DISCUSSION PAPER Nº IDB-DP-01028

COVID-19 Exodus:

Parent preferences for public schools in Peru

Gregory Elacqua
Nicolas Figueroa
Andrés Fontaine
Juan Margitic
Carolina Méndez Vargas

Inter-American Development Bank
Education Division

August 2023

COVID-19 Exodus:

Parent preferences for public schools in Peru

Gregory Elacqua
Nicolas Figueroa
Andrés Fontaine
Juan Margitic
Carolina Méndez Vargas

Inter-American Development Bank
Education Division

August 2023

JEL Codes: A20, D12, D83, I28

Keywords: School Choice, Centralized Assignment Systems, Private-Public Schooling, COVID-19

COVID-19 Exodus:

Parent preferences for public schools in Perú*

Gregory Elacqua† Nicolas Figueroa‡ Andrés Fontaine‡ Juan Margitic†

Carolina Méndez Vargas†

Abstract

In 2020 in Peru, the Ministry of Education (MINEDU), in response to the COVID-19 pandemic, organized a centralized assignment mechanism that allowed thousands of students at multiple levels of education to move from the private sector to the public sector due to an unprecedented rise in demand. Exploiting the randomness in the assignment of students to their new schools, we causally estimate which public school characteristics families that had decided to study in the private sector before COVID-19 value the most and how preferences for school attributes change after parents experience public schools. We find that families care about the distance to the assigned school and the relative academic and peer quality with respect to their school of origin. Parents weigh features such as distance to school and peer demographics differently when deciding whether or not to remain at the assigned school. These findings provide insights into how governments can strengthen the supply of public schooling.

JEL Classification: **A20, D12, D83, I28**

Keywords: School Choice, Centralized Assignment Systems, Private-Public Schooling, COVID-19.

*We thank the Ministerio de Educación of Perú for providing the data employed in this study. We want to thank Diana Vásquez Villanueva, Benjamín Madriaga and Jadira Sánchez Córdova for their support in implementing this project and performing the analysis. We gratefully acknowledge the Inter-American Development Bank for funding the research presented herein. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent. The authors have no conflicts of interest or financial and material interests in the results. All errors are our own.

†Inter-American Development Bank

‡Pontificia Universidad Católica de Chile

1 Introduction

In 2020 in Peru, the education sector was shocked by an unprecedented rise in the demand for spots in the public sector by students looking to migrate from the private sector. The exodus of more than a hundred thousand students into the public sector was likely fueled by a combination of factors: extended school closures due to Covid-19, lack of adaptability by the private schools to move to online education, and the unwavering cost of paying for private education during an economic crisis. This also follows similar trends observed in other countries in the region (Elacqua et al., 2022).

In response to this rapid shift in family preferences for public schooling, the Ministry of Education (MINEDU) designed a centralized assignment mechanism to enroll these students in public schools: the 2020 Exceptional Enrollment (*Matrícula Excepcional (ME)*) process, which created an opportunity to explore the relative weight parents give to different attributes of public schools when deciding whether or not to enroll and later remain in the assigned school. The 2020 ME process included a randomization component that we exploited to causally estimate family preferences over school characteristics at the moment of choosing to transfer to a public school. Moreover, for the families that migrate, we analyze the preferences when, one year later, families choose whether to remain on the public school or go back to a private one. This unique opportunity allows us to identify some dimensions of public education that this group of parents value, and governments could reinforce or dispel misconceptions about the public education system[1]. Strengthening public education has many potential benefits: functioning as a safety net for all members of society (Singh et al., 2014), as a large-scale adaptable apparatus that allows all students to cope with times of need (such as with COVID-19), as an equalizer of inequitable family backgrounds via reducing economic segregation (Murillo et al., 2018; Elacqua, 2012; Balarin, 2015).

Our paper first presents a theoretical framework that guides our empirical analysis

[1] *Este País*, May 1^{st}, 2014

on parent preferences on public education offering. The theoretical framework consists of a two-period model where students are offered a chance to accept or reject their public school assignment in period one; in period two, they decide whether or not to stay at their assigned school. Like our theoretical model, our empirical framework explores the decisions to accept public school allocations and remain at the assigned school in separate stages. This analysis allows us to understand the educational dimensions that families care about when accepting the assignment and deciding to stay after experiencing the school. Moreover, it sheds some light on the changes in the intensity of the preferences over some components.

Our identification strategy exploits the random component from the centralized assignment of students organized to allocate the unprecedented demand for public education and helps us relate decisions to accept and stay at the assigned schools with observable school characteristics. These observable attributes include information on the individual distance to the school, school size, average school characteristics such as class size, math standardized test scores, peer demographics, and teacher experience and contract type. In addition, we include information on the prices of the private schools of origin as a proxy for socioeconomic status.

We find that parents are more likely to accept their assigned public school at schools closer to their homes, with relatively better peer demographics (as measured by mean parents' education) and higher average math achievement compared to their private school of origin. We find that not only does the absolute distance to the school matter in accepting the assignment, but also the relative distance between the student's home and the assigned school compared to distance between the student's home and the private school of origin. This suggests that parents anchor their decisions using the previous school attributes as a benchmark. In addition, we find weak evidence that parents are more likely to accept larger schools and also ones with lower student-teacher ratios (STR).

We also find evidence that parents can positively change their perceptions about public schools once they experience them and have more information that allows them to compare the assigned school with their previous private school and other schools. Once parents are enrolled in a public school, features that seemed unimportant when deciding to accept their assigned school gain relevance in explaining their decision to stay or leave. Although we acknowledge that the sample of parents who decide to accept their school assignment is systematically different from parents in the first stage, this population of parents is of interest to policy-makers. As mentioned, distance to the school, relative peer demographics and school quality remain relevant. However, we also find that the relative difference in distance to the school does not affect the likelihood of remaining in the assigned school, suggesting a change in the “reference point" for the decision. We also find a higher coefficient associated with peer demographics. We also find that the difference in the share of novice teachers reduces the likelihood of remaining in the assigned school. At the same time, parents slightly prefer *larger* schools (with more students) compared to their previous private school.[2]

In the literature on parental preferences for schools, several studies have examined various factors that influence parents' choices. For instance,Hastings et al. (2005), Jacob and Lefgren (2007), and Hofflinger et al. (2020) have explored parents' preferences for school characteristics such as academic quality, proximity, and religious instruction. These studies found that parents from different socioeconomic backgrounds prioritize these factors differently. Specifically, parents with a lower socioeconomic status tend to prioritize proximity, while parents with a higher socioeconomic status prioritize quality and religious education. Similar findings were reported by Abdulkadiroğlu et al. (2020) and others, who found that parents generally prefer schools with higher test scores.

In terms of revealed preferences, Harris and Larsen (2017) conducted a study in New Orleans and found that parents preferred schools with higher test scores and value-added measures. However, they also considered non-academic factors such as after-school care,

[2] We include a literature review section in our Appendix **??** on the paper's broad connection to the literature on parent preferences for schooling dimensions.

extracurricular activities, and proximity to home. Similarly, Glazerman and Dotter (2017) emphasized the significance of commuting distance, school demographics, and academic indicators in school choice, highlighting the heterogeneity of preferences across parent demographics.

Our study aligns with previous research findings. We also find that families prefer higher-performing schools with favorable peer demographics and proximity to their residence, consistent with Jacob and Lefgren (2007) and Burgess et al. (2015). Furthermore, our results support the observations of Harris and Larsen (2017) regarding the consideration of both academic and non-academic attributes in school selection, as well as the findings of Glazerman and Dotter (2017) concerning the role of commuting distance, peer demographics, and academic indicators.

This paper contributes mainly to two different strands of the literature. Firstly, it adds to the body of evidence on exploiting information on parents' preferences and the randomization induced by centralized assignment mechanisms to learn about school choice, building on studies such as Abdulkadiroğlu et al. (2017), Glazerman and Dotter (2017), Kapor et al. (2016), and Agarwal and Somaini (2018). However, our study focuses on a unique case where preferences for the randomization algorithm implemented were imputed based on distance, contrasting with the option for parents to submit their choices.

The literature reveals that parental preferences for schools encompass multiple dimensions, including academic quality, proximity, socioeconomic composition, and non-academic factors. Our study contributes by examining preferences through a natural experiment, capturing the perspectives of parents who transitioned from private to public schools. By understanding how these parents perceive the public sector and what educational dimensions they value, policymakers can design effective strategies to strengthen public schooling and address potential stratification concerns.

Secondly, this paper offers a unique perspective for governments interested in strengthening public schooling by evaluating the preferences of parents who transitioned from private to public schools due to pandemic-related factors. These families, predominantly from middle and low-income backgrounds, represent a population that has shifted from the public sector over the last few decades. Gaining insights into how these parents perceive the public sector and what educational dimensions they value is crucial for governments to design and implement effective policies to strengthen public schooling. In particular, we can offer a glimpse into their preferences from a revealed preference perspective versus a stated preference perspective. Generally, the stated preference literature suggests that parents declare on surveys that they are most interested in a school's academic quality (e.g., test scores, teacher quality). However, the type and magnitude of these preferences may vary by the survey question and parent demographic characteristics. Previous studies, such as Schneider et al. (1998) and Kleitz et al. (2000), indicate that parents often prioritize teacher quality, high test scores, educational quality, class size, and safety when making school choices.

This paper is structured as follows: Section 2 describes the context and critical events that led to the 2020 ME process as well as the data used in this paper; Section 3 introduces the theoretical model that guides our empirical analysis; Section 4 describes our empirical analysis including the identification strategy we employed, the main results, and a heterogeneous analysis; finally, Section 5 discusses the findings and concludes.

2 Institutional Background

The pandemic caused by the COVID-19 virus led governments throughout Latin America to close schools due to the increased risk of contagion and lack of available vaccines. Countries in the region closed schools for an average of 158 days, in stark contrast with the OECD average of 57 days[3]. Clearly, the schools in the region were not equipped with the necessary resources to withstand protocols that could minimize the risk of contracting the disease. However, most students and schools were also well equipped for a transition

[3]Source: own elaboration based on UNICEF. Period: 11/3/2020 - 2/2/2021.

to a digital learning environment. There were critical differences in the schools' capacity to transition to this environment and the capacity of the students to adapt. Schools faced critical constraints such as the availability of technological resources as well as the capacity of the teachers to adapt to remote instruction[4]. On the other hand, students were limited in the resources they had access to at home to engage in online instruction, such as technological devices, internet access, a physical space to study, etc. (Abizanda et al., 2022).

In Peru, the pandemic had severe consequences. It has the world's highest COVID death rate, far higher than any of its neighbors and twice the rate of the United States. COVID officially caused nearly 6,470 deaths for every 1 million Peruvians[5]. In addition, unemployment increased considerably. During the second quarter of 2020, according to *Instituto Nacional de Estadística e Informática* (INEI), the categories of informal dependent and independent workers were the most affected, registering declines of 60% and 35%, respectively. In absolute terms, this meant the annual loss of approximately 4.8 million jobs in the informal segment of the labor market.

In the context of education, private schools were reticent to lower their fees[6] and even threatened to expel students in extreme cases[7,8]. Likely due to a combination of the factors mentioned above and economic hardships imposed by the economic impact of the pandemic, the public education sector saw an unprecedented increase in the number of students who requested to be transferred from their private schools to public schools. In response, the government took the initiative to design a system that could accommodate many transfer requests while preserving the principles of equity and efficiency. MINEDU, with the technical assistance from the Education Division at the Inter-American Development Bank, designed a centralized assignment mechanism using the Deferred Acceptance (DA) algorithm to allocate the students' to the available vacancy closest to their residence.

[4]Infobae, July 28 2022
[5]Our World Data
[6]El Comercio newspaper, April 26 2020
[7]Revista Matg, September 25, 2020
[8]Otras Voces en Educación, May 28, 2020

The government invited students to participate by registering in a web portal created during this process. Rather than eliciting the parents' preferences for schools, students were told that the process would accommodate them at the nearest public school with a vacancy. In other words, the system imputed linear distances between the family's reported address and the nearby schools (within 5 km) with vacancies, ranking the closest school as the highest preference. For instance, we will refer to *first preference* to the *closest* school to the student and progressively list closer schools as more preferred by students. As mentioned, families did not choose these vacancies; they were imputed during the process based on the distance from the students address. The *Matrícula Exepcional* (Exceptional enrollment) 2020 (ME 2020) program proceeded as follows: first, families registered and submitted their information on the platform, which included their home address [9]. The process would continue with the creation of rankings based on proximity to address for each student. Then running the DA algorithm would take both preferences and a list of school vacancies and assigned students. After receiving the offer, students were given the option of accepting or rejecting. If the reason for rejecting was due to an incorrect address, students were given the opportunity to correct it in the system and reenter the process in the second round. MINEDU would rerun the algorithm to compute the new allocation for those who rejected their first-round assignment. This process was iterated twice, resulting in a total of 3 rounds until every registered student was assigned to a school. Most of those who rejected their initial allocation in the first round and reentered the system reported issues such as misreported addresses in the registration process. In total, 125,295 students participated in ME 2020.

2.1 Data

Our starting sample includes 125,295 students who participated in the ME 2020 process. They belongs to a group of affected parents who responded to the shock imposed by

[9] Families were able to submit addresses that were different from their official ID, but it needed to be their permanent address

the COVID-19 pandemic and were mainly enrolled in lower-performing private schools in urban settings. These parents likely responded to the extended school closures, lack of adaptability by the private schools to online education, and continuing private school tuition fees when deciding to enter the process.

We use data provided by the MINEDU on student background, their choices, and allocations to investigate the possible factors that determined student and family decisions. We obtained individual records of students' enrollment status and teacher characteristics from 2019 to 2021 from the *Sistema de Información de Apoyo a la Gestión de la Institución Educativa* (SIAGIE) database, and data on average school characteristics from multiple databases: measures of parent education and schools' standardized academic performance from 2014, 2015, 2016, 2018, and 2019 from the *Evaluacion Censal de Estudiantes* (ECE) and information on staff and facilities from the School Directory (*Padron Escolar*) as well as the School Census (*Censo Escolar*) from 2014 to 2021.

We are interested in analyzing two stages of the process. The first stage is when the parents accept or reject the assigned public school. Second, when families decide to stay or leave the previously assigned and accepted school. In both stages, we exclude students who participated in rounds 2 and 3. We excluded them because they faced inherently different pools of potential schools for their assignment. Some of these students' preferences were incorrectly assumed, given incorrectly specified addresses. Also, we will consider only those students who come from private schools [10].

In the second stage, we filter out those students who did not accept their assignments from round 1. We do not incorporate these students in the study because they never experienced their assignment to the new schools. However, it is important to mention that the majority of these students were the ones who misreported their addresses during round 1.

We also define *accepting* the school assignment if the student accepted and virtually

[10] $5,761$ or 4.6% of our sample were enrolled in public schools

attended their assigned schools. In our sample, we observed some students who, although they accepted their assigned school, never enrolled in it [11]. For these cases, we assume that these students *de facto* rejected their assignment, and recode their responses in the first stage. In addition, we observed some students who accepted their assigned school, but appear to be last registered in 2020 at another school after the ME allocation took place. For instance, a student who accepts her assignment from the ME2020 process, but appears to be registered at another school in September, will be considered as treated. Specifically, we consider that student experienced some months at the assigned school given that they theoretically tried the assigned school in June and then decided to move in September.

Also, we will restrict our sample to consider students attending a school in 2021 within 50km of their assigned school by the 2020 ME process. We believe that students could switch schools for reasons not related to the schools' characteristics, such as moving. Therefore, to control for such scenarios, we only consider students registered in 2021 near their originally assigned schools.

After imposing all of our sampling conditions, we ended up with 61.43% of our original sample of students who participated in the 2020 ME process or 76,964 students in the first stage and 51.04% or 63,962 in the second stage [12]. We can observe in Figure 1 the flow of students from our first stage to our second stage. The dotted line separates the sample between our first and second stage analysis. The green arrows depict the decisions of students to continue in the assigned school from the ME2020 process; meanwhile, the red arrows depict exits. The first row depicts the decision of students to leave their assigned school during the ME2020 process. The second row of arrows defines those students who, although seemingly having accepted their assignment, never registered in their assigned

[11] We define these students as those who accepted their assignment but never formally enrolled in these schools

[12] We want to highlight that our regression samples differ with respect to the reported samples in this paragraph. This is due to two main reasons: First, at our preferred analysis level, many district by grade and level samples lack variation in the acceptance / rejection to be included in the analysis. About 16k students are dropped as a result. Second, some covariates in our analysis contain missing values for some students. Although the affected number of students remains small (about 3k), we're exploring missing value imputation methodologies to address this issue.

schools. We re-code their responses to *de-facto* rejections as mentioned. Rows three and four of the arrows define decisions to move out of the assigned school after having experienced at least some months in these schools. Finally, we can observe the geographical distribution of students and schools that form part of our analysis in Figure 2.

Figure 1: Student flow before and after 2020 ME

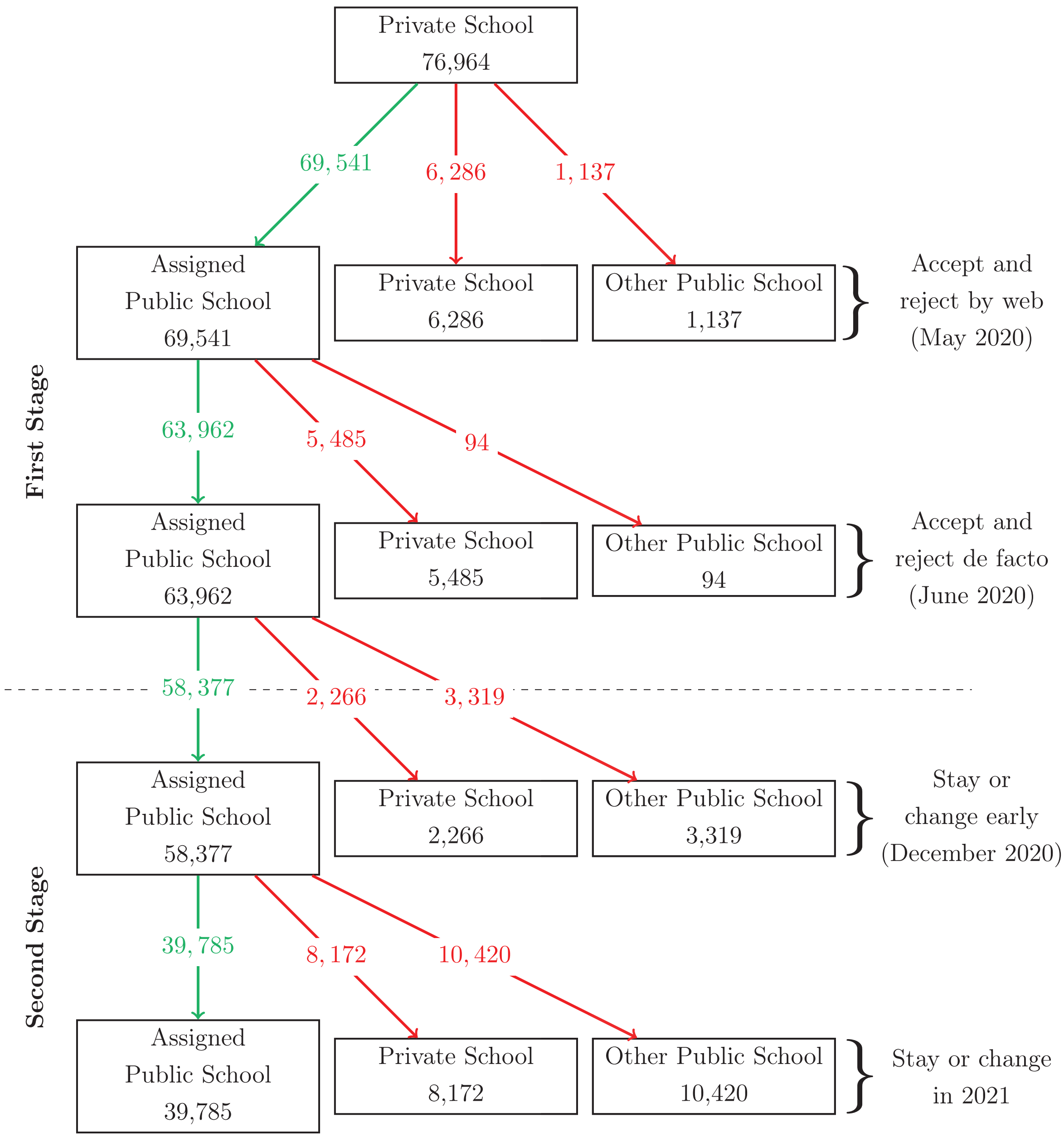
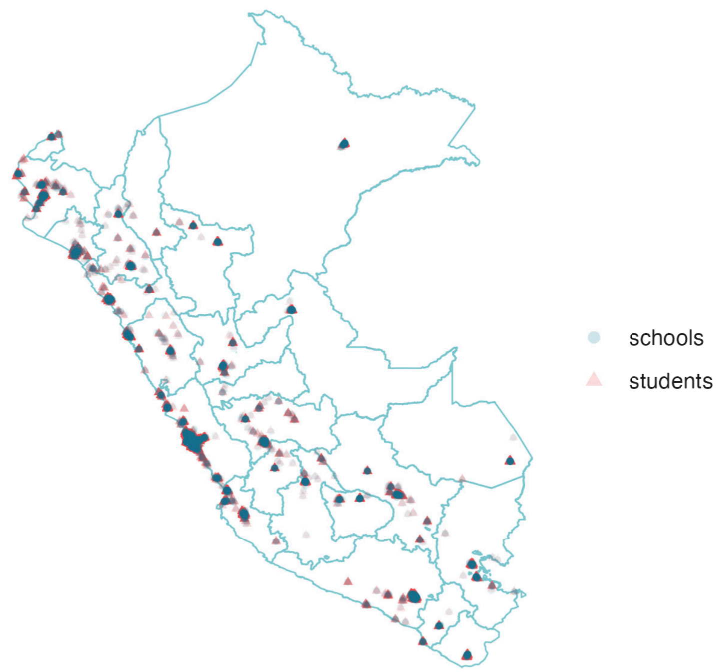


Figure 2: Geographic distribution of students and schools participating in ME2020



(a) Note: The higher color concentration indicates a higher density of students and schools.

3 Theoretical Framework

We introduce a simple two period model which allows us to rationalize the main empirical findings of section 4. Students choose first whether to switch to the public sector, based on the school they are offered. In this first decision, students are imperfectly informed about the quality of their match with the school. After attending school, they perfectly learn their valuation, and choose whether to remain or go back to the private sector. Students fully internalize the "experimentation value" of switching to a public school: if they like it, they can remain, if they do not, they have the possibility of going back to a private school.

The model is able to rationalize some key facts that emerge from the empirical analysis, namely the change in the intensity of preferences over key attributes like math test scores and distance. When a student chooses in the first period (that is whether to try out the public school or not), school characteristics are a priori more important, since they will affect a student over potentially two periods. This implies that first period decisions are more sensitive to school characteristics. On the other hand, in the second period decision (whether to remain or not) only a subset of students makes a choice, and they are more likely to care about these characteristics. This would imply higher sensitivity in the second period, which might seem counterintuitive at first.

We consider a two-period model, in which each student has two options, a public and a private school. A student payoff from attending the private school is $V_{PR}^{(1)} + \epsilon$ in the first period and $V_{PR}^{(2)} + \eta$ in the second. Here, we assume that $V_{PR}^{(2)} > V_{PR}^{(1)}$ to reflect that in period 1 households are experiencing a crisis due to COVID, that decrease their valuation for the private school, either because they are experiencing economic difficulties or because the quality of service has decreased. Moreover ϵ, η are idiosyncratic shocks distributed according to G which has a mean of 0. These shocks model the variation in the level of satisfaction with the current school (ϵ) and the strength of the recovery after the crisis (η). Agents know the realization of ϵ before their decision at $t = 1$ and the realization of η before $t = 2$. We assume that the the distribution G is the same in both periods and, moreover, that shocks are independently distributed, since they refer to different dimensions of the choice problem.

Students face significant uncertainty about the public school option, since it is unknown to them. We assume that the valuation for the public option is αV_{PU}, where V_{PU} is an objective measure of school attractiveness that captures observable attributes, like distance, parents education or schools' results on standardized tests. In its empirical counterpart V_{PU} is given by the vector of observable attributes of a given school.

On the other hand, α is an idiosyncratic shock that captures the quality of the school-student match, something not observable by the econometrician, but also unknown by students before attending a a school. Students are initially uninformed about α, which represents, among other things, the relationship to classmates, teachers and genera satisfaction with the learning environment. We assume that there are two types of students, $i = \{1, 2\}$ in proportions μ_0 and $1 - \mu_0$, and α is distributed according to F_i on [1,K], with $K < \infty$. We assume that $F_1 \succeq_{FOSD} F_2$, which captures that there is a fraction μ_0 of the population which is more likely to have a good experience in the public school system. Students know the group to which they belong, capturing the fact that some of them are more willing to try out the public sector.

The next assumption captures the fact that a bad match with public schools ($\alpha = 1$) makes it undesirable even in period 1, but a good match ($\alpha = K$) makes it the best option even in period 2, after the crisis has passed.

Assumption 3.0.1. *We assume that* $V_{PU} < V_{PR}^{(1)} < V_{PR}^{(2)} < KV_{PU}$.

First-period students choose whether to accept o reject the public school offered by the government. They do so knowing their valuations for the private school ($V_{PR}^{(1)} + \epsilon$ in the first period and $V_{PR}^{(2)}$ in expectation in the second), but knowing only the expectation of their valuation for the public school option ($\mathbb{E}_{F_i}(\alpha)V_{PU}$). Moreover, there is an "exploration value" of the public school option, since the parameter α becomes known after the first period if the student experiments with the public school. Thus, in $t = 1$, a student of type i accepts the public school offered if:

$$\underbrace{\int_{\alpha=1}^{K} \alpha V_{PU} dF_i(\alpha)}_{\text{First period expected payoff}} + \underbrace{\int_{\alpha=1}^{K} max\left\{\alpha V_{PU}, V_{PR}^{(2)}\right\} dF_i(\alpha)}_{\text{Second period expected payoff}} \geq \underbrace{V_{PR}^{(1)} + V_{PR}^{(2)} + \epsilon}_{\text{Private school payoff}} \tag{1}$$

where ϵ is distributed according to G. That is, the student decides to experiment if and only if the expected payoff in the first period plus the expected payoff in the second period, once α is known, is greater than the safe option of staying in the private sector. Experimentation is valuable, since the student will stay in the public sector if $\alpha \geq \frac{V_{PR}^{(2)}}{V_{PU}}$ and will return to the private sector otherwise.

Lemma 3.0.1. *In $t = 1$, a student of type i accepts the public school if and only if:*

$$H_i(V_{PU}) \equiv V_{PU} \cdot \left[\bar{\alpha}_i + \int_{\alpha=\frac{V_{PR}^{(2)}}{V_{PU}}}^{\alpha=K} \alpha dF_i(\alpha)\right] - \left[V_{PR}^{(1)} + V_{PR}^{(2)} \cdot (1 - F_i\left(\frac{V_{PR}^{(2)}}{V_{PU}}\right))\right] \geq \epsilon$$

therefore, the proportion of students that go to public schools is given by given by:

$$\mathbb{P}(try) = \mu_0 G(H_1(V_{PU})) + (1 - \mu_0)G(H_2(V_{PU}))$$

where $G(H_1(V_{PU})) > G(H_2(V_{PU}))$

Proof. Direct from algebraic manipulation and the fact that $F_1 \succeq_{FOSD} F_2$. □

In $t = 2$, students who tried out a public school choose whether to stay in the public system or not. In this period, they already know their match with the public school (α) and their valuation of going back to the private sector ($V_{PR}^{(2)} + \eta$). Thus, a student stays in their public school if:

$$\alpha_i V_{PU} \geq V_{PR}^{(2)} + \eta \tag{2}$$

Analogously to the previous lemma, characterize the proportion of students that stay in the public school.

Lemma 3.0.2. *At* $t = 2$*, a student stays in the public sector if:*

$$L(\alpha V_{PU}) \equiv \alpha V_{PU} - V_{PR}^{(2)} \geq \eta$$

therefore, the proportion of students that stay in the public system is given by:

$$\mathbb{P}(stay) = \mu_1 \int_{\alpha=1}^{K} G(L(\alpha V_{PU}))dF_1(\alpha) + (1 - \mu_1) \int_{\alpha=1}^{K} G(L(\alpha V_{PU}))dF_2(\alpha)$$

where $\mu_1 = \mathbb{P}(type1|try) = \frac{\mu_0 G(H_1(V_{PU}))}{\mu_0 G(H_1(V_{PU})) + (1-\mu_0)G(H_2(V_{PU}))} > \mu_0$

Proof. Direct. □

It is important to note that students of type 1, which are more likely to prefer the public school option, are over-represented in the sample at $t = 2$, that is $\mu_1 > \mu_0$. This is so because at $t = 1$, before knowing their actual match with the public system, they are more optimistic, so they are more likely to try it during the crisis. This *selection bias*, given by the difference $\mu_1 - \mu_0$ plays a crucial role in the results.

We are interested in analyzing the sensitivity of parents decisions to school characteristics, both at $t = 1$ and $t = 2$, and the changes in this sensitivity. In the model this corresponds to the quantities $\frac{\partial \mathbb{P}(try)}{\partial V_{PU}}$ and $\frac{\partial \mathbb{P}(stay)}{\partial V_{PU}}$, which indicate how the (expected) decision of a student changes if offered a marginally better school. The empirical counterpart of these quantities are given by the parameters β_i in the regressions, that indicate the increased likelihood of an option given a change in the observabke characteristic of a school.

As expected, the likelihood of trying the public school at $t = 1$ and staying in at $t = 2$ are increasing in V_{PU}. What is more interesting is that the sensibility to changes in V_{PU} can be greater or smaller in $t = 2$, depending on the strength of the selection bias.

Lemma 3.0.3. *Suppose that G is a uniform distribution on an interval $[-a, a]$. Then the sensibility in the second period will be greater if:*

$$(\mu_1 - \mu_0)(\bar{\alpha}_1 - \bar{\alpha}_2) > \left[(1-\mu_0)\int_{\alpha=\frac{V_{PR}^{(2)}}{V_{PU}}}^{K} \alpha \cdot dF_2(\alpha) + \mu_0 \int_{\alpha=\frac{V_{PR}^{(2)}}{V_{PU}}}^{K} \alpha \cdot dF_1(\alpha)\right] \quad (3)$$

Proof. If we assume that G is a uniform distribution, the derivate of the probability to experiment with respect to V_{PU} is

$$\frac{\partial \mathbb{P}(try)}{\partial V_{PU}} = g \cdot \left[\mu_0 H_1'(V_{PU}) + (1-\mu_0) H_2'(V_{PU})\right] \quad (4)$$

where $H_i'(V_{PU}) = \bar{\alpha}_i + \int_{\frac{V_{PR}^{(2)}}{V_{PU}}}^{K} \alpha \cdot dF_i(\alpha)$

and the derivate of the probability to stay with respect to V_{PU} is:

$$\frac{\partial \mathbb{P}(stay)}{\partial V_{PU}} = g\left[\mu_1 \int_{\alpha=1}^{K} \alpha dF_1(\alpha) + (1-\mu_1)\int_{\alpha=1}^{K} \alpha dF_2(\alpha)\right] = g\left[\mu_1 \bar{\alpha}_1 + (1-\mu_1)\bar{\alpha}_2\right] \quad (5)$$

If we substract (4) - (3) we have:

$$\underbrace{(\mu_1 - \mu_0)(\bar{\alpha}_1 - \bar{\alpha}_2)}_{\text{Selection effect}} - \underbrace{\left[(1-\mu_0)\int_{\alpha=\frac{V_{PR}^{(2)}}{V_{PU}}}^{\alpha=K} \alpha \cdot dF_2(\alpha) + \mu_0 \int_{\alpha=\frac{V_{PR}^{(2)}}{V_{PU}}}^{\alpha=K} \alpha \cdot dF_1(\alpha)\right]}_{\text{Horizon effect}}$$

The first element is the selection effect and the second element is the horizon effect. On the one hand, as the selection effect is greater, the sensitivity in the second stage will be greater. On the other hand, the horizon effect is mechanical. In the first stage, the student cares about two periods instead of just one. If the selection effect predominates, in the second stage we should observe a greater sensitivity.

Note that $\mu_1 > \mu_0$, due to in the second stage there are only those students who accepted their assignment. Since F_1 has first-order stochastic dominance over F_2, students who accepted are more likely to come from type 1. We call this the selection effect. □

The left hand side of condition (3) measures the strength of the selection effect. Since "optimistic" students, with higher levels of α, are over-represented at $t = 2$, the importance of the actual observable qualities of the school, V_{PU} becomes more acute. The right hand side of (3) represents the *horizon effect*. At $t = 1$, V_{PU} matters more simply because it potentially affects the student for two periods (always at $t = 1$, with some probability at $t = 2$).

The relative strength of the selection and horizon effects determines which decision, try the public school or stay enrolled in it, is more sensitive to the actual characteristics of the school offered. Moreover, it provides a rationale for the particular characteristics of the school (distance, standardized test scores, etc) being even more important for the decision of students *after* they have already stayed enrolled in the public school for a year. The results in section 4.2.2, in particular the comparison between tables 1 and 2, are consistent with a strong selection effect, and with parents that try the public sector being on average more willing to stay if everything goes right (type 1).

4 Empirical Analysis

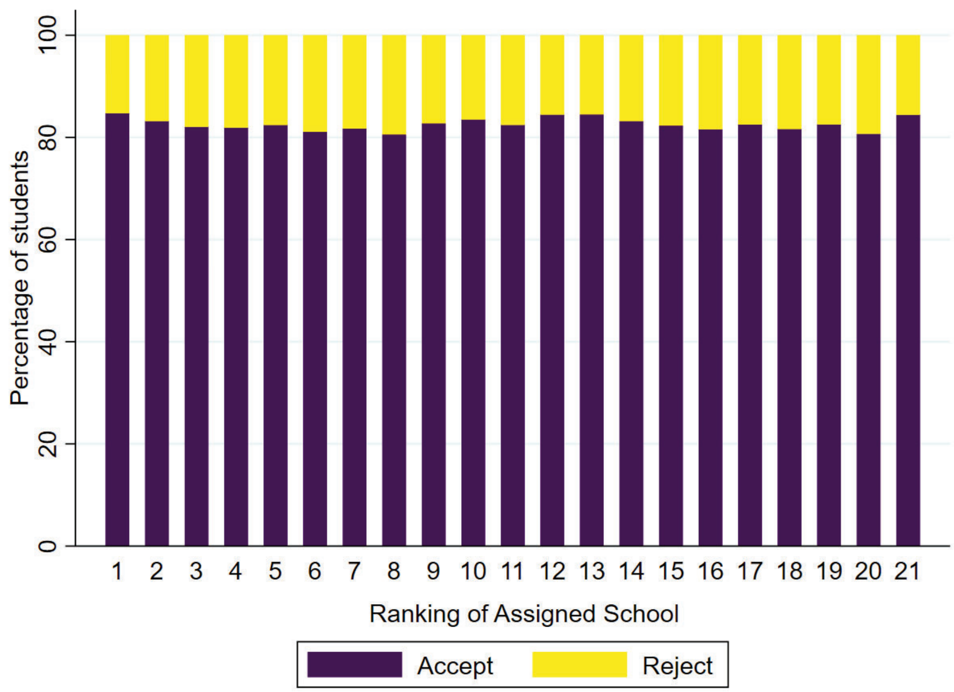
4.1 Identification Strategy

The *2020 Matricula Excepcional* (ME) algorithmic allocation process allows us to exploit the exogenous variation from the randomized assignment of students conditional on the local school offering to recover the causal estimates of the parents' preferences when choosing a school. The randomization process is part of the algorithmic mechanism of Deferred Acceptance (DA), which uses imputed preferences based on school proximity to the student's home (ranking the closest available school first and furthest school last) within a 5 km radius.

Conditional on students' preferences based on distance, students are randomly allocated to schools temporarily to maximize the likelihood of getting a spot at their most preferred school (closest in this context). Therefore, students with similar 'preferences,' in this case, close geographical proximity, achieve their final allocations by chance allowing us to compare their behavior.

Our identification strategy assumes exogenous randomness in the assignment. A possible threat to this assumption is observing that parents strategically listed their addresses close to their preferred school. However, we do not believe this is a reason for concern. As we can observe in Figure 3, the likelihood of accepting an assignment is roughly constant at around 85%. If parents were being strategic, we would observe a right-skewed distribution of acceptance where lower-ranked schools have lower rates of acceptance than higher ranked ones. In addition, public schools with high demand were already oversubscribed and unavailable to parents with this mechanism.

Figure 3: Distribution of Acceptances by preference ranking



This section will introduce a preferred specification that aims to estimate the determinants of parental preferences in their decision to accept or reject and to stay or leave their assigned school, conditional on having accepted the assignment. The first stage specification aims to causally estimate the parents' preferences over relative school characteristics in relation to accepting or rejecting their assignment offering. Then, conditional on accepting their assignment in the first stage, we explore the influence of observable characteristics on the decision to stay or leave the assigned school[13]. In particular, we run the following regression:

$$y_{isdgl} = \alpha_{dgl} + \psi \cdot \text{Dist}_{isd} + \mathcal{D}'_{isd}\Lambda + e_{isdgl} \tag{6}$$

Where y_{isdgl} takes two values: first, Accept_{isdgl} is an indicator variable at the student i, assigned school s, district d, level l, and grade g level that takes a value of 100 if the student decides to accept her assignment in the first stage and a value of 0 if the student rejects her assignment; and second Staying_{isdgl} which takes a value of 100 if the student decides to stay at her assigned school and a value of 0 if the student moves to another school conditional on having accepted the assignment. The independent variables used include Dist_{isd} is the distance between the assigned school s and the student i addresses in district d and $\mathcal{D}_{isd}$ capture differences in observable characteristics between assigned schools s and the previous school of student i in district d. To compute the differences, we use the information on school class sizes as measured by the student-to-teacher ratio, average parents' education, average measure of academic performance (math), and school size. We use average standardized math scores from the ECE assessment as a proxy for overall school quality. Also, we have included characteristics that give us an idea of the relevance of teaching quality, such as the proportion of novice teachers (with less than five years of experience) and the proportion of temporary teachers (teachers who did not meet the minimum standards). We include the prices of the school of origin to control for an approximate measure of income level and a dummy to control for the student's gender. We include district d by grade g and level l fixed effects denoted by α_{dgl} to control for the systematic differences in the feasible set of schools. Finally, e_{isdgl} represents the error term.

In our main results section, we will explore specification (6) as we are interested in

[13] As explained in the theoretical model, we acknowledge the selection issue in the second stage and its possible influence over our results; therefore, our claim of causality holds only for the first stage. However, we are working on adjusting our empirical framework to address these concerns in the second stage.

causally estimating ψ and Λ for the first stage and exploring the variation in these estimates in the second stage. Finally, we explore the robustness of these results by applying more granular fixed effects where we only compare among students who have the same top two schools listed in their set of preferences (based on proximity). We also refined our fixed effects at a more granular level to compare among those who have the same top 3 schools and include it in the appendix. Comparing students with the same schools listed among their top priorities closely mimics the randomization mechanism used by the algorithm. The level for the clustering of standard errors is the district by grade and level.

4.2 Main Results

4.2.1 Preferred specification

Table 1 presents the specification (6) results when the dependent variable indicates whether the parents accept or reject the assignment for columns (1) and (3). In contrast, columns (2) and (4) present the results when the dependent variable indicates whether the parents stay in their assigned school or leave. This last stage presents the results conditional on parents accepting their assignment in the first stage. As mentioned, specification (6) allows us to causally estimate the determinants of the first decision: accept the public school offered or reject it for another alternative (public or private). In contrast, columns (2) and (4) show the likelihood of remaining at the assigned schools after conditioning the sample of having accepted their assignment.

In Table 1, we consider the relative characteristics of the schools that students are assigned to compared to the characteristics that families experienced in their previous private schools. Columns (1) and (2) rely on district grade and level fixed effects, in contrast with columns (3) and (4), where we used a unique fixed effect per set of the same top 2 schools, which can be considered a stricter imposition on the distance between households, as only those living very close to each other will be assigned the same top two schools in their preference list. After conditioning by the more stringent set of fixed effects, we do not observe drastic differences with our preferred specification. There is a significant drop in the number of observations between our Figure 1 and column (1) from Table 1. The discrepancy occurs due to dropping clusters of observations where there needs to be more variation in the acceptance rate within our chosen observation level (district by grade and level).

Table 1: The effect of school characteristics on likelihood to accept and stay at assigned school

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay
Distance (km)	-1.058***	-2.205***	-1.164***	-2.987***
	(0.182)	(0.233)	(0.385)	(0.539)
Dif. Distance (km)	-0.328***	0.068	-0.574***	0.145
	(0.053)	(0.072)	(0.119)	(0.095)
Dif. Mean Math sc (s)	3.791***	3.404***	4.874***	5.560***
	(0.509)	(0.756)	(0.963)	(1.634)
Dif. Mean Parent Edu.	1.894***	3.887***	2.736***	5.705***
	(0.295)	(0.408)	(0.533)	(0.678)
Log(Ratio Num St.)	2.118***	4.403***	3.212***	4.310***
	(0.540)	(0.626)	(1.206)	(1.131)
Log(Ratio Num St.)^2	-0.118	-0.646***	-0.187	-0.910***
	(0.164)	(0.156)	(0.380)	(0.325)
Dif. STR	-0.056**	-0.011	-0.056	0.117*
	(0.022)	(0.038)	(0.052)	(0.059)
Dif. Prop. Novice Teachers	-1.196	-1.397**	-1.722	-1.814
	(0.766)	(0.657)	(1.303)	(1.171)
Dif. Prop. Temporary Teachers	-1.482	1.989	-3.348	0.467
	(1.331)	(2.139)	(3.318)	(3.405)
Price School of Origin	-0.008***	-0.007***	-0.006*	-0.009**
	(0.002)	(0.002)	(0.003)	(0.004)
Constant	83.690***	83.810***	73.788***	80.558***
	(1.164)	(1.873)	(2.876)	(2.854)
Observations	47,407	35,174	19,739	14,885
R-squared	0.072	0.082	0.151	0.170
Inc. Female	X	X	X	X
FE: Dist-Gr-Lev	X	X		
FE: Same Top 2			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(a) Note: Standard errors have been clustered at the district by level and grade level.

Distance to the school is highly relevant in the first decision. In Table 1, we find that both in columns (1) and (3), the effect of distance to school on the likelihood of accepting the assigned school varies from a decrease of 1.06 percentage points (pp.) to 1.16 pp. in the likelihood of accepting the assigned school for every 1 kilometer (km.) increase in the distance. This indicates a weak sensitivity of parents' choices to the distance it will require their children to travel daily to school in the first decision. We also find a negative impact of the distance to the school on the likelihood of remaining at their assigned school. In contrast with column (1), we observe an increase in the negative impact of distance from 1.06 to 2.21 p.p., implying a possible higher sensitivity to the distance at this stage. It is also important to highlight that all classes were still virtual when deciding to stay or switch schools. This fact is relevant because it might indicate that parents considered the move for the long run and not as a short-term solution. Given that distance to the school is irrelevant in online education, parents may have considered the move to public education as a long-term decision. We explore possible non-linear effects of distance in Appendix A.4 but find little evidence[14].

Interestingly, while the distance, in absolute terms, impacts the likelihood of accepting the assignment, the relative distance between the school of origin and the family's household (Dif. Distance) also negatively impacts the chance of accepting the assignment. In essence, families are unwilling to accept school assignments further away from their homes than their school of origin. This comparative framing is relevant at the first stage, but not for the decision to stay at the newly assigned school, as depicted in columns (2) and (4).

Beyond the role of distance, we observe that parents care about peer demographics. Parents value that the families in their newly assigned public school are relatively more educated than parents in their previous private school. An increase of a unit difference in the average level of education (an increase in the category towards receiving an academic degree) between the assigned school and the private school of origin affects the likelihood of accepting the assigned school by between 1.89 p.p. and 2.74 p.p. The results indicate that peer demographics matter when choosing which schools to attend. Moroever, when we analyze the likelihood of staying in the school, the coefficient increases to between 3.4 p.p. and 5.56 p.p., indicating the importance of this dimension when deciding to remain at the assigned school. This indicator can also be understood as a proxy or composite good of other features parents value in the school that are not necessarily related to peer demographics. For example, parents may believe that more educated parents are more

[14]In the iterations of the paper, we plan to explore heterogeneous effects of other dimensions such as families' income, using the information on the price of their original private school, as well as gender.

likely to attend improving schools (because more educated families may pressure for these improvements), safer schools, or other vital dimensions of school quality not directly related to student achievement. Since they may not have information on test scores, they may also use peer demographics as a proxy for student achievement.

In addition to caring about distance and peers, school quality as measured by student achievement is also relevant to parents. As mentioned above, we use the schools's mean standardized math score on the ECE exams as a proxy for school quality.In particular, we find that an increase of a standard deviation of the difference in mean math scores between the assigned and origin school increases the likelihood of accepting the assignment by between 3.79 and 4.87 p.p. and also increases the likelihood of staying in the assigned school between 3.4 p.p. and 5.56 p.p. The stability of student achievement in the first and second stage suggests that parents care about the quality of the schools their kids attend. This result is consistent with findings reported by Abdulkadiroğlu et al. (2020) and others that find that parents prefer schools with higher test scores. Regarding school size, we find that relative school size is also a factor that parents consider. The effect of an increase in a unit of the log ratio in school size (measured as the number of enrolled students) between the assigned and origin school raises the probability of accepting and staying at the assigned school. Parents may observe the large number of students enrolled at a school relative to their previous school as a possible signal of school quality (popularity). However, the relationship is not linear[15]. Accepting and staying is less likely if the assigned school is much larger than the private school of origin.

In addition, the results show that in the assignment stage, parents are less likely to accept being assigned to schools with larger teacher-student ratios (STR). However, this result does not hold if we consider our more stringent set of fixed effects.

The proportion of novice teachers has no relevance in the decision to accept the assigned school, probably because it is a characteristic unknown by the parents in the first stage.

[15] We have tested for joint significance of the quadratic terms: In the first column, the F-statistic is 16.24, while in the second column, it is 27.21, i.e., both are jointly significant.

Interestingly, when analyzing the likelihood of staying at their assigned school, parents are less likely to remain in schools with a higher proportion of novice teachers. Although this result does not hold in our more stringent fixed effect specification, it suggests that families are less likely to stay in schools with more inexperienced teachers. Following what we learn from our theoretical model, after parents experienced the new public school, they internalized the implications of this dimension and decided to leave.

Finally, higher tuition at the private school of origin is associated with a lower probability of accepting the assigned public school. We may interpret this variable as a family's purchasing power. Given the larger set of schools available, more affluent families were less likely to transition to the public sector. Another interpretation is that parents who initially paid higher private school fees perceive private schools to be higher quality than public schools and thus may be less likely to leave these schools.

Our findings are consistent with findings reported by Jacob and Lefgren (2007) and Burgess et al. (2015) that show that the families prefer higher-performing schools and better peer demographics, and are sensitive to school's proximity to their residence. Using data from a centralized matching system after eliminating school attendance zones in New Orleans post–Hurricane Katrina, Harris and Larsen (2017) find that parents prefer schools with higher test scores and higher value-added. However, parents also consider non-academic attributes such as after-school care, extracurricular activities, whether the child had a sibling in the school, and proximity to home. (Glazerman and Dotter, 2017) report that commuting distance, peer demographics, and academic indicators play essential roles in school choice and that there is considerable heterogeneity of preferences by parent demographics.

In Table 2a, we standardize same set of variables as in Table 1 in order to gain insight into the relative importance of each dimension on the decisions to accept and stay at the assigned schools. We find that most factors that affect the decision to accept the assigned schools are weighted similarly. However, in the second stage, school quality (difference in average math scores (s)) falls behind other factors such as distance to the school and difference in mean parents' education.

Table 2: The effect of school characteristics on the likelihood to accept and stay at assigned school - Standardized

	(1)	(2)	(3)	(4)
VARIABLES	accept	stay	accept	stay
Distance (km)	-2.071***	-4.317***	-2.278***	-5.847***
	(0.356)	(0.456)	(0.754)	(1.055)
Dif. Distance (km)	-1.722***	0.359	-3.020***	0.762
	(0.277)	(0.380)	(0.626)	(0.499)
Dif. Mean Math sc (s)	2.105***	1.890***	2.706***	3.087***
	(0.282)	(0.420)	(0.535)	(0.907)
Dif. Mean Parent Edu.	1.987***	4.077***	2.869***	5.983***
	(0.310)	(0.428)	(0.559)	(0.711)
Log(Ratio Num St.)	2.359***	4.904***	3.578***	4.801***
	(0.602)	(0.697)	(1.343)	(1.260)
Log(Ratio Num St.)^2	-0.350	-1.908***	-0.554	-2.690***
	(0.486)	(0.461)	(1.124)	(0.959)
Dif. STR	-0.523**	-0.106	-0.519	1.084*
	(0.206)	(0.351)	(0.480)	(0.551)
Dif. Prop. Novice Teachers	-0.451	-0.526**	-0.649	-0.684
	(0.289)	(0.248)	(0.491)	(0.441)
Dif. Prop. Temporary Teachers	-0.426	0.572	-0.962	0.134
	(0.382)	(0.615)	(0.953)	(0.979)
Price School of Origin	-1.484***	-1.374***	-1.110*	-1.751**
	(0.339)	(0.404)	(0.636)	(0.816)
Constant	80.728***	75.008***	72.298***	68.111***
	(0.137)	(0.212)	(0.316)	(0.357)
Observations	47,407	35,174	19,739	14,885
R-squared	0.072	0.082	0.151	0.170
Inc. Female	X	X	X	X
FE: Dist-Gr-Lev	X	X		
FE: Same Top 2			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(a) Note: Standard errors have been clustered at the district by level and grade level.

In contrast with our preferred specification, we could have used the observable characteristics of the assigned schools regarding the decisions to accept and stay. Table 12 (appendix levels results) presents the results of these specifications. We first find some similarities with our preferred specification, with the effect of distance and the importance of peer quality being relatively similar. However, we highlight some interesting differences. First, school quality does not seem to play a role in the decision to either accept the assignment or stay at the assigned school. The importance of school size only seems relevant in the second stage of remaining at the assigned school[16]. We also find that the proportion of

[16]We have tested the joint significance in the first stage and found null results for both coefficients of school size (p-value = 0.32)

inexperienced teachers appears relevant in the acceptance stage. It could be that younger teachers were perceived as a sign of quality, given that the educational offering was still remote and relied upon some tech expertise.

5 Discussion

During the first year of the COVID-19 pandemic, Peru experienced an unprecedented rise in the demand for public schools by students enrolled in private schools. More than a hundred thousand students requested a transfer to a public school due to factors such as the extended school closures during the first year of the pandemic, lack of adaptability by private schools to move to online education, and the cost of paying private school tuition during an economic crisis.

In response to the rapid rise in the demand for public schools, the Ministry of Education designed a centralized assignment system to absorb these students into the public sector. The inherent experimental variation from this system offered a unique opportunity to better understand what this specific group of parents, middle and lower middle class families who enroll their children in private schools, values in public schooling. Analyzing this setting offers insights for governments to strengthen the supply of public schooling and possibly reverse the pre-pandemic trends between public and private education.

This paper offers insights into what this group of parents value in a public school before and after experiencing the assigned school. We explore the role of school attributes such as distance, academic quality, and peer demographics, among other factors, both in the decisions of accepting and staying at the assigned public school. Parents value distance, academic quality, and peer demographics. When comparing these features between the decisions of accepting and remaining at the assigned public school, we find that absolute distance to the school and peer demographics play a more significant role than academic quality in the decision to stay. However, all three factors remain relevant in the decisions to accept and remain at the assigned school.

The results from this paper have important implications for policy [17]. First, consistent with previous evidence, controlling for other factors, parents care about student demo-

[17] However, one must be careful with extrapolating these results since, in this setting, parents only learned about the public school imperfectly through the online remote learning experience. Results could be different if parents had experienced the public school in-person.

graphics. In the context of Latin America and the Caribbean, public schools tend to be less diverse and serve the most disadvantaged students. Therefore, attracting a more diverse clientele into the public sector is essential to foster diversity and attract more middle class families. Otherwise, the segregation of students can become more severe as schools become less diverse, which can stigmatize public schools. Second, even after controlling for other relevant factors, the school's academic quality remains prevalent for parents in the decision to both choose and stay in the assigned school. The Peruvian government may continue to focus on school reforms that improve the quality of public schools. For instance, over the last decade, the government instituted reforms to modernize the teacher career path (Elacqua et al., 2018) and attract and select more effective teachers into the profession.

An important takeaway is that experimentation is essential for parents' decisions to consider enrolling their children in public schools. The policy implication is that short-term incentives to experience public services (like education) can have long-term consequences for a significant group of parents. Only 16%[18] of the students who experienced the assigned public school in 2020 returned to the private sector in 2021. Despite the evidence, that, all else equal, private schools do not have higher achievement than public schools (Balarin, 2015), parents in Peru have the perception that private schools are better [19]. Publishing objective school-level information on test scores or value-added measures may reduce misconceptions about public and private school quality. The Peruvian government may also run campaigns to increase citizen knowledge about the objective attributes of public schools.

[18]Only 10438 students out of the 63962 who experienced at least a few months in the assigned public school in 2020 returned to the private sector in 2021.

[19]*Este País*, May 1^{st}, 2014

References

A. Abdulkadiroğlu, N. Agarwal, and P. A. Pathak. The welfare effects of coordinated assignment: Evidence from the new york city high school match. *American Economic Review*, 107(12):3635–89, 2017.

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

B. Abizanda, G. Almeyda, E. Arias Ortiz, C. Berlanga, I. Bornacelly, M. S. Bos, E. Díaz, X. Dueñas, G. Elacqua, A. Elías, R. Fernández-Coto, V. Frisancho, V. A. García Moreno, J. C. Hernández Cardozo, D. Hincapie, J. F. Margitic, L. Marotta, M. M. Mateo-Berganza Díaz, A. Morduchowicz, F. Muñoz, E. Näslund-Hadley, M. Ruiz-Arranz, A. Thailinger, F. J. Valverde Rodríguez, E. Vezza, and P. Zoido. How to reboot education post-pandemic: Delivering on the promise of a better future for youth. *IDB Monograph ; 1012*, 2022.

N. Agarwal and P. Somaini. Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica*, 86(2):391–444, 2018.

M. Balarin. The default privatization of peruvian education and the rise of low-fee private schools: Better or worse opportunities for the poor? 2015.

S. Burgess, E. Greaves, A. Vignoles, and D. Wilson. What parents want: School preferences and school choice. *The Economic Journal*, 125(587):1262–1289, 2015.

G. Elacqua. The impact of school choice and public policy on segregation: Evidence from chile. *International Journal of Educational Development*, 32(3):444–453, 2012.

G. Elacqua, D. Hincapié, E. Vegas, M. Alfonso, V. Montalva, and D. Paredes. *Profesión: Profesor en América Latina:¿ Por qué se perdió el prestigio docente y cómo recuperarlo?* Inter-American Development Bank, 2018.

G. Elacqua, C. Méndez, and M. Navarro. Colegios privados en tiempos de covid-19 en américa latina y el caribe. *IADB Technical Note*, 2022. doi: 10.18235/0004434.

S. Glazerman and D. Dotter. Market signals: Evidence on the determinants and consequences of school choice from a citywide lottery. *Educational Evaluation and Policy Analysis*, 39(4):593–619, 2017.

D. N. Harris and M. F. Larsen. Demand, information, and the market for schooling: Evidence on revealed preferences from post-katrina new orleans. *Education Research Alliance for New Orleans*, 2017.

J. S. Hastings, T. J. Kane, and D. O. Staiger. Parental preferences and school competition: Evidence from a public school choice program, 2005.

A. Hofflinger, D. Gelber, and S. T. Cañas. School choice and parents' preferences for school attributes in chile. *Economics of Education Review*, 74:101946, 2020.

B. A. Jacob and L. Lefgren. What do parents value in education? an empirical investigation of parents' revealed preferences for teachers. *The Quarterly Journal of Economics*, 122 (4):1603–1637, 2007.

A. Kapor, C. Neilson, and S. Zimmerman. Heterogeneous beliefs and school choice. *American Economic Review*, 2016.

B. Kleitz, G. R. Weiher, K. Tedin, and R. Matland. Choice, charter schools, and household preferences. *Social science quarterly*, pages 846–854, 2000.

F. J. Murillo, G. Belavi, L. M. Pinilla, et al. Public-private school segregation in spain. *Universitat Autonoma de Barcelona Papers*, 2018.

M. Schneider, P. Teske, M. Marshall, and C. Roch. Shopping for schools: In the land of the blind, the one-eyed parent may be enough. *American Journal of Political Science*, pages 769–793, 1998.

A. Singh, A. Park, and S. Dercon. School meals as a safety net: An evaluation of the midday meal scheme in india. *Economic Development and Cultural Change*, 62(2):275–306, 2014. doi: 10.1086/674097. URL https://doi.org/10.1086/674097.

A Appendix

A.1 Sample details

Table 3: Covariates by Stages

	Sample			Accept			Stay		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Distance (km)	69373	2.11	1.40	57496	2.09	1.40	35639	2.05	1.40
STR	69344	22.47	4.96	57475	22.54	4.94	35630	22.55	5.02
Mean Math sc (s)	54331	-0.01	0.32	44533	-0.01	0.32	30993	0.01	0.31
Mean Parent Edu.	54321	4.91	0.63	44525	4.92	0.62	30992	4.94	0.61
Log(Num. of Students)	69344	6.01	0.77	57475	6.01	0.77	35630	6.14	0.71
Prop. Novice Teachers	69011	0.14	0.17	57203	0.14	0.17	35437	0.13	0.16
Prop. Temporary Teachers	69361	0.36	0.22	57486	0.36	0.22	35632	0.35	0.20
Price School of Origin	68543	232.85	192.05	56728	225.00	178.56	35175	220.06	178.88

Table 4: Difference in Covariates by Stages

	Sample			Accept			Stay		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Dif. Distance (km)	67933	-0.14	4.24	56195	-0.20	4.33	34773.00	-0.20	4.30
Dif. STR	67993	5.37	9.13	56313	5.53	9.17	34915.00	5.68	9.34
Dif. Mean Math sc (s)	50541	0.10	0.55	41161	0.12	0.55	28465.00	0.14	0.54
Dif. Mean Parent Edu.	49534	-1.46	1.02	40279	-1.39	0.99	27813.00	-1.31	0.95
Log(Ratio Num St.)	68228	1.22	1.04	56503	1.27	1.04	35042.00	1.37	1.04
Dif. Prop. Novice Teachers	65933	-0.20	0.38	54569	-0.20	0.38	33713.00	-0.20	0.38
Dif. Prop. Temporary Teachers	68028	-0.63	0.23	56273	-0.62	0.23	34884.00	-0.64	0.21

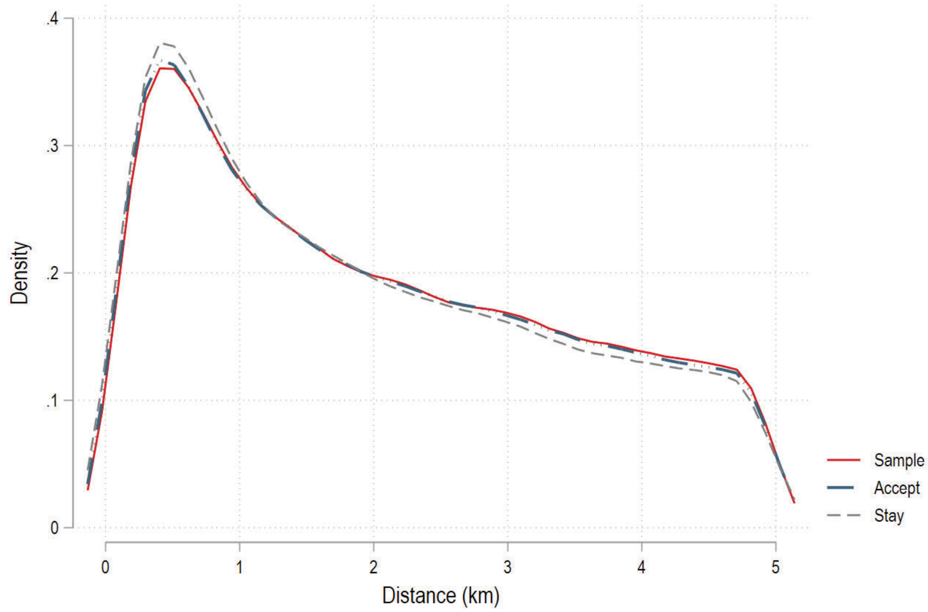


Figure 4: Distribution of Distance to Assigned School by Stages

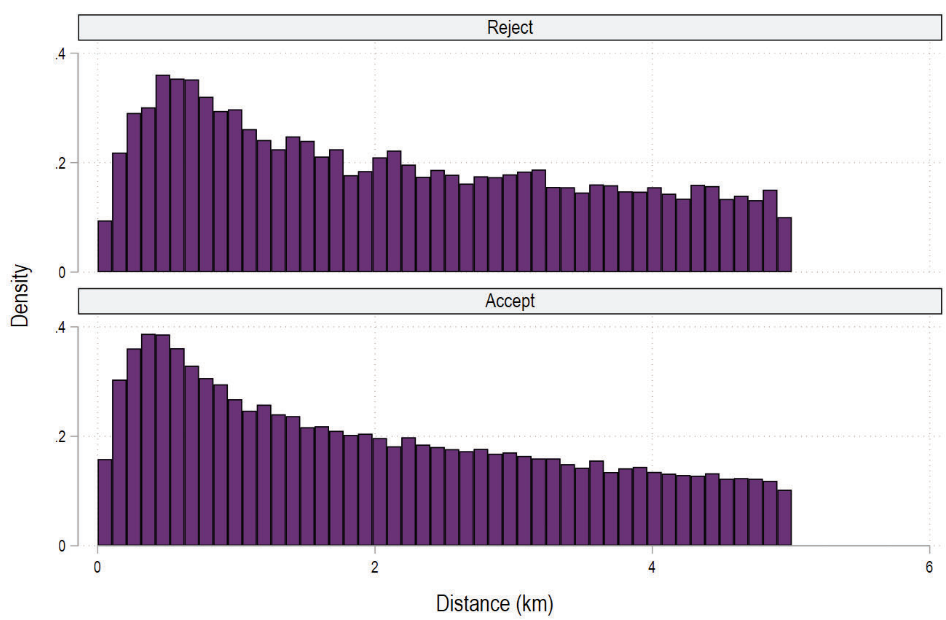


Figure 5: Histogram Of Distance to Assigned School Accept-Reject

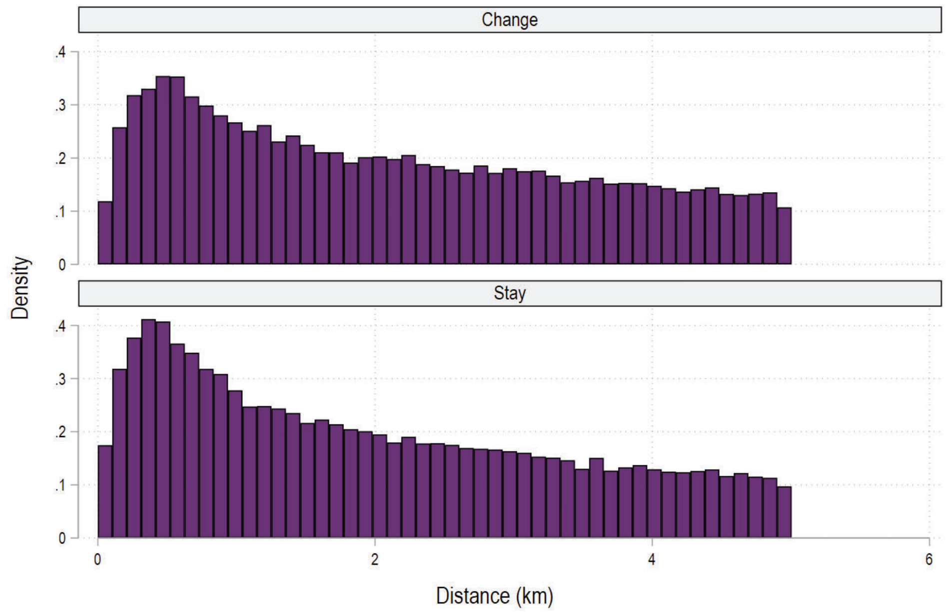


Figure 6: Histogram Of Distance to Assigned School Stay-Change

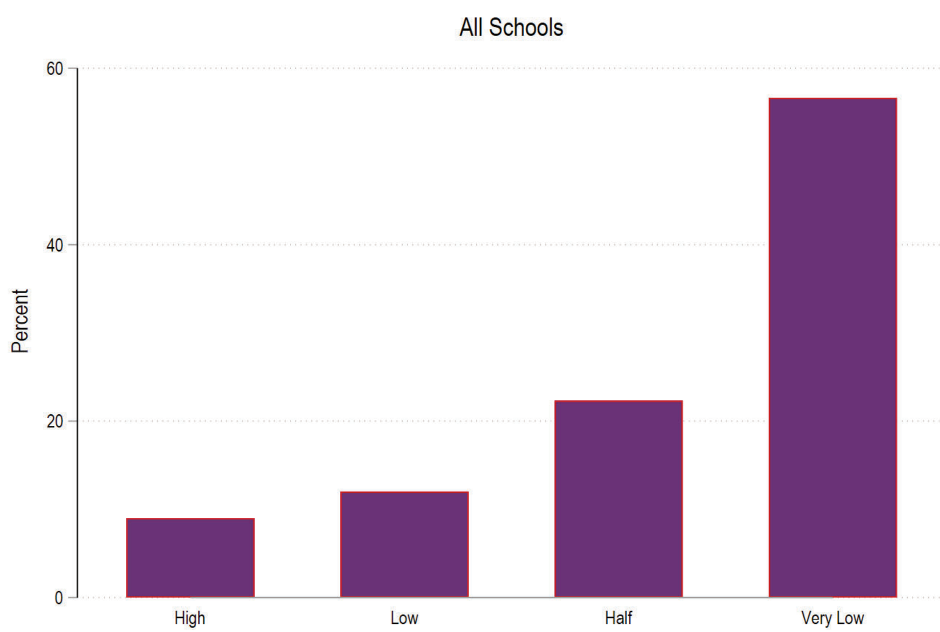


Figure 7: Distribution Socioeconomic Index All Schools

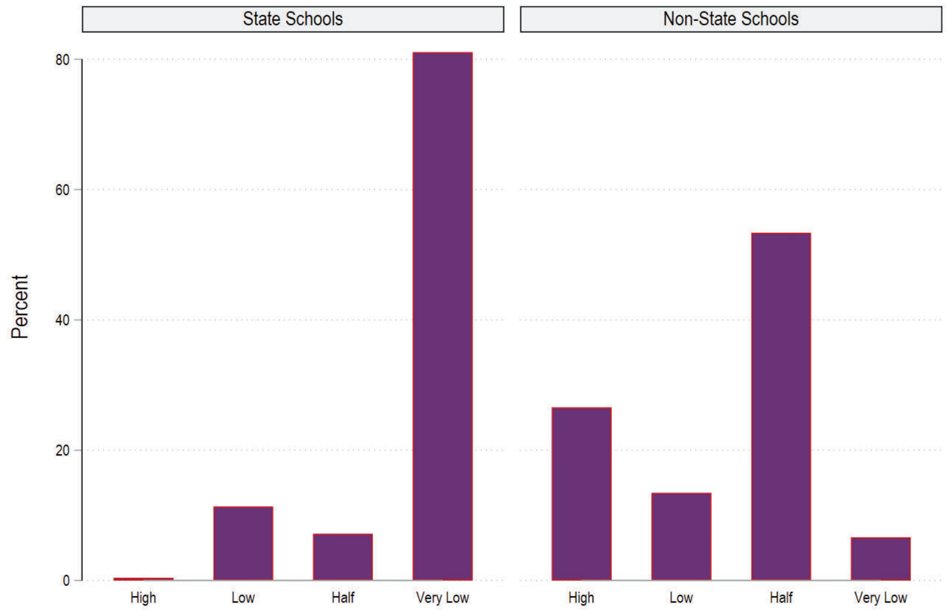


Figure 8: Distribution Socioeconomic Index All Schools State - Non-State

Figure 9: Distribution Socioeconomic Index Assigned Schools and Schools of Origin of Full Sample and Filtered School

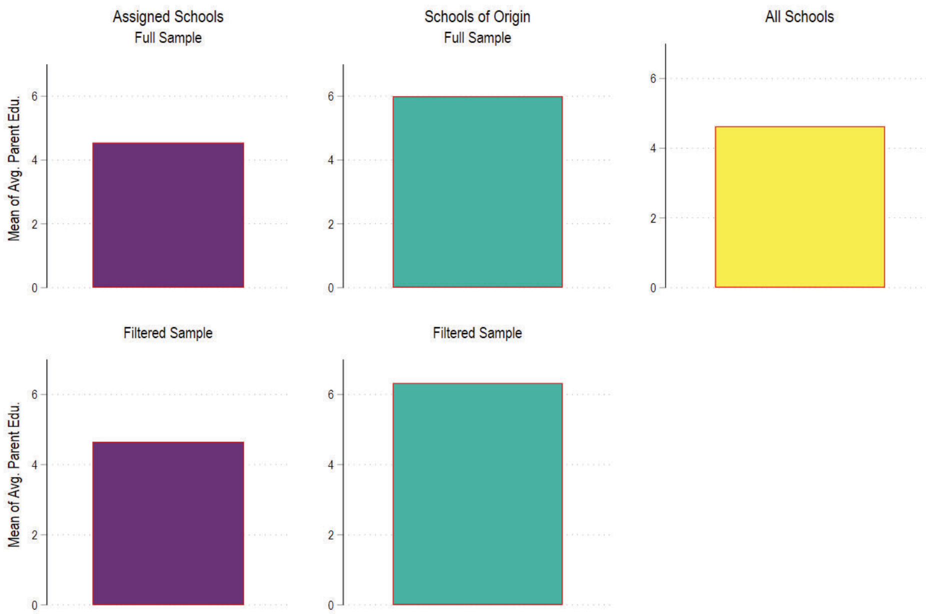


Figure 10: Mean of Average Parent Education of All Schools, Full Sample and Filtered Sample

A.2 Empirical Appendix

A.3 Specification with School Environment Index

In this new specification, school characteristics are incorporated into the School Environment Index. This allows us to address the fact that, even though school characteristics were randomly assigned, they were done jointly rather than individually, making it impossible to isolate the exogenous variation of each covariate.

However, we believe it is plausible to isolate the exogenous variation of distance, thus identifying a causal effect. The threat to this identification would be the presence of unobservable factors related to distance that are not observable to econometricians. Even when conditioning on other observable school characteristics, these factors may still influence the distance variable and have an impact on our dependent variable.

The School Environment Index is defined as follows:

$$\text{SE Index} = \frac{\begin{array}{r}\text{Mean Math sc} + \text{Mean Lang. sc.} + \text{Mean Parent Edu.} + \text{Log(Num Students)}+ \\ \text{STR} + \text{Prop. Non-Novice teachers} + \text{Prop. Permanent Teachers}\end{array}}{7}$$

Table 5: The effect of distance and School Environment Index on likelihood to accept and stay at assigned school

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay
Distance (km)	-1.100***	-2.164***	-1.453***	-3.044***
	(0.163)	(0.231)	(0.344)	(0.514)
Dif. Distance (km)	-0.325***	0.081	-0.540***	0.143
	(0.046)	(0.065)	(0.104)	(0.094)
School Environment Index	1.797***	6.056***	2.314**	8.109***
	(0.442)	(0.853)	(1.138)	(1.753)
Price school of origin	-0.018***	-0.025***	-0.021***	-0.034***
	(0.003)	(0.004)	(0.005)	(0.008)
Constant	87.485***	84.750***	80.662***	82.103***
	(0.617)	(1.067)	(1.100)	(1.982)
Observations	54,209	40,905	22,723	17,605
R-squared	0.060	0.066	0.128	0.135
FE: Dist-Gr-Lev	X	X		
FE: Female	X	X	X	X
FE: Same Top 2			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(a) Note: Standard errors have been clustered at the district by level and grade level.

Table 6: The effect of distance and School Environment Index on likelihood to accept and stay at assigned school

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay
Distance (km)	-1.118***	-2.158***	-1.315***	-2.975***
	(0.178)	(0.230)	(0.381)	(0.539)
Dif. Distance (km)	-0.317***	0.084	-0.576***	0.158
	(0.052)	(0.075)	(0.119)	(0.104)
School Environment Index	-6.456***	-2.660**	-9.845***	-4.346*
	(0.757)	(1.246)	(1.415)	(2.414)
Dif. SE Index	8.336***	9.191***	11.886***	12.855***
	(0.749)	(0.907)	(1.256)	(1.774)
Price school of origin	-0.008***	-0.013***	-0.007**	-0.017***
	(0.002)	(0.003)	(0.003)	(0.006)
Constant	84.787***	81.887***	76.837***	77.907***
	(0.507)	(0.882)	(1.083)	(1.697)
Observations	47,388	35,149	19,724	14,864
R-squared	0.072	0.078	0.151	0.166
FE: Dist-Gr-Lev	X	X		
FE: Female	X	X	X	X
FE: Same Top 2			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(a) Note: Standard errors have been clustered at the district by level and grade level.

Table 7: The effect of distance and School Environment Index on likelihood to accept and stay at assigned school - Standardized

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay
Distance (km)	-2.154***	-4.236***	-2.844***	-5.958***
	(0.319)	(0.453)	(0.674)	(1.007)
Dif. Distance (km)	-1.712***	0.425	-2.841***	0.754
	(0.240)	(0.342)	(0.548)	(0.495)
School Environment Index	1.797***	6.056***	2.314**	8.109***
	(0.442)	(0.853)	(1.138)	(1.753)
Price school of origin	-3.533***	-4.860***	-4.033***	-6.688***
	(0.504)	(0.786)	(0.951)	(1.612)
Constant	81.732***	75.159***	73.780***	68.748***
	(0.133)	(0.193)	(0.269)	(0.417)
Observations	54,209	40,905	22,723	17,605
R-squared	0.060	0.066	0.128	0.135
FE: Dist-Gr-Lev	X	X		
FE: Female	X	X	X	X
FE: Same Top 2			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(a) Note: Standard errors have been clustered at the district by level and grade level.

Table 8: The effect of distance and School Environment Index on likelihood to accept and stay at assigned school - Standardized

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay
Distance (km)	-2.189***	-4.225***	-2.573***	-5.823***
	(0.348)	(0.451)	(0.746)	(1.056)
Dif. Distance (km)	-1.669***	0.441	-3.030***	0.830
	(0.276)	(0.395)	(0.627)	(0.549)
School Environment Index	-6.456***	-2.660**	-9.845***	-4.346*
	(0.757)	(1.246)	(1.415)	(2.414)
Dif. SE Index	8.336***	9.191***	11.886***	12.855***
	(0.749)	(0.907)	(1.256)	(1.774)
Price school of origin	-1.658***	-2.600***	-1.389**	-3.358***
	(0.348)	(0.606)	(0.646)	(1.076)
Constant	80.917***	74.627***	72.981***	68.108***
	(0.112)	(0.194)	(0.266)	(0.393)
Observations	47,388	35,149	19,724	14,864
R-squared	0.072	0.078	0.151	0.166
FE: Dist-Gr-Lev	X	X		
FE: Female	X	X	X	X
FE: Same Top 2			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(a) Note: Standard errors have been clustered at the district by level and grade level.

To demonstrate that it is plausible to consider that the variation in the assigned distance to school is independent of the other covariates, the following table is presented, where the dependent variable is the assigned distance to school. It can be observed that none of the covariates are significant.

Table 9: Relation Distance and Covariates

VARIABLES	(1) Distance (km)
STR	0.001
	(0.007)
Mean Math sc (s)	0.092
	(0.166)
Mean Lang sc (s)	-0.127
	(0.214)
Mean parent edu	-0.046
	(0.061)
Log(Num Students)	-0.366
	(0.397)
Log(Num Students) cuad	0.023
	(0.033)
Prop. inexperienced teachers	0.175
	(0.122)
Prop. temporary teachers	0.279**
	(0.128)
Price school of origin	-0.000
	(0.000)
Constant	3.601***
	(1.198)
Observations	55,031
R-squared	0.356
FE: Dist-Gr-Lev	X
FE: Female	X
F-test:	6.84
Prob > F:	0.0000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(a) Note: Standard errors have been clustered at the district by level and grade level.

A.3.1 Main Specification - Robustness Check

In Tables 10 and 11, we present a robustness check of specification (6). The first four columns are the same as in Table 1. Columns (5) and (6) change the set of fixed effects we use. In these columns, we used a unique fixed effect per set of the same top 3 schools, which means an even stricter imposition on the closeness between households. After conditioning by the more stringent set of fixed effects, we do not observe drastic differences with our preferred specification. Notice that we experience a drop in the number of observations between our Figure 1 and column (1) from Table 1.

ov

Table 10: Main Specification - Robustness Check

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay	(5) accept	(6) stay
Distance (km)	-1.058***	-2.205***	-1.164***	-2.987***	-1.192**	-3.204***
	(0.182)	(0.233)	(0.385)	(0.539)	(0.506)	(0.699)
Dif. Distance (km)	-0.328***	0.068	-0.574***	0.145	-0.765***	0.083
	(0.053)	(0.072)	(0.119)	(0.095)	(0.209)	(0.155)
Dif. Mean Math sc (s)	3.791***	3.404***	4.874***	5.560***	4.695***	5.041***
	(0.509)	(0.756)	(0.963)	(1.634)	(1.039)	(1.925)
Dif. Mean Parent Edu.	1.894***	3.887***	2.736***	5.705***	3.450***	6.731***
	(0.295)	(0.408)	(0.533)	(0.678)	(0.774)	(0.896)
Log(Ratio Num St.)	2.118***	4.403***	3.212***	4.310***	3.802**	5.852***
	(0.540)	(0.626)	(1.206)	(1.131)	(1.463)	(1.420)
$Log(RatioNumSt.)^2$	-0.118	-0.646***	-0.187	-0.910***	-0.032	-1.246***
	(0.164)	(0.156)	(0.380)	(0.325)	(0.446)	(0.458)
Dif. STR	-0.056**	-0.011	-0.056	0.117*	-0.066	0.060
	(0.022)	(0.038)	(0.052)	(0.059)	(0.068)	(0.080)
Dif. Prop. Novice Teachers	-1.196	-1.397**	-1.722	-1.814	-2.429	-1.035
	(0.766)	(0.657)	(1.303)	(1.171)	(1.555)	(1.229)
Dif. Prop. Temporary Teachers	-1.482	1.989	-3.348	0.467	-0.763	-1.595
	(1.331)	(2.139)	(3.318)	(3.405)	(4.502)	(3.948)
Price School of Origin	-0.008***	-0.007***	-0.006*	-0.009**	-0.006*	-0.011**
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)
Constant	83.690***	83.810***	73.788***	80.558***	72.165***	77.968***
	(1.164)	(1.873)	(2.876)	(2.854)	(3.781)	(3.656)
Observations	47,407	35,174	19,739	14,885	13,838	10,261
R-squared	0.072	0.082	0.151	0.170	0.158	0.180
Inc. Female	X	X	X	X	X	X
FE: Dist-Gr-Lev	X	X				
FE: Same Top 2			X	X		
FE: Same Top 3					X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Main Specification - Robustness Check - Standardized

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay	(5) accept	(6) stay
Distance (km)	-2.071***	-4.317***	-2.278***	-5.847***	-2.333**	-6.271***
	(0.356)	(0.456)	(0.754)	(1.055)	(0.990)	(1.368)
Dif. Distance (km)	-1.722***	0.359	-3.020***	0.762	-4.024***	0.435
	(0.277)	(0.380)	(0.626)	(0.499)	(1.101)	(0.813)
Dif. Mean Math sc (s)	2.105***	1.890***	2.706***	3.087***	2.606***	2.799***
	(0.282)	(0.420)	(0.535)	(0.907)	(0.577)	(1.069)
Dif. Mean Parent Edu.	1.987***	4.077***	2.869***	5.983***	3.619***	7.059***
	(0.310)	(0.428)	(0.559)	(0.711)	(0.811)	(0.940)
Log(Ratio Num St.)	2.359***	4.904***	3.578***	4.801***	4.234**	6.518***
	(0.602)	(0.697)	(1.343)	(1.260)	(1.629)	(1.581)
$Log(RatioNumSt.)^2$	-0.350	-1.908***	-0.554	-2.690***	-0.095	-3.684***
	(0.486)	(0.461)	(1.124)	(0.959)	(1.319)	(1.353)
Dif. STR	-0.523**	-0.106	-0.519	1.084*	-0.614	0.556
	(0.206)	(0.351)	(0.480)	(0.551)	(0.636)	(0.742)
Dif. Prop. Novice Teachers	-0.451	-0.526**	-0.649	-0.684	-0.915	-0.390
	(0.289)	(0.248)	(0.491)	(0.441)	(0.586)	(0.463)
Dif. Prop. Temporary Teachers	-0.426	0.572	-0.962	0.134	-0.219	-0.458
	(0.382)	(0.615)	(0.953)	(0.979)	(1.294)	(1.134)
Price School of Origin	-1.484***	-1.374***	-1.110*	-1.751**	-1.241*	-2.142**
	(0.339)	(0.404)	(0.636)	(0.816)	(0.705)	(0.998)
Constant	80.728***	75.008***	72.298***	68.111***	69.233***	64.914***
	(0.137)	(0.212)	(0.316)	(0.357)	(0.374)	(0.462)
Observations	47,407	35,174	19,739	14,885	13,838	10,261
R-squared	0.072	0.082	0.151	0.170	0.158	0.180
Inc. Female	X	X	X	X	X	X
FE: Dist-Gr-Lev	X	X				
FE: Same Top 2			X	X		
FE: Same Top 3					X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

A.3.2 Alternative Specification

This section will explore an alternative specification similar to the main one. We will also use two dependent variables: first, Accept_{isdgl} is an indicator variable at the student i, school s, district d, level l, and grade g level that takes a value of 100 if the student decides to accept her assignment in the first stage and a value of 0 if the student rejects her assignment; and second Staying_{isdgl} which takes a value of 100 if the student decides to stay at her assigned school and a value of 0 if the student moves to another school conditional on having accepted the assignment.

The independent variables of interest $\mathcal{X}_{sd}$ aim to capture school s at district d characteristics of parents' potential interest when they make a decision. First, we calculate the distance Dist_{isd} between the school s and the student i addresses in district d. In addition, we find information on school class sizes as measured by the student-to-teacher ratio, average parents' education, average measures of academic performance (math), and school size. Also, we have included characteristics that give us an idea of the relevance of teaching quality, such as the proportion of novice teachers (with less than five years of experience) and the proportion of temporary teachers (teachers who did not meet the minimum standards). We include the school of origin's prices to capture an approximate measure of income level and control for students' gender. We include district d by grade g and level l fixed effects denoted by α_{dgl} to control for the systematic differences in the feasible offer sets of schools. Finally, e_{isdgl} represents the error term.

In summary, we will use the same variables as in specification (6), but in levels instead of the differences between the assigned and original schools. The results are in Table 12 and 13

$$y_{isdgl} = \alpha_{dgl} + \eta \cdot \text{Dist}_{isd} + \mathcal{X}'_{sd}\Omega + e_{isdgl} \tag{7}$$

Table 12: Alternative Results - Levels

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay	(5) accept	(6) stay
Distance (km)	-1.388***	-2.162***	-2.003***	-2.855***	-2.470***	-3.313***
	(0.156)	(0.220)	(0.311)	(0.468)	(0.371)	(0.592)
Mean Math sc (s)	0.866	2.265	0.126	2.516	-0.318	1.899
	(0.766)	(1.492)	(1.453)	(2.961)	(1.662)	(3.707)
Mean Parent Edu.	1.447***	2.324***	3.370***	7.470***	4.844***	9.310***
	(0.484)	(0.811)	(1.028)	(1.400)	(1.385)	(2.163)
Log(Num Students)	6.605	19.710***	1.850	49.520***	14.823	63.642***
	(4.918)	(6.617)	(10.866)	(10.434)	(12.400)	(11.694)
$Log(NumStudents)^2$	-0.495	-1.300**	-0.138	-3.837***	-1.254	-4.999***
	(0.394)	(0.518)	(0.875)	(0.793)	(0.996)	(0.897)
STR	-0.048	0.174*	-0.033	0.317	-0.065	0.193
	(0.054)	(0.089)	(0.114)	(0.194)	(0.155)	(0.218)
Prop. Novice Teachers	2.662**	-2.219	6.260**	-0.476	8.352**	-0.263
	(1.225)	(1.646)	(2.591)	(2.823)	(3.394)	(3.356)
Prop. Temporary Teachers	-1.545	2.506	-5.281**	0.588	-5.667*	-3.473
	(1.212)	(2.714)	(2.459)	(3.852)	(3.356)	(4.079)
Price School of Origin	-0.018***	-0.025***	-0.020***	-0.035***	-0.024***	-0.039***
	(0.003)	(0.004)	(0.005)	(0.008)	(0.006)	(0.011)
Constant	60.926***	-2.290	61.155*	-119.640***	15.889	-168.051***
	(15.125)	(21.754)	(34.630)	(35.699)	(39.261)	(39.953)
Observations	55,031	41,742	22,956	17,841	16,226	12,478
R-squared	0.058	0.068	0.127	0.137	0.132	0.142
Inc. Female	X	X	X	X	X	X
FE: Dist-Gr-Lev	X	X				
FE: Same Top 2			X	X		
FE: Same Top 3					X	X

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 13: Alternative Results - Levels (Standardized)

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay	(5) accept	(6) stay
Distance (km)	-2.718***	-4.232***	-3.921***	-5.590***	-4.835***	-6.485***
	(0.306)	(0.431)	(0.609)	(0.917)	(0.726)	(1.158)
Mean Math sc (s)	0.290	0.759	0.042	0.843	-0.107	0.636
	(0.257)	(0.500)	(0.487)	(0.992)	(0.557)	(1.242)
Mean Parent Edu.	0.960***	1.541***	2.235***	4.953***	3.212***	6.173***
	(0.321)	(0.538)	(0.682)	(0.928)	(0.918)	(1.434)
Log(Num Students)	5.516	16.460***	1.545	41.354***	12.379	53.147***
	(4.107)	(5.526)	(9.075)	(8.713)	(10.355)	(9.766)
Log(Num Students)^2	-4.700	-12.335**	-1.306	-36.416***	-11.897	-47.442***
	(3.742)	(4.921)	(8.303)	(7.530)	(9.457)	(8.514)
STR	-0.259	0.939*	-0.179	1.707	-0.351	1.040
	(0.289)	(0.480)	(0.613)	(1.047)	(0.838)	(1.176)
Prop. Novice Teachers	0.470**	-0.392	1.105**	-0.084	1.475**	-0.046
	(0.216)	(0.291)	(0.457)	(0.498)	(0.599)	(0.593)
Prop. Temporary Teachers	-0.347	0.562	-1.185**	0.132	-1.272*	-0.780
	(0.272)	(0.609)	(0.552)	(0.865)	(0.753)	(0.916)
Price School of Origin	-3.571***	-4.818***	-3.964***	-6.829***	-4.596***	-7.555***
	(0.513)	(0.770)	(0.929)	(1.620)	(1.199)	(2.159)
Constant	81.501***	74.095***	73.469***	66.640***	70.102***	63.353***
	(0.276)	(0.354)	(0.695)	(0.743)	(0.763)	(0.833)
Observations	55,031	41,742	22,956	17,841	16,226	12,478
R-squared	0.058	0.068	0.127	0.137	0.132	0.142
Inc. Female	X	X	X	X	X	X
FE: Dist-Gr-Lev	X	X				
FE: Same Top 2			X	X		
FE: Same Top 3					X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

A.4 Heterogeneous Effects

A.4.1 Heterogeneous effect: Distance

Next, we discuss the non-linear effects of distance on the likelihood of accepting and staying at the assigned school:

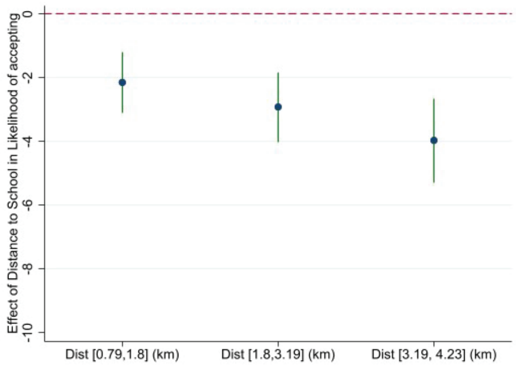
$$y_{isdgl} = \alpha_{dgl} + \sum_{n=2}^{n=4} \eta_r \cdot I\{Dist_{isd} \in Q_r\} + \mathcal{D}'_{sd}\Lambda + e_{isdgl} \tag{8}$$

We explore the heterogeneity of the effects of distance using the same set of variables as specification (6) with the exception of the distance variable. We replace the distance variable with an indicator that categorizes distance into their respective quartiles in the distribution resulting in equation (8). We denote each quartile of the distribution of distance to school as Q_r where r = 1 denotes the group with the smallest distances, and r = 4 denotes the group with the largest distances. The group ranges are [0, 0.79), [0.79, 1.8), [1.8, 3.19), and [3.19, 4.23].

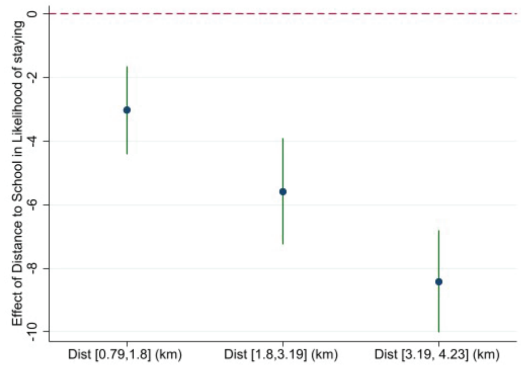
We include the regression tables in Appendix Table 14 and plot the estimates for η_r for $r \in 2, 3, 4$ from column (1) from each specification. In all specifications, we exclude the smallest distance category ($r = 1$); therefore, the coefficients are meant to be interpreted as differences between each group and the group in the first category. We also run a similar specification observed in Figure 12 where, instead of using distance quartiles, we bin distances by km to observe if a more granular analysis of distance yields different conclusions. This approach yields the following group ranges [0,1), [1,2), [2,3), [3,4), and [4,5). Similarly to specification 8, we drop the smallest distance category ([0,1)) and plot the coefficients, which are meant to be interpreted as differences between each group and the group in the first category.

Figure 11 (both panels) shows a surprisingly linear relationship between the distance between an increase in distance to the school and the likelihood of accepting and staying at the assigned school when allowing for non-linear effects in each group category of distance. We find that the coefficients of the effect on the likelihood of accepting the assigned school for each group, when compared to the smallest quartile of distance, vary between -2.13 p.p. up to -4 p.p. We cannot rule out that the effect sizes for all these categories concerning the first group are the same as we observe overlapping confidence intervals. It is important to note that in the first stage, we incorporate the difference in distances between the origin and the assigned school as an additional regressor. This variable is significant in the first stage and may tamper the overall effect of distance on the likelihood of accepting the school assignment. Also, the coefficients of the effect on the likelihood of staying at the assigned school for each group, when compared to the smallest quartile of distance, vary between -3.06 p.p. up to -8.42 p.p. (the effects can reach up to -11.18 considering different sets of fixed effects). These effects are significant considering the size of the rest of the coefficients in the Table and the economic implication of these estimates. We find a drop of about 8 p.p. in the likelihood of staying at their assigned school if attending a school that is more than 3 km from the parent's house compared to being assigned to a school that is within 1 km from a parent's house implies that distance is a very relevant factor when assessing the desired qualities in a school for a family. In contrast to panel (a), we observe that the effect sizes of the largest quartile group are statistically different when compared to that

Figure 11: Heterogeneous effects of distance by quartiles

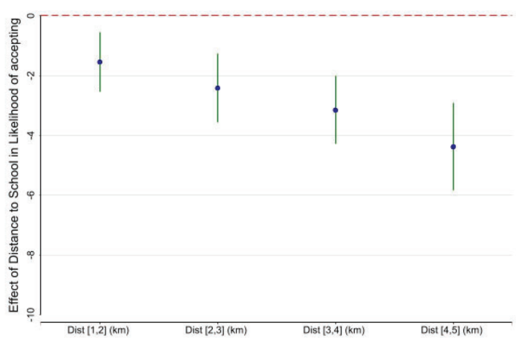


(a) On likelihood of Accepting

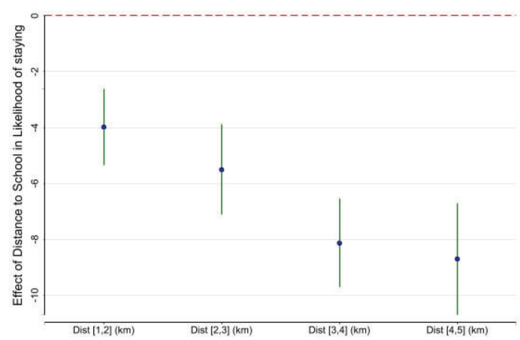


(b) On likelihood of Staying

Figure 12: Heterogeneous effects of distance by sections



(a) On likelihood of Accepting



(b) On likelihood of Staying

of the second quartile ([0.79,1.8)).

In contrast to Figure 11, although we still observe a linear pattern in the coefficients from panel (a), we now find that the point estimates of the coefficients of panel (b) seem to remain constant after 3km. However, we can only differentiate the coefficients in the second category [1,2) from those in categories [3,4) and [4,5). The effects of distance on the likelihood of staying at the assigned school grow from -3.98 in the second category to -8.7 in the last category, more than doubling. As discussed in the text, parents in the second stage seem very sensitive to distance..

In the Table 14 we show the results of the specifications (8), which explore the non-linear effects of distance on the likelihood of accepting and staying at the assigned school.

Table 14: Heterogeneous effects of distance by quartiles

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay	(5) accept	(6) stay
Dist. [0.79,1.8] (km)	-2.125***	-3.059***	-2.564*	-3.573**	-2.710	-2.933
	(0.480)	(0.703)	(1.325)	(1.547)	(1.889)	(2.072)
Dist. [1.8,3.19] (km)	-2.925***	-5.586***	-5.279***	-6.541***	-5.626***	-6.229***
	(0.557)	(0.844)	(1.353)	(1.737)	(1.742)	(2.377)
Dist. [3.19,4.23] (km)	-4.003***	-8.416***	-4.457***	-11.180***	-4.580**	-11.218***
	(0.672)	(0.815)	(1.492)	(1.757)	(2.055)	(2.292)
Dif. Distance (km)	-0.344***	0.051	-0.604***	0.115	-0.802***	0.040
	(0.052)	(0.070)	(0.115)	(0.093)	(0.200)	(0.151)
Dif. Mean Math sc (s)	3.772***	3.390***	4.827***	5.516***	4.661***	4.992**
	(0.507)	(0.759)	(0.975)	(1.641)	(1.055)	(1.935)
Dif. Mean Parent Edu.	1.890***	3.891***	2.767***	5.746***	3.496***	6.788***
	(0.296)	(0.407)	(0.532)	(0.680)	(0.776)	(0.898)
Log(Ratio Num St.)	2.118***	4.396***	3.215***	4.327***	3.809**	5.854***
	(0.541)	(0.624)	(1.207)	(1.130)	(1.470)	(1.417)
Log(Ratio Num St.)^2	-0.114	-0.639***	-0.190	-0.910***	-0.040	-1.246***
	(0.164)	(0.155)	(0.379)	(0.319)	(0.447)	(0.453)
Dif. STR	-0.056**	-0.011	-0.053	0.117**	-0.063	0.062
	(0.022)	(0.038)	(0.051)	(0.059)	(0.068)	(0.080)
Dif. Prop. Novice Teachers	-1.201	-1.392**	-1.744	-1.805	-2.464	-0.991
	(0.766)	(0.657)	(1.303)	(1.174)	(1.556)	(1.228)
Dif. Prop. Permanent Teachers	1.456	-2.017	3.127	-0.550	0.528	1.524
	(1.322)	(2.137)	(3.283)	(3.420)	(4.453)	(3.968)
Price School of Origin	-0.008***	-0.007***	-0.006*	-0.009**	-0.006*	-0.011**
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)
Constant	83.776***	83.551***	74.662***	79.807***	73.165***	76.517***
	(1.155)	(1.970)	(2.921)	(2.847)	(3.881)	(3.659)
Observations	47,407	35,174	19,739	14,885	13,838	10,261
R-squared	0.072	0.082	0.151	0.170	0.159	0.179
Inc. Female	X	X	X	X	X	X
FE: Dist-Gr-Lev	X	X				
FE: Same Top 2			X	X		
FE: Same Top 3					X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 15: Heterogeneous effects of distance by sections

VARIABLES	(1) accept	(2) stay	(3) accept	(4) stay	(5) accept	(6) stay
Dist. [1,2] (km)	-1.547***	-3.983***	-3.200***	-5.800***	-2.957**	-5.146***
	(0.505)	(0.688)	(1.141)	(1.296)	(1.426)	(1.722)
Dist. [2,3] (km)	-2.420***	-5.506***	-5.605***	-7.998***	-5.595***	-7.844***
	(0.585)	(0.821)	(1.345)	(1.636)	(1.873)	(2.193)
Dist. [3,4] (km)	-3.151***	-8.130***	-4.464***	-11.681***	-4.343**	-11.305***
	(0.573)	(0.807)	(1.374)	(1.416)	(1.899)	(1.976)
Dist. [4,5] (km)	-4.386***	-8.701***	-5.351***	-12.339***	-5.236**	-13.221***
	(0.741)	(1.016)	(1.649)	(2.175)	(2.029)	(3.013)
Dif. Distance (km)	-0.333***	0.059	-0.579***	0.137	-0.777***	0.073
	(0.053)	(0.072)	(0.117)	(0.095)	(0.204)	(0.157)
Dif. Mean Math sc (s)	3.796***	3.403***	4.840***	5.533***	4.684***	5.020**
	(0.509)	(0.756)	(0.965)	(1.650)	(1.041)	(1.932)
Dif. Mean Parent Edu	1.887***	3.891***	2.746***	5.735***	3.469***	6.754***
	(0.295)	(0.406)	(0.533)	(0.686)	(0.770)	(0.901)
Log(Ratio Num St.)	2.117***	4.401***	3.198***	4.298***	3.807**	5.859***
	(0.540)	(0.623)	(1.205)	(1.124)	(1.464)	(1.408)
Log(Ratio Num St.)^2	-0.117	-0.644***	-0.184	-0.906***	-0.038	-1.246***
	(0.164)	(0.155)	(0.379)	(0.319)	(0.446)	(0.453)
Dif. STR	-0.056**	-0.012	-0.054	0.117*	-0.064	0.061
	(0.022)	(0.038)	(0.051)	(0.059)	(0.068)	(0.081)
Dif. prop. Novice Teachers	-1.206	-1.418**	-1.738	-1.794	-2.458	-1.012
	(0.769)	(0.656)	(1.313)	(1.177)	(1.570)	(1.233)
Dif. Prop. Permanent Teachers	1.445	-2.077	3.159	-0.473	0.631	1.577
	(1.322)	(2.129)	(3.296)	(3.402)	(4.465)	(3.937)
Price School of Origin	-0.008***	-0.007***	-0.006*	-0.009**	-0.006*	-0.011**
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)
Constant	83.356***	83.577***	74.695***	80.657***	72.871***	77.438***
	(1.136)	(1.972)	(2.802)	(2.756)	(3.670)	(3.494)
Observations	47,398	35,167	19,735	14,883	13,837	10,259
R-squared	0.072	0.082	0.152	0.171	0.159	0.180
Inc. Female	X	X	X	X	X	X
FE: Dist-Gr-Lev	X	X				
FE: Same Top 2			X	X		
FE: Same Top 3					X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1