV: Market Structure										
Number of Schools	-0.053	(0.018)	-0.063	(0.028)	0.037	(0.229)	0.014	(0.034)	0.004	(0.002)
Private (Mean Others)	-0.015	(0.007)	-0.027	(0.011)	0.003	(0.001)	0.007	(0.006)	0.008	(0.006)
For-profit (Mean Others)	0.088	(0.027)	0.012	(0.004)	0.005	(0.002)	0.011	(0.005)	0.023	(0.013)
Religious (Mean Others)	0.014	(0.006)	-0.036	(0.024)	0.004	(0.002)	0.013	(0.009)	0.012	(0.006)
HS Integrated (Mean Others)	-0.008	(0.005)	-0.010	(0.005)	0.009	(0.005)	0.037	(0.018)	0.023	(0.015)
Morning Shift (Mean Others)	0.023	(0.013)	0.024	(0.014)	0.000	(0.000)	0.010	(0.005)	0.009	(0.005)
Year FE	$\times$		$\times$		$\times$		$\times$		$\times$	
School District FE	$\times$		$\times$		$\times$		$\times$		$\times$	
R^2										
F-statistic of the Excluded Instruments	23.645		19.887		28.112		15.294		39.767	
N obs	23,543		23,543		23,543		23,543		23,543	

Note: This tables presents the first stage for each of the endogenous school characteristics on exogenous school characteristics and the four sets of excluded instruments. Some interactions are omited from this table. Prices are in dolars and divided by 1,000. Market shares are multiplied by 1,000. The F-statistic is reporte on the exluded instruments only.

Table 4: IV Estimates for Peer Effects

	VA-I (Full Period)		VA-I (Full Period)		VA-I (2018 only)	
	Coef.	StdErr	Coef.	StdErr	Coef.	StdErr
Share of High SES (Income)	0.078	(0.030)	0.081	(0.032)	0.085	(0.038)
Share of High SES (Education)	0.107	(0.040)	0.117	(0.053)	0.112	(0.048)
Share of High SES (Income $\times$ Education)	0.023	(0.009)	0.029	(0.010)	0.031	(0.014)
Year FE	$\times$					
School District FE	$\times$		$\times$		$\times$	
R^2	0.137		0.923		0.735	
N obs	23,543		5,734		5,734	

Note:

Table 5: Effect of the Private Voucher on Choices

	Private	Value-added	Price		Peers	
			Sticker	OOP	Income	Educ
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect of Being Over the Academic Cutoff (Reduced Form)						
Score $\geq$ Cutoff	0.135	0.212	29.591	-13.213	0.109	0.124
	(0.063)	(0.099)	(13.866)	(-6.190)	(0.051)	(0.053)
Mean	0.568	0.041	80.61	80.61	0.326	0.371
Obs left	2981	3034	2995	3006	3055	2857
Obs right	2150	2105	2085	2106	2155	2083
Panel B: Effect of being Elegible (IV)						
Scholarship	0.409	0.642	89.673	-40.032	0.330	0.375
	(0.189)	(0.288)	(41.478)	(-18.395)	(0.154)	(0.155)
Mean	0.568	0.041	80.61	80.61	0.326	0.371
Obs left	3032	3012	2936	3007	2916	2920
Obs right	2148	2169	2091	2183	2197	2120

Notes: Estimates are computed using optimal bandwidths and robust confidence intervals proposed by Calonico et al. (2014).

Table 6: Demand Model Estimates

Panel A: Linear parameters	Parameter	Coeff.	Std Err.
Quality	$\bar{\beta}^q$	0.734†	(0.341)
Price	$\bar{\beta}^p$	-2.397†	(1.114)
Peers (Income)	$\bar{\beta}^{zy}$	0.448†	(0.207)
Peers (Educ.)	$\bar{\beta}^{ze}$	0.603†	(0.274)
Distance	$\bar{\beta}^d$	-1.288†	(0.651)
Panel B: Observed Heterogeneity			
Quality			
Family Income	β^q_y	0.657†	(0.331)
Mothers' Educ	β^q_e	0.464†	(0.156)
Price			
Family Income	α_y	0.567†	(0.275)
Mothers' Educ	α_e	1.321†	(0.652)
Peers (Income)			
Family Income	β^{zy}_y	0.641†	(0.326)
Mothers' Educ	β^{zy}_e	0.706†	(0.275)
Peers (Educ.)			
Family Income	β^{ze}_y	0.389†	(0.176)
Mothers' Educ	β^{ze}_e	0.310 †	(0.107)
Distance			
Family Income	β^d_y	0.234†	(0.109)
Mothers' Educ	β^d_e	0.554†	(0.282)
Panel C: Unobserved Heterogeneity			
Means			
Price	m^p_v	0.013†	(0.006)
Peers (Educ.)	m^{ze}_v	0.007†	(0.003)
Peers (Poor)	m^{zy}_v	0.019†	(0.009)
Variance - Covariance			
Price	σ_p	0.132†	(0.064)
Peers (Income)	σ_{zy}	0.179	(0.098)
Peers (Educ.)	σ_{ze}	0.211†	(0.101)
Peers (Corr.)	$\rho_{zy,ze}$	0.019	(0.010)
Scholarship Network	σ_s	0.399	(0.237)
Panel D: Application Costs			
Fixed - Constant	$\bar{\eta}$	2.424†	(0.064)
Fixed - Educ.	η_e	-1.345†	(0.101)
Distance - Constant	$\bar{\eta}^d$	1.9290	(0.010)
Distance - Educ.	η^d_e	-0.432	(0.098)

Notes: † indicates significance at 0.05 confidence level. The table presents results from the estimation using thirty-nine markets. Random coefficients have a lognormal distribution, and the simulation is done using the sparse grid rule of Heiss and Winschel (2008) for dimension 4 with the Gauss-Hermite rule.

Table 7: Supply Model Estimates

Panel A: Observable Marginal Cost Shifters	Parameter	Coeff.	Std Err.
Quality	γ_μ	$0.214^\dagger$	(0.197)
Other Characteristics	γ_l		
For-profit		$0.072^\dagger$	(0.035)
Religious		$-0.026^\dagger$	(0.011)
HS Integrated		$0.089^\dagger$	(0.043)
Morning Shift		0.231	(0.120)
Chain A (Innova)		$-0.233^\dagger$	(0.097)
Chain B (Saco Oliveros)		$-0.189^\dagger$	(0.104)
Chain C (Pamer)		$-0.173^\dagger$	(0.074)
Chain D (Trilce)		-0.071	(0.039)
Chain Other		$-0.031^\dagger$	(0.015)
Constant (Mean Market FE)		0.472	-
Panel B: Quality Component of the Fixed Cost			
Constant (Mean Firm FE)	ψ_j	1.917	-

Notes: † indicates significance at 0.05 confidence level. Market FE and Mean Firm Effects are shown to give a sense of the magnitude

Table 8: Counterfactuals

N	Policy	VA-Low SES			VA-High SES			Value Added		Cost	
		p25	p50	p75	p25	p50	p75	Δ Gap	Mean	Ex-ante	Ex-post
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
I	Flat Vouchers	-0.24	-0.04	0.24	-0.03	0.08	0.21	-0.072	0.066	1.000	0.880
II	Targeted Vouchers	-0.07	0.08	0.29	-0.09	-0.02	0.04	0.158	0.051	1.000	0.790
III	Scholarships	-0.28	-0.09	0.17	-0.15	0.05	0.33	-0.144	0.046	1.000	0.910
IV	Tuition tax deduct.	-0.09	-0.03	0.03	-0.03	0.13	0.34	-0.197	0.079	1.000	0.953
V-a	Public school invest. A	-0.10	0.04	0.26	-0.03	0.05	0.11	0.074	0.069	1.000	1.000
V-b	Public school invest. B	-0.13	0.03	0.24	-0.03	0.06	0.11	0.048	0.054	1.000	1.000
VI	Land lease subsidies	-0.14	-0.09	0.06	-0.01	0.14	0.35	-0.218	0.087	1.000	1.000
II+VI-a	Combination (2/3-1/3)	0.02	0.08	0.19	-0.09	-0.02	0.11	0.093	0.072	1.000	0.910
II+VI-b	Combination (1/3-2/3)	-0.04	-0.00	0.10	0.02	0.09	0.19	-0.070	0.078	1.000	0.970

Notes: This table shows the results for the simulations of the counterfactuals descrived in Table 1. Columns (1)-(6) report the distribution of the changes in the value-added of the schools chosen by families of low and high income, column (7) shows the reduction in the value-added income gap, column (8) shows the mean quality provided by the schools chosen by families in the market, column (9) shows the ex-ante cost of the policy (as a fraction of the original budget for policy I), and column 10 shows the ex-post cost of the policy. The last two columns may differ due to incomplete policy take-up.

Appendix

A1 Additional Details on the Setting

A1.1 More Details on the Law for the Promotion of Private Investments in Education (Law N-882)

Law N-882 also introduced tax incentives for private schools. Firms that offer education services receive credit for 30% of the reinvested utilities, in addition to import and sales tax exemptions.Tax incentives are particularly attractive for multi-market firms under the prevailing integrated tax system, as tax credits in one firm can be used to reduce taxes for other firms in the holding. An interesting fact in this setting is that many schools have entered as for-profit chains. There are around 15 chains of schools,[38] most of which entered the market in the recent years, with aggressive expansion plans. Together, they represent a 12% of the share of elementary schools in Lima, but most of them have schools in other cities.

A1.2 More Details on the Law for the Teachers' Payment Reform (Law N-29944)

To enter into a tenure-track position and to get promoted to a higher performance group, teachers go through a centralized process. In the first stage, they take a National Knowledge Test for teachers. If they are above a minimum cutoff, they participate in the centralized application for tenure-track positions, for which teachers submit a rank order list of 5 schools and are then assigned to two schools based on their score. In the second stage, teachers are evaluated in a decentralized process by the schools in terms of their "didactic capacity, training, merits, and experience". The schools make offers to candidates they like. For teachers and vacancies that are not matched in this process, there is an aftermarket, in which teachers are assigned to schools using a serial dictatorship algorithm. The aftermarket contracts are non-tenure track and last for only one year.

These changes generated substantial variation in wages across time and across schools. The opening for the contract vacancies is school-specific and is decided at the school district level. Due to budget constraints and planned staged implementation of the policy, only a fraction of teaching positions in the public sector were subject to these higher paying contracts. The way the first stage is designed also generates variation, as rank-ordered lists are short, and several positions end up not being filled with tenure-track contracts.

A final comment on the reform. Private schools strongly opposed the increases in wages, arguing that the rise in wages in the public sector was unfair competition with private schools as

[38] See Scotiabank (2017) financial report in which they analyze the education industry in Peru and discuss the horizontal integration of schools.

it led to ”substantial increases in their cost of hiring”.[39]

A1.3 More Details on the New Regulation for Primary Schools (Directiva N 014-2012-MINEDU)

The Ministry of Education introduced a new regulation for schools in 2012. These rules include two elements:

1. **Price discrimination**: the new regulation forces schools to make their enrollment conditions public and uniform for all students. This includes application dates, fees, and other payments to the school.

2. **Selection from the school side**: schools are not allowed to engage into active selection or discrimination of students. In particular, the new regulation establishes that:

   i. Forbids any type of evaluation processes for children enrolling in preschool and first grade in both private and public schools.

   ii. In the case of oversubscribed schools, the order forbids prioritization criteria that involves the abilities or characteristics of the child, except for children with obvious special needs that are not able to attend that school.

   iii. These criteria must respond to objective aspects that do not entail discrimination or social exclusion. The criteria should be publicly announced before the annual application process for new students, and published in a visible place to guarantee transparency and equal opportunities for all applicants.

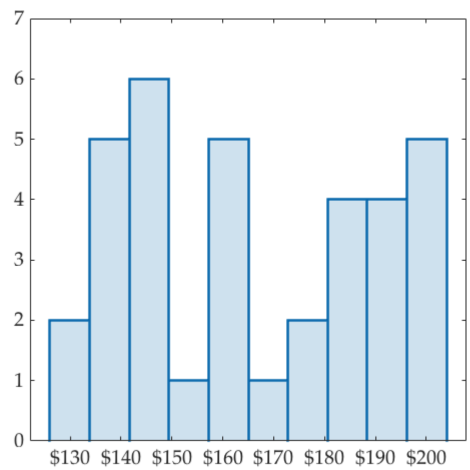
   iv. In the case of schools that have a special characteristic that is intrinsic to their institutional objectives, or within its mission and vision, this can be part of the prioritization criteria, to the extent to which it does not violate the previous requirements.

It is hard to verify to what extent this regulation is enforced in practice. The Ministry of Education has a platform in which parents can report violations to the schools' regulations, and they have the legal power to give a fine to the schools. In private conversations with one of the largest chains of schools, they claimed that they follow this order closely and work on a first come first based basis (and that is how they describe the admission process on their website). Figure A1 provides evidence for the pricing strategy for that firm. Panel (a) shows the price dispersion across schools in the chain. Panel (b) shows daily transaction data for the first monthly due for 5 of the schools in the chain. While there is clear price discrimination across schools, there is no evidence

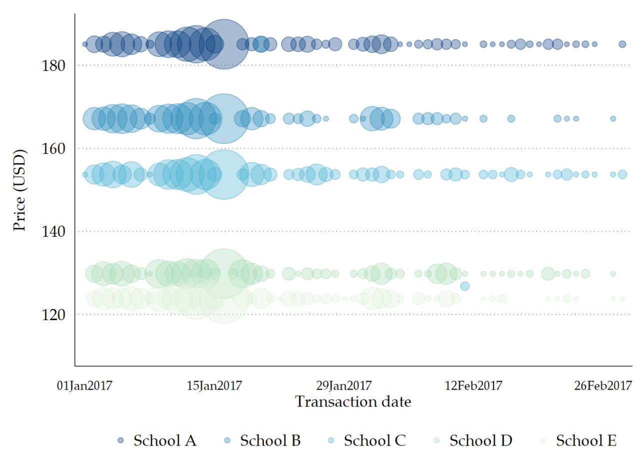
[39]This claim was made by the director of Grupo Educación al Futuro (GEF), a consulting company focused in private schools in Peru. He also argued that private school tuitions will go up on 2018 ”due to last year salary increase for public education teachers which have been transferred to private education”. See `https://rpp.pe/economia/economia/gef-colegios-privados-subirian-sus-pensiones-para-incrementar-el-sueldo-de-sus-docentes-noticia-1102091`.

of price discrimination between students in the same school. Regarding the selection from the school side, Figure A2 shows the answers to the 2016 principals' survey. Panel (c) asks for admission requirements, to which almost all schools reply that they do not test kids or ask for other requirements that could be potentially discriminatory. Panel (d) asks for schools for what they do if oversubscribed, and they reply that they operate on a first-come, first-served basis and do not use screening exams or interviews.

Figure A1: Pricing Strategy for Large Chain



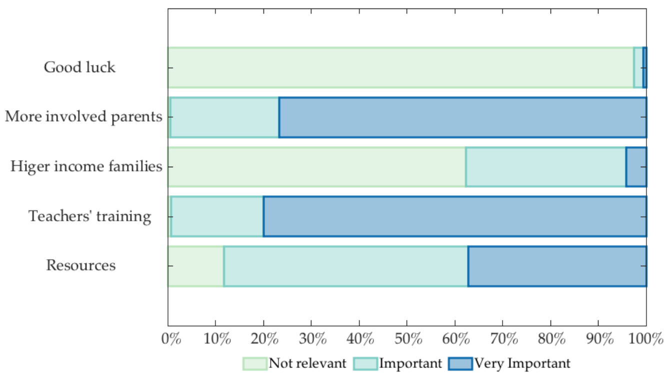
(a) Price dispersion across schools (Lima 2017)



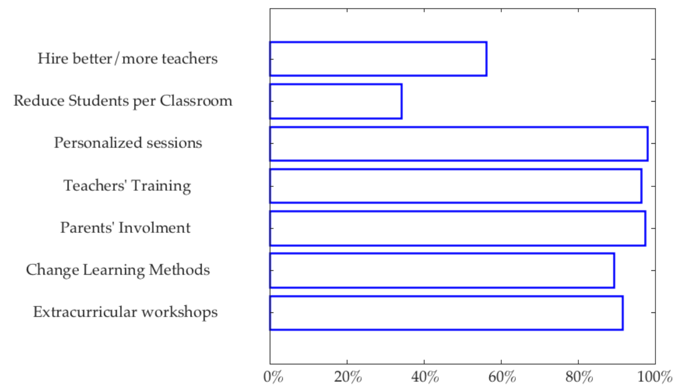
(b) Price dispersion within schools

Notes: The data in this figure comes from one of the largest chains of schools in Peru. Panel (a) shows the monthly tuition charged by schools in the city of Lima in 2017. All these schools operate under the same brand, have standardized buildings, and centralized management for curriculum development and teachers' training. Panel (b) shows daily transactions for the first monthly dues for the academic year for a sample of 5 schools from the ones in panel (a). The size of the circle is proportional to the number of transactions at each price, in each day, for each school.

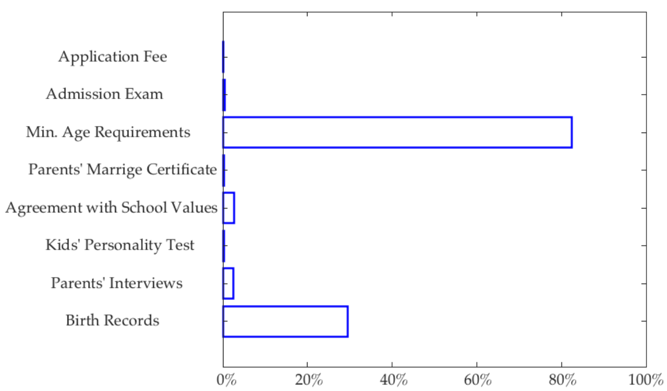
Figure A2: School Principals' Survey



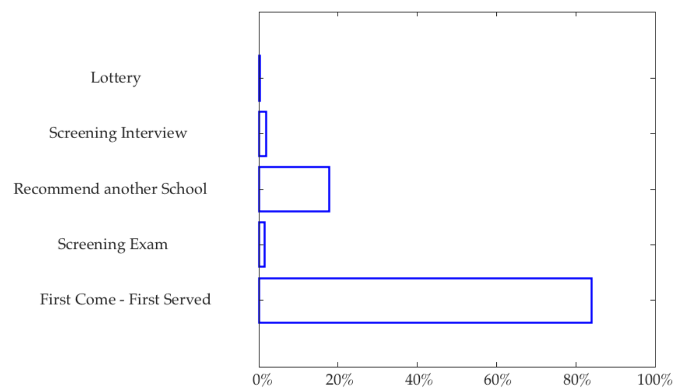
(a) Factors that explain tests scores



(b) Actions to improve performance



(c) Admission requirements



(d) Screening mechanism

Notes: These figures show the answers to survey questions for 621 principals in 2016. Each panel corresponds to the following question: Panel (a) - What factors explain that some schools have higher test scores? Panel (b) - What would you do to improve the performance of the students in your school? Panel (c) - What are the requirements for the admission of a child in your primary school? Panel (d) - How is the process of admission to this school when the number of students applying to the school exceeds the number of vacancies?

A1.4 More Details on the 2017 Teachers' Strike

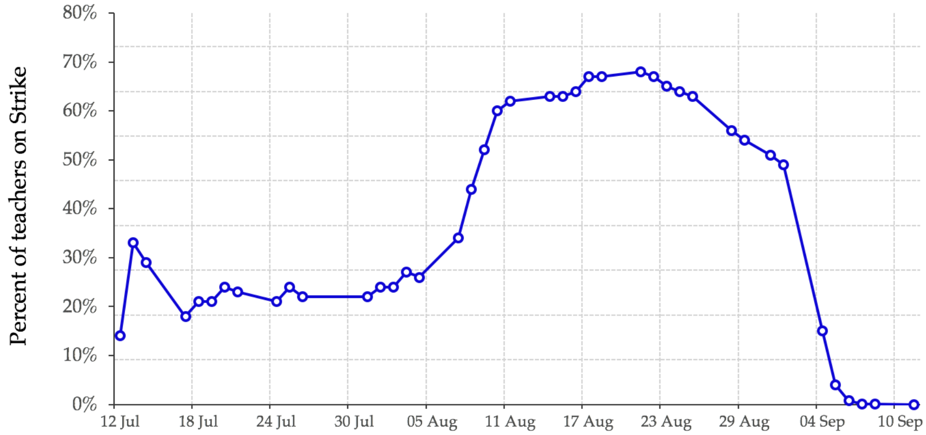
To have an idea of the proportions of the strike, panel (a) Figure A3 shows the percentage of public school teachers that were on strike each day during the strike period. At its peak, this number was close to 70% at the national level. However, the intensity and duration of the strike had substantial variation by the school district, as shown for the the city of Lima in Figure A3.

According to the the Ministry of Education the strike had an important political component, and generated a crisis that ended with the impeachment of the Minister. As mention in section 3, the decision of the teachers to join the strike was heavily influenced by the teachers' union membership, political affiliation, and political networks. There are several teacher unions to which public school districts have been historically subscribed to, and there is variation in whether the political parties supported the strike or not. As a result, there is a substantial geographical variation in the intensity of the strike in different schools, both in terms of duration and number of teachers involved. This information was collected daily through surveys in each school district by the Intelligence and Crisis Management team at the Ministry of Education. They also collected data on the union leaders, their political networks and maps for the unions' internal geographical divisions within a market.

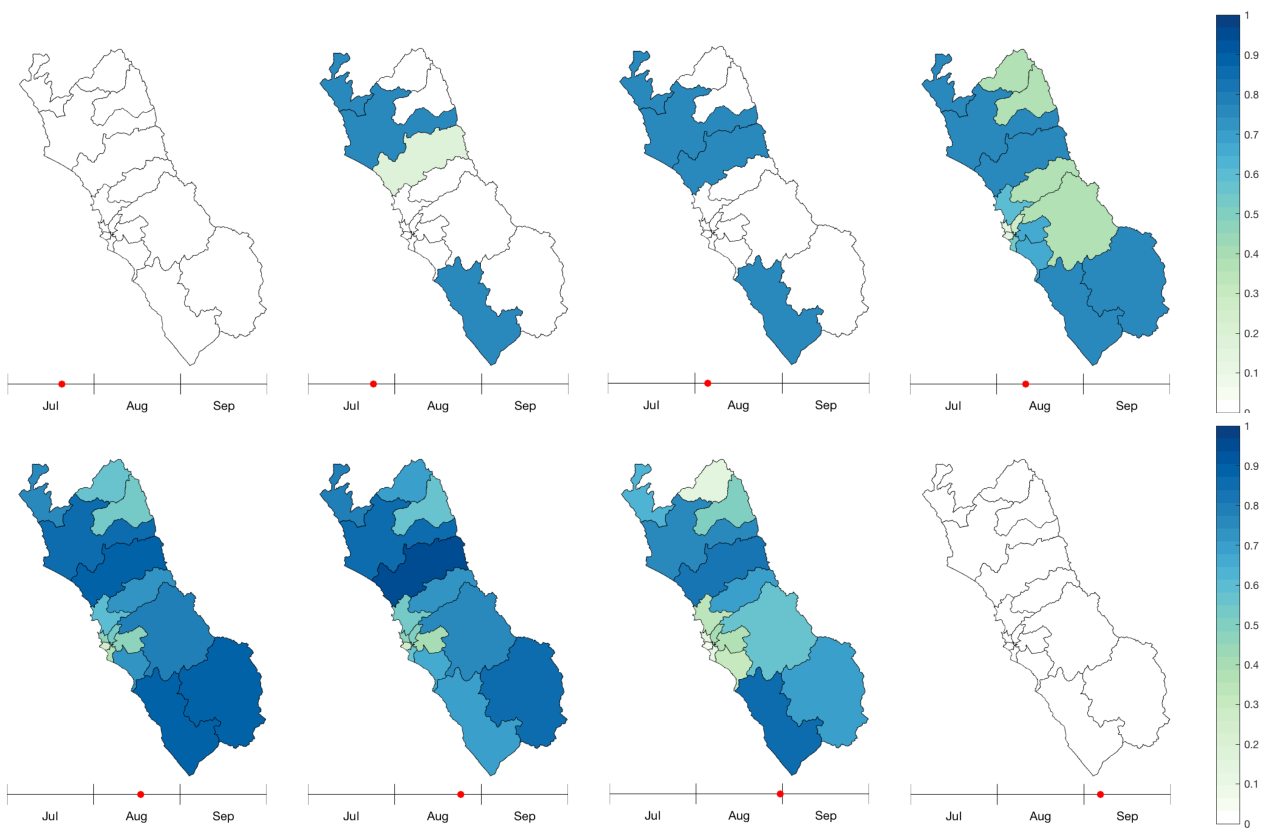
There are three final considerations regarding the strike. First, At the time of the strike, both the Government and experts were worried that many public school families that could afford a private school would switch to the private sector, leaving the lower SES families behind.[40]. Second, the timing of the strike is important. Panel (a) in Figure A4 shows a timeline for the school year, school choice decisions, standardized tests and the period in which the strike took place. Finally, it is important to note that strikes are not uncommon in Peru. This is shown in panel (c) in Figure A4.

[40]See `https://elcomercio.pe/economia/dia-1/educacion-migracion-alumnado-huelga-noticia-453629`. The article comments on this and cites the owner of a big chain of schools who claims that he is ”getting prepared receive several new strudents” after the strike.

Figure A3: Evolution and Intensity of the Strike



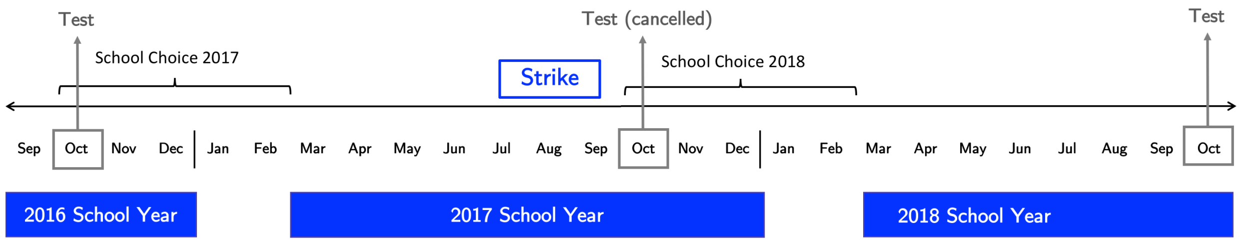
(a) Intensity of the 2017 Teachers' Strike (National Level)



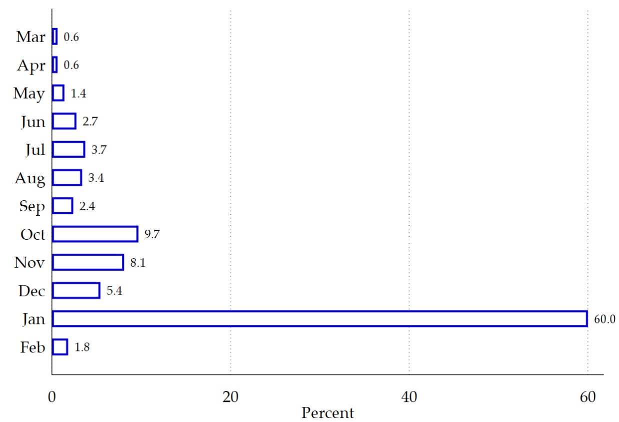
(b) Evolution of the strike in Lima

Notes: These plots display the evolution and intensity of the 2017 teachers' strike for the school districts in Lima. The intensity goes from 0 to 1, where one means that 100% of the public schools in the district were on strike that day. Dates correspond to 2017 and are represented with a red dot in the timeline below each graph.

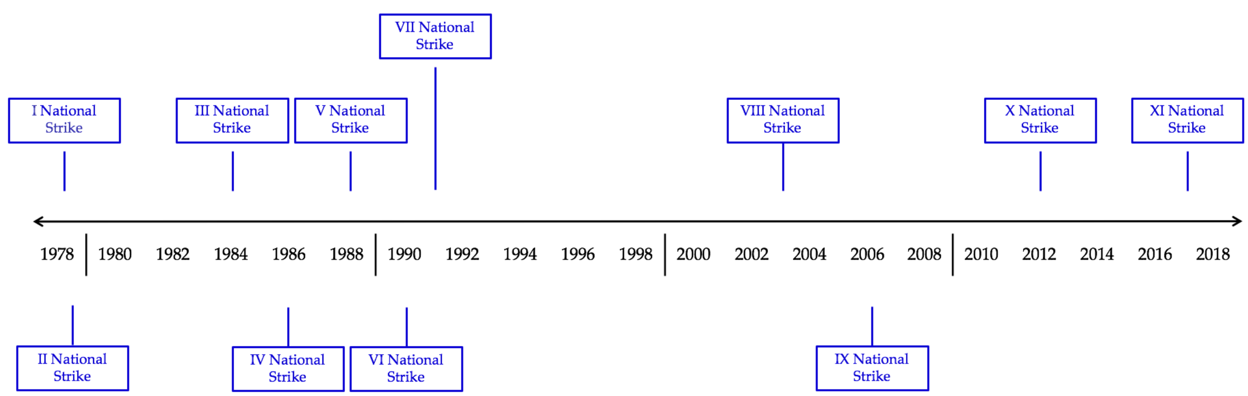
Figure A4: Teachers' Strike



(a) Timeline of the 2017 Teachers' Strike



(b) School Survey: Application Date



(c) Historic Strikes in Peru

Notes:

A2 More Details on the Scholarship Program

Scholarships are assigned using a two-stage discontinuous rule[41]. In the first stage, students take the academic test, which has two parts: language and math. Then, based on the average score, students who are over a cutoff are eligible for the second stage[42], in which students take a leadership test, evaluated one-by-one by professional psychologists. Students who pass the second cutoff are eligible for the scholarship in a school of their choice among the ones with which the NGO has an agreement.

Panel (a) in Figure A5 plots the first step of the scholarship assignment as a function of the academic scores. Blue lines show a first order polynomial fit for units below and above the cutoff separately, and grey circles represent the sample average for each bin. Not all of these students above the cutoff receive the scholarship, as there is a second step, depicted in panel (b), in which all the ones that qualified in the first step take a leadership test, which also has a cutoff. Only students that are above both cutoffs are eligible for the scholarship. However, as not all the students take the leadership test, I cannot use a two dimension RD.

To identify the price parameter I need an instrument that generates exogenous variation in prices, while trying to keep the sample as representative as possible. That is why I will focus on the fist cutoff for the scholarship assignment.

Table A1 presents estimates of the change in the probability of being eligible for the scholrship for students who are above the cutoff, where $Score \geq Cutoff$ is a dummy variable that equals one if the students' academic score is above the cutoff for scholarship eligibility. I use optimal bandwidths and robust confidence intervals proposed by Calonico et al. (2014). Being above the cutoff increases the probability of receiving the scholarship by 37% percentage points. Panels (c)-(b) present robustness checks for the regression discontinuity design: the McCrary(2008) density test and falsification tests that look at the effect of the scholarship on pre-treatment covariates.

A3 Total Derivatives of the Share under Social Interactions

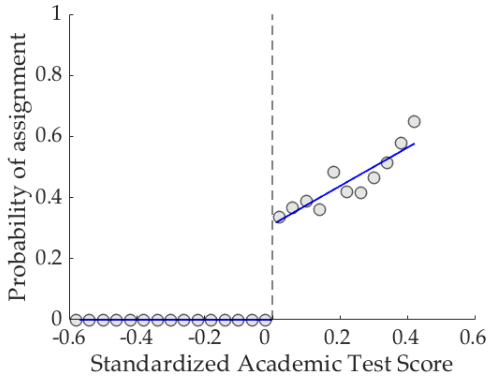
A3.1 With respect to Price

Equation 13 shows the total derivative of the individual choice probabilities for school j with respect to changes in the price for j. Terms (a), (b), and (d) have simple expressions that are functions of demand parameters and individual choice probabilities:

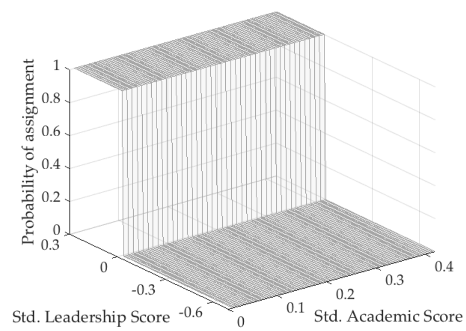
[41] Except for the slots reserved for students with unique talents (mainly featured athletes and artists), which are identified at the beginning of the process and are subject to different rules

[42] Cutoffs may differ by location/date

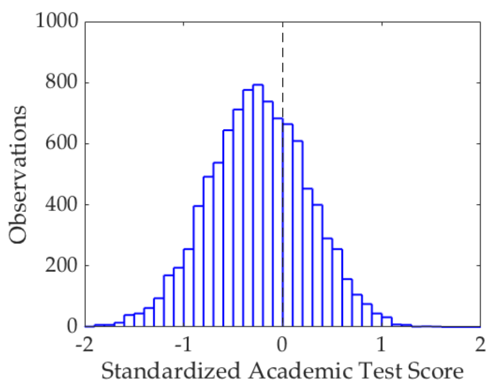
Figure A5: Regression Discontinuity: Design and Falsification Tests



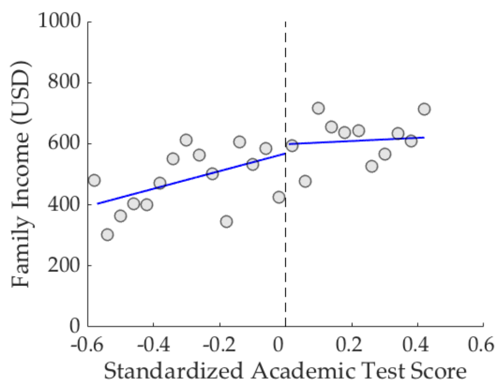
(a) First Cutoff



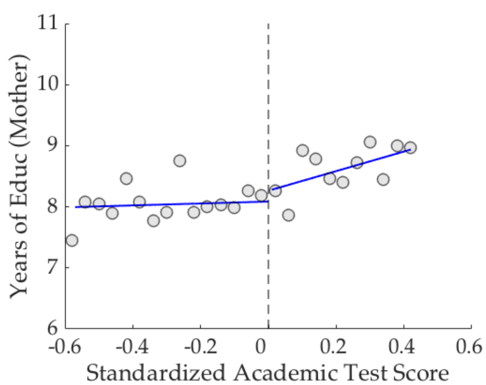
(b) Second Cutoff



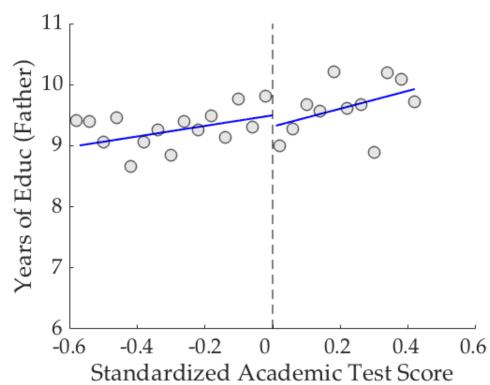
(c) Density Visualization



(d) Pre-treatment: Family Income



(e) Pre-treatment: Mother's Education



(f) Pre-treatment: Father's Education

Notes: White circles present the average in the outcome variable for individuals in equally spaced disjoint bins. Blue lines depict a first-order polynomial fit, separate for units at each side of the threshold.

Table A1: First Stage - Scholarship

	Elegible
Score $\geq$ Cutoff	0.3681**
	(0.1823)
Obs left	2990
Obs right	2126

Notes: Estimates are computed using optimal bandwidths and robust confidence intervals proposed by Calonico et al. (2014).

$$\frac{\partial s_{ijt}}{\partial p_{jt}} = -\alpha_i s_{ijt}(1 - s_{ijt})$$
$$\frac{\partial s_{ijt}}{\partial z^x_{jt}} = \tilde{\beta}_i^{z^x} s_{ijt}(1 - s_{ijt}), \quad \text{for } x = \{y, e\}$$
$$\frac{\partial s_{ijt}}{\partial z^x_{rt}} = \tilde{\beta}_i^{z^x} s_{ijt} s_{irt}, \quad \text{for } x = \{y, e\} \text{ and } r \neq j.$$

The expressions for terms (c) and (e) cannot be obtained directly from demand parameters, as they are functions of the left side of Equation 13, the total derivatives of the shares with respect to p_{jt}. Then, I can aggregate Equation 13 at the type x and stack expressions by type and school, to form a linear system from which we can solve for the whole vector of total derivatives with respect to price.

First, note that z^x_{jt} (where x can be either y or e) is the fraction of families of type x, so I rewrite:

$$z^x_{jt} = \frac{s^x_{jt}\pi^x N}{s_{jt}N} = \frac{\pi^x s^x_{jt}}{\pi^x s^x_{jt} + (1 - \pi^x)s^{-x}_{jt}}$$

In the previous expression z^x_{jt} is a function of the shares $f(s^x_{jt}, s^{-x}_{jt})$. Then, I express the total derivative as:

$$\frac{dz^x_{jt}}{dp_{jt}} = \frac{\partial z^x_{jt}}{\partial s^x_{jt}}\frac{ds^x_{jt}}{dp_{jt}} + \frac{\partial z^x_{jt}}{\partial s^{-x}_{jt}}\frac{ds^{-x}_{jt}}{dp_{jt}}$$

Where

$$\frac{\partial z^x_{jt}}{\partial s^x_{jt}} = \underbrace{\frac{\pi^x}{s_{jt}}(1-\pi^x s^x_{jt})}_{\kappa^x_j} \text{ and } \frac{\partial z^x_{jt}}{\partial s^{-x}_{jt}} = \underbrace{-\frac{\pi^x}{s_{jt}} s^x_{jt}(1-\pi^x)}_{\kappa^{-x}_j}$$

I write:

$$\frac{dz^x_{jt}}{dp_{jt}} = \kappa^x_j \frac{ds^x_{jt}}{dp_{jt}} + \kappa^{-x}_j \frac{ds^{-x}_{jt}}{dp_{jt}} \tag{30}$$

Then, I add Equation 13 at the type x level:

$$\frac{ds^x_{jt}}{dp_{jt}} = \frac{\partial s^x_{jt}}{\partial p_{jt}} + \sum_l \frac{\partial s^x_{jt}}{\partial z^y_{lt}} \frac{\partial z^y_{lt}}{\partial p_{jt}} + \sum_l \frac{\partial s^x_{jt}}{\partial z^e_{lt}} \frac{\partial z^e_{lt}}{\partial p_{jt}} \tag{31}$$

And replace Equation 30 in Equation 31. I get 4 equations for the derivatives of each school with respect to p_{jt}:

$$\left[\frac{ds^y_{jt}}{dp_{jt}}\left(1-\kappa^y_j \frac{\partial s^y_{jt}}{\partial z^y_{jt}}\right) - \sum_{l\neq j} \frac{ds^y_{lt}}{dp_{jt}}\left(\kappa^y_l \frac{\partial s^y_{lt}}{\partial z^y_{lt}}\right)\right] - \sum_l \frac{ds^{-y}_{lt}}{dp_{jt}}\left(\kappa^{-y}_l \frac{\partial s^y_{lt}}{\partial z^y_{lt}}\right) - \sum_l \frac{ds^e_{lt}}{dp_{jt}}\left(\kappa^e_l \frac{\partial s^y_{lt}}{\partial z^e_{lt}}\right) - \sum_l \frac{ds^{-e}_{lt}}{dp_{jt}}\left(\kappa^{-e}_l \frac{\partial s^y_{lt}}{\partial z^e_{lt}}\right) = \frac{\partial s^y_{lt}}{\partial p_{jt}} \tag{32a}$$

$$-\sum_l \frac{ds^y_{jt}}{dp_{jt}}\left(\kappa^y_l \frac{\partial s^{-y}_{lt}}{\partial z^y_{lt}}\right) + \left[\frac{ds^{-y}_{jt}}{dp_{jt}}\left(1-\kappa^{-y}_l \frac{\partial s^{-y}_{jt}}{\partial z^y_{jt}}\right) - \sum_{l\neq j} \frac{ds^{-y}_{lt}}{dp_{jt}}\left(\kappa^{-y}_j \frac{\partial s^{-y}_{lt}}{\partial z^y_{lt}}\right)\right] - \sum_l \frac{ds^e_{lt}}{dp_{jt}}\left(\kappa^e_l \frac{\partial s^{-y}_{lt}}{\partial z^e_{lt}}\right) - \sum_l \frac{ds^{-e}_{lt}}{dp_{jt}}\left(\kappa^{-e}_l \frac{\partial s^{-y}_{lt}}{\partial z^e_{lt}}\right) = \frac{\partial s^{-y}_{lt}}{\partial p_{jt}} \tag{32b}$$

$$-\sum_l \frac{ds^y_{lt}}{dp_{jt}}\left(\kappa^y_l \frac{\partial s^e_{lt}}{\partial z^y_{lt}}\right) - \sum_l \frac{ds^{-y}_{lt}}{dp_{jt}}\left(\kappa^{-y}_l \frac{\partial s^e_{lt}}{\partial z^y_{lt}}\right) - \left[\frac{ds^e_{jt}}{dp_{jt}}\left(1-\kappa^e_l \frac{\partial s^e_{jt}}{\partial z^e_{jt}}\right) - \sum_{l\neq j} \frac{ds^e_{jt}}{dp_{jt}}\left(\kappa^e_j \frac{\partial s^e_{lt}}{\partial z^e_{lt}}\right)\right] - \sum_l \frac{ds^{-e}_{lt}}{dp_{jt}}\left(\kappa^{-e}_l \frac{\partial s^e_{lt}}{\partial z^e_{lt}}\right) = \frac{\partial s^e_{jt}}{\partial p_{jt}} \tag{32c}$$

$$-\sum_l \frac{ds^y_{lt}}{dp_{jt}}\left(\kappa^y_l \frac{\partial s^{-e}_{lt}}{\partial z^y_{lt}}\right) - \sum_l \frac{ds^{-y}_{lt}}{dp_{jt}}\left(\kappa^{-y}_l \frac{\partial s^{-e}_{lt}}{\partial z^y_{lt}}\right) - \sum_l \frac{ds^e_{lt}}{dp_{jt}}\left(\kappa^e_l \frac{\partial s^{-e}_{lt}}{\partial z^e_{lt}}\right) + \left[\frac{ds^{-e}_{jt}}{dp_{jt}}\left(1-\kappa^{-e}_j \frac{\partial s^y_{jt}}{\partial z^e_{jt}}\right) - \sum_{l\neq j} \frac{ds^{-e}_{lt}}{dp_{jt}}\left(\kappa^{-e}_l \frac{\partial s^{-e}_{lt}}{\partial z^e_{lt}}\right)\right] = \frac{\partial s^{-e}_{lt}}{\partial p_{jt}} \tag{32d}$$

I stack the set of equations for each school and rewrite them to form a system of linear equations to solve for $\frac{ds^x_{jt}}{dp_{jt}}$. The, I can recover the object of interest, $\frac{ds_{jt}}{dp_{jt}}$. The system of equations has the form:

$$A_j X_j = B_j$$

Where:

- X is a $(N_j \times 4) \times 1$ vector of total derivatives of s^x_{jt} with respect to price p_{jt}

$$X'_j = \left[x'_{1j}, x'_{2j}, x'_{3j}, ..., x'_{Jj}\right] \text{, where } x'_{jj} = \left[\frac{ds^y_{jt}}{dp_{jt}}, \frac{ds^{-y}_{jt}}{dp_{jt}}, \frac{ds^e_{jt}}{dp_{jt}}, \frac{ds^{-e}_{jt}}{dp_{jt}}\right]$$

- B_j is a $(N_j \times 4) \times 1$ vector of partial derivatives of s^x_{jt} with respect to price p_{jt}, which are functions of demand parameters and individual choice probabilities.

$$B'_j = \left[b'_{1j}, b'_{2j}, b'_{3j}, ..., b'_{Jj}\right] \text{, where } b'_{jj} = \left[\frac{\partial s^y_{jt}}{\partial p_{jt}}, \frac{\partial s^{-y}_{jt}}{\partial p_{jt}}, \frac{\partial s^e_{jt}}{\partial p_{jt}}, \frac{\partial s^{-e}_{jt}}{\partial p_{jt}}\right]$$

- A_j is a $(N_j \times 4) \times (N_j \times 4)$ matrix:

$$A_j = \begin{bmatrix} A^{11}_j & \cdots & A^{1j}_j & \cdots & A^{1J}_j \\ \vdots & \ddots & \vdots & & \vdots \\ A^{j1}_j & \cdots & A^{jj}_j & \cdots & A^{jJ}_j \\ \vdots & & \vdots & \ddots & \vdots \\ A^{J1}_j & \cdots & A^{Jj}_j & \cdots & A^{JJ}_j \end{bmatrix}$$

  Where:

$$A^{jj}_j = \begin{bmatrix} 1 - \kappa^y_j \frac{\partial s^y_{jt}}{\partial z^y_{jt}} & \kappa^{-y}_j \frac{\partial s^y_{jt}}{\partial z^y_{jt}} & \kappa^e_j \frac{\partial s^y_{jt}}{\partial z^e_{jt}} & \kappa^{-e}_j \frac{\partial s^y_{jt}}{\partial z^e_{jt}} \\ \kappa^y_j \frac{\partial s^{-y}_{jt}}{\partial z^y_{jt}} & 1 - \kappa^{-y}_j \frac{\partial s^{-y}_{jt}}{\partial z^y_{jt}} & \kappa^e_j \frac{\partial s^{-y}_{jt}}{\partial z^e_{jt}} & \kappa^{-e}_j \frac{\partial s^{-y}_{jt}}{\partial z^e_{jt}} \\ \kappa^y_j \frac{\partial s^e_{jt}}{\partial z^y_{jt}} & \kappa^{-y}_j \frac{\partial s^e_{jt}}{\partial z^y_{jt}} & 1 - \kappa^e_j \frac{\partial s^e_{jt}}{\partial z^e_{jt}} & \kappa^{-e}_j \frac{\partial s^e_{jt}}{\partial z^e_{jt}} \\ \kappa^y_j \frac{\partial s^{-e}_{jt}}{\partial z^y_{jt}} & \kappa^{-y}_j \frac{\partial s^{-e}_{jt}}{\partial z^y_{jt}} & \kappa^e_j \frac{\partial s^{-e}_{jt}}{\partial z^e_{jt}} & 1 - \kappa^{-e}_j \frac{\partial s^{-e}_{jt}}{\partial z^e_{jt}} \end{bmatrix} \quad A^{jl}_j = \begin{bmatrix} \kappa^y_l \frac{\partial s^y_{jt}}{\partial z^y_{jt}} & \kappa^{-y}_l \frac{\partial s^y_{lt}}{\partial z^y_{lt}} & \kappa^e_l \frac{\partial s^y_{lt}}{\partial z^e_{lt}} & \kappa^{-e}_l \frac{\partial s^y_{lt}}{\partial z^e_{lt}} \\ \kappa^y_l \frac{\partial s^{-y}_{lt}}{\partial z^y_{lt}} & \kappa^{-y}_j \frac{\partial s^{-y}_{jt}}{\partial z^y_{jt}} & \kappa^e_l \frac{\partial s^{-y}_{lt}}{\partial z^e_{lt}} & \kappa^{-e}_l \frac{\partial s^e_{lt}}{\partial z^e_{lt}} \\ \kappa^y_l \frac{\partial s^e_{lt}}{\partial z^y_{lt}} & \kappa^{-y}_l \frac{\partial s^e_{lt}}{\partial z^y_{lt}} & \kappa^e_j \frac{\partial s^e_{jt}}{\partial z^e_{jt}} & \kappa^{-e}_l \frac{\partial s^{-y}_{lt}}{\partial z^e_{lt}} \\ \kappa^y_l \frac{\partial s^{-e}_{lt}}{\partial z^y_{lt}} & \kappa^{-y}_l \frac{\partial s^{-e}_{lt}}{\partial z^y_{lt}} & \kappa^e_l \frac{\partial s^{-e}_{lt}}{\partial z^e_{lt}} & \kappa^{-e}_j \frac{\partial s^{-e}_{jt}}{\partial z^e_{jt}} \end{bmatrix}$$

I get the vector of type-specific total derivatives from:

$$X = A^{-1}B$$

Finally, I recover the total derivatives of each school with respect to p_jt:

$$\frac{ds_{jt}}{dp_{jt}} = \pi^x \frac{ds_{jt}^x}{dp_{jt}} + (1-\pi^x)\frac{ds_{jt}^{-x}}{dp_{jt}}$$

I repeat the procedure for the price of each school.

A3.2 With respect to Quality

Additional elements (f), (g), (h), (i), (j), and (k) need to be computed to recover the total derivative of the shares with respect to quality. These elements are individualized in Equation 18. Terms (a)-(e) are described in the previous section of the Appendix.

Term (f) is a function of demand parameters and individual choice probabilities:

$$\frac{\partial s_{ijt}}{\partial \mu_{jt}} = \beta_i^{\mu} s_{ijt}(1-s_{ijt})$$

In practice, (i) is the sum of (g) and (k). I separate these two elements in the text to show that there is a direct and an indirect change in second stage prices when quality changes but, in practice, I only need the total effect (i). The decomposition into (g) and (k) can be done trivially from the proposed method. Terms (j) and (h) are also obtained in the process.

$$\frac{dp_{lt}^*}{d\mu_{jt}} = \gamma_\mu + \frac{ds_{jt}}{d\mu_{jt}} \times \left(\frac{ds_{jt}}{dp_{jt}}\right)^{-1} - s_{jt} \times \left(\frac{ds_{jt}}{dp_{jt}}\right)^{-2} \times \left(\frac{d\frac{ds_{jt}}{dp_{jt}}}{d\mu_{jt}}\right)$$

where:

$\frac{ds_{ijt}}{d\mu_{jt}}$: is calculated using the analogous to the one in Appendix A3.1, taking the total derivative with respect to quality instead of price. This is the sensititivy of demand to quality, considering the spillover effects, but holding prices constant. Terms (j) and (h) are also obtained in this process.

$\frac{ds_{jt}}{dp_{jt}}$: Is calculated in Appendix A3.1.

$\frac{d\frac{ds_{jt}}{dp_{jt}}}{d\mu_{jt}}$: I obtained using an approach based on the same principle as the one in Appendix A3.1.

A4 Interpretation of the parameters under uncertainty about school characteristics

The linear and the non-linear parameters in the utility function $\theta_1 = (\beta_0, \eta)$ and $\theta_2 = (\alpha_k, \beta_k, \lambda_k, \beta^u)$ define the weights that the family place on the observed school characteristics. I present a parsimoniuos model of belief formation to interpret these weights under uncertainty about school characteristics. I assume that the true distribution for quality is $q_j \sim N(0, \sigma_q^2)$, but families only observe a noisy signal which correponds to the true quality plus an error distributed $v_{ij} \sim N(0, \sigma_{\epsilon k}^2)$. Note that the variance of the error can differ by family type k. The expected quality would be: $\tilde{q}_{ij} = E[q_{ij}] = \rho_k(q_{ij} + v_{ij})$, where $\rho_k = \dfrac{\sigma_q^2}{\sigma_q^2 + \sigma_{\epsilon k}^2}$. If we instert the expression for the expected quality in **??** we can rewrite the utility problem of the family as:

$$max_{j \in \Omega_i} \; U_{ijt} = \pi_k q_{jt} - \alpha_k p_{jt} + \bar{z}'_{jt} \tilde{\beta}_i^z + \lambda_k d_{ij} + \sum_r \eta^r x_{jt}^r + \xi_{jt} + \tilde{\epsilon}_{ijt} \tag{33}$$

Where $\tilde{\epsilon}_{ijt} = \pi_k v_{ij} + \epsilon_{ijt}$ and $\pi_k = \beta_{ik} \rho_k$. Then, the estimated parameter on quality, π_k, is the preference for quality scaled by a measure of how precise the beliefs on quality are. If $\sigma_{\epsilon k}^2 = 0$, then $\rho_k = 1$, implying that families observe quality perfectly and $\pi_k = \beta_{ik}$, so the preference weight is not shrunk. As $\sigma_{\epsilon k}^2$ increases, the signals get noisier, and families put less weight on quality.

A5 Information on Data Sources and Their Descriptions

A5.1 Student Level

Student records. I use the SIAGIE (*Sistema de Información de Apoyo a la Gestión Educativa*) dataset from the Ministry of Education. SIAGIE is the system that keeps the records for every student in the education system from 2012 to 2018. The information includes the school, grade, and classroom they attend, native language, parents' education, other demographic and personal information, and basic academic performance, including grades by subject and final situation. Student identifiers allow tracking individual students across years and merging this information with other datasets.

Student test scores. ECE (*Evaluación Censal de Estudiantes*) data has individual-level results for Peru's standardized test. All private and public schools' students in specific grades take this test every year since 2006.[43] The test is low-stakes for the students and measures their basic competencies in math and language at the end of the school year.

Birth and health records. Observations in the previous datasets are matched with the CVN (*Cer-*

[43] Except for 2017, year in which the test was canceled due to *El Niño* naturally occurring phenomenon and the theachers' strike.

tificado de Nacido Vivo) and PdN (*Padrón de niños*) data from the Ministry of Health. CVN are the birth records kept by the Ministry of Health, which contain information on the health conditions of a child at birth (birth weight, height, gestation age, type of birth, type of professional serving, hospital) and information about the family (parents' education level, occupation, marriage status, address type of health insurance and the possibility ofidentifying of the mother and siblings of each student). PdN is a dynamic record of all children of ages 0 to 6 that has information on health outcomes, health and social programs in which the student is a beneficiary, and the family's updated addresses, which are geocoded to generate data on spatial locations.

National Household Survey. The student records data is matched with PGH (*Padrón General de Hogares*) data from the Ministry of Social Development. PGH is the data for the socioeconomic characterization of households for the focalization of social programs. It is a detailed household survey that covers 90% of the population (except very isolated non-urban areas and very wealthy locations). This Survey is used to give an SES-class to the household.

National Census. Several of the datasets described above have student addresses (SIAGIE, PdN, and CVN). However, addresses in Peru are not standardized, difficulting a massive geocoding process.[44] The *Instituto Nacional de Estadstica e Informtica (INEI)* is the agency responsible for the Census of Peru. Through an agreement between INEI and the Ministry of Finance, we were able to associate each student in the education system with the census block in which their household is located for the 2017 Census. The match was done by INEI using their National ID Numbers, names and birth dates.[45]

Scholarship application. An additional source of data at the student level is the Peru Champs Scholarship application records for the years 2014-2018. The NGO assigns an internal identifier to all the students who apply to the program. These identifiers allow me to match the parents' Survey, academic test results, and leadership test results. The NGO also collects the students' and parents' National ID Numbers for all applicants, which I use to merge the scholarship application data with administrative records.

A5.2 School Level

School directory. The first source is the data on every school in Peru (both private and public) collected by the Ministry of Education using two systems -ESCALE and SIAGIE- and a unique school identifier. ESCALE includes three main datasets, available from 2000 to date for both private and public schools. The School Directory (*Padrón Escolar*) has data on the type of school management, ownership, whether it is coeducational, religious orientation, focus (vocational, for students with

[44] There are several formats and no unified source to map them into longitude and latitude cooordinates

[45] In Peru, the Census includes a non-mandatory National ID Number question. The rate of response is high (close to 90%). When the ID Number was missing, the matching was made using two names, last name and birth dates

special talents or for students with cognitive limitations), addresses and geocoded location, grades offered, and date in which the school was registered in the official system for the first time. The School Census (*Censo Escolar*) includes data on total enrollment (by grade, sex, age, and native language), the number of classrooms, and staff (by position, experience, type of education). Finally, the School Infrastructure section of the Census contains detailed information about the labs, libraries, sports facilities, computers, and other equipment, and access to the internet and sanitary services. Finally, chains of private schools and the network of schools they own have been identified using the information provided by a consulting company and self-collected by accessing the chain's website.

The school directory also includes information on the teachers that work in the school (for both public and private), including their national ID numbers. This data is matched with the 2017 Census to obtain the census block in which graduates live.

Quality and student body. The second source is variables at the school level that result from either aggregating the individual level data to generate variables that characterize the student body composition (this includes mother's education and income) or using the individual-level data to estimating elements at the school level, like the value-added measures discussed in section 5.3.

Tuition. The third group of data sources are the ones used to construct a price variable for each school over time. IDENTICOLE is a system that the Ministry of Education uses to collect data on schools to inform parents and has data on the monthly tuition for private schools for the years 2015-2018. This data is complemented with other sources that are also used to verify and get information on prices for other years: i) survey for prices in the city of Lima conducted by researchers at the Catholic University for the year 2015-2016, ii) data from a consulting company collects data on private schools and publishes a guide for parents since 1999 (information contained in the 2011 to 2017 guides was digitalized and matched to the school ID), and iii) data from the NGO that provides students with a scholarship and also collects data on private school prices since the year 2013.

Teachers' salaries . The fourth data source is payroll data for teachers in the public sector (SUP *Sistema Único de Planillas*). Using the theachers' ID number, I can link teacher payments and type of contracts to schools using NEXUS (*Sistema de Administración y Control de Plazas*), a dataset that has the information on all the teachers in public schools in Peru.

Teachers' recruitment process. In the years 2015, 2017 and 2018 the Ministry of Education held contests to allocate teaching positions. The data includes information on the vacancies, teachers' test scores, ranked order lists for applications, and final assignment.

Teacher unions' data. dataset on the union leaders, their political networks and maps for the unions' internal geographical divisions within a market.

Teachers' political affiliations. The Government of Peru has a website called *Infogob*[46], which has the objective of promoting governance and transparency in the elections. This website has information on every person ever affiliated to a political party in Peru. I web scrapped this data and matched it with NEXUS using the teachers' ID numbers. Then, I have a measure of whether the teacher is affiliated to a political party, and to which party they are affiliated.

Survey to school principals (2016). MineduLab (part of the Ministry of Education of Peru) implemented this Survey as part of the DFM project. [47] The Survey asks two sets of questions to school principals. The first set is related to the admission process, and the second one is related to how they perceive competition from other schools. A summary of the main answers for this survey is shown in Figure A2.

Price survey to private schools (2016). This survey collects price data for 2,380 private schools in Lima and was implemented by Gallego et al. (2016) in 2016.

Survey to private schools (2019). I implemented a new price survey in 2019 for 3,150 private schools in all the markets used in the empirical section. The objective of the survey was to verify current price data collected by the Government and to recover missing price data from 2014. Other questions related to financial aid and enrollment requirements were also included.

A5.3 Market Level

National Census of Peru I use the Census to get the joint distribution of family SES characteristics (income, mother's education, etc.) for kids about to enter primary school in each node for each market. The two recent Censuses were in 2007 and 2017. The Census is used to define the main school markets in the country. I restrict my attention to urban areas. Boundaries of the markets are defined as in (**?**). Maps for each market are shown in Appendix O-1.

Nodes are defined using an hexagonal grid of equally spaced points over the area of the market. Then, census blocks are assigned to their closest node. Each node is an aggregation of census blocks. This is a detailed geographic disaggregation relative to the ones usually used in the literature (zip codes in the U.S.). Then, for each node, I calculate the number of families with kids in the age of choosing schools (5 years old) and the percentage for each discrete SES family type. To capture changes in SES across time, I take the number of families and percentages by type in each node for 2007 and 2017, and use a linear interpolation to get values for the years in between.

College graduates. Finally, I have data on all the college gradutes registered in SUNEDU, including the program, and year of graduation. This data is matched with the 2017 Census to obtain the census block in which the graduates live.

[46] `https://infogob.jne.gob.pe/Politico`

[47] See Gallego et al. (2016) for more details.

A6 Description of the Instruments

A6.1 Student body characteristics

Lagged strike exposure indexes. These instruments are constructed as follows. The Ministry of Education collected daily data on the schools and the number of teachers that were on strike through the local school districts, covering a large sample of schools. I use LASSO to predict the total number of hours lost per student for the schools in the sample. Explanatory variables include the political affiliations of the teachers in the school, the teacher union affiliation, unions' political affiliation (which vary by school district), and several other demographic variables for teachers. The results are presented in Appendix A6. I use the estimated statistical model to get a prediction for the strike intensity (number of hours lost per student) for the rest of the public schools in the market.

Exclusion restriction ($Cov(Z_{inst}, \xi_j) = 0$). The use of local demographics and market structure as instruments for the student body relies on the assumption that these characteristics do not respond to current demand and cost unobservables. Residential sorting and school entry decisions can be thought of as decisions made in previous stages of the game, determined by past shocks not related to short-run school decisions. These previous stages require several periods for adjustment and thus are not modeled. Taking them as given seems reasonable for the time frame under analysis.

The strike instrument exploits the political component and the interaction with teachers' affiliation before the conflict, which has significant within-market school level variation and is likely to be uncorrelated with school demand and cost unobservables. The exclusion restriction for the strike instrument relies on the idea that competing schools' decisions to participate are unrelated to their own and others' demand shock for the next year. Recent exposure to school closure has effects on next year's enrollment due to frictions (not included in the model), but there are no persistent effects on the schools' unobservables.

This assumption can be violated if the effects of the strike are not limited to parents switching schools.[48] The strike can have long-lasting effects on school quality if, for example, high performing teachers leave the affected schools. Panel (b) in Figure 5 shows no evidence of negative effects on teachers' scores in the next academic year, while panel (a) shows significant effects in enrollment.

The timing of the strike is consistent with this assumption. Panel (a) in Figure A4 shows a timeline for the school year, school choice decisions, standardized tests, and the period in which

[48]Current quality can be affected by the strike if kids get fewer hours of instruction. A report from the Ministry of Education calculates that, on average, kids ended up losing 200 hours of lectures. The Government was aware of this possibility, so they suspended the standardized test in the year 2017. To address these potential problems, I drop the year 2017 from the estimation.

the strike took place. Because of how fast the movement gained force and the months in which the strike took place(mid-July to mid-August), families choosing schools for children entering first grade in the 2017 academic year should not have been able to predict the events. On the other hand, most families chose the school for 2018 after the strike was settled and the political situation had decompressed.[49]

Finally, it is important to note that strikes are not uncommon in Peru. This is shown in Panel (c) in Figure A4. I assume that parents hold beliefs on the probability that there will be a strike in public schools each year. Parents take this into account when they make their choice, shifting local demand. Families that were exposed to a more intense strike may update their beliefs about the probability of strike in the public schools. However, given the fact that the origin of the strike was unrelated to relevant school characteristics, I assume that they do not update the probability that a specific school will go on strike for the coming year.

Relevance Condition ($Cov(\bar{z}_{jt}, Z_{inst}) \neq 0$). The instruments should be correlated with the fraction of high SES families in the school. This is natural for the local demographics instrument in a context in which families do not travel far to go to school. All else equal, a school located in a neighborhood with a high concentration of high SES families should enroll a higher share of that group than a school located in a neighborhood with a lower concentration of high SES families.

The fundamental assumption for the strike to be a relevant shifter for the student body composition is related to differences in price sensitivities by SES. A school that is not operating for an extended period is likely to generate a relevant negative shock on parents' utility, and can potentially induce families to switch to the private sector. This was, indeed, one of the concerns that experts and the Government had after the extended strike. Differences in preferences will generate different response patterns to the strike, changing the student body composition of the schools from which the students leave and the ones that receive them.

These ideas can be appreciated in the event studies presented in Figure 5. Panel (a) shows that public schools that had a very high (high) baseline predicted probability of experiencing an intense strike (using Strike Exposure Index II) strike lost 17 (9) students on average. Panel (c) and (d) show that private schools located close to public schools that had a very high (high) baseline predicted probability of experiencing an intense strike (Strike Exposure Index II) increased the shared of high SES students enrolled in their schools.

A6.2 Instruments for price and quality

Relevance Condition ($Cov(\bar{p}_{jt}, Z_{inst}) \neq 0$ and $Cov(\bar{\mu}_{jt}, Z_{inst}) \neq 0$). Figure 4 shows that the wages for public school teachers almost doubled in the period 2012-2018. This can have a large effect on

[49]Panel (b) in Figure A4 shows school survey data for the dates for the school applications, most of which take place after the strike was over (October to January).

both quality and prices. With respect to quality, this reform significantly increased the recruiting standards and incentives in the hiring process, both due to higher average wages and merit payments based on a holistic evaluation process. Previous research has shown that one standard deviation increase in teacher quality raises math achievement by 0.15 to 0.24 standard deviations and English by 0.15 to 0.20 standard deviations (Rivkin et al., 2005; Rockoff, 2004; Kane and Staiger, 2008). Bobba et al. (2019) study part of this reform, taking advantage of arbitrary cutoff rules that determine teacher wages in rural areas. They find that offering higher wages in Peru increases the number of applications, the recruited teachers' scores in teachers' evaluation tests, and teacher retention. They also find significant improvements in students' test scores, suggesting that the reform was effective at increasing the quality of instruction.

This evidence suggests that the reform can have a positive effect on the quality of public schools, and a potentially negative effect on the quality of the private ones, which may now find more competition to hire effective teachers. Given that teacher wages are one of the most important components of the schools' expenditures in Peru, higher wages in the public sector should shift costs for private schools if they compete to hire teachers in the same labor market.

Table 4 presents the first stage for the share of high SES students in the school, using the three types of instruments and their interactions. The F test rejects the null hypothesis that the first-stage coefficients are zero.

Exclusion restriction ($Cov(Z_{inst}, \xi_j) = 0$). For instruments (i)-(ii), the exclusion restriction is given by the design of the teacher payment reform. The initial increase in base wages was unexpected and made effective immediately. Then, there is time and geographic variation in the opening of vacancies for the new contracts (therefore, higher wages). Contracts were rationed for a gradual implementation of the policy. Requests for new contract vacancies were decided by the school district. There are differences in the timing of vacancies openings. Appendix **??** provides evidence that this timing within the school district is unrelated to school characteristics, suggesting that this is unlikely to be correlated with school unobservables. As mentioned in Section 3, the design of the centralized application system also adds variation to whether the vacancies are filled. To avoid the problem of high-quality teachers self selecting into schools with certain unobservables, I use a measure of the average payments close to the school, excluding their own wages.

For instruments (iii)-(vi), the argument for the exclusion restriction is similar to the one in the previous section: residential sorting, school entry, and, in this case, workers sorting decisions are taken previous stages and take time to adjust. Then, they do not respond fast to current local school shocks.

A6.3 Using RDs to identify price parameters

Relevance Condition ($Cov(p, Z_{inst}) \neq 0$). To understand how this approach works, let me start by defining a simplified version of the choice model[50]:

$$u_{ij} = x_j\beta - \alpha p_{ij} + \xi_j + \epsilon_{ij}$$

Among the applicants of the scholarship, the price varies with a discontinuous function of the student test scores, where c_1 and c_2 are the tests cutoffs:

$$p_{ij} = \begin{cases} p_j/4 & \text{if } x_i^{sc1} \geq c_1, x_i^{sc2} \geq c_2. \\ p_j & \text{otherwise} \end{cases} \tag{34}$$

I assign applicants to bins of students with similar tests, using bandwidth h, and number of bins K_1 and K_2 for each test. To do so, I define:

$$b^L_{k_1} = c_1 - (K_1 - k_1 + 1)h, \text{ and} \quad b^L_{k_2} = c_2 - (K_2 - k_2 + 1)h$$

Each bin $\boldsymbol{k} = (k_1, k_2)$ includes students with scores:

$$b_{k_1} < x_i^{sc1} \leq b_{k_1+1}, \quad b_{k_2} < x_i^{sc2} \leq b_{k2+1}$$

Using the bins, I can build market shares conditional on bins for students with similar tests $\boldsymbol{k} = (k_1, k_2)$, given by ($b_{k_j} < x_{scj} \leq b_{k_j+1}$):

$$s_{j\boldsymbol{k}} = \frac{exp(x_j\beta - \alpha p_{j\boldsymbol{k}} + \xi_j)}{1 + \sum_l exp(x_l\beta - \alpha p_{l\boldsymbol{k}} + \xi_l)} \tag{35}$$

Following Berry et al. (1995), the shares for each bin $\boldsymbol{k}$ can be inverted and expressed as[51]:

$$\ln s_{j\boldsymbol{k}} - \ln s_0 = x_j\beta - \alpha p_{j\boldsymbol{k}} + \xi_j \tag{36}$$

Given this expression for the difference of shares, I can use the indicator for being over the

[50] Part of the identification argument follows Bucarey (2018)

[51] Here I am presenting the case for a discrete choice problem with an outside option. There is no outside option in my model. However, both problems are analogous in terms if the identification argument.

cutoffs $z_{\boldsymbol{k}} = 1(x_i^{sc1} \geq c_1, x_i^{sc2} \geq c_2)$ as a valid instrument for α, the parameter on price. In terms of the instrumental variables literature, this instrument satisfies the properties for a valid instrument. This follows directly from Equation 34, as crossing the threshold reduces the cost of tuition in 3/4.

Exclusion restriction ($Cov(Z_{inst}, \xi_j) = 0$**).** The argument for the exclusion restriction follows from the discontinuous design of the instrument. While there can be a correlation between the unobserved school quality (ξ_j) and the test scores (for example, better-performing students may come from families with a stronger preference for the school unobserved quality), the treatment assignment at the threshold should be "as good as random". The cutoff scores are not revealed to students, and they cannot perfectly manipulate their scores as there is probably some randomness of student performance. Students are not allowed to retake the test if they failed to pass the first stage of the process.

In section I show how to use this instrument as an additional moment for identification of the parameters.

A7 Estimation Details

A7.1 Dealing with Selection in the Scholarship Application Decisions

I propose a simple model for selection into de scholarship application to implement a correction when I use the RD sample. The binary variable Z_i indicates if the student is offered a scholarship. Expected utility conditional on the scholarsip, but before the realization of ϵ_{ijt} is given by[52]:

$$w(Z_i \mid x_i, \boldsymbol{v_i}, loc_i) = E\left[\max_{j \in \Omega_{it}} U_{ijt}(\boldsymbol{p_i}|Z_i) \mid x_i, \boldsymbol{v_i}, loc_i\right] \tag{37}$$

Where $\boldsymbol{p_i}$ is a vector of prices in which each element is:

$$p_{ij} = \begin{cases} p_j/4 & \text{if } Z_i = 1 \text{ and } j \in \Omega_{network} \\ p_j & \text{otherwise} \end{cases} \tag{38}$$

Where $\Omega_{network}$ is the set of schools for which the scholarship can be used.

Scholarship assignment. The scholarship assignment rule is a discontinuous function of the student test scores, where x_i^{sc1} and x_i^{sc2} are the scores for the academic and leadership tests and c_1 and

[52]The expression for the max is given by: $E\left[\max_{j \in \Omega_{it}} U_{ijt}\right] = \log\left(\sum_{j \in \Omega_{it}} \exp(U_{ijt})\right)$

c_2 are the cutoffs.

$$Z_i = \begin{cases} 1 & \text{if } x_i^{sc1} \geq c_1, x_i^{sc2} \geq c_2 \\ 0 & \text{otherwise} \end{cases} \tag{39}$$

I assume that parents do not know the potential results of their kids in these tests, so they do not select into applying for the scholarship based on their individual probabilities of getting it. Parents' share the prior that their children can get the scholarship with probability π if they apply and 0 otherwise.

$$Pr(Z_i = 1) = \begin{cases} \pi & \text{if } a_i = 1 \\ 0 & \text{if } a_i = 0 \end{cases} \tag{40}$$

Application decision. I assume that there is an application cost c_i that I parametrize as a type-specific fixed cost and a linear function of D_i^a, the distance from the student house to the location in which they have to take the tests for getting the scholarships. D_i^a is excluded from Equation 37, implying that distance is a valid instrument for scholarship application.

$$c_i(a, x_i, D_i^a) = a \times \left(\bar{\eta} + \eta_e x_i^e + D_i^a \times (\bar{\eta}^d + \eta_e^d x_i^e)\right) \tag{41}$$

Then, the application decision is given by:

$$A_i = \arg\max_{a \in \{0,1\}} V_i(\theta_2, \eta, x_i, \boldsymbol{v_i}, loc_i) \tag{42}$$

Where

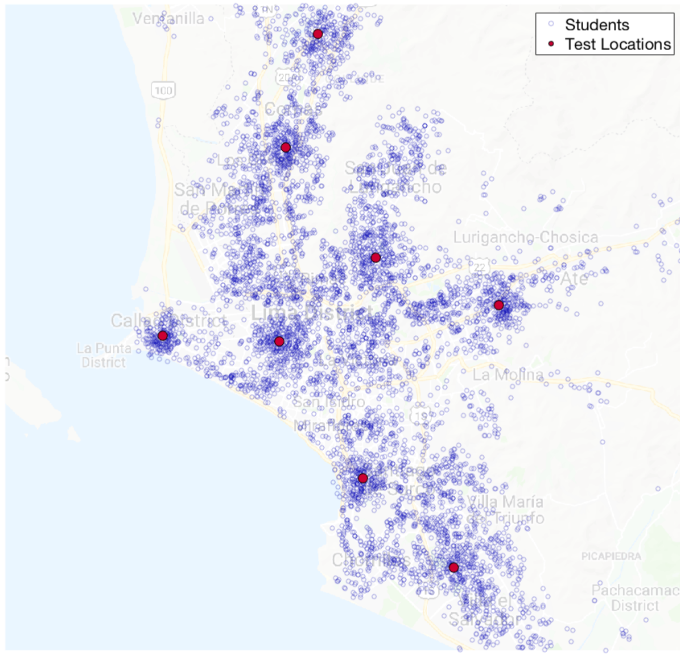
$$V_i(\theta_2, \eta, x_i, \boldsymbol{v_i}, loc_i) = \pi \times \left[w(Z_i = 1 \mid x_i, \boldsymbol{v_i}, loc_i) - w(Z_i = 0 \mid x_i, \boldsymbol{v_i}, loc_i)\right] - c_i(a, x_i, D_i^a) \tag{43}$$

Likelihood function. The likelihood of the application decision and choices is given by:

$$\ell_i(\theta_2, \eta, x_i, \boldsymbol{v_i}, loc_i) = Pr\left(\arg\max_{a \in \{0,1\}} V_i(\theta_2, \eta, x_i, \boldsymbol{v_i}, loc_i)\right)$$

Simulation with selection correction. For each student, the simulated characteristic is obtained by integrating over the individual choice probabilities. These probabilities are functions of the demand parameters, the students' characteristics, the students' test scores, which define eligibility for the scholarship, and as a consequence, the prices that the student faces:

Figure A6: Applicants' to the Scholarship and Tests' Location



$$y_{l,i}^{sim}(\theta_2, x_i, x_i^{sc1}, x_i^{sc2}, \bar{v}_{i,s}^*) = \sum_j^{N^f} s_{ijt}^{nk}(\theta_2, x_i^{sc1}, x_i^{sc2}, \bar{v}_{i,s}^*) \times y_{jt}$$

The simulation is made taking 500 draws from the unobservables. The unobservable preferences for quality and peers are drawn from the lognormal distribution, conditional on the parameters. The selection correction is implemented using a student-specific individual posterior distribution for the preferences for schools in the network, instead of the normal distribution.

This posterior is based on Equation 42, and is computed conditional on the parameters as a function of the students' application decision, type, and location. More details on how the posterior is calculated are provided in Appendix **??**. The intuition for doing this is simple: if a student decided to apply to the scholarship when the predicted probability if doing so was low given the student's observables, the student likely had a strong preference for the schools in the network. I correct for this type of selection by updating the distribution of preferences with the information from the selection equation.

A8 Demand Estimation of Geographically Differentiated Products in Large Markets

There are several markets in which geography plays an important role in consumer demand (healthcare, education, retail stores, restaurants, etc.). In these markets, it is important to allow for preferences over the distance from the consumers' location to the point of consumption. This implies that there is a product characteristic that varies by product-consumer pair (distance as a function of the pair locations). If we also allow tastes to be heterogeneous over observed and unobserved characteristics, we need to integrate choice probabilities over three dimensions to simulate market shares. While the dimensionality of the observable and unobservable heterogeneity is likely to be fixed across markets, the number of possible locations will grow as markets get larger. This implies that the number of product-consumer pairs' locations grows exponentially, and so does the required memory allocation.

To overcome this problem, I divide the largest market (Lima) into four sub-markets, as depicted in Figure A7. I use structures that naturally separate sections of the city into the northern, western, eastern, and southern sub-markets.

However, these structures do not fully isolate the markets, and families can actually cross these imaginary boundaries to attend schools in a different sub-market. This can be problematic for the estimation of the model. For example, take a school that is just to the right of the boundary with another sub-market. If we just split the schools and consumers and assign them to the corresponding sub-markets, then when we calculate the market shares we would miss an important fraction the potential consumers that could have gone to that school, while also ignoring potential competitors (the schools to which the students close to the boundary could have gone). These two problems will not necessarily compensate each other for every section of the border.

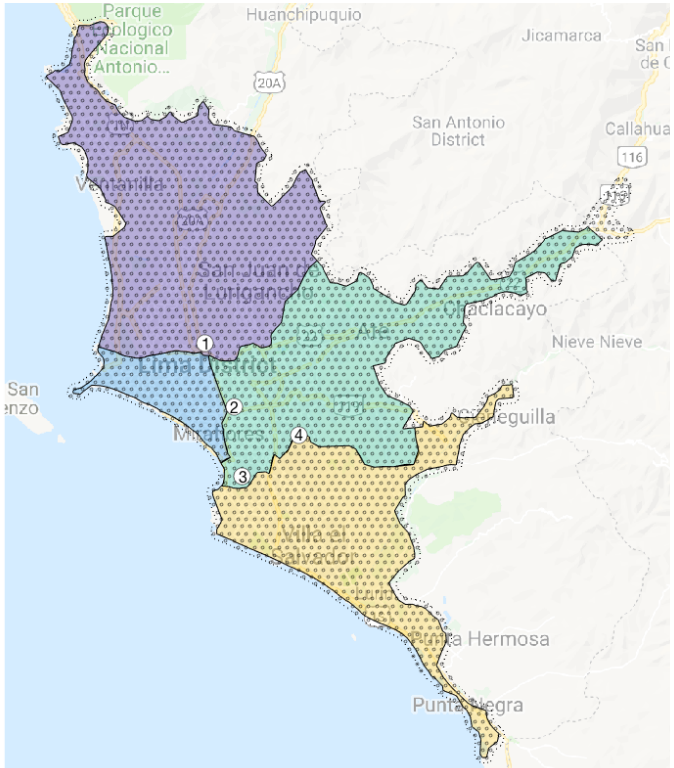
To deal with this problem, I correct the market shares and the weights for the distributions of families of different types across the grid of nodes that I define for each market. This would be trivial if I had all the students in the market geocode. Even though I have microdata (including address) for all the students in the market, I only have geocoded locations for a sample that corresponds to 81% of the students.

Recall that in the model families differ in three dimensions: their observable type (k), their unobservable type (v_i), and the node in which they are located (N_i). Nodes for Lima are shown in figure 1. Simulated and observed market shares are calculated in the following way:

- Simulated Shares ($\mathcal{S}_{jt}(\theta_2, \xi)$)

$$U_{ijt} = \tilde{\beta}_i^{\mu} \mu_{jt} - \tilde{\alpha}_i p_{jt} + \bar{z}'_{jt} \tilde{\beta}_i^{z} + \tilde{\beta}_i^{d} D_{ij} + r'_{jt} \beta^{r} + \xi_{jt} + \epsilon_{ijt}, \qquad \text{where } \tilde{\beta}_i = \bar{\beta} + \sum_k \beta_k + \beta_u^q v_i$$

Figure A7: Sub-markets for Lima



Note: Natural structures that split the city include: 1) Rimac River, 2) Via Expresa Urban Highway, 3) Military Base, and 4) Villa Maria Hills. The small circles show the equidistant location nodes in the city, which are 1 km appart from each other.

Where

$$\mathcal{S}_{jt}^{ink}(q,p,z,\xi) = \left(\frac{\exp(U_{ijt}(q,p,z,\xi))}{\sum_{l \in \Omega_{it}} \exp(U_{ilt}(q,p,z,\xi))} \right)$$

$$\mathcal{S}_{jt}(q,p,z,\xi) = \sum_{k}^{K} \sum_{n}^{N_m} \sum_{i}^{v_i} \mathcal{S}_{jt}^{ink} \cdot w_i^v \cdot w_{nk}^m \cdot \pi_k^m \tag{44}$$

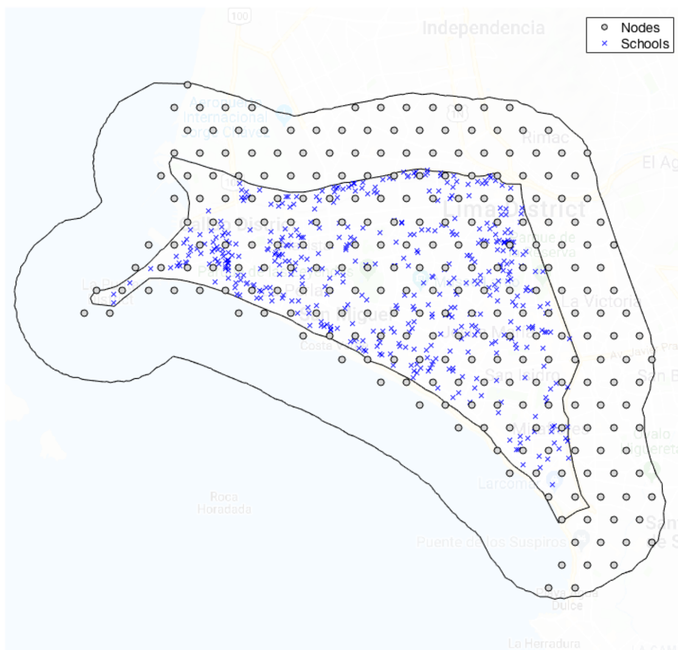
- Observed Shares (S_{jt})

$$S_{jt} = \frac{N_{jt}}{\sum_{l \in \Omega_t} N_{lt}} \tag{45}$$

I make the following adjustments to the basic structure:

1. **Buffer:** First, I generate a 3 km buffer around each sub-market. Figure A8 shows an example for one of the sub-markets, Lima West. For each sub-market, the set of schools will be the ones strictly within the boundaries of the sub-market. For the nodes, I will consider the ones within the sub-market in addition to the ones that are within the buffer zone. Figure A9 shows how the sparsity structure of the simulation for the market shares for Lima after the division.

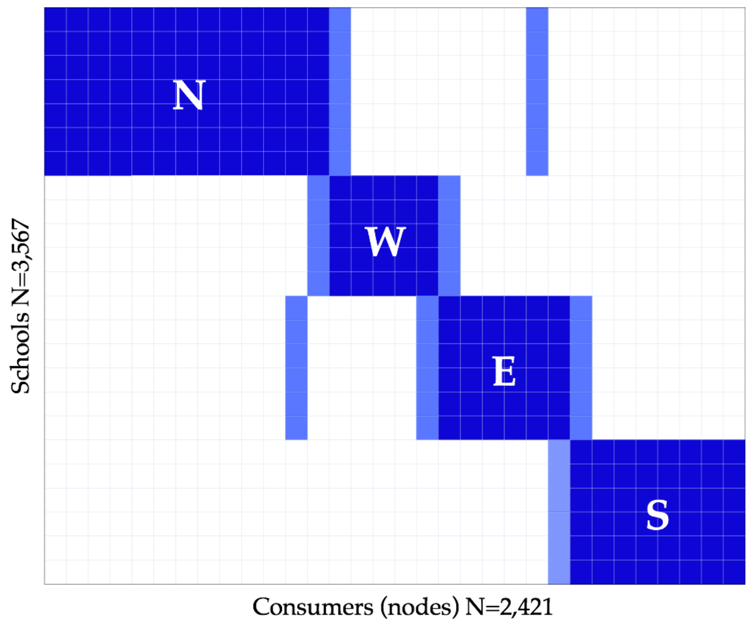
Figure A8: Sub-market: West Lima



2. **Adjustment to the Observed Market Shares**: For each school in the sub-markets, I calculate the share of families enrolled in that school that live within the boundaries of the market+buffer. If the number of families enrolled in that school in the geocoded sample, represent less than 25% of the total enrollment, I impute this number based on the distance to the boundary and whether the school is public or private. I multiply this share for the total enrollment from administrative data. The resulting numbers will be the school enrollment considered to calculate the adjusted market shares at the sub-market level.

$$N_{jt}^{h*} = \frac{\textit{N of students in school j from submarket h (geo sample)}}{\textit{Total N of students that go to school j in Lima (geo sample)}} \times N_{jt}\textit{(school records)}$$

Figure A9: Sparcity structure for sub-marekts in Lima



Note: This figure shows the school-node pairs for Lima. Simulated market shares and moments are obtained by taking a weighted average of individual choice probabilities over location nodes, types and unobservables. The dark areas show the schools and nodes in each sub-market. The light areas show the nodes that belong to the buffer for each market, which are also included for the simulations.

$$S_{jt}^{*} = \frac{N_{jt}^{h*}}{\sum_{l \in \Omega_t} N_{lt}^{h*}}$$

3. **Adjustments to the Weights to integrate over the individual choice probabilities:** I adjust the weight for the following distributions:

   - Distribution of types over the nodes:

$$w_{nk}^{h*} = \frac{N \text{ students of type } k \text{ in node } n \text{ that go to school in } h \text{ (geo sample)}}{N \text{ students of type } k \text{ in node } n \text{ (geo sample)}} \times w_{nk}^{m}\text{(Census)}$$

   - Agregate distribution of types in the market:

$$\pi_{k}^{h*} = \frac{N \text{ students of type } k \text{ that go to school in } h \text{ (geo sample)}}{N \text{ students of type } k \text{ in Lima (geo sample)}} \times \pi_{k}^{m}\text{(Census)}$$

These adjustments rely on two assumptions:

- **Most families do not travel more than 3 km** In the microdata, a 3 km distance between the family's location and the chosen school corresponds to the 96th percentile of the geocoded sample.

- **The geocoded sample is a random sample of the population:** To geocode the location of the families, I match their National IDs with the census microdata, and obtain the latitude and longitude of the census block in which the family lives. I was able to do this for 81% of the students. Reporting the ID in the census is not mandatory, but most people answer that question (close to 93%). This does not appear to vary by SES. However, the ID is not verified, so apart for the 7% that do not report their ID, some observations do not match because of typos, incomplete numbers, etc. I don't expect this to vary by SES.

A9 Counterfactual Policies Budget Assumptions

It is estimated that the optimal Operational Cost (as a % of Sales) is $\leq$ 63% (or EBITDA Margin of $\geq$37%), which is reached at approximately 90 schools. As of 2019, there are 54 schools and by 2023, the Company expects to reach 90 schools. In order to reach 90 schools, Innova requires to invest $\sim$\$540 million in CapEx:

- Each school is built on a land area of $\sim$ 5,000 m2, whose cost is US\$ 3 million on average.

- The construction and equipment are carried out in 4 stages, being the first stage the most expensive (US\$ 1.8 million) and the remaining stages are less expensive (US\$ 500 thousand each).

- In the first stage, the common spaces and a large part of the classrooms are built. In the rest of the stages, only new classrooms are added.

- A school in its fourth stage has an accumulated investment of US\$ 6.3 million (average), 40 classrooms and capacity for 1,200 students.

- The accumulated investment by 2026 is expected to be US\$ 542 million.

A10 Other Outcomes for Counterfactual Simulations

A11 Market Construction and Description

??

Appendix Online Appendix

O-1 Summary Statistics by Market

Link to the online appendix