A Dynamic Model of Elementary School Choice

Autores:

Nicolas Grau

Santiago, Enero de 2016

A Dynamic Model of Elementary School Choice

Nicolás Grau[*]

June 2015

Abstract

This paper builds and estimates a dynamic model of elementary school choice using detailed Chilean administrative data. In the model, parents care about different features of primary schools: school's socioeconomic composition, quality (measured as the school's contribution to standardized test scores), religiosity, location, type of administration, tuition fee and GPA standard. Parents are heterogeneous in two dimensions: whether they have the skills needed to understand public information about quality (standardized tests), and their involvement in their child's school. The results suggest that: (1) Parents care about school quality, but to a moderate degree. (2) Parents have an important misperception about school quality, which results in a less favorable opinion about the quality of public schools, relative to private schools. (3) If parents were only concerned about quality, they would choose public schools more often. (4) Admission restrictions play a relevant role; otherwise, parents would choose private school more frequently.

JEL Classification(s): I24, C35, D03

---

[*]ngrau@fen.uchile.cl. Department of Economics, University of Chile and Centre for Social Conflict and Cohesion Studies (COES)

1 Introduction

A frequent topic in policy debates is what should be the role – if any – of market incentives in education provision. Given that parents' choice is the critical mechanism to increase school quality in a market-oriented educational system, the literature has focused on the extent to which parents consider school quality when they make their decisions, and how this consideration is heterogeneous across parents. To understand parents' school choice, and the potential heterogeneity in their preferences, one must separate the effects of differences in their preferences, in perceptions about quality, and in choice sets. Distinguishing these three elements is a complex task given that, in general, these determinants of parents' choice are not observable.

In this paper, I build and estimate a dynamic model of elementary school choice. To this end, I use detailed Chilean administrative data for the students who entered 1st grade in 2004. As many authors have emphasized (e.g., Gallego and Hernando (2008) and Hsieh and Urquiola (2006)), the Chilean system is probably the most massive school choice program in the world, hence the importance of studying the determinants of school choice in this context.

I model elementary school choice at the end of each academic year, allowing for parental heterogeneity along several dimensions: their ability to understand public information about quality (standardized tests), how much they care about school quality (measured as the school's contribution to standardized test scores), their involvement in the school attended by their child, and their choice set.[1] By estimating the structural parameters of the model, I am able to assess the empirical relevance of these components in explaining both the observed preference for private over public schools and the unequal access to high quality schools.

In the model, parents care about different characteristics of primary schools, such as the school's socioeconomic composition, quality, religious affiliation, location, type of administration (i.e., public, subsidized private and non-subsidized private), tuition fee, and GPA standard (i.e., grading standard). Parents do not perfectly observe school quality. To estimate the quality of each school, they can access two different sources of information. First, every year they observe the performance of each school on a standardized test, which is made public with a one year lag. Parents can have different levels of misperception in processing this information; because test scores depend on school quality and on the socioeconomic status (SES) of the school, parents can confound these two effects, confusing high quality schools with schools that have higher SES students. Second, parents also differ in their exogenous level of involvement in the schooling process of their child, which implies that those who are involved in their child's school observe the quality of that particular school without misperception.

I estimate the parameters of the model by simulated maximum likelihood, using the Monte Carlo integration and interpolation method (Keane and Wolpin (1994)). To build

---

[1]In Chile, at least in my sample period, schools were allowed to select students based on academic and non-academic characteristics (e.g., parents' marital status).

the database of students, I use the administrative panel data from 2004 to 2011, which includes the school attended by each student in each year, their average grade, the municipality where they live and where the school is located, and some basic demographic information. Because the sample of students entering 1st grade in 2004 took the SIMCE test in 4th grade (2007) and in 8th grade (2011), I merge this panel with information from parent surveys associated with those rounds of SIMCE administration, including mother's and father's information. To build the database of schools, I use the test scores from the SIMCE test and the information collected from SIMCE parents' surveys for the years 2002, 2005-2011.[2] Test scores are used to estimate school quality for every year. The surveys include questions about school tuition fees, and information about the elements considered by the school in the admission process. Furthermore, from administrative data of the Ministry of Education, I collect information about schools' religious affiliation, if any.

The results show that parents do care –but in a moderate way– about school quality, that more involved parents care marginally more about school quality, and that parents' decisions are not sensitive to quality after the first decision (1st grade). Moreover, the results also suggest that parents have an important misperception about school quality, which results in a less favorable opinion about the quality of public schools, relative to private schools. This result supports the idea that parents may have difficulties in isolating a school's quality from its socioeconomic composition when they observe test scores. However, given that quality is not very relevant for their decision, such a misperception only partially affects parents' choices.

Regarding the debate about why parents choose private schools over public schools, the results show that, if parents were only concerned about quality, they would choose public schools more often. The same would be true if they did not have a misperception about quality. However, the results suggest that admission rules are binding restrictions and that relaxing them would increase the demand for private schools. The simulations also show that schools' admission rules and household location are both important in explaining the rise in the achievement gap between students from different SES.

The paper has three main contributions. First, to the best of my knowledge, this is the first paper that structurally estimates a dynamic model of elementary school choice. The dynamic nature of school choice has particular relevance in the Chilean context, where around 30% of the students switch schools at least once between 1st and 8th grade (excluding those who moved to other municipalities and those who fail at least one year). Second, the structural approach followed in this paper allows me to quantify different causes of unequal access to high quality schools and of the higher demand for private schools than for public schools. Finally, the model considers the difficulties that parents may have in processing and understanding information about school quality, which contributes to the scarce literature on structural estimation with bounded rationality, as well as to the literature that uses observed choices to infer agents' information.

---

[2]4th grade SIMCE for years 2002, 2005-2009 and 8th grade SIMCE for years 2004, 2007, 2009 and 2011.

The structure of the paper is as follows: Section 2 presents a review of the related literature; Section 3 briefly describes the Chilean educational system; Section 4 introduces the model; Section 5 discusses the data and the procedure to estimate the model; and Section 6 presents the results and the analysis of the counterfactual experiments. Section 7 concludes.

2   Literature Review

This paper is related to several strands in the literature. First of all, it is related to the papers that evaluate the role of competitive market incentives in education provision. On one side, there are theoretical papers, such as Epple and Romano (1998), and McMillan (2004), which debate the potential for those incentives, specifically tuition vouchers, to increase schools' quality and to make significant improvements for poor families.[3] On the other side, there are empirical studies that show mixed evidence regarding this debate; see, for example, Hsieh and Urquiola (2006), and Angrist et al. (2006).[4]

Authors have studied the determinants of parents' school choice because this is one of the important mechanisms that could explain the shortcomings in the implementation of market-oriented policies in education. In particular, they have studied whether and to what extent parents consider school quality when they make their choice.[5] For instance, in an interesting paper, Hastings and Weinstein (2008) used a natural experiment and a field experiment that provided direct information on school test scores to lower-income families in a public school choice plan, finding a significant increase in the probability that those families would choose higher-performing schools. In an alternative strategy, several studies have focused on estimating the value that parents place on school quality by calculating how much more people pay for houses located in areas with better schools (e.g., Black (1999) and Kane et al. (2006)).

One caveat about this literature is that, in general, it measures school quality using average school test scores. The problem with this approach is that, from the point of view of the parents, this average is not relevant; what is relevant is what their child's performance would be if she were to attend a particular school. Given that sorting is a common feature in education, these two conditional expectations should not coincide. Exceptions to this general problem are shown in Mizala and Urquiola (2013), Neilson (2013), and Rothstein (2006). For instance, Mizala and Urquiola (2013) use a sharp regression discontinuity to estimate the effect that being identified as a SNED winner (a program which seeks to identify effective schools, controlling for schools' SES) has on

---

[3]In a survey of this literature, Epple and Romano (2012) conclude: *Research taking account of distinctive features of the education "market" has shown that early arguments touting the virtues of laissez-faire flat-rate vouchers were overly optimistic. However, the research does not vindicate voucher opponents who use shortcomings of the laissez-faire voucher to justify the wholesale dismissal of vouchers.*

[4]Bettinger (2011) reviews the cases of Chile, Colombia, and Sweden, emphasizing the context-specific nature of the results.

[5]See for example Alderman et al. (2001), and Bast and Walberg (2004).

schools' enrollment, finding no consistent evidence that winning a SNED award affects this outcome.[6]

Several authors study the determinants of school choice in the Chilean context.[7] For instance, Gallego and Hernando (2008), using a semi-structural approach, find results that suggest that the school choice implemented in Chile increased overall student welfare, but they also find that there is a lot of heterogeneity in the size and even the sign of the welfare change. Along the same lines, Chumacero et al. (2011), using a database that accurately estimates the distance between the household and school, find that both quality and distance are highly valued by households. In a recent and novel paper, Neilson (2013) studies the effects of targeted school vouchers on the outcomes of poor children in Chile; his findings suggest that this program effectively raised competition in poor neighborhoods, pushing schools to improve their academic quality. Finally, Mark et al. (2006) study how parents construct their school choice sets and comparing this to what they say they are seeking in choosing schools. Their results indicate that parental decisions are influenced by demographics.[8]

This paper is also related to the literature that models individuals' economic decisions incorporating bounded rationality. In general, this literature follows the idea that, as Simon (1986) points out, cognitive effort is a scarce resource, and the knowledge and computational power of the decision-maker are always limited. In an interesting paper, which is one of the few papers that perform a structural estimation with bounded rationality, Houser et al. (2004) develop a Bayesian procedure for classification of subjects into decision rule types in choice experiments, finding that, in a very difficult dynamic problem, more than a third of the experimental subjects followed a rule very close to the optimal (expected wealth maximizing) rule.

Because school choice is a complex task, which involves gathering and processing information, different authors have studied the presence of bounded rationality in that context. For instance, Schneider et al. (1998) find that, on average, low-income parents have very little accurate information about objective conditions in the schools.[9] However, even though levels of objective information held by parents are low, their actual choice of schools reflects their preferences in education. Along the same lines, Azmat and Garcia-Montalvo (2012) conclude that, as well as parents' education, information gathering and information processing are important determinants for the quality of school

---

[6]Mizala et al. (2007) present evidence indicating that, in the case of Chile, once we control for the students' socioeconomic status, the remaining part of the test scores are very volatile from year to year. Hence, they argue that producing a meaningful ranking of schools that may inform parents and policymakers may be harder than is commonly assumed.

[7]Chile's school choice policies will be described in the next section.

[8]There are several studies that try to study the effect of the voucher system implementation in Chile. Although the evidence is mixed regarding its effect on school quality, there is more agreement on the negative effect of this policy on student socioeconomic segregation (Auguste and Valenzuela (2006); Gauri (1999); and Hsieh and Urquiola (2006)). In a different approach, Bravo et al. (2010) find that educational vouchers increased educational attainment, high school graduation, college attendance and graduation, and wages.

[9]The same is found by Henig (1996).

choice.

Finally, this paper is also related to the literature that attempts to infer agents' information using observed choices, such as: Carneiro et al. (2003); Cunha et al. (2005); and Navarro (2011).

3   The Chilean Educational System

In 1981, the Chilean military government created a voucher market in the educational system, which was part of a broader reform that also included the decentralization of public schools (which were transferred to municipalities) and the introduction of flexibility in teachers' contracts.[10] This reform transformed the way schools were funded by the government, establishing a system where private and public schools were paid per student, with a flat voucher, on the basis of attendance.

Since then, the allocation of public resources has been mainly determined by parents' decisions. However, in practice, this decision has had several restrictions: schools can select students based on their previous performance, tests, and the characteristics of their parents (e.g., marital status and religion). On top of that, since 1994, when a co-payment law was passed, schools that are eligible for public funding can also charge a tuition fee; in that case, depending on the amount charged, there is a discount to the school's subsidy.[11]

This reform consolidated a system of mixed provision of education, with three types of schools; municipal (public), private subsidized (voucher-private), and entirely private (non voucher-private). The first two receive most of their funds from state vouchers, and, since 1994, privately subsidized schools may additionally charge a tuition fee. In 2013, over 90% of the Chilean students received funding via vouchers.[12]

In order to guide parents' decisions and to measure the student learning process, a new testing system, SIMCE, came into existence in 1988. The SIMCE is an annual nationwide standardized test. Its results have been public information for more than two decades, publicized in part by listings in major newspapers of individual schools' performance. The government also uses SIMCE scores to allocate resources.[13]

More than 30 years after the reform, there are several clear stylized facts. First, there has been a massive migration from public to private schools. Indeed, the student fraction in the public system went from 78%, in 1981, to 38% in 2012.[14] Secondly, enrollment in

---

[10]For a summary of these reforms, see Gauri (1999) and Mizala and Romaguera (2000).

[11]Epple and Romano (2008) emphasize the consequences of this selection mechanism for the outcomes of an educational voucher system.

[12]Source: Ministry of Education, Chile.

[13]Meckes and Carrasco (2010) describe SIMCE's main features, purposes, institutional framework, and strategies for communicating results.

[14]There is a debate, and mixed evidence, about whether voucher private schools have higher quality than public schools. In a meta-analysis, Drago and Paredes (2011) find that voucher-private schools have a small advantage over public schools. On the contrary, Bellei (2009) finds that voucher-private schools are no more effective than public schools, and that they may be less effective.

voucher-private schools was accelerated after passage of the co-payment law.[15] Thirdly, the magnitude of socioeconomic school segregation is very high (and higher than the geographical segregation), and has increased slightly over the last decade (Valenzuela et al. (2014)). Finally, despite important increases in the public budget allocated to education, Chile's performance is relatively poor when compared with similar countries (Chumacero et al. (2011)).

Another salient feature of the Chilean system, which is consistent with its "free choice" design, is that a fairly large number of parents switch schools at some point during primary school. In this regard, Table 10 of the Appendix A shows that, in any grade, around 4-7% of the parents change their child's school, and that more educated parents are more likely to do so.[16] Moreover, Table 9 of the Appendix A shows that more than 30% of parents changed their child's school at least one time during primary school.[17]

4   The Model

I consider a model in which each family $i \in \{1, 2, ..., I\}$ decides among their possible elementary school alternatives in each of T (finite) discrete periods of time, where T is the end of the elementary cycle. The educational market is composed of J schools. The parents' decision is restricted in two ways. First, each parent i, has a specific choice set $\Lambda_i \subseteq \{1, 2, ..., J\}$. The cardinality of Λ_i is denoted by $S(\Lambda_i)$. Second, each school $j \in \Lambda_i$ may or may not admit the student i based on a rule that will be described below.

4.1   Parents' Utility

Let $D_{it} \in \Lambda_i$ be the school chosen by parent i at time t. The flow utility of parents i when their child is attending school j at time t is given by:

$$u_{ijt} = \beta_y \mathbf{Y}_{jt} + \beta_z \mathbf{Z}_{ij} + \beta_g G_{ijt} + C \, 1(D_{it-1} \neq j) + \beta_{eg}[G_{ijt-1} - \widehat{G}_{ijt-1}] \qquad (1)$$
$$1(D_{it-1} = j) + \epsilon^u_{ijt}$$

where $\mathbf{Y}_{jt}$ is a vector of characteristics of the school j, including socioeconomic composition (dummies for 5 tiers), tuition fee, and type of administration (public, voucher-private, and non voucher-private);[18] $\mathbf{Z}_{ij}$ is a vector of variables which are determined by the relationship between the school j and the individual i, including religion (a dummy variable which takes one if the school is religious and parents profess a religion, and zero

---

[15]See Larrañaga (2004).

[16]These figures do not include parents who change the municipality where they live, or students who repeat a grade. If one considers those cases, this fraction rises to around 11% (Zamora (2011)).

[17]These levels of student mobility are similar to what is observed in other countries. For instance, Hanushek et al. (2004) show that, in Texas' public schools, one-third of all children switch schools at least once between grades 4 and 7, excluding changes due to the transition from elementary to middle school.

[18]Given that some prestigious public schools begin at 7th Grade, I also include a dummy variable for those schools.

otherwise), years that student i has been attending school j, and location (a dummy variable which takes one if the school and the family house are located in different municipalities, and zero otherwise), G_{ijt} is the GPA obtained by the student, C is the direct cost of changing school, $1(A)$ is a function that takes 1 when A is true, $\widehat{G}_{ijt-1}$ is the expected GPA given the information at $t-2$, and ϵ^u_{ijt} is an iid shock. The term $[G_{ijt-1} - \widehat{G}_{ijt-1}]$ is included in the utility function to capture that students may prefer to stay in their current school when their performance in the previous grade was above their prediction.

This model requires a cost of switching schools, C, in order to fit the patterns of the data. Otherwise, the model would predict a higher probability of changing schools than actually exists. However, and beyond this practical consideration, it is reasonable to assume such a cost. Indeed, in the estimation, this parameter will capture an average of different costs that parents face when they change schools, namely, the student's adaptation cost,[19] and the monetary cost (many schools charge an enrollment fee).

In the final period, there is a utility that also captures all future payoffs, such that

$$u_{ijT} = \beta^T_{Ki} K_{ijT} + \beta_y \mathbf{Y}_{jT} + \beta_z \mathbf{Z}_{ij} + C\, 1(D_{iT-1} \neq j) + \beta_{eg}[G_{ijT-1} - \widehat{G}_{ijT-1}] \quad (2)$$
$$1(D_{iT-1} = j) + \beta^T_s SEC_j + \beta^T_{ag} \overline{G}_{ijT} + \beta^T_{ta} TA_{iT} + \epsilon^u_{ijT},$$

where K_{ijT} is the knowledge achieved by student i in school j at time T, TA_{iT} represents the time, in years, that student i has been attending the current school and SEC_j takes one when school j also offers secondary level grades (from 9th to 12th) and zero otherwise. Including the latter in the terminal utility captures the changing costs that parents are forced to incur in $T+1$ when their child attends a school that does not offer the secondary class level. Furthermore, $\overline{G}_{ijT}$ and TA_{iT} are included because, as will be noted, they determine the future chances of being admitted in the desired high school.

4.2 Student knowledge

I model student knowledge as a cumulative process. In particular, let q_{jt} be the quality of school j at time t and q_{ijt} the quality of the school attended by student i at time t, such that $q_{ijt} = \sum_{j=1}^{J} q_{jt} 1(D_{it} = j)$; thus, the learning process is given by:

$$K_{i0} = \alpha_0 X_i, \quad (3)$$
$$K_{ijt} = K_{ijt-1} + \alpha_1 q_{ijt}.$$

---

[19]The model, as is explained below, also has an explicit cost of switching schools, which is given by the fact that the GPA in school j is, among other things, a function of the time that the child has attended that particular school j. The economic literature has found evidence of this pedagogical cost. See, for example, Hanushek et al. (2004). Thus, C captures the adaptation costs, which come in addition to this cost in performance.

Therefore, the knowledge achieved by student i in school j at time t, K_{ijt}, is a function of student i's previous knowledge and the quality of the school she attends that year, q_{ijt}. In addition, the initial knowledge only depends on student i's characteristics X_i (i.e., parents' education). It should be noted that in the model, school quality q_{ijt} may change over time, this possibility is also considered in the estimation part.

4.3   GPA function

Grades in elementary school are determined by the following production function:

$$G_{ijt} = \lambda_{0j}^t + \lambda_{1j}^t K_{ijt} + \lambda_2^t TA_{it} + \varepsilon_{ijt}^g, \tag{4}$$

This specification captures the idea that each school may have a particular way to map knowledge onto grades. In particular, the higher the value of λ_{0j}^t, the more likely it is that students perform well in school j. Moreover, even conditioning on student knowledge, TA_{it} has an effect on grades. This accounts for the fact that it may take time for new students to learn the characteristics of the evaluation system of each school. As noted above, this implies an explicit cost of switching, given that moving to a new school will imply a cost in GPA performance.

Regarding parents' information, I assume that parents know $\{\lambda_{0j}^t, \lambda_{1j}^t\} \ \forall \ t, j$, which is necessary to forecast GPA performance of their children. Intuitively, and this means that parents know the level of difficulty of any school.

4.4   Probability of admittance

Parents are restricted in their choices to the extent that schools have the right of admittance. Let AD_{ijt} be a binary variable, which is unobservable for the econometrician, that equals one if student i can enter school j at time t and zero otherwise, such that:

$$AD_{ijt} = \begin{cases} 1 & \text{if } \varrho_{ijt} - \varepsilon_{ijt}^{ad} \geq 0 \\ 0 & \text{if } \varrho_{ijt} - \varepsilon_{ijt}^{ad} < 0 \end{cases} \tag{5}$$

where

$$\begin{aligned}
\varrho_{ijt} = {} & \varphi_0 + \varphi_q q_{jt} Sel_{jt} + \varphi_{0k} Sel_{jt}^k + \varphi_{1k} K_{it-1} Sel_{jt}^k + \varphi_{0g} Sel_{jt}^g + \varphi_{1g} \overline{G}_{it-1} Sel_{jt}^g \\
& + \varphi_{0mr} Sel_{jt}^m + \varphi_{1mr} MR_{ijt-1} Sel_{jt}^m + \varphi_{0r} Sel_{jt}^r + \varphi_{1r} REL_i Sel_{jt}^r + \varphi_s Sel_{jt}^o \\
& + \varphi_{0ns} New_{jt} + \varphi_{1ns} New_{jt} * Size_{jt} + \varphi_{fe} X_i * fee_{jt} + \varphi_{sx} X_i * Sel_{jt}.
\end{aligned} \tag{6}$$

and Sel_{jt}^k takes one when school j selects students based on academic tests and zero otherwise; Sel_{jt}^g takes one when school j selects students based on previous grades and zero otherwise; Sel_{jt}^r takes one when school j selects students based on students' religion and zero otherwise; takes one when school j selects students based on parents' marital

status and zero otherwise Sel_{jt}^m; Sel_{jt}^o takes one when school j selects students based on other reasons and zero otherwise; Sel_{jt} is an index to measure how selective is school j at time t.[20][21] These variables may all equal one at the same time. New_{jt} takes one when the school is new (or doesn't offer the previous grades), $Size_{jt}$ is the size of this new school, and fee_{jt} denotes the tuition fee. Finally, MR_{ijt} takes one if parents are married and zero otherwise, and REL_i takes one if parents are religious and zero otherwise.[22]

Then, assuming that ε_{ijt}^{ad} are iid, following a logistic distribution, the probability of admission is described by:

$$Pr(AD_{ijt} = 1) = \frac{\exp(\varrho_{ijt})}{1 + \exp(\varrho_{ijt})} \tag{7}$$

To have a tractable likelihood calculation, I assume that ε_{ijt}^{ad} is realized before parents make the D_{it} decision.[23] Moreover, to simplify the solution of the model, I assume that, for any student, there is always at least one school willing to admit her.[24] In particular,

$$h \in argmax_{j \in \Lambda_i}(\varrho_{ijt}) \Rightarrow AD_{iht} = 1. \tag{8}$$

I chose this specification for the random process of AD_{ijt} for two reasons: given the rich information that I have about the admission rules of each school, and considering the challenge in separately identifying the parameters of this process and the parameters of the utility function, when AD_{ijt} is latent. In this regard, it should be noted that there is no variable that enters in the same way in the admission probability function and in the utility function.[25] For instance, the admission probability is also a function of the interaction between the tuition and the parents' socioeconomic status. This is included to capture the fact that some tuition is unaffordable to some families. Note that tuition is also present in the utility function, but this variable only interacts with parents' socioeconomic status in the admissions probability. Thus, the idea of the first element is to capture the fact that parents do not like to pay high prices. The idea of the second is to capture the fact that some prices are impossible for some families to pay.

---

[20]In the empirical implementation, all these variables are proportions, instead of binary variables. This is because I construct these variables from parents' surveys, and in each school their answers are not always the same. Thus, for instance, in the empirical implementation, Sel_{jt}^k is the fraction of parents in school j who affirm that school j selects students based on an academic test.

[21]In the empirical implementation of this model, $Sel_{jt} = (Sel_{jt}^k + Sel_{jt}^g + Sel_{jt}^o) * \frac{1}{3}$.

[22]Contreras et al. (2010) present evidence indicating that student selection is a widespread practice among private subsidized schools.

[23]This means parents do not apply to schools. Instead, at the end of each period they know their feasible set for the next period and they pick the feasible school that maximizes their expected utility.

[24]Given that in the data the $argmax_{j \in \Lambda_i}(\varrho_{ijt})$ is almost certainly a public school (because most of those schools do not select students), this assumption is equivalent to assuming that there is one public school that admits the student i without uncertainty. Which is actually the way that it works in reality.

[25]The few variables that are in both functions are interacting with other variables in the function that determine the probability of admission.

Moreover, this specification takes advantage of the interaction between the features of the model and data availability. For instance, if the school selects students based on an academic test (with $\varphi_{1k} > 0$, as expected), then the higher K_{it-1}, the higher the probability of i being admitted at j.

Note that students who decide to stay in their current school do not have to complete an admissions process. If they want, they can stay in their current school with probability 1.

4.5 Parents' information and perception about quality

Parents have two sources of information about school quality. Firstly, they observe the results of the standardized tests for all the schools, which is public information. Secondly, they may observe the quality of the school which their child is attending, which is private information.

Regarding public information, it is assumed that standardized tests are measures of school quality, whose values also depend on the characteristics of the student. Thus, in this model, school quality (q) is defined as a school's contribution to learning, such that:

$$ST_{ijt}^\chi = q_{jt} + \theta_2 X_i + \varepsilon_{ijt}^\chi, \tag{9}$$

where ST_{ijt}^χ denotes the standardized test score in subject χ, such that

$$\chi \in \{Math, Spanish, Natural\ Science, Social\ Science\}.$$

In this context, I define $\tilde{q}_{jt}$ as the fixed effect estimation of the expected quality of school j at time t, given the public information $\mathbf{ST}_{jt}$.[26] Hence, $\tilde{q}_{jt}$ is the unbiased estimate of q_{jt} given public information at t.[27] To the extent that the demographic composition of the schools students' matter for test scores, these peer effect will also be included in q_{jt}. Thus, $\tilde{q}_{jt}$ is in practice the average test score in the school that is not explained by the individual characteristics of the students (teachers, infrastructure, school manager, etc.).

Although school quality may vary over time, I assume that parents have no information about the evolution of quality. Thus, for them the best estimate of future quality is their estimate of current quality.[28]

The model is flexible in terms of how parents access and understand the information about schools' quality. In the first place, there are different types of parents with regard to their ability to distinguish the school's contribution from the students' contribution to test scores. In the second place, there are different types of parents with regard to their involvement in the schooling process of their child, which determines whether

---

[26]The same approach to estimate school quality is followed by Neilson (2013).

[27]X is a matrix with the characteristics of all the students. In practice, to estimate $\tilde{q}_{jt}$, we only need the characteristics of parents who send their child to school j.

[28]When I estimate the model, I estimate school quality of year t using only information for that year. Thus, the estimation approach is totally agnostic to any kind of dynamics.

they observe the quality of that school. Namely, only involved parents have access to private information. The first is denoted *parent's cognitive skill*, whereas the second is denoted *parent's school involvement*. The school involvement type of parent i is given by $\psi_i \in \{0, 1\}$, where 1 means involved.[29]

To present how parents access and understand the information about the quality of schools, I divide the analysis into three cases: (1) their perception about the quality of the schools not attended by their child; (2) the involved parents' perception about the quality of the school attended by their child; and (3) the non-involved parent's perception about the quality of the school attended by their child. In all three cases, what matters is parents' perception at the end of $t-1$ (when they make the choice of school for period t), about school quality at time t, given their information at time $t-1$, i.e. $E_{t-1}[q_{jt}|D_{it-1}, \psi_i]$.[30]

Case 1: Schools not attended by their child.

$$E_{t-1}[q_{jt}|D_{it-1} \neq j] = \tilde{q}_{jt-2} + \sum_{\chi \in A} \eta_\chi^i (\overline{ST}_{jt-2}^\chi - q_{jt-2}), \tag{10}$$

where, $\eta_\chi^i = \eta_{1,\chi}^1 X_i + \eta_{2,\chi} S(\Lambda_i)$ and $A = \{m, s, n, sc\}$. Thus, η depends on the parents' education and the size of the choice set ($S(\Lambda_i)$), where the latter is motivated by the bounded rationality literature.[31] Indeed, the way in which the complexity of the information processing to make decisions increases the probability of making mistakes is well-established in that literature. Moreover, in the context of my model, it is rather natural to think that the larger the choice set, the higher the complexity of the information processing. In short, in this model, *bounded rationality* implies that parents differ in their ability to perceive school quality information correctly, and that their misperception is a function of their educational level and the size of their choice set.

In this case, if as expected $\eta_\chi^i \geq 0$, then parents will overestimate the quality for schools whose students have, on average, highly educated parents.[32] Moreover, given the fact that in the Chilean educational system standardized tests are published one year after taken, even if parents did not have a misperception about quality (i.e., $\eta = 0$), when they choose the school for time t (at the end of $t-1$), they would estimate school qualities using the public information at $t-2$.

In sum, the use of public information presents two potential drawbacks: it is published after a delay, and parents may have difficulties in interpreting it, namely, when they ob-

---

[29]This is an exogenous, time invariant, parents' characteristic and therefore it does not depend on the school's characteristic.

[30]I am assuming that parents do not use past grades as a potential source of information to infer school quality.

[31]A survey can be found in Conlisk (1996).

[32]Given the functional form of the standardized tests, $\overline{ST}_{jt-2}^\chi - \tilde{q}_{jt-2}$ is the part of the average test, of subject χ, that is not explained by school quality. Hence, this is the part explained by the socioeconomic composition of the school.

serve the tests, they can have problems in isolating school quality from the socioeconomic composition of its students.

Case 2: School attended by their child, when parents are involved in that school ($\psi_i = 1$).[33]

$$E_{t-1}[q_{jt}|D_{it-1} = j, \psi_i = 1] = q_{jt-1} \tag{11}$$

Thus, parents who are involved in their child's school observe the quality of that school without distortion and without lag.[34]

Case 3: School attended by their child, when parents are not involved in that school ($\psi_i = 0$).

$$E_{t-1}[q_{jt}|D_{it-1} = j, \psi_i = 0] = \tilde{q}_{jt-2} + \sum_{\chi \in A} \eta_\chi^i (\overline{ST}_{jt-2}^\chi - q_{jt-2}), \tag{12}$$

$A = \{m, s, n, sc\}$.

Thus, parents who are not involved in their child's school have, for that school, the same information that they have for all the other schools (public information).

In this context, parents' perception about their child's knowledge is given by the following expressions:

- $E_0[K_{i0}|\psi_i] = K_{i0}$.

- $E_t[K_{it}|D_{it}, \psi_i] = E_t[K_{it-1}|D_{it}, \psi_i] + \alpha_1 \sum_{j=1}^J E_t[q_{jt}|D_{it}, \psi_i]1(D_{it} = j)$.

- $E_{t-1}[K_{it}|D_{it-1}, \psi_i] = E_{t-1}[K_{it-1}|D_{it-1}, \psi_i] + \alpha_1 \sum_{j=1}^J E_{t-1}[q_{jt}|D_{it-1}, \psi_i]1(D_{it} = j)$.

I denote $\tilde{K}_{it}^a = E_a[K_{it}|D_{ia}, \psi_i], \quad a = \{t-1, t\}$.

4.6 Decision Timing and Solution of the Model

At the end of period $t-1$, the following random variables are realized: (1) Utility idiosyncratic shocks: $\epsilon_{ijt}^u \ \forall i, j$; (2) the right of admittance shocks: ε_{ijt}^{ad} (hence, AD_{ijt}) $\forall i, j$; and (3) test scores, published with lag: $\{\mathbf{ST}_{t-2}^m, \mathbf{ST}_{t-2}^l, \mathbf{ST}_{t-2}^n, \mathbf{ST}_{t-2}^{sc}\}$. Given this

---

[33]I allow for β_{ki} being different for this type of parent.

[34]This assumption is supported by the evidence presented in Azmat and Garcia-Montalvo (2012), who find that knowing about and/or visiting more schools is related to more accurately assessing local schools.

information, parents decide D_{it}, taking into consideration the expected flow utility at t and the expected future payoff associated with each school.[35]

The model is solved by backward recursion, where the dynamic decision is driven by the state variables (Ω_{it}).[36]

$$\Omega_{it} = \begin{cases} \{D_{it}, TA_{it}, \tilde{K}_{ijt}^t, \overline{G}_{ijt}, q_{ijt}, \mathbf{ST}_{t-1}, \varepsilon_{\mathbf{it}}^{\mathbf{ad}}, \epsilon_{\mathbf{it}}^{\mathbf{u}}\} & \text{if } \psi_i = 1 \\ \\ \{D_{it}, TA_{it}, \tilde{K}_{ijt}^t, \overline{G}_{ijt}, \mathbf{ST}_{t-1}, \varepsilon_{\mathbf{it}}^{\mathbf{ad}}, \epsilon_{\mathbf{it}}^{\mathbf{u}}\} & \text{if } \psi_i = 0 \end{cases} \tag{13}$$

I define Ω_{it}^- as the state variables which are observed by the econometrician, such that:[37]

$$\Omega_{it}^- = \begin{cases} \{D_{it}, TA_{it}, \tilde{K}_{ijt}^t, \overline{G}_{it}, q_{ijt}, \mathbf{ST}_{t-1}\} & \text{if } \psi_i = 1 \\ \\ \{D_{it}, TA_{it}, \tilde{K}_{ijt}^t, \overline{G}_{it}, \mathbf{ST}_{t-1}\} & \text{if } \psi_i = 0 \end{cases} \tag{14}$$

To consider the school's right of admittance, I redefine the flow utility as:

$$\bar{u}_{ijt} = \tilde{u}_{ijt}(\epsilon_{ijt}^{ad}) + \epsilon_{ijt}^u \tag{15}$$

where,

$$\tilde{u}_{ijt}(\epsilon_{ijt}^{ad}) = \begin{cases} \beta_x \mathbf{Y}_{jt} + \beta_z \mathbf{Z}_{ij} + \beta_g G_{ijt} + C\ 1(D_{it-1} \neq j) \\ +\beta_{eg}[G_{ijt-1} - \widehat{G}_{ijt-1}]1(D_{it-1} = j) & \text{if } AD_{ijt}(\Omega_{it-1}^-, \epsilon_{ijt}^{ad}) = 1 \\ \\ -\infty & \text{if } AD_{ijt}(\Omega_{it-1}^-, \epsilon_{ijt}^{ad}) = 0 \end{cases} \tag{16}$$

The solution to this dynamic problem is fully characterized by the integrated value function, $\overline{V}(\Omega_{it-1}^-)$, such that:[38]

$$\overline{V}(\Omega_{it-1}^-) = \int \max_{j \in \Lambda_i} \left\{ E_{t-1}\tilde{u}_{ijt}(\epsilon_{ijt}^{ad}) + \epsilon_{ijt}^u + \delta E_{t-1}[\overline{V}(\Omega_{it}^-)|\Omega_{it-1}^-, D_{it} = j] \right\} dG_\varepsilon(\varepsilon_{it}), \quad \varepsilon_{it} = [\boldsymbol{\epsilon}_{it}^u\ \boldsymbol{\epsilon}_{it}^{ad}]'. \tag{17}$$

Then, defining the auxiliary function $v(\Omega_{it-1}, D_{it} = h)$ as:

---

[35] ε_{ijt}^g is realized after the decision of D_{it} is made.

[36] $TA_{it} = 1 + 1(D_{it} = D_{it-1})TA_{it-1}$ and $\overline{G}_{ijt} = \frac{\overline{G}_{ijt-1}*(t-1)+G_{ijt}}{t}$.

[37] Observed conditional on types.

[38] $\boldsymbol{\epsilon}_{it}^u = \{\epsilon_{ijt}^u\}_{j \in \Lambda_i}$ and $\boldsymbol{\epsilon}_{it}^{ad} = \{\epsilon_{ijt}^{ad}\}_{j \in \Lambda_i}$.

14

$$v(\Omega_{it-1}, D_{it} = h) = E_{t-1}\tilde{u}_{iht}(\epsilon_{it}^{ad}) + \epsilon_{iht}^u + \delta E_{t-1}[\overline{V}(\Omega_{it}^-)|\Omega_{it-1}^-, D_{it} = h] \quad (18)$$

$$\Rightarrow D_{it} \in argmax_{j\in\Lambda_i}\{v(\Omega_{it-1}, D_{it} = j)\} \quad \forall t \in \{1, 2, ..., T-1\}$$

At the end of $T-1$:

$$\tilde{u}_{ijT}(\epsilon_{ijT}^{ad}) = \begin{cases} \beta_K K_{ijT} + \beta_x \mathbf{Y}_{jT} + \beta_z \mathbf{Z}_{ij} + \beta_{eg}[G_{ijT-1} - \overline{G}_{ijT-1}]1(D_{iT-1} = j) \\ +C\ 1(D_{iT-1} \neq j) + \beta_s^T SEC_j + \beta_{ag}^T \overline{G}_{ijT} + \beta_{ta}^T TA_{iT} & \text{if } AD_{ijt}(\Omega_{iT-1}^-, \epsilon_{ijT}^{ad}) = 1 \\ \\ -\infty & \text{if } AD_{ijt}(\Omega_{iT-1}^-, \epsilon_{ijT}^{ad}) = 0 \end{cases}$$
$$(19)$$

$$v(\Omega_{iT-1}, D_{iT} = j) = E_{T-1}\tilde{u}_{ijT}(\epsilon_{iT}^{ad}) + \epsilon_{ijT}^u, \quad (20)$$

$$\Rightarrow D_{iT} \in argmax_{j\in\Lambda_i}\{v(\Omega_{iT-1}, D_{iT} = j)\}.$$

Thus, using equation 20, it is possible to solve the maximization in the last period. Then, using Monte Carlo integration and interpolation method (Keane and Wolpin (1994)), it is possible to approximate $\overline{V}(\Omega_{iT}^-)$ for any value of vector Ω_{iT}^-, which allows me to solve the maximization in the period $T-1$ (see equation 18). Following this procedure for all periods, the model can be solved by backward recursion.

5   Data and Empirical Implementation

5.1   Data Description

The main source of information in this paper is the administrative panel data from 2004 to 2011 on all students in the country from the Ministry of Education of the government of Chile. This panel includes the school attended every year, the average grade, the municipality where the student lives and where the school is located, and some basic demographic information. As mentioned, the sample of students entering 1st grade in 2004 took the SIMCE test in 4th grade (2007) and 8th grade (2011), and I merge this panel with information from parent surveys that are carried out during the SIMCE process. These contain mother's and father's education, whether they care about school religion, and their marital status.

In order to characterize schools, I use SIMCE test scores and the information collected from SIMCE parents' surveys for the years 2002, 2005-2011.[39] Test scores are used to

---

[39]4th grade SIMCE for years 2002, 2005-2009 and 8th grade SIMCE for years 2004, 2007, 2009 and 2011.

estimate schools' quality for every year. The surveys include questions about school tuition fees, and whether the school considered some of the following elements in the admission process: a student test, previous GPA, parents' marital status (and whether they had a religious wedding), and a general category to account for any other information considered in the admission. Furthermore, from administrative data of the Ministry of Education, I collect information about each school's religious affiliation.

Finally, from the SIMCE of 2011, 8th grade for my cohort, I use the answers to two types of questions as determinants of parent involvement. First, I use the questions to parents:

1. *How often do you attend the periodic parents' meeting of your child's class?*

2. *Name the first three reasons why you chose your child's current school.*

Second, I use the questions to students, *How often does one of your parents do each of the following activities?*:

1. *She or he explains to me the class material that I don't understand.*

2. *She or he helps me to study.*

5.2 Empirical Implementation

Two inputs are needed to estimate the model, namely, the measures of school quality and parents' choice set. Moreover, to gain in speed, and given the detailed information that I have, I estimate the parameters of the knowledge production function and the parameters of the grade production function outside of the model.

5.2.1 Estimating Measures of Quality

As presented above, the observable test scores have the following functional form:

$$ST_{ijt}^{\chi} = q_{jt} + \theta_2 X_i + \varepsilon_{ijt}^{\chi}, \quad \chi \in \{m, s, n, sc\}. \tag{21}$$

I estimate the values of q_{jt} by fixed effect regressions, the prediction from that estimation procedure is denoted $\tilde{q}_{jt}$. To be clear, in this estimation I only use students who have attended the same school during the first four years of primary school, and X_i is a vector or parents education.[40]

Figure 8 (Appendix B.1), shows the distribution of estimated school quality by school type in 2004, which is consistent with Bellei (2009), in the sense that, when one does not control for peer effects, voucher-private schools have higher quality than public schools.

---

[40]Test scores are available in 4th grade at the elementary school level, and in 8th and 10th grade in alternating years. This precludes including student fixed effects to estimate the school quality in every year of my sample.

Because I want to understand parents' decisions and to what extent they base such decisions on school quality, it makes sense to consider peer effects as part of the definition of school quality.[41] Furthermore, to have an idea about how schools' qualities are estimated, in Figure 9 (Appendix B.1), I show the distribution of school quality by school type based only on test scores (i.e., the raw measure of quality).[42] The comparison between the estimated quality versus the raw measure of quality (Figures 8 and 9) shows, as expected, a reduction in the variance of school quality and in the quality gap between schools of different types of administration.

To estimate the parameters of the knowledge production function, I run the following OLS regression:[43]

$$\tilde{K}_{ijT} = \alpha_0 X_i + \alpha_1 \sum_{t=1}^{T} \tilde{q}_{ijt} + \vartheta_{ijT} \tag{22}$$

To the extent that ϑ_{ijT} may include unobserved variables correlated to $\tilde{q}_{ijt}$, it is not accurate to interpret α_1 as the causal effect. However, there are two considerations to include here. First, in this regression I control for parents' education, which should help to attenuate the potential problem of bias. Second, and much more important, my approach does not need α_1 to be the causal effect. Since the utility is a linear function of K and K is a linear function of q, what I am really estimating when I find β_k is the relevance of school quality (mediated by α_1 and β_k) in parents decision. Thus, if α_1 is estimated with bias, β_k is going to correct that, because what matters is the estimation –and identification– of $\beta_k \alpha_1$. To illustrate this point, imagine that α_1 and β_k are the correct values, then if the former is estimated with bias, where the estimated value is $\hat{\alpha_1}$, and $E[\hat{\alpha_1}] = \alpha_1 \nu$, then the estimated β_k, $\hat{\beta_k}$, will correct that, such that $E[\hat{\beta_k}] = \frac{\beta_k}{\nu}$.

5.2.2 Parents' Choice Set

As opposed to other educational systems, in Chile parents are allowed to choose any public (or voucher-private) school, regardless of the municipality where they live. Given that, it is a hard empirical problem to define the choice set Λ_{it}. To do so, I classify families in G groups, grouped by their home location (municipality) and their level of education, then:

$$G_g = \{i, \ s.t. \ (ed_i, loc_i) = (ed_g, loc_g)\},$$

$$g(i) \Leftrightarrow i \in G_g.$$

---

[41]However, this should be kept in mind in the analysis of the counterfactual experiments.

[42]By test scores I mean the simple average between standardized Math and standardized Spanish test scores.

[43]Where K_{ijT} is estimated by EM algorithm, assuming that K_{ijT} is a latent variable measured by the SIMCE tests at time T. Further, $\tilde{q}_{ijt} = \sum_{l=1}^{3} \hat{\pi}_l \hat{\mu}_{jt}^l$.

$$\Lambda_{it} = \{j, \ s.t. \ \exists \ i' \in G_{g(i)} \mid \sum_{t=0}^{T} 1(D_{i't} = j) > 0\} \tag{23}$$

This means that, by definition, for each pair of parents, the chosen school belongs to their choice set.

Having a large number of families belonging to each group implies that, if no family belonging to group G_g has chosen a particular school, it is because that school is not feasible for that group of families. Figure 7 (Appendix B.1) shows the distribution of the size of parents' choice set, which indicates that, if anything, this approach is overestimating that size.

5.2.3 Estimating the Parameters of the Grade Production Function

To estimate the parameters in the grade production function, λ_{0j}^t and $\lambda_{1j}^t \ \forall j, t$, I use the math test to replace K_{ijt} by $ST_{ijt}^m - \varepsilon_{ijt}^m$ in the grade production function, such that:

$$G_{ijt} = \lambda_{0t}^t + \lambda_{1j}^t ST_{ijt}^m + \lambda_2^t TA_{ijt} - \lambda_{1j}\varepsilon_{ijt}^m + \varepsilon_{ijt}^g \tag{24}$$

Then, I estimate the parameters of interest by Two Stage Least Squares, using ST_{ijt}^l, ST_{ijt}^n and ST_{ijt}^c as instruments of ST_{ijt}^m.

5.2.4 Estimating the Parameters of the Utility Function

The specification of the utility function includes (in vector $\mathbf{Y}_j$) school socioeconomic composition dummies, tuition fee, school type dummies (public, voucher-private, and non voucher-private);[44] (in vector $\mathbf{Z}_j$) a dummy variable that takes one if both the school and parents are religious and zero otherwise, and a dummy variable that takes one if parents live in the municipality where the school is located and zero otherwise.

I estimate the parameters of the utility function by simulated maximum likelihood, using the Monte Carlo integration and interpolation method (Keane and Wolpin (1994)).[45] Given that this –and any method that solves the dynamic problem in each parameter iteration– is time consuming, I select a sample in the following way: I sort the municipalities belonging to Santiago City in descending order, in terms of their total student population, and I use the students living in the municipalities ranked 1, 3, 5 (odd numbers) ..., 19.[46] As a result, the final sample for the estimation has $9,752$ families and

---

[44]Because the public system has two types of schools, one from 1st to 6th and the other one from 7th to 12th, in the case of public schools, I allow for a different dummy for each type.

[45]The discount parameter δ is not estimated, but it is assumed equal to 0.95.

[46]The considered municipalities are Estación Central, Huechuraba, La Granja, La Reina, Macul, Melipilla, Pedro Aguirre Cerda, Recoleta, San Miguel, and Ñuñoa. I used 10 of the 33 municipalities of Santiago city.

856 schools. I also drop the students who fail one or more years and those who change the municipalities where they live. The former is because I use the student information collected in the 2011 SIMCE (8th grade), information that is obviously missing for those who enter 1st grade in 2004 and fail at least one year between 2004 and 2010. The latter is because the dynamic problem is solved for each student type, where a type is defined by the student location, among other things. Therefore, if I considered people who change their location, I would have to solve the dynamic problem for all the combinations of locations observed in the data, which would dramatically increase the estimation time. Table 8, Appendix A shows how much students who fail at least one grade level differ from students who never fail a grade. In short, students who fail at least one grade switch schools more often, and when they change schools, they do so for different reasons. For example, they have a higher probability of switching schools due to low performance (i.e. they are forced by the school).

Given the solution to the dynamic problem, which is fully characterized by the integrated value function $\overline{V}(\Omega_{it-1}^-)$, and assuming that ϵ_{ijt}^u is iid, following a standard type-1 extreme value distribution, then:

$$P(D_{it} = h | \boldsymbol{\epsilon}_{it}^{ad}, \psi_i) = \tag{25}$$
$$\frac{\exp(E_{t-1}[\tilde{u}_{iht}(\epsilon_{iht}^{ad}) + \delta\bar{V}(\Omega_{it}^-)|\Omega_{it-1}^-, D_{it} = h, \psi_i]) AD_{iht}(\Omega_{it-1}^-, \epsilon_{iht}^{ad})}{\sum_{j \in \Lambda_i} \exp(E_{t-1}[\tilde{u}_{ijt}(\epsilon_{ijt}^{ad}) + \delta\bar{V}(\Omega_{it}^-)|\Omega_{it-1}^-, D_{it} = j, \psi_i]) AD_{ijt}(\Omega_{it-1}^-, \epsilon_{ijt}^{ad})}.$$

Therefore, the probability of a sequence of schools chosen by parents i, $\mathbf{D}_i$, is given by:[47]

$$P(\mathbf{D}_i) = \prod_{t=1}^{T} P(D_{it}) = \sum_{n \in \{0,1\}} \pi_n^i \int \sum_{t=1}^{T} P(D_{it} | \boldsymbol{\epsilon}_{it}^{ad}, \psi_i = n) dG_\varepsilon \boldsymbol{\epsilon}_{it}^{ad}. \tag{26}$$

where the log-likelihood function L, is given by:[48]

$$L = \sum_{i=1}^{I} log(P(\mathbf{D}_i)). \tag{27}$$

Given that AD_{ijt} is a latent variable, I approximate $P(D_{it})$ by:

$$P(D_{it} = h) \approx \tag{28}$$
$$\sum_{n \in \{0,1\}} \pi_n^i \frac{1}{N_s} \sum_{\kappa=1}^{N_s} \frac{\exp(E_{t-1}[\tilde{u}_{iht}(\epsilon_{iht}^{ad}) + \delta\bar{V}(\Omega_{it}^-)|\Omega_{it-1}^-, D_{it} = h, \psi_i = n]) AD_{iht}^\kappa}{\sum_{j \in \Lambda_i} \exp(E_{t-1}[\tilde{u}_{ijt}(\epsilon_{ijt}^{ad}) + \delta\bar{V}(\Omega_{it}^-)|\Omega_{it-1}^-, D_{it} = j, \psi_i = n]) AD_{ijt}^\kappa}.$$

---

[47]$\pi_n^i = P(\psi_i = n | X_i)$.

[48]Because each likelihood calculation takes around 5 minutes, I use HOPSPACK (Hybrid Optimization Parallel Search PACKage) to optimize the likelihood function. This program is a derivative-free optimization solver.

where N_s is the number of simulations and the values of AD_{ijt}^{κ} are drawn from G_{ε}.[49]

5.2.5 Identification

The identification of this model faces two challenges not commonly present in any standard discrete choice dynamic programming models of individual behavior.[50] First, there is a challenge in separately identifying the parameters of the utility function and the parameters of the admission probabilities, without observing parents' applications. Second, there is a challenge in identifying the parameters that determine parents' perception about quality (i.e., η).

The former challenge is overcome through exclusion restrictions, which are naturally developed given the available data and the features of the model. Specifically, there is no variable (nor interaction of variables) that is simultaneously present in the utility function and in the admission probability function. For instance, parents care about school quality through its effect on student knowledge, a variable that also enters in the admission probability function. However, the admission probability is affected by student knowledge only for schools that select students based on academic tests. Furthermore, given the limitation that the model imposes on the heterogeneity of parents' preferences for quality, the fact that parents do not choose some schools that would give them higher utilities is rationalized by the model as if those schools did not admit such a student.[51]

The intuition behind the solution for the latter challenge is the following: if those parents who choose schools, not for the estimated quality, but for their average test scores (which is also determined by the socioeconomic composition of the school) have a higher probability of belonging to a particular education group or live in higher proportions in municipalities with a particular pattern in terms of the choice set size, then one can use those correlations to identify the parameters that determine η. Technically speaking, given the fact that the socioeconomic composition of schools enters directly in the utility function, the parameters of η_i are identified given the variation than comes from the interaction – in the utility function – of the socioeconomic composition of the school and the socioeconomic composition of parents i.[52]

6 Results

In the Appendix B.1, I show the estimated parameters of the knowledge production function and the production function of grades. In short, almost all the signs are as expected and the magnitudes are, in around two third of the cases, statistically significant. An interesting result, presented in Table 13, is that, even controlling for student

---

[49]In the estimation, I consider 50 simulations for each individual-time data point.

[50]For a survey, see Aguirregabiria and Mira (2010), Eckstein and Wolpin (1989) or Rust (1994).

[51]This is what identified the constant of the admission probability function. Another possible approach could be the one developed by Geyer and Sieg (2013).

[52]The difference between test scores and quality is, on average, equal to the contribution of the school's SES to test scores.

knowledge, the number of years a student stayed in a particular school positively impacts grades which, as was described above, constitutes a heterogeneous switching cost. Moreover, Figure 10 shows how schools have different standards by which they evaluate their students.[53]

The results of the estimation by simulated maximum likelihood are shown in Tables 1-4. These tables contain the parameters of the utility function, the admission probability function, parents involvement probability function, and the parameters that determine parents' quality perception. As can be seen, most of them have the expected sign and more than a half are statistically significant.

In terms of the utility function, everything is as expected, namely, besides having an important switching cost,[54] parents prefer schools with high quality, whose students come from high income families, with low tuition fees, private (without subsidy), with religious orientation (for those parents who state that they profess a religion), close to home, and with high GPAs (low difficulty conditional on quality). In terms of the admission process, in both the first and the other grades' admission probability functions, all of the signs go in the expected direction and many are statistically significant. For instance, the estimated parameters show that it is more difficult to be admitted to a school which considers admission tests, but this negative effect on admission probabilities is attenuated for students with high level of knowledge. Finally, the results related to parents' quality perception and parents' involvement are less robust. In regard to the former, the only variable that is statistically significant is the size of the choice set. However, and to the extent that the size of the choice set is a variable that is clearly related to the complexity of the decision, as opposed to the other considered variables, this is not necessary a bad result. For the latter, three of the four parameters have the expected sign. For example, parents who attend school meetings more often have a higher probability of being a more involved type. However, all of these parameters have high standard errors, to the extent that they are not statistically significant.

Table 1: Parameters of the Utility Function

	Coeff	Std Err.		Coeff	Std Err.
Student Knowledge in T (β_{kT})	0.0271	0.0035	$\Delta\beta_{kT}$ of Involved Parents	0.0024	0.0090
Medium Low SES School	0.0376	0.0232	Medium SES School	0.0959	0.0243
Medium High SES School	0.1185	0.0267	High SES School	0.1305	0.0356
School Tuition Fee	-0.0140	0.0071	Voucher-Private School	-0.0006	0.0054
Non Voucher-Private School	0.0608	0.0219	Religious School $*$ Religious Parents	0.2512	0.0085
Student GPA	0.0311	0.0497	GPA Correctiion $G_{ijt-1} - \widehat{G}_{ijt-1}$	0.0485	0.0081
Home and School in Different Municipalities	-0.3396	0.0057	School Offers Secondary Education	0.0403	0.0353
GPA at time T	0.0600	0.4521	Years Enrroled in Current School (at T)	0.0794	0.0090
Switch Cost	-5.9648	0.0281	Public School (2nd cycle)	1.4770	0.0487

[53]In this context, an easy school is one where the constant is big and the slope is small, hence all the students have good grades and their achieved knowledge has an irrelevant impact on their performance.

[54]Below I present some counterfactual exercises that show the relevance of the switching cost.

21

Table 2: Parameters of the Adsmission Probability Function

	Coeff	Std Err.		Coeff	Std Err.
Constant (1st Grade)	-0.3141	0.0137	School Quality $* Sel^k$ (1st Gr.)	-0.1000	0.0288
School Quality $* Sel^g$ (1st Gr.)	-0.2110	0.0361	School Quality $* Sel^o$ (1st Gr.)	-0.0964	0.0358
Sel^k (1st Gr.)	-0.0614	0.0099	$K * Sel^k$ (1st Gr.)	0.0033	0.0001
Sel^m (1st Gr.)	-0.0050	0.0076	$MR * Sel^m$ (1st Gr.)	0.0006	0.0188
Sel^r (1st Gr.)	-0.0280	0.0127	Religious$*Sel^r$ (1st Gr.)	0.0009	0.0191
Sel^o (1st Gr.)	-0.0221	0.0095	Tuition Fee $*$ Low Education (1st Gr.)	-1.9681	0.0905
Tuition Fee $*$ Medium Education (1st Gr.)	-1.0137	0.0352	Tuition Fee $*$ High Education (1st Gr.)	-0.1746	0.0142
School selectivity(+) $*$ Low Ed. (1st Gr.)	-2.2056	0.1270	School selectivity $*$ Medium Ed. (1st Gr.)	-0.9976	0.0701
School selectivity $*$ High Ed. (1st Gr.)	0.0017	0.0092	Constant (2nd-8th Grade)	-0.7468	0.0188
School Quality $* Sel^k$ (2nd-8th)	-0.1000	0.0271	School Quality $* Sel^g$ (2nd-8th)	-0.1886	0.0305
School Quality $* Sel^o$ (2nd-8th)	-0.0999	0.0438	Sel^k (2nd-8th)	-0.0525	0.0370
$K * Sel^k$ (2nd-8th)	0.0002	0.0001	Sel^g (2nd-8th)	-0.0200	0.0079
$GPA * Sel^g$ (2nd-8th)	0.0010	0.0008	Sel^m (2nd-8th)	-0.0103	0.0373
$MR * Sel^m$ (2nd-8th)	0.0020	0.0090	Sel^r (2nd-8th)	-0.0200	0.0181
Religious$*Sel^r$ (2nd-8th)	0.0019	0.0101	Sel^o (2nd-8th)	-0.0100	0.0091
New School (2nd-8th)	1.0097	0.3442	New School $*$ School Size (2nd-8th)	0.0025	0.0023
Tuition Fee $*$ Low Education (2nd-8th)	-0.7028	0.0761	Tuition Fee $*$ Medium Education (2nd-8th)	-0.5036	0.0287
Tuition Fee $*$ High Education (2nd-8th)	-0.1872	0.0112	School selectivity(++) $*$ Low Ed. (2nd-8th)	0.4190	0.0785
School selectivity $*$ Medium Ed. (2nd-8th)	0.1222	0.0063	School selectivity $*$ High Ed. (2nd-8th)	-0.0058	0.0060

(+) School Selectivity at first grade is defined as $(Sel^k + Sel^o)/2$.
(++) School Selectivity between 2nd and 8th grade is defined as $(Sel^k + Sel^o + Sel^g)/3$.

Table 3: Parameters that Determine Parents' Quality Perception

	Coeff	Std Err.		Coeff	Std Err.
Low Education (Math SIMCE)	0.0000	0.4322	Medium Education (Math SIMCE)	0.0046	0.3436
High Education (Math SIMCE)	0.0023	0.3456	Size of the Choice Set (Math SIMCE)	0.0104	0.0020
Low Education (Spanish SIMCE)	0.0000	0.3607	Medium Education (Spanish SIMCE)	0.0000	0.2579
High Education (Spanish SIMCE)	0.0003	0.0479	Size of the Choice Set (Spanish SIMCE)	0.0000	0.0003

Table 4: Parameters of the Parents Involment Probability Function

	Coeff	Std Err.		Coeff	Std Err.
Constant	-0.7476	0.1660	Parents Attend School Meetings	0.6817	0.5469
Parents Explain Class Material	0.0151	0.2140	Parents Help to Study	-0.0470	0.2201
Parents Care about Quality	0.2191	0.1980			

Besides the analysis of signs and significance, given the non-linear relationship between the parameters and parents' decision, the best way to assess the relevance of parameter magnitudes is through model fit analysis and counterfactual experiments.

6.1 Model Fit

I present the fit of the model under two scenarios. In the first case, I consider the sample used for the estimation. In the second case, whose figures are presented in Appendix

C.1, I use the complete sample (all the students in Santiago City).[55] I discuss the fit of the model under these two scenarios, since these are also the two samples that I consider in the counterfactual experiments, and because showing that the fit is similar in the two cases reinforces the point that the model is capturing the main mechanisms that determine parent decision, without overfitting the data.

As Figure 1 shows, the simulation of the model overall fits the pattern of the students who switch school by grade. However, the model has difficulty in generating the increase in school switching that occurs at the end of 6th grade. This increase is mainly driven by the entry of new public schools in 7th grade (for the new cycle), something that the model can only partially generate. Moreover, as Figure 2 shows, the model does a good job of predicting the 8-year total school changes, by parents' education.[56]

Figure 1: Fraction of students changing their school by grade

---
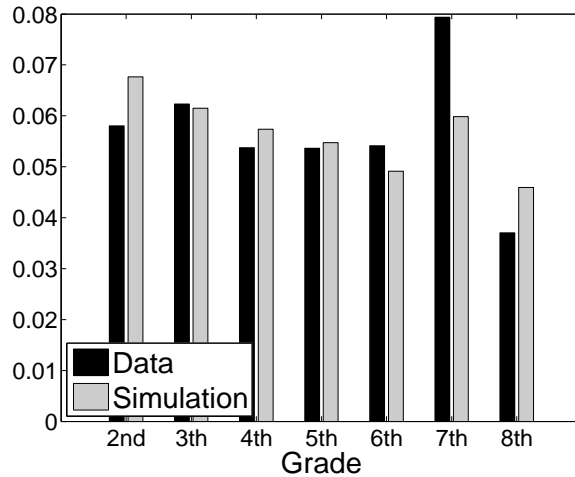
[55]In Appendix A, Table 11 shows descriptive statistics for these two samples.

[56]In the estimation and in the simulation, I collapse the information of parents' education into three categories: (1) both parents did not complete secondary education (low education); (2) One of the parents completed secondary education, but both parents did not attend higher education (medium education); and (3) at least one of the parents attended higher education (high education).

23

Figure 2: Average total change by parents' education
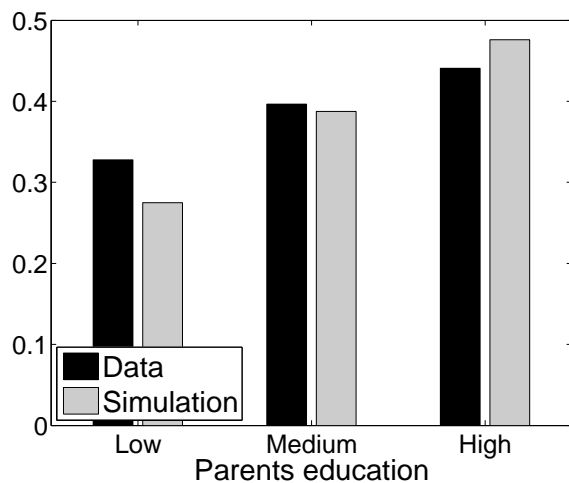
Given the kind of counterfactual experiments that I perform, it is relevant to assess how the model fits the data with regard to some patterns of the decision of parents. For instance, Figure 3 shows how the model fits parents' choice in terms of school type, namely, their decision about attending public, voucher-private or non voucher-private schools.

Figure 3: Student fraction by school type
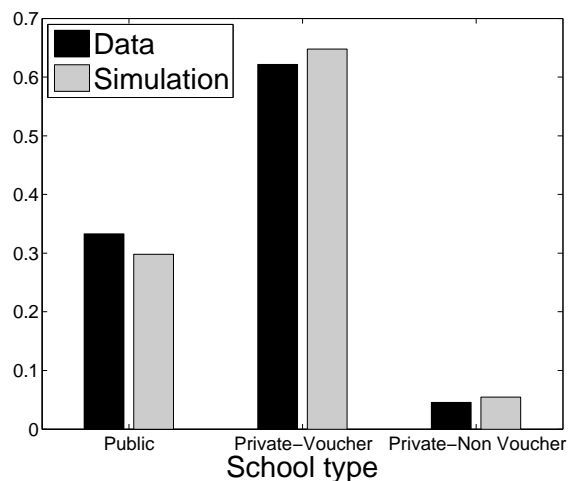
One common feature of many educational systems, which is extremely problematic in

the Chilean case,[57] is the fact that students' access to different schools in terms of quality depends on their income. In the context of the model, this means that the initial knowledge gap $K_{0i} - K_{0i'}$, is increased by $K_{Ti} - K_{Ti'} - (K_{0i} - K_{0i'}) = (K_{Ti} - K_{0i}) - (K_{Ti'} - K_{0i'})$.

The model has several channels that can generate this correlation: parents can have differences in preferences about quality, differences in cognitive skills to understand information about schools, differences in involvement in the child's school, and differences in their choice restrictions. Figure 4 shows how, in the data and in the model, the knowledge gain is positively correlated with parents' education. This figure also says that the model overpredicts the gain for students with parents of low or medium education, and underpredicts this gain for students with highly educated parents.

Figure 4: Gain in Knowledge by parents education ($K_T - K_0$)
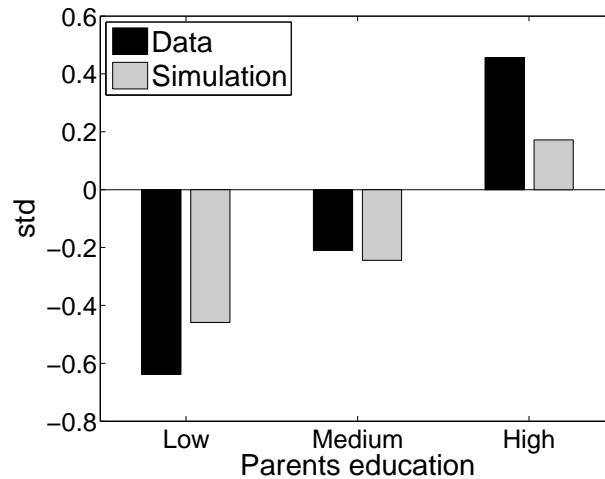
Overall, the fit of the model when the sample considers all the students in Santiago City is similar to the model fit when the estimation sample is used (Appendix C.1). The most important difference is that the former underestimates the frequency of school switches.

6.2   Parents Perception about Quality

Given the estimated parameters, it is possible to calculate the differences between parents' perception about school quality (which is determined by $\hat{\eta}_i$) and the effective quality of each school ($\tilde{q}_{jt}$, i.e., the schools' fixed effects of test score regressions). Moreover, it is interesting to see how the distance between perception and reality affects the three school types differently.

---

[57]Valenzuela et al. (2014).

To this end, I take two prototypical parents, both of the same educational level, but one with a small choice set (20 schools), the other with a big choice set (100 schools),[58] and then calculate which would be their quality perception for each school of the sample. To conclude, I calculate the distance between perception and reality for each school.[59] Figure 5 shows the results of this exercise. In the first place, there is an important distance between perception and reality. In the second place, this misperception is less severe for parents facing an smaller choice set. Finally, this misperception biases parents' preferences toward private schools. This bias is driven by the fact that voucher-private schools have more educated parents than public schools, whereas the same is true between non voucher-private and voucher-private schools. As discussed in the model section, because $\eta_\chi^i \geq 0$, parents overestimate the quality for schools whose students have, on average, highly educated parents.

Figure 5: Quality misperception by school administration type (2004)

(a) Parents with small choice set ($S(\Lambda) = 20$)   (b) Parents with big choice set ($S(\Lambda = 100)$)
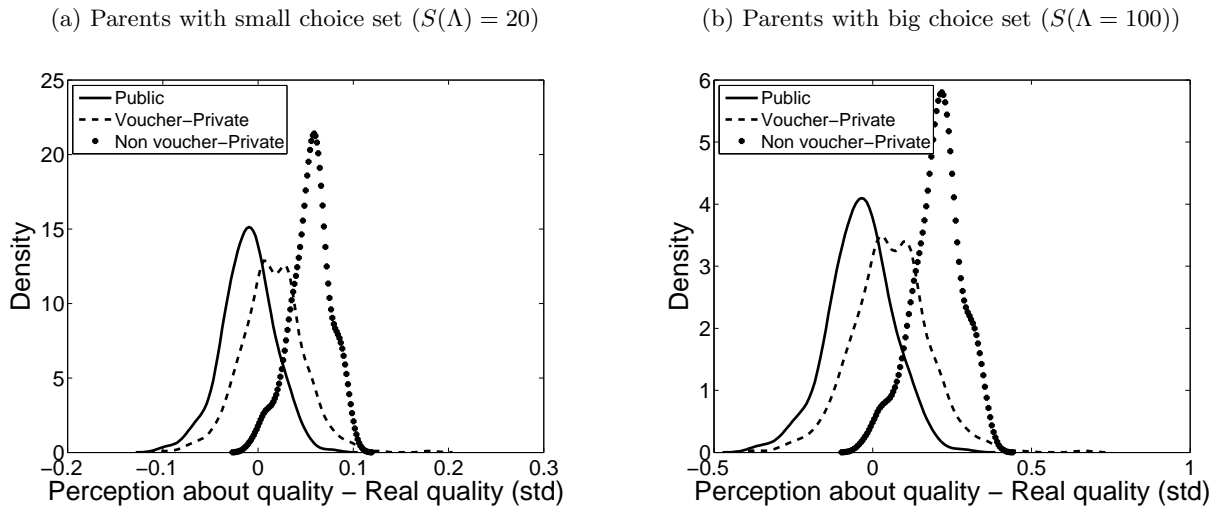
6.3 Counterfactual Experiments

To assess how important school quality is in parents' decisions, I simulate the model, randomly picking half of the schools and increasing their quality by 0.5 std, while decreasing the quality of the rest by the same amount. Then, I calculate the increase in the fraction of parents sending their children to the former schools. To see how relevant quality is in the first decision (first grade), vis-a-vis later decisions, I do this exercise by increasing schools' quality in different periods. For instance, I do not affect school quality until t, and I perform these quality changes from $t+1$ to T.

Figure 6 shows the results of these exercises. On one hand, there is a moderate increase, of 8-10 percentage points, in the demand for schools that increase their quality since the

---

[58]As can be seen from Figure 7, a choice set of 100 schools is not an outlier.

[59]In practice, what I calculate is $E_{t-1}[q_{jt}|D_{it-1} \neq j] - \tilde{q}_{jt}$.

first period. On the other hand, the effect is irrelevant when schools change their quality after the first decision is made (1st grade). These results highlight the relevance of the estimated switching cost.

Figure 6: Increase in the fraction of students in schools with higher quality
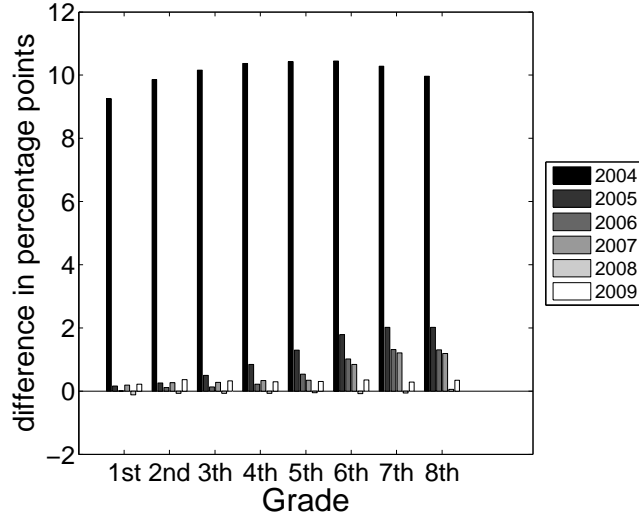
To study the mechanisms in parents' demand that explain the frequency of switching schools, the allocation of students across school types, and the correlation between the gain in knowledge and parents' educational level, I simulate the model under the following scenarios:[60]

- **Scenario (1):** There is no misperception ($\eta = 0$), thus parents correctly estimate school quality from standardized tests.

- **Scenario (2):** Parents only cares about quality, which means that $U = \beta * K + C * (D_{it} \neq D_{it-1}) + \epsilon$.

- **Scenario (3):** All students are admitted, thus $P(AD_{itj} = 1|X_i) = 1 \; \forall i, j, t$.

- **Scenario (4):** There is a random admissions process, i.e., $AD_{ij} \perp\!\!\!\perp X_i$.

- **Scenario (5):** the cost of changing school is reduced by 10% ($C * 0.9$).

---

[60]It should be noticed that many of these policies may affect the choice set definition. Thus, given that a choice set is fixed in all these simulations, the effects of these policies are underestimated in this analysis.

- **Scenario (6):** Parents with the lowest education are relocated to the municipality with highest average quality. Parents with the highest education are relocated to the municipality with lowest average quality.

- **Scenario (7):** All the students have the same knowledge endowment (K_0). In particular, $ED_i = 3 \; \forall i$.

Table 5 shows the fractions of parents who switch schools by grade in the baseline simulation (first column), and the differences in percentage points – compared to the baseline – under each of the counterfactual experiments. I include the confidence intervals for these estimations at a 90% confidence level.[61] From this table, it follows that, if parents were just concerned about quality, they would switch more often, which is explained by the fact that the other schools' characteristics are more stable across the years (SES, price in std, and type). Admissions restrictions play a relevant role in attenuating the frequency of switches. Finally, and more obviously, this frequency is also attenuated by the switch cost.

Table 5: Fraction of students changing school by grade (with respect to baseline in percentage points)

| | Baseline | Counterfactual scenarios | | | | | | |
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
2nd	6.85%	-0.20	0.37	11.18	-0.02	4.75	-1.81	-0.38
	(6.1 7.3)	(-0.58 0.58)	(-0.04 1.21)	(10.62 12.24)	(-0.59 0.69)	(4.18 5.54)	(-2.35 -1.23)	(-0.80 0.40)
3rd	6.18%	-0.00	1.10	11.25	0.19	4.64	-1.64	-0.18
	(5.6 6.8)	(-0.68 0.59)	(0.46 1.75)	(10.27 12.08)	(-0.62 0.61)	(3.96 5.40)	(-2.20 -1.03)	(-0.76 0.43)
4th	5.80%	-0.08	1.41	10.96	0.04	4.38	-1.46	-0.20
	(5.1 6.3)	(-0.57 0.55)	(0.96 2.16)	(10.07 11.97)	(-0.51 0.66)	(3.72 5.05)	(-2.01 -0.98)	(-0.76 0.41)
5th	5.33%	-0.04	2.03	10.89	0.03	4.20	-1.36	-0.17
	(4.7 5.9)	(-0.59 0.55)	(1.40 2.58)	(10.14 11.71)	(-0.42 0.75)	(3.64 4.96)	(-1.82 -0.84)	(-0.69 0.47)
6th	5.00%	-0.01	2.20	10.46	0.05	3.90	-1.28	-0.20
	(4.5 5.4)	(-0.61 0.52)	(1.70 3.00)	(9.58 11.30)	(-0.40 0.63)	(3.46 4.74)	(-1.69 -0.80)	(-0.62 0.39)
7th	6.03%	0.06	1.71	9.91	0.07	4.85	-1.86	-0.12
	(5.4 6.8)	(-0.66 0.47)	(0.79 2.34)	(8.93 10.65)	(-0.74 0.46)	(3.93 5.45)	(-2.46 -1.39)	(-0.81 0.34)
8th	4.47%	-0.00	2.92	9.66	-0.09	3.65	-1.20	-0.13
	(4.0 5.2)	(-0.55 0.46)	(2.13 2.34)	(8.56 10.37)	(-0.62 0.34)	(3.11 4.32)	(-1.67 -0.82)	(-0.76 0.23)

The confident intervals are in parenthesis.

Table 6 shows the fraction of parents by school type in the baseline simulation (first column), and the differences in percentage points – compared to the baseline – under each of the counterfactual experiments. As above, I include the confidence intervals

---

[61]The confidence intervals, for this and the following tables, are calculated using a bootstrap procedure, with the following steps in each simulation: (1) Draw the set of parameters from a normal distribution with mean and standard deviation equal to the estimated parameters and standard errors; (2) Simulate the model under the different counterfactual scenarios. With these values, I find the the respective percentiles.

for these estimations at a 90% confidence level. Even though parents' perception is importantly biased in favor of private schools (with and without vouchers), when they decide based on the real quality, the fraction of parents attending public schools increases by a moderate 1.2 percentage points (with the confidence interval including 0). This small effect, relative to the size of the misperception, is explained by the fact that parents do not care too much about quality. In fact, if parents were only concerned with school quality, there would be an increase of 2.6 percentage points in the fraction of parents choosing public schools, while this figure would decrease by 3.4 percentage points for voucher-private schools. This basically reflects the fact that the other elements of the utility function (SES of the school, the preference for its type, etc.) lead parents to apply to private schools.[62] Finally, these simulations allow us to see what would happen if less educated parents had a more relaxed choice set constraint. Scenarios 3, 4, and 7, all tell the same story: less choice set restrictions would lead (less educated) parents to choose private schools more often, though not necessary because of their higher quality.

Table 6: Student fraction by school type at first grade (with respect to baseline in percentage points)

| | Baseline | Counterfactual scenarios | | | | | | |
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Public	29.26%	1.20	2.60	-2.87	-2.99	0.71	0.45	-3.39
	(26.6 31.7)	(-0.37 2.00)	(-0.02 4.98)	(-4.11 -1.84)	(-4.30 -2.13)	(-0.46 1.91)	(-1.27 1.05)	(-5.02 -2.37)
Voucher Private	65.35%	-0.76	-3.44	-0.29	-0.26	-0.76	2.53	1.96
	(62.2 68.1)	(-1.61 0.85)	(-6.06 -0.47)	(-1.38 1.05)	(-1.15 1.28)	(-2.00 0.63)	(1.68 4.35)	(0.67 3.30)
Non voucher Private	5.38%	-0.44	0.84	3.16	3.25	0.05	-2.98	1.44
	(4.0 7.0)	(-1.14 0.12)	(-1.07 2.15)	(2.45 3.75)	(2.12 4.08)	(-0.73 0.42)	(-3.96 -2.13)	(0.91 2.32)

Table 7 shows the knowledge that students gained between 2004 and 2011, by parents' education. In addition, in the last row I present the change in the overall quality. Again, I include the confidence intervals for these estimations at a 90% confidence level. While the numbers of the baseline simulation are presented in the first column, the numbers in the other columns are the differences in standard deviations – compared to the baseline – under each of the counterfactual experiments. The first result to notice is that, while an exclusive focus on quality would increase the knowledge gained by students whose parents have medium or high education, this shift in preferences would not have a relevant effect for students whose parents have a low level of education. This confirms the relevance of choice restrictions: for some parents, even if they put more weight on quality, they cannot find a better school for their child. A second element to notice is that both prohibiting schools from making admission decisions based on student characteristics and

---

[62]It should be noticed that, in this model, peer effects are part of the school quality, which is constant in all the policy experiments. Therefore, in this model it is not possible to study a potential self-fulfilling prophecy, in which parents think that private schools are better, and therefore apply to those schools; those schools select the best students (those who have more educated parents); and, because of that pattern of admissions decisions, and given the peer effect, private schools end up being better than the public ones.

reallocating the poor families to better municipalities are effective measures to reduce the gap between students with parents with different levels of education.[63]

Table 7: Gain of knowledge ($KT - K0$) by parents education (with respect to baseline in standard deviations).

| | Baseline | Counterfactual scenarios | | | | | | |
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Incompleted High School	-52.79% (-61.1 -43.1)	0.00 (-0.07 0.04)	0.03 (-0.07 0.11)	0.10 (0.04 0.16)	0.08 (0.02 0.13)	0.00 (-0.05 0.04)	-0.09 (-0.14 -0.03)	0.10 (0.05 0.17)
Completed High School	-27.32% (-35.2 -18.5)	-0.02 (-0.07 0.02)	0.08 (0.00 0.16)	0.14 (0.09 0.18)	0.11 (0.07 0.16)	0.00 (-0.03 0.03)	-0.00 (-0.03 0.03)	0.12 (0.07 0.17)
With college Studies	18.76% (7.4 31.2)	-0.06 (-0.11 -0.00)	0.17 (0.06 0.27)	0.09 (0.04 0.13)	0.05 (0.00 0.09)	0.00 (-0.05 0.03)	0.08 (0.00 0.16)	0.00 (-0.04 0.04)
All	-15.90% (-23.1 -7.0)	-0.03 (-0.07 -0.01)	0.10 (0.03 0.16)	0.12 (0.08 0.14)	0.08 (0.05 0.11)	0.00 (-0.03 0.02)	0.01 (-0.02 0.04)	0.07 (0.04 0.10)

Finally, regarding the gain in the overall quality, as expected, an exclusive focus on quality increases the gain in quality. If the parents have a more relaxed choice set constraint, there is an increase in such a gain. Surprisingly, if parents have no quality misperception, there is a small decrease in the overall gain in quality. This is due to the fact that when parents exaggerate the quality of some schools due to misperception, implicitly they also care more about school quality relative to other characteristics of the schools, which leads parents to choose better schools in terms of quality. In other words, parents' misperception generates two distortions. On the one hand, parents fail to properly assess the real ranking of schools in terms of quality. This effect decreases the overall gain in quality. On the other hand, there is a change in school' quality distribution, which is equivalent to a change in the weights that parents place on school quality in their utility function. The latter effect could decrease or increase the overall gain in quality. However, in this case, it increases it.

Appendix C.2 contains the tables that show the results of the counterfactual experiments when using the complete sample, which includes all the students of Santiago City who entered first grade in 2003. Although the complete sample incorporates all the small municipalities that were not part of the estimation sample, the main conclusions (elaborated from Tables 5, 6, and 7) are not affected.

---

[63]This can be concluded by looking at columns 5, 8, and 7 of Table 7. Notice that in all these counterfactual experiments, the choice set (Λ) is fixed, in the sense that the set is invariant conditional on the municipality where parents live and their educational level. Thus, when a family of parents with low education is relocated from municipality A to municipality B, their new choice set (Λ) is going be the choice set of a family with low-educated parents who live in municipality B.

7   Conclusions

This paper estimates a dynamic model of elementary school choice. To this end, I use detailed Chilean administrative data for the students who entered 1st grade of the elementary cycle in 2003, following them until 8th grade (2011), which in Chile is the end of the elementary cycle. The estimated model considers several elements that are relevant to explain parents' decisions, namely, how much do they care about school quality (and other school characteristics), parents' skill in understanding information about quality (national standardized tests), parents' involvement in the school attended by their children, and their choice set.

Assessing the relevance of these different components contributes to a better understanding of the demand for schools and the role that markets with competitive incentives can have in education. In particular, the structural approach followed in this paper allows me to quantify different sources of unequal access to high quality schools and of the higher demand for private schools than for public schools. In doing so, this paper also contributes to the scarce literature that estimates structural models with bounded rationality, as well as to the literature which uses observed choices to infer agents' information.

Regarding the debate about the extent to which parents base their decisions on school quality, I find that parents do care about school quality, but only to a moderate degree. Moreover, the simulations show that parents' decisions are not sensitive to changes in quality after the first decision (1st grade). I also find that more involved parents care marginally more about school quality.

The results show that parents have an important misperception about school quality, which causes them to have a less favorable opinion about public schools, relative to private schools. This result supports the idea that parents may have difficulty in isolating a school's quality from its socioeconomic composition when they observe test scores. However, given that quality is not very relevant for their decision, such a misperception has only a limited effect on parents' decisions.

Concerning the question of why parents choose private schools over public schools, the results show that, if parents were only concerned about quality, they would choose public schools more often. The result would be the same if they did not have a misperception about quality. However, if parents had more freedom in terms of the schools their children could attend, they would choose private schools more often. This last result suggests that admission rules are binding restrictions and that relaxing them would increase the demand for private schools.

Regarding the causes of the increase in the knowledge gap between students from different socioeconomic backgrounds, simulations show that schools' admission rules and household location are relevant in explaining the rise in this gap. This result supports the papers which argue that Chilean SES school segregation cannot be explained only by geographical segregation.[64]

---

[64]See, for example, Valenzuela et al. (2014), Elacqua (2012) and Hsieh and Urquiola (2006).

Finally, it should be noticed that, even though these counterfactual exercises are very useful to compare the effects of different policies on relevant outcomes (e.g., inequality), these are in general small effects. The latter can be partially explained by the fact that, in all the simulations, the choice set is fixed conditional on parents' education and home location. This limitation is something that should be addressed in future research.

References

Aguirregabiria, V. and P. Mira (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics 156*(1), 38 – 67. Structural Models of Optimization Behavior in Labor, Aging, and Health.

Alderman, H., P. F. Orazem, and E. M. Paterno (2001). School quality, school cost, and the public/private school choices of low-income households in pakistan. *The Journal of Human Resources 36*(2), pp. 304–326.

Angrist, J., E. Bettinger, and M. Kremer (2006). Long-term educational consequences of secondary school vouchers: Evidence from administrative records in colombia. *American Economic Review 96*(3), 847–862.

Auguste, S. and J. P. Valenzuela (2006). Is it just cream skimming? school vouchers in chile. Technical report.

Azmat, G. and J. Garcia-Montalvo (2012, June). The role of awareness, information gathering and processing in school choice. Working Papers 639, Barcelona Graduate School of Economics.

Bast, J. L. and H. J. Walberg (2004). Can parents choose the best schools for their children? *Economics of Education Review 23*(4), 431 – 440. Special Issue In Honor of Lewis C. Solman.

Bellei, C. (2009). The private-public school controversy: The case of chile. In P. Peterson and R. Chakrabarti (Eds.), *School Choice International*, pp. 165–192. Cambridge, MA: MIT Press.

Bettinger, E. (2011). Chapter 7 - educational vouchers in international contexts. In S. M. Eric A. Hanushek and L. Woessmann (Eds.), *Handbook of The Economics of Education*, Volume 4 of *Handbook of the Economics of Education*, pp. 551 – 572. Elsevier.

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

Bravo, D., S. Mukhopadhyay, and P. E. Todd (2010). Effects of school reform on education and labor market performance: Evidence from chile's universal voucher system. *Quantitative Economics 1*(1), 47–95.

Carneiro, P., K. T. Hansen, and J. J. Heckman (2003). 2001 lawrence r. klein lecture: Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review 44*(2), pp. 361–422.

Chumacero, R. A., D. Gómez, and R. D. Paredes (2011, October). I would walk 500 miles (if it paid): Vouchers and school choice in chile. *Economics of Education Review 30*(5), 1103–1114.

Conlisk, J. (1996). Why bounded rationality? *Journal of Economic Literature 34*(2), 669–700.

Contreras, D., P. Sepulveda, and S. Bustos (2010). When schools are the ones that choose: The effects of screening in chile. *Social Science Quarterly 91*(s1), 1349–1368.

Cunha, F., J. Heckman, and S. Navarro (2005, April). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers 57*(2), 191–261.

Drago, J. L. and R. Paredes (2011). La brecha de calidad en la educación chilena. *Cepal Review 104*, 167–180.

Eckstein, Z. and K. I. Wolpin (1989). The specification and estimation of dynamic stochastic discrete choice models: A survey. *The Journal of Human Resources 24*(4), pp. 562–598.

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

Epple, D. and R. Romano (2008). Educational vouchers and cream skimming. *International Economic Review 49*(4), 1395–1435.

Epple, D. and R. Romano (2012, 07). Economic modeling and analysis of educational vouchers. *Annual Review of Economics 4*(1), 159–183.

Epple, D. and R. E. Romano (1998). Competition between private and public schools, vouchers, and peer-group effects. *The American Economic Review 88*(1), pp. 33–62.

Gallego, F. A. and A. E. Hernando (2008). On the determinants and implications of school choice: Semi-structural simulations for chile [with comments]. *Economía 9*(1), pp. 197–244.

Gauri, V. (1999). *School Choice In Chile: Two Decades of Educational Refor*. Pitt Latin American Studies. University of Pittsburgh Press.

Geyer, J. and H. Sieg (2013, November). Estimating a model of excess demand for public housing. *Quantitative Economics 4*(3), 483–513.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (2004). Disruption versus tiebout improvement: the costs and benefits of switching schools. *Journal of Public Economics 88*(9?10), 1721 – 1746.

Hastings, J. S. and J. M. Weinstein (2008, November). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics 123*(4), 1373–1414.

Henig, J. (1996). The local dynamics of choice: Ethnic preferences and institutional responses. In R. E. B. Fuller and G. Orfield (Eds.), *Who chooses? Who loses? Culture, institutions, and the unequal effects of school choice*, pp. 93–117. New York: Teachers College Press.

Houser, D., M. Keane, and K. McCabe (2004). Behavior in a dynamic decision problem: An analysis of experimental evidence using a bayesian type classification algorithm. *Econometrica 72*(3), 781–822.

Hsieh, C.-T. and M. Urquiola (2006, September). The effects of generalized school choice on achievement and stratification: Evidence from chile's voucher program. *Journal of Public Economics 90*(8-9), 1477–1503.

Kane, T. J., S. K. Riegg, and D. O. Staiger (2006). School quality, neighborhoods, and housing prices. *American Law and Economics Review 8*(2), 183–212.

Keane, M. P. and K. I. Wolpin (1994, November). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. *The Review of Economics and Statistics 76*(4), 648–72.

Larrañaga, O. (2004, March). Competencia y participación privada: La experiencia chilena en educación. Working Papers wp208, University of Chile, Department of Economics.

Mark, G. Elacqua, and J. Buckley (2006). School choice in chile: Is it class or the classroom? *Journal of Policy Analysis and Management 25*(3), 577–601.

McMillan, R. (2004). Competition, incentives, and public school productivity. *Journal of Public Economics 88*(9?10), 1871 – 1892.

Meckes, L. and R. Carrasco (2010). Two decades of simce: an overview of the national assessment system in chile. *Assessment in Education: Principles, Policy and Practice 17*(2), 233–248.

Mizala, A. and P. Romaguera (2000). School performance and choice: The chilean experience. *The Journal of Human Resources 35*(2), pp. 392–417.

Mizala, A., P. Romaguera, and M. Urquiola (2007). Socioeconomic status or noise? tradeoffs in the generation of school quality information. *Journal of Development Economics 84*(1), 61–75.

Mizala, A. and M. Urquiola (2013). School markets: The impact of information approximating schools' effectiveness. *Journal of Development Economics 103*(0), 313 – 335.

Navarro, S. (2011). Using observed choices to infer agent's information: Reconsidering the importance of borrowing constraints, uncertainty and preferences in college attendance. Technical report.

Neilson, C. (2013). Targeted vouchers, competition among schools, and the academic achievement of poor students. Technical report.

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

Rust, J. (1994, January). Structural estimation of markov decision processes. In R. F. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume 4 of *Handbook of Econometrics*, Chapter 51, pp. 3081–3143. Elsevier.

Schneider, M., P. Teske, M. Marshall, and C. Roch (1998). Shopping for schools: In the land of the blind, the one-eyed parent may be enough. *American Journal of Political Science 42*(3), pp. 769–793.

Simon, H. A. (1986). Rationality in psychology and economics. In R. M. Hogarth and M. W. Reader (Eds.), *In Rational Choice: The Contrast Between Economics and Psychology*, pp. 25–40. Chicago: University of Chicago Press.

Valenzuela, J. P., C. Bellei, and D. d. l. Ríos (2014). Socioeconomic school segregation in a market-oriented educational system. the case of chile. *Journal of Education Policy 29*(2), 217–241.

Zamora, G. (2011, 00). Movilidad escolar en Chile: Análisis de las Implicancias para la Calidad y Equidad de la Educación. *Estudios pedagógicos (Valdivia) 37*, 53 – 69.

A Figures and Tables

Table 8: Descriptive statistics of students with and without grade retention history

Variable	Students who did not repeat a grade	Students who repeated a grade	Difference	z-statistic	P-Value
Changed school at least once	0.50	0.70	-0.201	-67.37	0.00
Moved to another house	0.09	0.12	-0.025	-14.11	0.00
Can not afford the tuition fee	0.03	0.05	-0.020	-19.27	0.00
Forced to switch given low performance	0.01	0.05	-0.041	-63.72	0.00
Students did not like the school	0.06	0.11	-0.054	-35.82	0.00
Parents found better school	0.15	0.10	0.046	22.32	0.00
School does not offer secondary level	0.04	0.04	0.000	0.18	0.86
Another reason	0.11	0.17	-0.063	-32.85	0.00
Number of observations	165606	34327			

The sample contains all the students who took 2011 SIMCE for 8th grade.

Table 9: Fraction of students changing school

	2nd	3th	4th	5th	6th	7th	8th	N
All	0.06	0.06	0.05	0.05	0.05	0.07	0.04	77432
Mother's education								
Incomplete elementary school	0.04	0.04	0.04	0.04	0.03	0.04	0.03	6962
Complete elementary school	0.06	0.05	0.04	0.05	0.03	0.05	0.03	8182
Incomplete secondary school	0.06	0.05	0.05	0.05	0.05	0.06	0.04	11085
Complete secondary school	0.06	0.06	0.05	0.05	0.05	0.08	0.04	29667
Complete or incomplete college	0.06	0.07	0.06	0.06	0.06	0.09	0.06	21536

Note: I drop the students who fail at least one class between 1st and 8th grade, and I also drop the students whose families switch the municipality where they live.

Table 10: Total school change

	0	1	2	>2	N
All	0.69	0.24	0.05	0.01	77432
Mother's education					
Incomplete elementary school	0.78	0.19	0.03	0.01	6962
Complete elementary school	0.75	0.20	0.04	0.01	8182
Incomplete secondary school	0.72	0.22	0.05	0.01	11085
Complete secondary school	0.68	0.26	0.05	0.01	29667
Complete or incomplete college	0.65	0.27	0.07	0.02	21536

Note: I drop the students who fail at least one class between 1st and 8th grade, and I also drop the students whose families switch the municipality where they live.

Table 11: Descriptive Statistics of the Two Samples

	Estimation Sample	Complete Sample
Parents with Low Education	0.209	0.203
Parents with Medium Education	0.428	0.409
Parents with High Education	0.363	0.389
Religious Family	0.212	0.243
Parents Attend School Meeting*	0.946	0.946
Parents Explain Class Material**	0.469	0.470
Parents Help Student Study***	0.402	0.392
Parents Care about Quality****	0.377	0.386
N	9752	18280

All the figures are proportions.
(*) From the question: How often do you attend the periodic parents' meeting of your child's class? I construct a dummy variable that takes one if parenst always or almost always attend, and zero otherwise.
(**) From the question: How often does one of your parents explain you the class material that you don't understand? I construct a dummy variable that takes one if the answer is always or almost always attend, and zero otherwise.
(***) From the question: How often does one of your parents help you to study? I construct a dummy variable that takes one if the answer is always or almost always attend, and zero otherwise.
(****) From the question: Name the first three reasons why you chose your child's current school.; I construct a dummy variable that takes one if parenst name school quality as one of the reasons, and zero otherwise.

B Estimated Parameters and Model's Inputs

B.1 Model's Inputs and Parameters Estimated Outside the Model
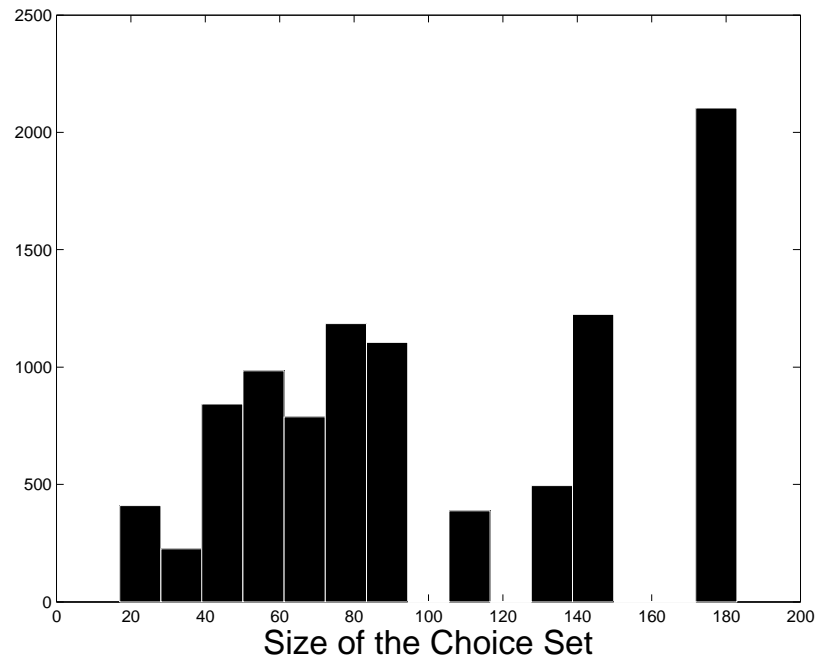
Figure 7: Histogram of the Size of the Choice Sets

Figure 8: Distribution of Estimated Quality by School Type in 2004
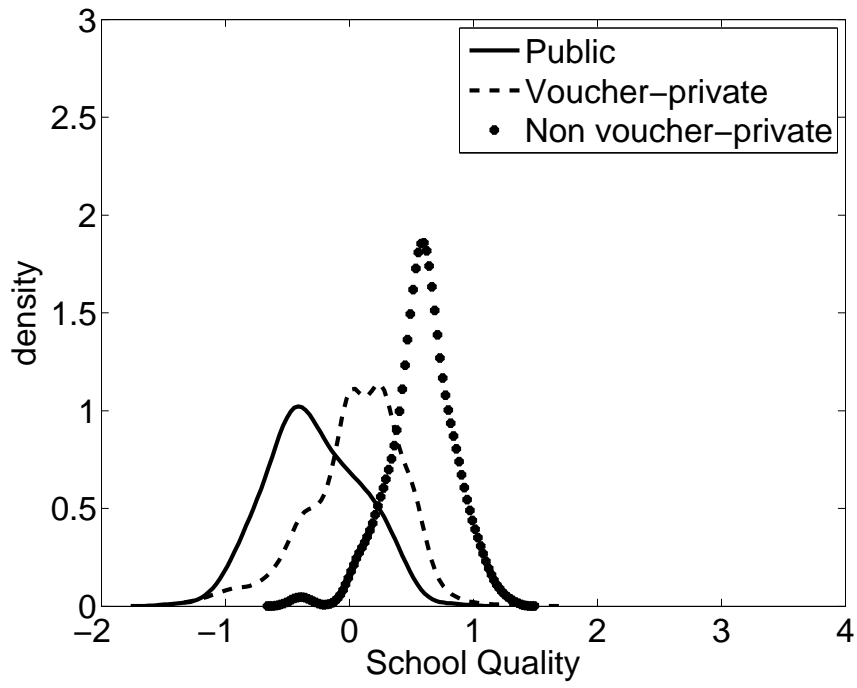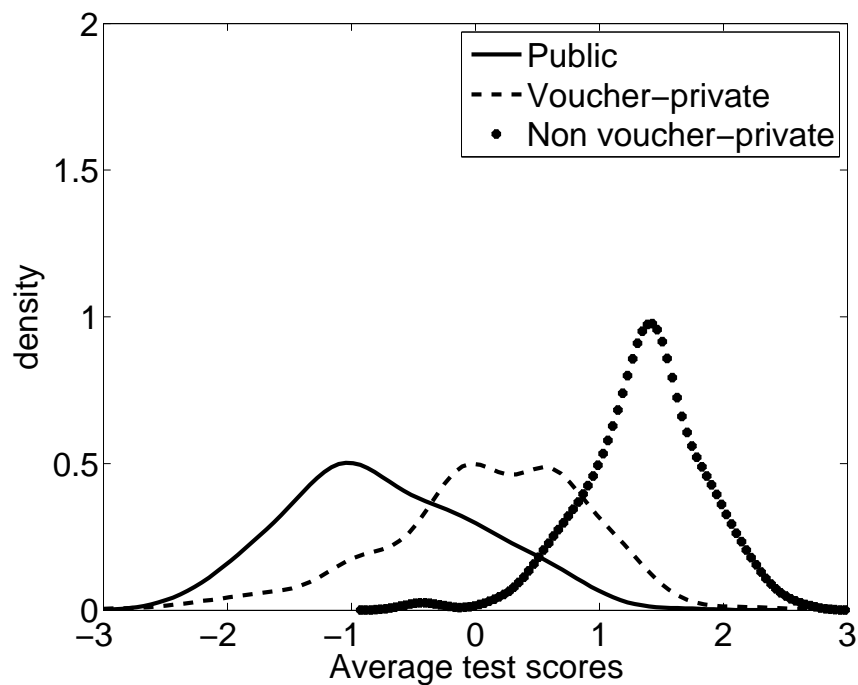
Figure 9: Distribution of test scores by School Type in 2004

??

Table 12: Knowledge Production Function

|	Coeff	Std Err.
Accumulated School Quality at 8th Grade	3.69	0.0314
Parents with Medium Level of Education	5.49	0.2396
Parents with High Level of Education	9.21	0.2646
Constant	256.59	0.1947

$N = 55,421$
$R^2 = 0.2113$

Note: This is the result of estimating the Knowledge Production Function $\hat{K}_{ijT} = \alpha_0 X_i + \alpha_1 \sum_{t=1}^{T} \tilde{q}_{ijt} + \vartheta_{ijT}$.
The omiited dummy variable is Parents with Low Level of Education.

41

Table 13: GPA Production Function

	4th 2005		4th 2006		4th 2007		4th 2008		4th 2009		8th 2011	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Entered in 2nd	-0.07	0.0034	-0.06	0.0034	-0.06	0.0038	-0.06	0.0039	-0.07	0.0034	-0.05	0.0054
Entered in 3th	-0.08	0.0033	-0.08	0.0033	-0.08	0.0036	-0.07	0.0038	-0.07	0.0033	-0.04	0.0050
Entered in 4th	-0.11	0.0030	-0.11	0.0030	-0.12	0.0035	-0.10	0.0037	-0.11	0.0032	-0.06	0.0048
Entered in 5th											-0.06	0.0045
Entered in 6th											-0.08	0.0045
Entered in 7th											-0.10	0.0040
Entered in 8th											-0.17	0.0046
Constant	3.63	0.0063	3.72	0.0062	3.74	0.0066	3.75	0.0066	3.72	0.0065	3.02	0.0097
N	214,661		210,177		162,552		169,500		178,685		161,920	
Adjusted R-squared	.549		.545		.559		.537		.558		.465	

Note: This is the result of estimating the GPA production function $G_{ijt} = \lambda_{0j}^t + \lambda_{1j}^t K_{ijt} + \lambda_2^t TA_{it} + \varepsilon_{ijt}^g$, where I only present the values for λ_2^t. The omitted dummy variable is Entered in 1st grade.
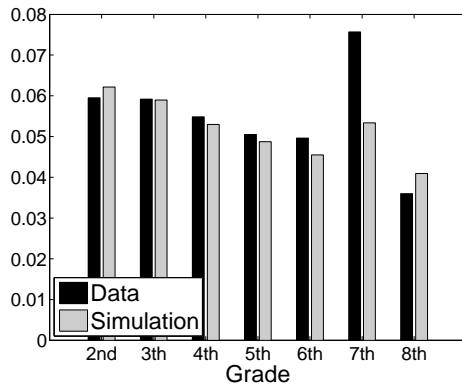
Figure 10: GPA Standards

(a) 2005 (4th Grade)   (b) 2011 (8th Grade)
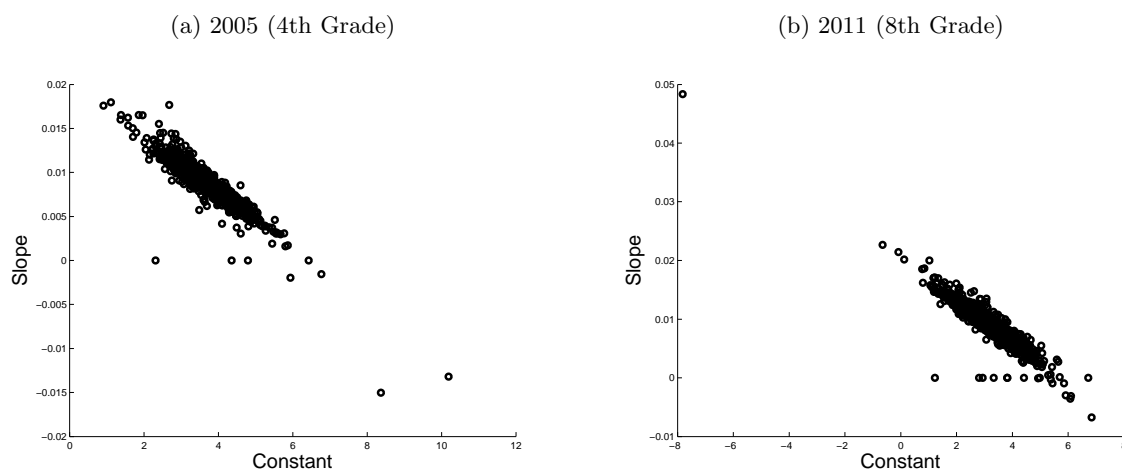
Note: Giving the GPA production function $G_{ijt} = \lambda_{0j}^t + \lambda_{1j}^t K_{ijt} + \lambda_2^t TA_{it} + \varepsilon_{ijt}^g$; λ_{0j} is the constant and λ_{0j} the slope.

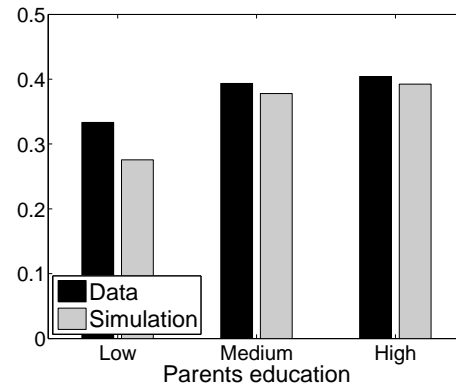# C Model Fit and Counterfactual experiments (complete sample)

C.1 Model Fit
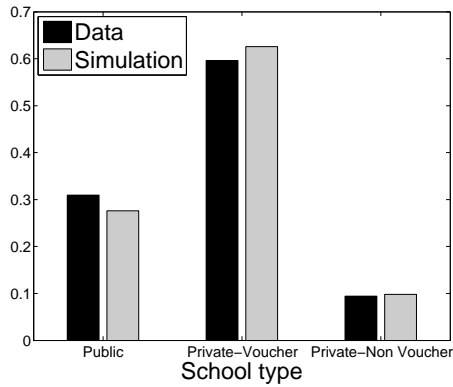
Figure 11: Model fit for the complete sample

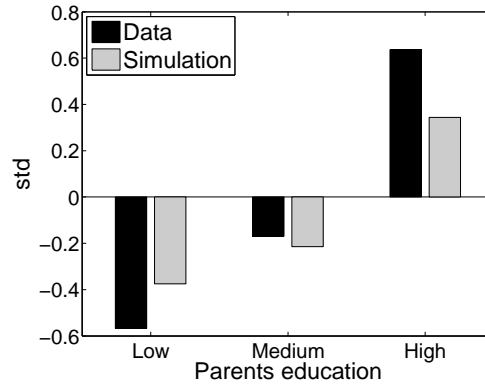(a) Fraction of student changing their school by grade

(b) Average total change by parents education

(c) Student fraction by school type

(d) Gain in Knowledge by parents education

C.2 Counterfactual Experiments

Table 14: Fraction of students changing school by grade (with respect to baseline in percentage points), complete sample.

	Baseline	No misperception	Only Q maters	All admitted	Random admission	$C*0.9$	New locations	All $ED=3$
2nd	6.2%	0.1	0.6	10.9	0.1	4.6	-1.3	-0.1
3th	5.8%	-0.0	1.1	10.8	0.0	4.1	-1.4	-0.1
4th	5.3%	-0.0	1.6	10.4	0.0	4.1	-1.2	-0.1
5th	4.9%	0.0	1.9	10.3	0.2	4.0	-1.0	-0.1
6th	4.5%	0.0	2.2	9.9	0.2	3.8	-1.0	-0.1
7th	5.2%	-0.0	2.0	9.1	-0.0	4.1	-1.4	-0.2
8th	4.2%	-0.0	2.6	8.8	-0.1	3.4	-1.1	-0.2

Table 15: Student fraction by school type at first grade (with respect to baseline in percentage points), complete sample.

	Baseline	No misperception	Only Q maters	All admitted	Random admission	$C*0.9$	New locations	All $ED=3$
Public	27.7%	0.8	2.7	-2.7	-2.8	0.4	0.7	-3.1
Voucher Private	62.6%	-0.3	-3.0	-0.3	-0.3	-0.3	6.6	1.8
Non voucher Private	9.7%	-0.6	0.3	3.0	3.1	-0.1	-7.4	1.3

Table 16: Gain of knowledge $(KT - K0)$ by parents education (with respect to baseline in standard deviations), complete sample.

	Baseline	No misperception	Only Q maters	All admitted	Random admission	$C*0.9$	New locations	All $ED=3$
Incompleted High school	-0.429	-0.002	0.010	0.084	0.059	-0.002	-0.118	0.085
Completed High school	-0.233	-0.028	0.038	0.111	0.087	-0.009	-0.004	0.112
With college studies	0.386	-0.047	0.137	0.075	0.045	-0.001	-0.129	0.011
All	-0.032	-0.030	0.071	0.091	0.065	-0.005	-0.076	0.067