Preference Estimation in Deferred Acceptance with Partial School Rankings

(JOB MARKET PAPER)

Shanjukta Nath[*]

Department of Economics, University of Maryland, College Park

January 16, 2021

Abstract

The Deferred Acceptance algorithm is a popular school allocation mechanism thanks to its strategy proofness. However, with application costs, strategy proofness fails, leading to an identification problem. In this paper, I address this identification problem by developing a new Threshold Rank setting that models the entire rank order list as a one-step utility maximization problem allowing for both ranked as well as non-ranked alternatives. I apply this framework to study student assignments in Chile. There are three critical contributions of the paper. I develop a recursive algorithm to compute the likelihood of my one-step decision model. Partial identification is addressed by incorporating the outside value and the expected probability of admission into an additive cost framework. The empirical application reveals that although school proximity is a vital variable in school choice, student ability is critical for ranking high academic score schools. The results suggest that policy interventions such as tutoring aimed at improving student ability can help increase the representation of low-income low-ability students in better quality schools in Chile.

Keywords: Education, Centralized Algorithm, Segregation.

JEL classification: I20, I24, I28.

1 Introduction

Policymakers are using centralized allocation algorithms to match students to schools all across the globe. These algorithms require parents to submit a ranking over schools, which is a critical input to the matching process.[1] Among such algorithms, the Deferred Acceptance (DA) algorithm is extensively used as it is strategy-proof (Abdulkadiroğlu & Sönmez, 2003). Parents are expected to report school rankings according to their true underlying preferences and not manipulate their behavior to achieve favorable allocations.[2]

However, strategy proofness of DA fails if there are positive application costs (Fack, Grenet, & He, 2019). Empirical studies studying parental rankings in DA find that parents often report a partial list. The partial list can result from hard limits on the length of the list(Haeringer & Klijn, 2009; Luflade, 2017). But, it is observed that parents do not reveal the complete ranking over schools, even without any hard limits. Multiple factors can contribute to such behavior. Often parents have access to a guaranteed school, and they do not want to list schools below this outside option. Parents are also aware of the oversubscribed schools and would like to skip reporting such schools if the likelihood of admission is very low and does not compensate for the application cost. Partial ranking due to these factors can lead to non-strategy proof equilibria and pose an issue with identification.

This paper addresses identification under partial student rank ordered list (ROL) in DA. I incorporate the value of the outside option and the probability of admission in an additive cost framework in the school choice model to identify the determinants of true preferences over schools (Kuersteiner, Nath, & Urzua, 2020). The school choice framework developed here differentiates between ranked and non-ranked alternatives using the rank cut-off. Ranked schools are assumed to provide higher utility (accounting for the probability adjusted cost) than the cut-off, and non-ranked do not offer as much utility as the cut-off. Notably, the cut-off is determined by the outside option value. I contribute to the methodology of rank-ordered choices by modeling the student's decision to rank schools as a single step process. In my model, the random error component of latent utilities is known to the parents and is drawn once and assumed to be fixed for the entire school choice process. The one-step process is computationally intractable as the likelihood involves an m-dimensional definite integral. For instance, if a parent ranks m schools, the likelihood involves $O(2^m)$ terms, which becomes computationally prohibitive if m is large. Therefore, I develop a novel recursive algorithm to compute the likelihood efficiently. The new model predicts the length of student ROL along with its determinants by allowing for variation in rank cut-offs. This is applied to Chilean DA for high school assignments. I find that student ability plays a critical role in ranking behavior in Chile that may dominate the costs imposed due to low household income or travel distance to schools.

---

[1]Variants of centralized student assignment algorithms have been used in New York City, Boston, Chicago, Mecklenburg County in the United States (Abdulkadiroğlu, Agarwal, & Pathak, 2017; Abdulkadiroğlu, Pathak, Roth, & Sönmez, 2005; Hastings, Kane, & Staiger, 2005; Pathak & Sönmez, 2013), Finland (Salonen et al., 2014), Hungary (Biró, 2012), Ghana (Ajayi, 2013), Tunisia (Luflade, 2017) and several other countries.

[2]Another common assignment algorithm is the Boston mechanism (Ergin & Sönmez, 2006; Pathak & Sönmez, 2008). However, research indicates that it gives sophisticated parents an advantage as there are gains associated with strategic behavior.

There are several unique features of the Chilean system that leads to partial ROL and variations in the rank cut-off. Parents submit partial ROLs even though there are no monetary application fees or limits on length of ROL. About 60% of parents rank at most three schools for high school admission. Although there are no direct monetary costs, there might be implicit costs. Such implicit costs are likely to be a function of the level of sophistication of the parents.

The outside option is an essential component of the cut-off at which the parents decide to stop ranking other schools. In Chile's case, the outside option depends on two features for middle to high school transition. If the student's pre-DA school offers the grade at which the student seeks admission, the student is guaranteed admission at pre-DA school when the algorithm fails to allocate the student at any of the listed schools in ROL. On the contrary, if the pre-DA school does not offer a higher grade, which is the case for many middle schools, the student gets allocated to the nearest public school with a vacancy. Moreover, there is a high negative correlation between vacancy and school academic quality. Students enrolled in high score schools pre-DA are less likely to participate in DA. This correlation makes it critical to account for indicators of the probability of admission explicitly in the school choice model.

In this paper's school choice estimator, identification can be achieved for Chile with full support. If there is a positive density of all types of students around different school types, this variation in student location of a similar type can be critical for identification. In other words, let's assume one can use student income and the ability to define student type. School types are defined by academic quality as a simplifying assumption. If similar typed low income and low ability students are located around high and low academic quality schools, then this distance variation based on location can be used for identification. The student location provides variation in the outside option. If there is variation in outside options for similar type students, this can be used to explain the length of ROL for these students. Empirical analysis reveals a notable variation in the placement of similar type students around different schools and the outside value due to this geographic location.

The parametric framework allows me to account for the outside option value and the additive costs adjusted by the assignment probability. There are likely systematic differences in costs based on unobserved characteristics. I account for unobserved heterogeneity using the correlated random effect. Correlated random effects are a commonly used technique to account for unobserved heterogeneity, particularly in situations with a small number of observations per agent, making fixed effect estimation inconsistent. I use the EM algorithm to account for the unobserved component in the underlying preferences.

A crucial source of variation for identification in the Chilean context is the distance to school as it modifies the outside option. Consequently, it is critical to computing this variable precisely. Most of the current work in related literature uses crude proxies of this distance measure either due to the precise student address's unavailability or due to the computational costs of calculating such travel distances for large administrative data-sets (Burgess & Briggs, 2010; Burgess, Greaves, Vignoles, & Wilson, 2015; Laverde, 2020). The Chilean government has provided precise student and school geo-coordinates to the researchers to analyze the new system. I take advantage of this information and employ the Open Source Routing Machine (OSRM) API to compute travel distances for each student school combination for the

set of all schools in a student's school market to use as an input into the rank order model.[3]

A decentralized system preceded the centralized system with vouchers for students belonging to poor income families. However, school segregation in Chile based on socioeconomic status has increased in the last decade.[4] This is due to the disproportionate flight of students belonging to middle-income households from the public to voucher schools. Consequently, the Chilean government's key motivation behind introducing centralized student assignments is to reduce the existing levels of school segregation. However, Kutscher, Nath, and Urzua (2020) does not find any evidence suggesting an unambiguous decline in school segregation due to the government's new policy. This result is crucial and necessitates the study of parental ranking preferences.

My results illustrate that travel distance is a critical element in the decision-making process. Nevertheless, parents also care about the extent of the match between student ability and school academic rigor. The impact of distance on the ROL varies substantially by student ability and income. I observe two critical results. Higher-income households might have an advantage in overcoming the travel cost to good quality schools. However, student ability proxied by pre-centralization test scores is the most crucial determinant of listing the best quality schools in the students' school market. Once I condition the marginal effects on student income, the results showcase that higher ability students do end up applying to good quality schools irrespective of income levels. In other words, the ability can compensate for income levels and induce parents to incur the additional travel cost as parents care about the ability match between the student and school.

The above result is critical as it suggests that reducing travel costs alone might not improve students' representation from low socioeconomic status in higher-quality schools. Policies geared towards improving student ability, especially for students belonging to low-income households, can go a long way in improving their representation in high-quality schools. Optimal reallocation can help in improving student outcomes. It can help reduce absenteeism, drop out rates (Hanushek, Lavy, & Hitomi, 2008) and improve academic performance (Glewwe & Jacoby, 1994; Hastings & Weinstein, 2008; Kirabo Jackson, 2010).

This paper is organized as follows. In section 2, I discuss the theoretical framework for the school choice model. I present the estimation strategy and the recursive algorithm in section 3. Section 4 discusses Chile's institutional setting, critical features of centralized allocation, followed by a detailed description of the analysis's data-sets. I also provide some reduced-form evidence on the impact of school quality on student attendance and the extent of school diversity before and after the reform. I apply the new estimator to student data in Chile and present the results in section 5. I discuss alternative simulations and make policy recommendations in section 6. Lastly, I conclude in section 7.

---

[3]These computations were done on secure servers in collaboration with the Ministry of Education in Chile. The resulting data-sets are owned by the Ministry of Education, Chile.

[4]The voucher system led to a proliferation of voucher schools in Chile funded partially by government vouchers but managed by private entities

2 Theoretical Framework

I follow the school choice model provided in Fack et al. (2019) and Kuersteiner et al. (2020). I assume that every student $i \in \{1, 2, .., n\}$ chooses from a set $\mathcal{J} = \{1, 2, ..., J\}$ schools.[5] The three components of DA comprises of student preferences, priority indices for each student school pair and the set of vacancies $q = \{q_1, ..., q_J\}$ at the DA participating schools. Given these three components, the student submits a ROL $L_i = \{l_{1,i}, l_{2,i}...., l_{K_i,i}\} \in \mathcal{L}$, which are manifested ranks over latent utilities. Here, $l_{1,i}$ is the top ranked school, $l_{2,i}$ is the second choice and so on and so forth till the student ranks the least preferred school $l_{K_i,i}$. I allow the cut-off where the student decides to stop ranking schools to vary across individuals and is denoted by K_i.

I use a linear index for the student utilities from schools. This linear index is the sum of a component V_{ij} that is observed to the researcher and a random component ϵ_{ij} not observed by the researcher. Consequently, the set of student utilities is described as $u_i = \{u_{1,i}, ..., u_{J,i}\}$, where $u_{ij} = V_{ij} + \epsilon_{ij}$.

V_{ij} comprises of three types of explanatory variables i) covariates varying for each student school pair, ii) student specific characteristics and iii) school specific characteristics. Finally, V_{ij} also comprises of any type of advantage that student i enjoys for school j. Such advantages for school j can be a function of factors such as sibling enrolled in school j, parents employed in school j, former students or for students with special needs. Conditional on these factors, DA uses a random lottery to generate priority indices for each student school pair used for tie breaking in each iteration in DA. In other words, students with an advantage at school j will be assigned a higher lottery number relative to a student without an advantage all else equal.

The popularity of DA is associated with its property of being strategy-proof. This mainly depends on parents revealing a complete ordering over schools and no costs associated with adding additional schools to ROL. However, the ROL observed by the researcher in real applications is a partial list instead of a complete ordering, and second, there are often some positive costs of the application.

First, the partial list in DA can be attributed to the guaranteed seat or outside option. Since students have a positive probability of getting assigned to one of the listed schools, the ranked schools must offer at least as much utility as the guaranteed school. For instance, in the Chilean DA, the value of the outside option is determined by one of the following two components. Student i, while making the transition from middle to high school through DA, has a guaranteed seat in the middle school if it offers the high school grade. On the contrary, if the middle school does not offer high school grades, the student gets automatically allocated to the nearest public school with vacancies.[6] This generates the outside option $u_{i,0}$ for student i. The outside option is observed to the researcher in the Chilean DA. Consequently, all schools in $\mathcal{J}$ that do not offer at least $u_{i,0}$ will not be ranked in ROL ($u_{i,j} - u_{i,0} < 0$).

Second, DA is often associated with some costs of adding schools to ROL. The costs have multiple interpretations in student assignment. In some DA applications such as Hungary, there is an application

---

[5] $\mathcal{J}$ corresponds to the set of the total number of participating schools in DA, and this is likely different from the total available schools in the schooling market. For instance, only public and voucher schools participated in DA and not the non-voucher schools in the Chilean schooling market.

[6] In Chile, this cut-off for the nearest public school with vacancies is 17 km from the student's place of residence.

fee associated with additional schools. On the contrary, Fack et al. (2019) interestingly suggests that there could be an implicit cost of application even without an explicit fee. Often this could correspond to the mental cost of obtaining information about several aspects of the school and the effort of listing additional schools on ROL.

Variation in the value of an outside option can often result in partial lists. However, the critical question is even when I account for $u_{i,0}$ in the school choice model, do I observe a partial list based on true ranking for the schools strictly preferred over the outside option. This might not hold if there are costs associated with listing additional schools or parents are trying to exclude schools that are impossible to get into due to low vacancies. Proposition 1 in Fack et al. (2019) illustrates that the truth-telling property is no longer the equilibrium strategy for all under a positive application cost. Moreover, the authors highlight that the cost's magnitude need not be considerable to deviate from truth-telling. Even if this cost component is minuscule, the marginal benefit of adding school can be meager if the probability of admission to the additional school is close to 0 or there is a high chance of admission to a higher ranked school. All these factors poses difficulty for identification of student preferences in DA.

Kuersteiner et al. (2020) extends the results in Fack et al. (2019) and illustrates that observed ROL's are a subset of actual preferences by parameterizing the reasons for leaving out certain schools. This parameterization can be used for identification. We show that under no costs or a linear cost specification where the probability of admission enters the utility function in a additive manner one can identify the parameters of the student preferences over schools.

In Chilean DA, cost comprises of the implicit mental cost of application as there is no application fee. Such cost are less likely to be school specific and hence I assume a linear cost function. Using the linear index parameterization for utilities and proposition 2 in Kuersteiner et al. (2020), the utilities from a ranked alternative has to compensate for the value of outside option and the cost component adjusted by the probability of admission. In other words,

$$
u_{i,j} - \frac{c_i}{p_j^s} = \begin{cases} X_{ij}\beta + \epsilon_{ij} - \frac{c_i}{p_j^s} > u_{i,0}, & \text{if } j \text{ is ranked} \\ X_{ij}\beta + \epsilon_{ij} - \frac{c_i}{p_j^s} \leq u_{i,0}, & \text{otherwise} \end{cases}
$$

Further, I impose the following assumptions.

Assumption 1 *The school choice decision is made in a single step where student i draws from the error component ϵ_{ij} from an i.i.d Type 1 extreme value distribution.*

According to Assumption 1 the unobserved component of the utility is known to the student and I assume it to be fixed for the school choice decision process. This deviates from the sequential choice process used in the urn model where the agent is assumed to draw the unobserved component repeatedly in each step.

Assumption 2 *The individual specific component c_i is modeled as a correlated random effect as follows*

$$
c_i = Z_i\gamma + w_i
$$

6

where Z_i is a student specific observed characteristic such as the outside option value and w_i is drawn from $N(0, \sigma^2)$.

Identification in data using the above set up can be achieved with the assumption of full support. Full support implies that there should exist variations in the placement of all types of students across different school types. This geographic variation will generate variation in the outside value or the quality of the guaranteed school, which will help to pin down the parameters of parental preferences in the presence of partial lists.

I explain the argument using a simple example. To keep the illustration simple, I assume two students type low income and high income. I also assume two school types of high and low ability. The student and school types are part of the observed covariates vector X_{ij}. Besides, this vector contains information for travel distance for every school student pair. The placement of every student type around all types of schools will make the outside option school for each type of student accessible for students of other types. This variation in the outside option for similar types of students such a low income due to the geospatial location will reveal the preference ordering of low-income students over all types of schools.

3   Estimation: Threshold Rank Order Model

The school choice model outlined in section 2 provides a framework to identify the determinants of true parental preferences with the partial ROL. In this section I lay out the steps used for estimation.

The individual i ranks $|L_i| = K_i$ schools out of a choice set of $\mathcal{J}$ schools. Also the individual draws the random component ϵ_{ij} once and then it is fixed for the school choice process. This deviates significantly from the assumption on errors in a sequential decision making process where the random components are drawn in every draw from the urn (Glazerman & Dotter, 2017).[7] The likelihood for an individual given the unobserved cost c and unknown parameter β is as follows

$$g(L|c;\beta) = P(u_1 - \frac{c}{p_1^s} > u_2 - \frac{c}{p_2^s} > .. > u_K - \frac{c}{p_K^s} > u_0, u_{K+1} - \frac{c}{p_{K+1}^s} < u_0, .., u_J - \frac{c}{p_J^s} < u_0)$$

$$= P(u_1 - \frac{c}{p_1^s} > .. > u_K - \frac{c}{p_K^s} > u_0 | u_{K+1} - \frac{c}{p_{K+1}^s} < u_0, .., u_J - \frac{c}{p_J^s} < u_0) P(u_{K+1} - \frac{c}{p_{K+1}^s} < u_0, .., u_J - \frac{c}{p_J^s} < u_0)$$

$$= P(u_1 - \frac{c}{p_1^s} > .. > u_K - \frac{c}{p_K^s} > u_0) \prod_{l=K+1}^{J} P(u_l - \frac{c}{p_l^s} < u_0)$$

$$= \left[\int_{-W_1}^{\infty} \int_{-W_2}^{W_1+\epsilon_1-W_2} ...... \int_{-W_K}^{W_{K-1}+\epsilon_{K-1}-W_K} f(\epsilon_K) d\epsilon_K ..... f(\epsilon_1) d\epsilon_1 \right] \prod_{l=K+1}^{J} \int_{-\infty}^{-W_l} f(\epsilon_l) d\epsilon_l$$

The derivation from the second to the third equality holds as the events u_j for $j \in \{K+1, ..., J\}$ are independent conditional on the covariates and c. Additionally, the movement from step three to four works as the event $u_1 - \frac{c_i}{p_1^s} > u_2 - \frac{c_i}{p_2^s} > .... > u_K - \frac{c_i}{p_K^s} > -u_0$ is independent of $u_{K+1} - \frac{c_i}{p_{K+1}^s} < -c, u_{K+2} - \frac{c_i}{p_{K+2}^s} < -u_0, .., u_J - \frac{c_i}{p_J^s} < -u_0$.

I compute the marginal distribution that student i ranked K_i schools by integrating over the distribution

---

[7]I drop the i subscript for the following derivation to keep the notation simple.

of the unobserved individual correlated effect c.

$$f_l(L_i) = \int g(L_i|c_i; \beta) f_w(c; \gamma, \sigma^2) dc$$

where f_w denotes the distribution of unobserved cost c with unknown parameters γ and σ^2.

Recursive Algorithm: With the single step decision, the likelihood becomes highly complicated and computationally expensive as now every integral in the above density has two bounds. For instance, the above problem with K ranked alternatives has a complexity of 2^K terms. In the Chilean data, there are parents ranking as many as 44 schools. This results in 17 trillion calculations in the likelihood. Due to this complexity, the literature on ROL models has been agnostic on the rank cut-off heterogeneity. I solve this computational challenge by developing a recursive algorithm for the Threshold Rank Order estimator.

Theorem 1 *If there are a total of K ranked alternatives from a set of J schools such that $K+1,..,J$ are non-ranked, then the likelihood of the observed ranks conditioned on covariates and the latent variable is given as*

$$P(u_1 - \frac{c}{p_1^s} > u_2 - \frac{c}{p_2^s} > .. > u_K - \frac{c}{p_K^s} > u_0, u_{K+1} - \frac{c}{p_{K+1}^s} < u_0, .., u_J - \frac{c}{p_J^s} < u_0)$$

$$= \left[\kappa_1 \kappa_2 ... \kappa_{K-2} \kappa_{K-1}(1 - F(-W_K)...F(-W_1)) - \left\{ \kappa_2 .. \kappa_{K-1} F(-W_K) F(-W_{K-1}) .. F(-W_2) \boldsymbol{I(1)} \right\} \right.$$

$$\left. - ...... - \left\{ \kappa_{K-1} F(-W_K) F(-W_{K-1}) \boldsymbol{I(K-2)} \right\} - \left\{ F(-W_K) \boldsymbol{I(K-1)} \right\} \right] \prod_{l=K+1}^{J} F(-W_l)$$

where $\boldsymbol{I(k)} = \int_{-W_1}^{\infty} \int_{-W_2}^{W_1+u_1-W_2} ...... \int_{-W_k}^{W_{k-1}+u_{k-1}-W_k} f(\epsilon_k) d\epsilon_k ..... f(\epsilon_1) d\epsilon_1$, $\kappa_{n-j} = \frac{1}{\sum_{j=n-j}^{n} e^{-(W_{n-j}-W_j)}}$ and $F(-W_j) = e^{-e^{W_j}}$.

The proof of theorem 1 is provided in Appendix A.

The urn model is the most popular model used to analyze ranked data in the literature. According to Plackett and Luce, the ranking process for J items can be thought of as an aggregation of J independent iterations. In the first iteration, the individual chooses the most preferred item from the set of J items. In the next iteration, individual chooses the best from the left over items after the top choice is removed. This process continues till all the items are ranked. The likelihood of the event $u_1 > ..... > u_J$ under a logit specification is

$$P(U_1 > ..... > U_J) = \frac{e^{W_1}}{\sum_{j=1}^{J} e^{W_j}} \times \frac{e^{W_2}}{\sum_{j=2}^{J} e^{W_j}} \times ..... \times \frac{e^{W_{J-1}}}{\sum_{j=J-1}^{J} e^{W_j}} \tag{1}$$

$$= \kappa_1 \kappa_2 ... \kappa_{n-2} \kappa_{n-1}$$

where $\kappa_{n-j} = \frac{1}{\sum_{j=n-j}^{n} e^{-(W_{n-j}-W_j)}}$. I use the same notation as the threshold rank order model for comparison (see Theorem 1). The Plackett-Luce (urn model) multiple stage process for ranking data did not explicitly discuss impartial rankings. Some discussion on ties and partial rankings is done in Allison and Christakis (1994); Guiver and Snelson (2009); Skrondal and Rabe-Hesketh (2003). For

instance, Allison and Christakis (1994) suggested that one can assume an underlying rank over non-ranked alternatives, which is not observed by the researcher. However, one can account for all permutations of possible rank orderings over the non-ranked alternatives in the likelihood. Suppose, the individual ranks two alternatives out of 4 and last two alternatives are non-ranked. I follow the model in Allison and Christakis (1994) and modify the likelihood of the urn model as

$$P(u_1 > ..... > u_4) = \frac{e^{W_1}}{\sum_{j=1}^{4} e^{W_j}} \times \frac{e^{W_2}}{\sum_{j=2}^{4} e^{W_j}} \times \left[\frac{e^{W_3}}{e^{W_3} + e^{W_4}} + \frac{e^{W_4}}{e^{W_3} + e^{W_4}} \right]$$
$$= \frac{e^{W_1}}{\sum_{j=1}^{4} e^{W_j}} \times \frac{e^{W_2}}{\sum_{j=2}^{4} e^{W_j}}$$

The urn model likelihood (equation 1) has several differences with the assumptions used for the school choice model proposed in this paper. First, this method does not differentiate between ranked and non-ranked alternatives in terms of the underlying utility. Second, splitting the decision process into multiple stages assumes that the individual cares about the most preferred alternative for that stage and this decision is completely independent of the decision process during other stages. On the contrary, I work with the assumption that the individual ranks all the ranked schools in the same step.

Next, I discuss the computation of the parameters for the manifested variable ranks and those determining c with unobserved individual heterogeneity. I use the Monte Carlo EM algorithm to estimate the parameter (Dempster, Laird, and Rubin (1977) and Sammel, Ryan, and Legler (1997)). The parameter space for the manifested ranks consist of β (manifested ranks) and individual specific effect consist of γ, σ^2. If I were to observe the c_i, the log-likelihood for the complete data is

$$log f_{rc}(L_i, c_i) = \sum_{i=1}^{N} \{ log[g(L_i|c_i; \beta)] + log[f_w(c_i; \gamma, \sigma^2)] \}, \tag{2}$$

However, the latent cost is unobserved and the computation of the observed likelihood ($f_l(L_i) = \int g(L_i|c_i; \beta) f_w(c_i; \gamma, \sigma^2) dc$) of listed ROL requires to integrate over the unobserved cost. EM algorithm provides an iterative solution to maximum likelihood estimation. In the E step, the algorithm computes the expectation of the observed likelihood using the initial guess of the parameters, and then the initial guess is updated in the M step by optimization of the expectation obtained in the E step.[8]

Equation 2 is critical as it informs that the parameter space can be split into two parts. The first part can be used to estimate β, and the second component can be used to obtain estimates of γ, σ^2 corresponding to the correlated latent variable c. Since c is unobserved, so instead of the above likelihood (equation 2), I use the EM algorithm and compute the expectation of the gradient of the log-likelihood function (score) corresponding to γ and σ^2. I follow the methodology in Sammel et al. (1997) very closely

---

[8]There exists an extensive work on the use of the EM algorithm in the context of latent class structure Croon and Luijkx (1993), Francis, Dittrich, and Hatzinger (2010) and Marden (2014)[9]. However, for my setting, c is a continuous latent variable, and I need an added approximation in the E step to compute the integral (Booth and Hobert (1999); Ibrahim, Chen, and Lipsitz (1999)). Such approximation is not required with finite latent classes as the integral becomes a weighted sum of the conditional likelihood over the posterior distribution of discrete latent classes.

Algorithm 1 Estimate school choice model

Input: ROL L, Explanatory variables X, Z
Result: Parameter estimates for manifested ranks β, latent variable parameter estimates γ, σ

1: Start with initial guess of parameters $\{\beta^0, \gamma^0, \sigma^0\}$
2: Draw T Monte Carlo samples from the prior distribution of latent variable c, which is $f_w(c_i; \gamma, \sigma^2)$.
3: Using the recursive algorithm of theorem 1 compute $g(L|c; \beta)$ and calculate expectation of the score function
4: Solve for γ, σ^2 using equation (2) and (3). Update γ, σ^2.
5: Use the $f_w(c_i; \gamma, \sigma^2)$ and $g(L|c; \beta)$ obtained in step 2 and 3 respectively to formulate the Q as derived in equation 4.
6: Maximize and update β.
7: Repeat steps 1 to 6 till all the parameters converge.
8: **return** $\hat{\beta}, \hat{\gamma}, \hat{\sigma}$

---

for the EM algorithm. The methodology in Sammel et al. (1997) has been modified for my problem of rank-ordered ROL as Sammel et al. (1997) developed the EM method for multinomial choice. Therefore, the distributional assumptions on error term is different in Sammel et al. (1997). Moreover, this paper did not solve for the variance of the distribution of the unobserved cost. I modify their proof used for the mean of the latent variable to additionally solve for the variance. The posterior distribution $h(c_i|L_i)$ is used to form the expectation of the score function.

Step 1, solving for $\hat{\gamma}, \hat{\sigma}$: The expectation of the score function w.r.t γ is given below. I can equate $E_c S_i(\gamma) = 0$ and solve for $\hat{\gamma}$. The proof for the mean is provided in Sammel et al. (1997), section 3.3, page 671.

$$
\hat{\gamma} = \sum_{i=1}^{N}(z_i'z_i)^{-1}\left(\sum_{i=1}^{N} z_i'\left[\frac{\sum_{t=1}^{T} c_i g(L_i|c; \beta)}{\sum_{t=1}^{T} g(L_i|c; \beta)}\right]\right)
$$
$$
\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{N}\left[\frac{\sum_{t=1}^{T}(c_i - z_i\gamma)^2 f(L_i|c; \beta)}{\sum_{t=1}^{T} f(L_i|c; \beta)}\right]
$$

(3)

The proof for σ^2 is provided in Appendix 7.

Step 2, solving for $\hat{\beta}$: In the E step, I compute the expectation of the likelihood w.r.t the distribution of the latent variable c_i conditional on the observed data L_i. This is based on the same steps as Sammel et al. (1997) on how Monte Carlo samples can be used for approximation of the integral. The steps for computation are given below. These are comparable with equation six and the next (unnumbered) equation on page 671 in Sammel et al. (1997) modified for the likelihood in this paper.

$$
E_{h(c|L)}g(L_i, c) = \int f_{lc}(L_i, c)h(c|L; \gamma, \sigma^2)dc
$$
$$
= \frac{\int f_{lc}(L_i, c)g(L_i|c; \beta)f_w(c)dc}{\int g(L_i|c; \beta)f_w(c)dc}
$$

I use Monte Carlo approximation at this step. I draw a sample of c generated from the distribution

$f_w(c)$. $f_w(c)$ follows a normal distribution as $w \sim N(0, \sigma^2)$, but the means are adjusted by $Z\gamma$. Using this sample the Monte Carlo approximation of the integrals are in equation 4. In the M step, I maximize $Q(\beta, \gamma, \sigma^2 | \beta^{(t-1)}, \gamma^{(t-1)}, \sigma^{2(t-1)})$ and update the estimate to $\beta^{(t)}, \gamma^{(t)}, \sigma^{2(t)}$. The optimization over β and γ, σ^2 takes place in two steps within the maximization step. I keep iterating between the E and M step till the estimates converge. The steps of the algorithm needed to compute the estimates are provided as Algorithm 1.

$$E_{h(c|L)}g(L_i, c) = \frac{\sum_{t=1}^{T} f_{rc}(L_i, c_t) f(L_i | c_t)}{\sum_{t=1}^{T} g(L_i | c_t; \beta)} \quad , \quad Q(\beta, \gamma, \sigma^2 | \beta^{(0)}, \gamma^{(0)}) = \frac{\sum_{t=1}^{T} g(L_i | c_t; \beta)^2 f_w(c)}{\sum_{t=1}^{T} g(L_i | c_t; \beta)} \quad (4)$$

Monte Carlo Simulations: I evaluate the performance of the estimator for a model with multiple covariates (four covariates) and no latent individual heterogeneity. First, I generate the observed covariates X_1, X_2 as random draws from uniform distribution $U[-3, 1]$, X_3 is drawn from $U[-1, 3]$ and lastly, X_4 from $U[-1, 2]$. The unobserved component ϵ follows a Type 1 extreme value distribution. Using the utility model, I generate the matrix of observed ranks. For an individual, I observe a partial ordering of ranks over schools from which the individual obtains a positive net benefit. An individual does not rank a school if the net benefit from that school is zero. The observed data for the econometrician consists of

Table 1: Simulation: Multiple covariates and without latent variable

n	J	β	MSE	MAE	Median bias	Mean bias
100	15	0.5	0.0030	0.0410	0.0071	0.0061
1000	15	0.5	0.0003	0.0129	0.0005	-0.0004
2000	15	0.5	0.0001	0.0085	0.0029	0.0011
100	15	0.8	0.0041	0.0502	0.0003	0.0017
1000	15	0.8	0.0004	0.0160	0.0025	0.0031
2000	15	0.8	0.0002	0.0121	0.0013	-0.0002
100	15	-0.9	0.0049	0.0569	0.0054	0.0051
1000	15	-0.9	0.0004	0.0168	-0.0033	0.0000
2000	15	-0.9	0.0002	0.0109	-0.0005	-0.0021
100	15	-0.8	0.0059	0.0611	-0.0123	-0.0058
1000	15	-0.8	0.0007	0.0217	-0.0061	-0.0068
2000	15	-0.8	0.0004	0.0152	-0.0002	-0.0007

Notes: This table illustrates the error distribution for the parameter estimates using the recursive model. The measures of error distribution are shown for four parameters. The number of schools, J, has been kept constant at 15, and n increases from 100 to 1000 to 2000.

$L, X1, X2, X3, X4$. The likelihood for the observed ranks can be obtained using the recursive solution provided in Theorem 1. The data is generated to match closely with the actual ranking data that I examine in section 4. Figure B1 shows the frequency distribution of ranks for one random simulation in this analysis (see (e) in figure B1). The figure shows that the frequency distribution tapers before the maximum ranked schools which is 15 in this case. The parameters to be estimated are $[0.5, 0.8, -0.9, -0.8]$. The sample size $n = [100, 1000, 2000]$ and the number of simulations are 100. I use multiple starting values for optimization where starting values are random draws from $U[-1, 1]$. Table 1 provides a summary of the simulations. The mean squared error as illustrated in column 4 in Table 1 drops as I increase the

sample size from 100 to 2000. Moreover, I also find better concentration of the estimates as the sample size increases, shown in panel (A) to (D) in figure B1.

Next I introduce latent heterogeneity to incorporate the correlated random effect. I discuss the performance of estimator for the parameters $\theta = \{\beta_1, \beta_2, \beta_3, \beta_4, \gamma\}$. I generate the student ranks over J in the following way. First, I generate the observed covariates X_{ij} and Z_i. $X_{ij}'s$ are generated from uniform distribution $U[-3, 3]$ and Z_i is drawn from a standard Gaussian distribution. Since the correlated effect c_i varies at the level of individual, the covariate Z_i is average characteristic over schools for an individual. The unobserved component ϵ_{ij} follows a Type 1 extreme value distribution and $w_i \sim N(0, 0.25)$. The latent variable c is generated as a linear function of observed component $Z\gamma$ and the unobserved random component w. Using the utility model as described in section 2.1, I obtain the matrix of observed ranks L. The observed data comprises of R, X, Z. I use the recursive solution derived above for density of

Table 2: Simulation: Multiple covariates with latent variable

n	J	β	MSE	MAE	Median bias	Mean bias
100	10	0.8	0.0062	0.0643	0.0014	0.0005
1000	10	0.8	0.0009	0.0237	-0.0138	-0.0118
100	10	0.4	0.0052	0.0530	0.0091	0.0066
1000	10	0.4	0.0003	0.0140	0.0018	0.0017
100	10	-0.8	0.0067	0.0651	0.0242	0.0143
1000	10	-0.8	0.0009	0.0233	0.0101	0.0131
100	10	-0.9	0.0079	0.0668	-0.0016	-0.0114
1000	10	-0.9	0.0006	0.0201	0.0070	0.0070
100	10	0.5	0.0127	0.0869	-0.0076	-0.0139
1000	10	0.5	0.0029	0.0331	-0.0113	-0.0115

Notes: This table illustrates the error distribution for the parameter estimates using the recursive model. The measures of error distribution are shown for four parameters. The number of schools, J, has been kept constant at 10, and n increases from 100 to 1000.

ranks conditioned on the vector of latent variable c. Since c is the latent component in the utility, I use EM algorithm for estimation as described in section 2. For the posterior distribution of the latent vector $h(c|L; \gamma, \sigma) = \frac{g(L|c;\beta)f_w(c;\gamma,\sigma^2)}{\int g(L|c;\beta)f_w(c;\gamma,\sigma^2)dc}$, I need to integrate over the latent vector c. I approximate for the integral using monte carlo simulations. I draw $T = 500$ random samples of w from $N(0, 0.25)$ and obtain the latent variable $c = Z\gamma + w$. The denominator $\int g(L|c;\beta)f_w(c;\gamma,\sigma^2)dc = \frac{1}{T}\sum_{t=1}^{T} g(L|c;\beta)$. Similarly, I use the average over Monte Carlo samples for the numerator in Q. The initial values of all parameters are random draws from a $U[-1, 1]$.

The maximization step is divided into two components. First, I solve for the parameters determining the distribution of the latent cost. Second, I obtain the MLE for β. Based on parameter estimates obtained in this step, I update it for the next iteration. I keep on iterating alternatively between the E and M step till the estimates converge.

Table 2 illustrates the simulation results for the parameters of this model. For all the parameters, I observe a decline in mean squared error as I increase the sample size from 100 to 1000. I also illustrate the concentration of the parameters as I increase the sample size in figure B2. The estimated values for

the parameters get closer to the population value as the number of individuals increase in the simulations.

Comparison with the urn model: Here, I compare the performance of the new preference model estimator with the existing estimators used in the literature to study rank data. I use the urn model for comparison. Figure B3 shows that the threshold rank order model estimates are closer to the actual parameter than the estimates obtained using the urn estimator for each of the four parameters in the model.[10] Table 3 provides a detailed comparison of the mean squared error as I increase the sample size for the two models. The mean squared error for every parameter and N is lower for the recursive estimator as compared to the urn model.

Table 3: Simulation: Comparison of urn model and recursive estimator

n	J	β	MSE (recursive)	MSE (urn)
100	15	0.5	0.0033	0.0052
500	15	0.5	0.0005	0.0009
1000	15	0.5	0.0002	0.0005
100	15	0.8	0.0043	0.0053
500	15	0.8	0.0007	0.0010
1000	15	0.8	0.0003	0.0006
100	15	-0.9	0.0053	0.0064
500	15	-0.9	0.0008	0.0011
1000	15	-0.9	0.0004	0.0006
100	15	-0.8	0.0061	0.0078
500	15	-0.8	0.0014	0.0017
1000	15	-0.8	0.0006	0.0009

Notes: This table illustrates the error distribution comparison for the underlying agent utility parameters between the recursive and the urn estimator. The comparison has been made keeping J fixed at 15, and n varies from 100 to 500 to 1000.

A critical feature of the recursive estimator is that it can calculate the probability of a school getting listed. I compute these probabilities for each student school pair in the simulated data using $\hat{\beta}$. I compare these predicted probabilities with the actual probability using the population parameters for every student school pair. An essential limitation of the urn model gets highlighted in this comparison. The urn estimator does not differentiate between ranked and non-ranked alternatives. Every student ranks every school. Consequently, any exercise that intends to compute a school's popularity using the sum of predicted probabilities across students will be possible using the recursive estimator but will always put a probability 1 under the urn model. Moreover, using the recursive model I can predict the expected ROL for each student. Such an exercise is not possible using the urn estimator as there is no cut-off to distinguish between the ranked and non-ranked alternatives.

I plot the error distribution between the actual and predicted total probability of ranking each school in panel (A) in figure 1. This can be interpreted as a measure of expected school popularity. The size of

---

[10]Observed covariates X_1, X_2 are random draws from $U[-3,1]$, X_3 is drawn from $U[-1,3]$ and lastly, X_4 from $U[-1,2]$.

Figure 1: Error distribution

A. School popularity   B. Expected ROL
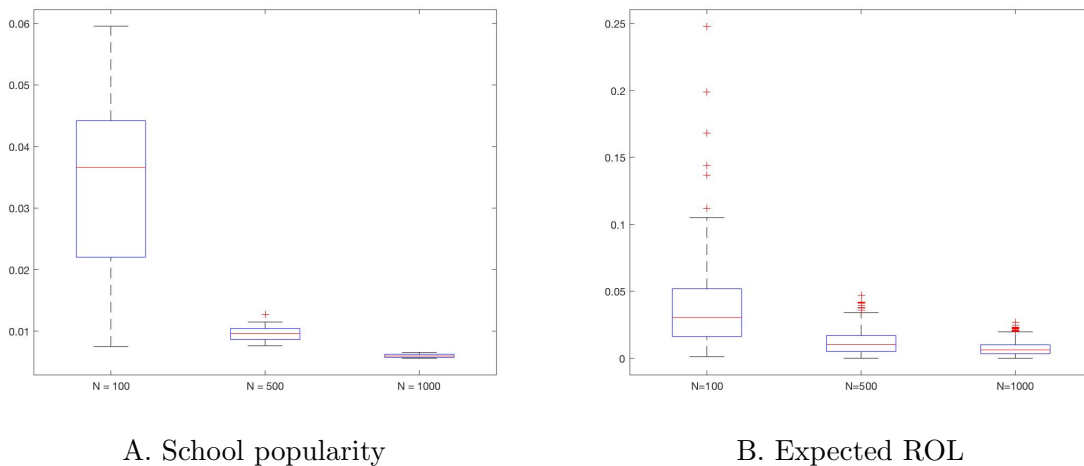
Notes: Panel (A) depicts the distribution of error in the total predicted probability of ranking schools in the simulated data by students. Panel (B) shows the error in the length of expected ROL. The error distribution diminishes as sample size increases from 100, 500 to 1000.

this error consistently shrinks as I increase the sample size. Panel (B) shows the error between the true and predicted expected ROL.

4 Application: Chile's Schooling System

4.1 Background

Chile has undertaken several significant education reforms. The first round of reforms happened in the early 1980s, followed by another reform in 1993. The last two rounds of reform happened in 2008 and 2015, respectively. Back in the 1980s, the government decided on the decentralization of primary and secondary education in Chile. Consequently, they transferred the public school system from the jurisdiction of the central government to local municipalities (school districts).[11] This transfer was complemented with the introduction of a school voucher system.

Due to this reform public, private voucher, and private non-voucher schools were created. The government financed and administered public schools. On the contrary, private voucher schools were managed privately but received government vouchers, and lastly, private non-voucher schools had no financial or administrative intervention from the government.

The second round of reforms happened in 1993, where the private voucher schools were allowed to charge partial tuition from students in return for some reduction in government vouchers. Following this, in 2008, the government introduced additional vouchers to schools if they enrolled students from poor socio-economic backgrounds (also known as priority students).

---

[11]School admissions in the municipalities (school districts) in Chile works differently than the United States as students are allowed to apply to schools outside their municipality of residence.

Although the government intended to make the education system more inclusive through these reforms, education research in Chile has documented that school segregation has been on the rise in the last couple of decades (Valenzuela, Bellei, & Ríos, 2014). Hsieh and Urquiola (2006) has shown that the voucher system resulted in the disproportionate flight of students belonging to the middle class to the private sector from public schools (10-15 year old for years 1982-1996).

Figure 2: Time-line of Schooling Reforms in Chile

The Congress in Chile introduced the Inclusion law in 2015 to address the ongoing concerns over school segregation and to promote inclusiveness in the schooling system. Under this law, the centralized system of school assignment was launched. The government transferred all the schools from the jurisdiction of the municipalities to that of the central government. Parents were required to apply for school admission through a common web portal, and admissions were no longer decentralized.

Chile has adopted the deferred acceptance (DA) algorithm (Abdulkadiroğlu & Sönmez, 2003; Gale & Shapley, 1962) for student assignment. The first region that underwent the reallocation under DA in 2016 is Magallanes. In 2017, the new system was implemented in four other regions, namely Tarapaca, Coquimbo, O'Higgins, and Los Lagos. In 2018, the new system was extended to the rest of the country except Metropolitana. Finally, in 2019, the Ministry extended it to Metropolitana, and therefore, students all over the country could participate in the new system.

4.2 Data

I compile data-sets from multiple sources for the empirical application. There are three critical inputs to the algorithm-student ranks over schools, student priorities, and school vacancy. All these variables are obtained from the DA files. Students who want to change school (>pre-K) or enter the public schooling system (pre-K) can list all their preferred schools (ranks) in the common application. If the student decides to continue in the existing school, there is no requirement to participate in the centralized system. The vacancy at each grade for a school is the difference between the capacity for that grade and the number of students who get promoted to that grade and decide to continue in the same school. In other words, the vacancy at ninth grade is computed as the difference between ninth grade capacity and the number

of previous year's eighth-graders who get promoted and decide not to change school. Lastly, the special priorities include applicants who belong to a lower socioeconomic status based on priority index, sibling studying in the same school, school officials' children, and previous alumni of the school[12].[13]

Although I observe DA for all grades, the highest participation is for two grades-pre-K and ninth grade. The primary reason for the highest student application in these two grades is that they are the entry points for the primary and secondary school in Chile, respectively.[14] I focus on the ninth grade cohort for my analysis due to the availability of background variables. I obtain the student background characteristics by matching unique student identifiers in DA files with the standardized test scores in Chile, also known as SIMCE. Such files can be obtained for students already in the education system (ninth-grade) and not for students entering primary education through DA (pre-K). Additionally, the Chilean education system has witnessed more pronounced SES segregation in high schools than primary schools (Torche, 2005; Valenzuela et al., 2014). This makes it compelling to study the schooling choices in the transition from middle school to high school.

I illustrate the participation for ninth-grade admission by region in Table B1. The number of regions vary by year, as DA was sequentially implemented. The first three columns summarize the number of high schools that participated in the new system. It is critical to note that only public and private voucher schools participated in DA. Private non-voucher schools did not participate in DA. Moreover, columns 2 and 3 display substantial variation in the distribution of participating schools by type. This difference emanates from variation in local schooling structure across regions in Chile. There are regions such as Tarapacá and Coquimbo, which had a much higher supply of private voucher schools relative to other regions such as Los Lagos and Aysén, which had an almost balanced availability of both public and private voucher schools. Overall, the fraction of private voucher schools is higher among the participating schools indicating a higher presence of such schools in most of the regions. Earlier work in Chile has shown that the former government schooling reforms aimed at the school voucher led to the proliferation of private voucher schools. Such heterogeneity in school supply can have important implications for the parental decisions on school ranks. The type of school is strongly correlated with school fees. School fee is likely a critical component in the school choice decision, particularly for low income households. Table B3 illustrates that less than 1% of public schools were charging any fee during the DA implementation. On the contrary, around 50% private voucher schools were charging fee. Columns 5 and 6 in Table B1 displays the number and the percent of total ninth graders who participated in DA by region. The participation of students vary between [32%,67%] with an average around 50%. [15]

---

[12]The last special priority for previous alumni of the school excludes students who were expelled from the school.

[13]In the existing literature on school choices in Chile such as Chumacero, Gómez, & Paredes, 2011; Gallego, Hernando, Flabbi, & Tartari, 2008, the researcher can mostly observe only the final school choices, and the student ranks over schools are unobserved. Here, however, I observe the ranking of schools in addition to the final allocation. In this regard, my work relates to Hastings et al. (2005); Hastings and Weinstein (2007), Ajayi (2013) and Fack et al. (2019) where the researchers observe parents' preferences. Estimating a parametric model on the preferences provides useful information on demand-side heterogeneity on school ranks, which is not possible to capture if researchers can merely observe the final allocation.

[14]It is mandatory for parents who are seeking admission for their children in public/voucher schools in pre-K to participate in DA. Students who intend to switch schools between primary and secondary must participate in DA unless they are seeking admission in private non-voucher schools.

[15]All the descriptive statistics corresponds to the information in the regular allocation files.

I distinguish between the ranked and non-ranked alternatives in my school choice framework. The Chilean data display specific characteristics that make this distinction critical. Some of those features are; i) substantial variation in the number of ranked alternatives and cut-off varies across individuals, ii) significant fraction lists three or fewer schools in ROL, iii) a sizable number of students do not end up in their top choice, iv) high disparity in the academic quality of guaranteed school and v) strong negative association between vacancies and school academic quality.

I display the distribution of total ranks for ninth-graders in panel (A) in figure 3. At least 54% of families listed three or fewer ranks in their application in 2016. The corresponding figures for 2017 and 2018 are 52% and 60%, respectively. This suggests that the cut-off where families stop listing schools can be extremely critical in determining their final school assignment through the centralized algorithm. Moreover, I illustrate in panel (B) in figure 3, the fraction of students who got allocated to their top choice. About two-thirds of students get allocated to their top choice. Nonetheless, one-third get allocated to their second, third, or latter choices. The key takeaway is that there is a possibility that students end up in lower-ranked schools on their list. This compels the need to study factors that determine both the ranking order as well as the cut-off of ranks.

Figure 3: Distribution of ranks: outcome variable for choice model
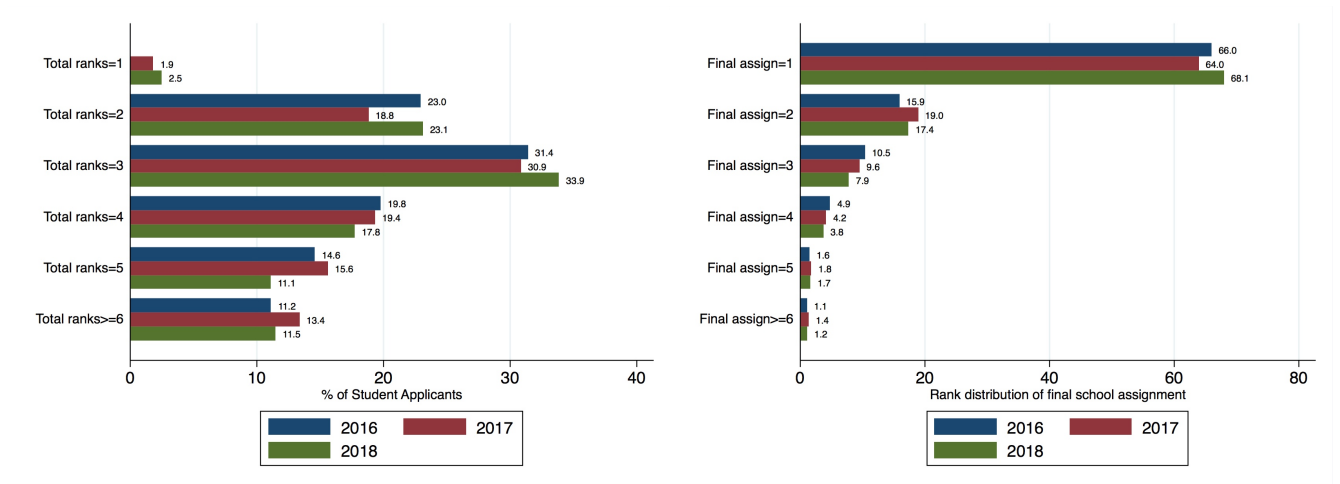
A. Rank dist. for ninth-graders   B. Rank of final assignment

Notes: The figure in panel (A) illustrates that there is significant variation in reported number of ranked schools in student applications. This figure uses the data for ninth-graders for 2016, 2017 and 2018. In panel (B), I illustrate the association between final assignment and the rank for the assigned school in student application. The sample for this analysis consist of participants who were allocated to one of the ranked schools through the algorithm. For students for whom the algorithm failed to assign a school in ROL, they were assigned the closest school based on their place of residence. Consequently, the sample for this analysis is slightly lower than the actual number of participants.

Student cut-offs in ROL are likely to vary by the quality of the guaranteed school. There is a positive probability associated with the DA algorithm in Chile of not allocating a student to any of the schools on ROL. The Ministry of Education in Chile provides detailed guidance on allocation in this scenario. There

are two possibilities: first, if the student's old school, the school in which the student is enrolled before DA reallocation, offers the grade to which the student seeks admission, then the student is guaranteed a seat there DA fails to allocate. Second, if the prior school does not offer the grade, the student is guaranteed admission to the nearest public school with vacancies. This rule creates a significant variation in the value of the outside option for the participating student. I plot the distribution of pre-DA test scores for the guaranteed school in figure 4. I observe substantial differences in outside value across students. Moreover, on an average high income students (Panel (A)) have a higher outside value than the low income students (Panel (B)).

Figure 4: Variation in outside value/guaranteed school quality by student income

A. High income   B. Low income
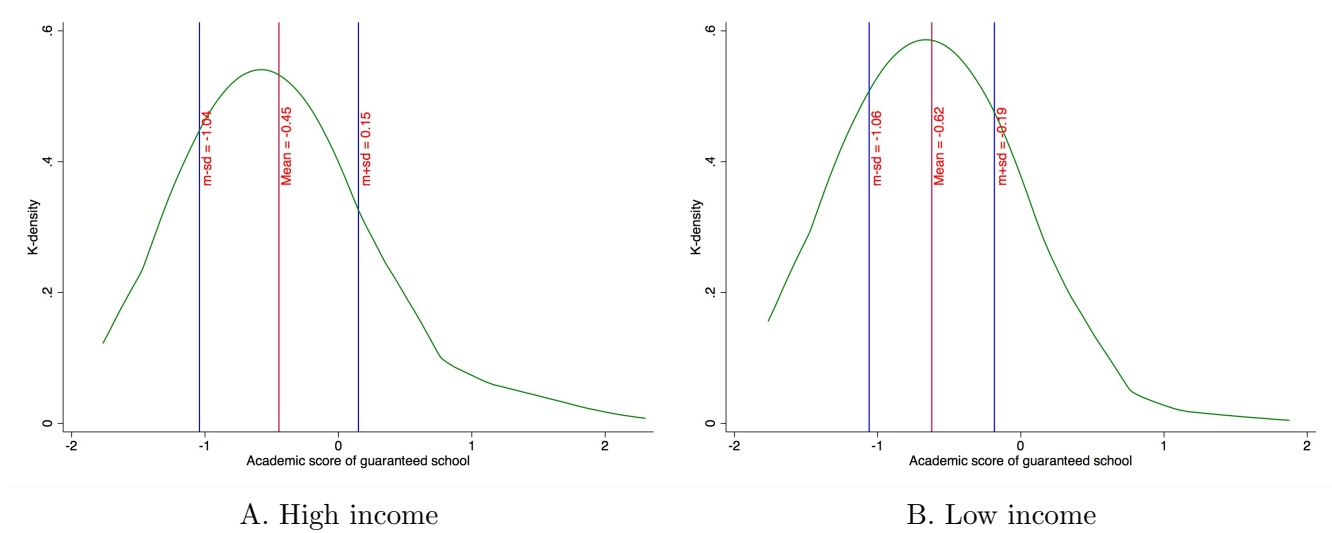
Notes: These graphs display the kernel density plots for academic quality of the guaranteed school. The academic quality has been obtained as an average of math and language SIMCE scores measured before DA. These test scores have been adjusted by the mean and standard deviation. The resulting test score distribution has $\mu = 0$ and $\sigma = 1$.

The second important component that can result in parents revealing a partial set of rankings is their expectation on the probability of acceptance at different schools. It is often seen that high academic quality schools are oversubscribed, and this behavior impacts the likelihood of admission. In DA, the likelihood of admission is never zero as admission at every iteration of the algorithm is tentative, and ties are resolved by lotteries. However, parents, even under DA, can modify their behavior if they expect the chances of admission to a high-quality school are low due to fewer vacancies.

Since I focus on the transition between the middle and high school, the vacancies are directly a function of the fraction of eighth-graders participating in DA to switch schools. Figure 5 illustrates that there is a strong negative correlation between vacancies and school academic quality. This is likely as students enrolled in good academic quality schools are less likely to switch schools in ninth-grade. Consequently, high academic quality schools post fewer openings for ninth-grade admissions.

Figure 5: Vacancies and pre DA test score

A. 2016

B. 2017
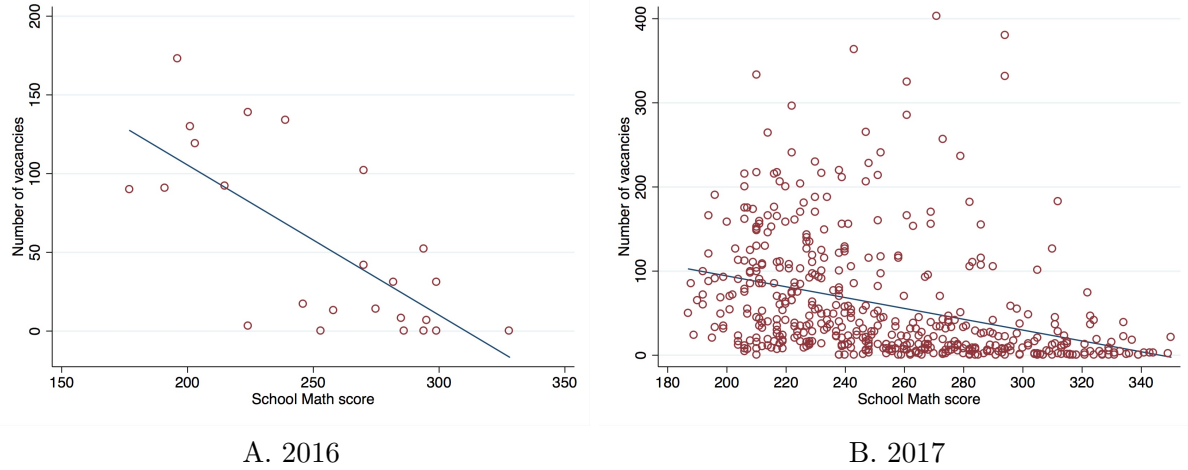
Notes: These graphs show the relationship between school quality and vacancies in DA. I study the transition from middle to high schools and some students can choose to continue in their pre-DA school if it offers high school grades. The expected probability of admission takes advantage of this relationship in the empirical specification.

I provide suggestive reduced-form evidence on the relationship between the length of ROL reported in DA and the value of the outside option and school vacancies. For this exercise, I use the eighth-grade cohort in 2017 who participated in DA for ninth grade admissions. I construct the measure of outside options using the pre-DA school test scores. Next, I account whether parents consider the expected probability of admission in their ranking process by computing the average vacancies in schools to which the student did not apply but was part of the student's choice set. Additionally, I account for various background characteristics of the student. Table 4 illustrates the results of this association. Once I condition on student characteristics, the value of the outside option is negatively associated with the length of ROL (Model (2)), which is in line with expectation. But I do not observe a significant association between length of ROL and vacancies at non-listed schools. This might be because the expected likelihood of admission might be less relevant for the Chilean parents on average. Nevertheless, I do account for the likelihood of admission in my school choice model as there might be parents at the margin accounting for such probabilities.

To determine the factors explaining the student rank order list (ROL), I closely follow the literature on school choice. I incorporate determinants for both benefits and costs associated with an application to a school. On the benefits side, parents care about the academic quality of the school. Black (1999); Hanushek, Kain, Rivkin, and Branch (2007); Hastings, Kane, and Staiger (2009); Hastings and Weinstein (2008); Reback (2008) show that school average test scores are important determinants of school choice. Moreover, it is often the case that value attached to academic quality varies by parental income. Parents with higher incomes tend to put a higher value on school test scores than parents with lower levels of income (Burgess et al., 2015). Since I focus on ninth-graders in this analysis, I use the tenth-grade average (school) math and language SIMCE test scores as primary measures of school quality (see Table

Table 4: Length of ROL, outside value and vacancies

VARIABLES	ROL length (1)	ROL length (2)
Vacancy to capacity in non-applied schools	0.416	0.612
	[1.014]	[1.015]
Outside option value	0.001	-0.046*
	[0.021]	[0.023]
Student pre-DA score		0.055**
		[0.022]
Mother's education		0.034***
		[0.008]
Income		0.070***
		[0.007]
Total school availability	0.011***	0.011***
	[0.003]	[0.003]
Constant	3.059***	2.352***
	[0.578]	[0.593]
Observations	15,125	10,558
R-squared	0.030	0.050

Notes: Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. This analysis uses the data on eighth grade cohort that participated in DA in 2017 in five regions-Tarapaca, Coquimbo, O'Higgins, Los Lagos and Magallanes.

B2 for details).[16] I observe significant variation in school test score distribution across all three years displayed in Table B2. I use the SIMCE data from 2015, 2016, and 2017 as the parents need to observe the tenth-grade test scores when making school choice decisions in 2016, 2017, and 2018 respectively, and the SIMCE tenth grade scores for the same year will not be reported at the time of application.

School fees can be a major barrier to private school enrollment (Alderman, Orazem, & Paterno, 2001; Glick & Sahn, 2006). As discussed in section 4 school fee structure is closely associated with the school type in Chile. Table B3 shows the fee structure for high schools in Chile. As of 2018, secondary education in most of the public schools was free. On the contrary, 47% private voucher schools charge an add on fee to parents. This is a critical difference between the two types of schools that participated in DA (public and private voucher).

School choice literature such as Hastings and Weinstein (2008), Gallego and Hernando (2010) illustrate school proximity is a key determinant of parental preferences. For computing the travel time to school, I use precise student residential and school addresses, provided by the Ministry of Education. I use open street maps API to calculate commuting time to schools. Travel time or distance using actual road network provides a much more accurate measure than any geodetic or straight-line measures used in related literature for similar analysis (Chumacero et al., 2011; Frenette, 2006; Laverde, 2020).

Figure 6 displays the commuting route and travel distance using a car for two ninth graders who applied in DA in Magallanes in 2016. The travel distance for Student A shown in panel (A) is 2.4 km to the preferred school. The geodetic distance computes this as 1.95 km. Similarly, travel distance for student 2 is 8.3 km (Panel (B)). However, the corresponding geodetic measure is 6.3 km. Next, I also illustrated the kernel density plots for all the student school rankings of ninth graders in Magallanes

---

[16]Unique school identifier in the DA files can be matched with SIMCE files to obtain the school academic quality variable.

in 2016 (Panel (C)). I observe significant variation in the two densities. Moreover, geodetic distances consistently underestimate the distance to school.

Figure 6: Travel time and route to ranked school

| A. Student A | B. Student B | C. Travel vs. geodetic |
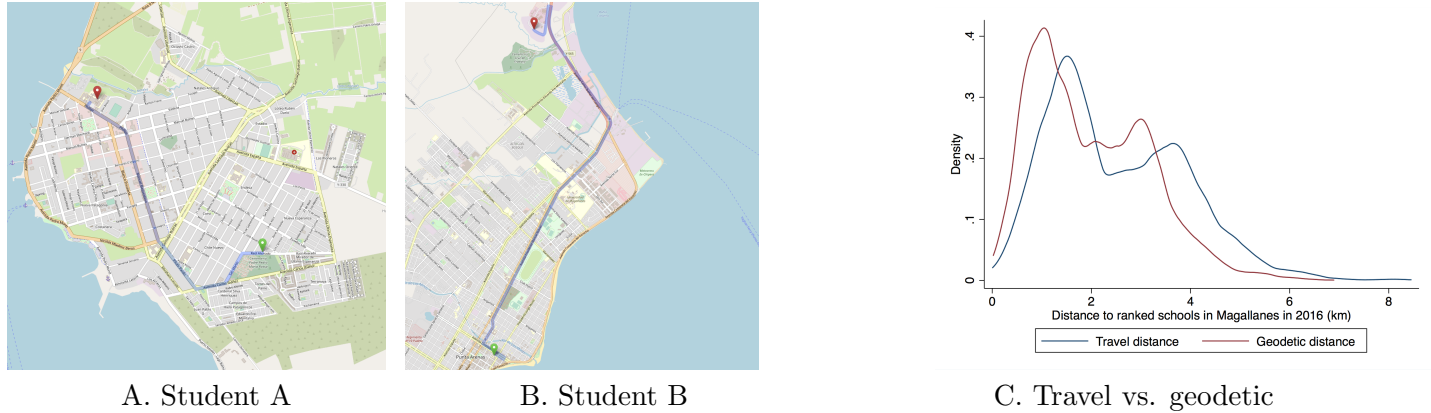
Notes: The figure in panel (A) and (B) illustrate the travel distance by car computed using the actual road network between the student's residence and school. Panel (C) depicts the kernel density plots of travel and geodetic distance to the schools listed in ROL.

Beyond the above variables, there are several other determinants of applying to a school. Fack et al. (2019) and Hastings et al. (2005) suggest that parents care about the socio-economic make up of a school. Particularly, parents seek schools that have students coming from a similar socio-economic background. Besides, parents might also favor schools where the academic standards match the student's academic ability (Fuller, Manski, & Wise, 1982; Light & Strayer, 2000). I do control for ability and SES match in my school choice model.

Lastly, parents' decision to stop ranking schools critically hinges on three critical variables in the Chilean context. First, if students do not get allocated to any of the schools listed in their ROL, they are guaranteed a seat in their old school, conditional it offers ninth-grade. Else, they are allocated to the nearest public school with a vacancy. Second, there could be heterogeneity in the psychological cost of listing additional schools, which can be strongly correlated with the extent of parent sophistication. Lastly, parents might modify their behavior based on the expected likelihood of admission, and school academic quality variables can account for these differences. I account for such components in my school choice model.

The identification argument for partial lists illustrated in Kuersteiner et al. (2020) requires two critical components. First, empirically one needs to illustrate the existence of full support. In other words, there should exist variations in the placement of all types of students across different school types. This geographic variation will generate variation in the outside value or the quality of the guaranteed school, which will help to pin down the parameters of parental preferences in the presence of partial lists. I explain the argument using a simple example. To keep the illustration simple, I assume two students types low income and high income. I also assume two school types high and low ability. Figure 7 shows the type of variation necessary for identification in case of partial lists.

Figure 7: Geographic variation in placement of school types and student types

A. Identification fails   B. Identification possible
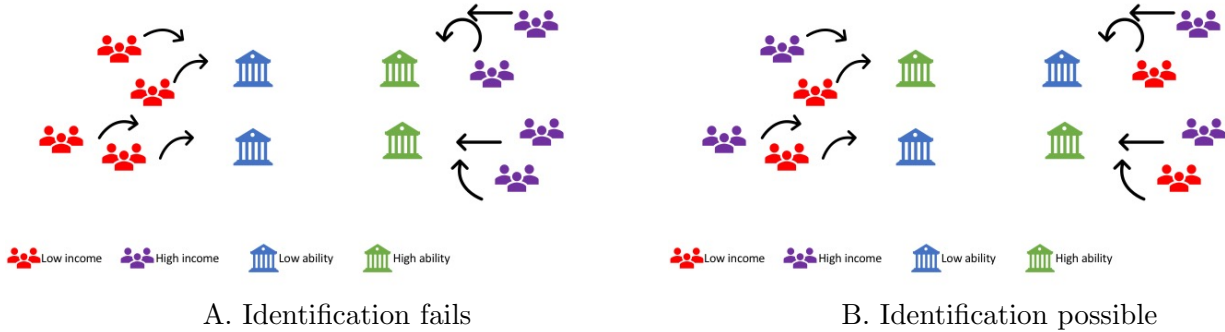
Notes: Panel (A) and (B) illustrate the variation required in the data to satisfy the assumption of full support. Students of the same type should be placed geographically around every school type for the required variation in the value of guaranteed school (outside value).

Assume that parents care about school quality and distance while listing schools. In panel A, I observe that low-income students are all geographically clustered around low ability schools, and consequently, these parents list only the low-income schools. On the contrary, since the high-income students are clustered around the high ability school, only the high ability school features on their ROL. One of the primary reasons driving this partial listing for the low-income group could be that the outside option (closest school) is higher than the utility provided by listing a faraway high ability school. For the high-income students, there is less reason to list a low ability school further away from residence. In this context, it is hard to pin down determinants of parental ranks for the complete set of schools, so identification fails.

In panel B, I reshuffle this set-up, and now there is variation in the placement of different student types around high and low ability schools. Under this set-up, even with partial ranking, I observe the lists of low-income parents for both high and low-income schools due to differences in placement and variation in the quality of the guaranteed school. Using this variation and the partial ROL, I can back out the complete ordering of parental preference parameters in a school choice set-up.

It is important to identify the source of the above variation in data. For this, I show the geospatial makeup of student and school types in the regions that participated in DA in 2017. Figure 8 provides suggestive evidence on full support. I observe a significant mix of low and high-income students around every school type in each of the regions. This variation is critical for identification in DA with partial lists. The dimension of student and school types are kept at two for the simplicity of illustration. However, I expand the dimension of student types and condition on both income and ability later in the paper.

Figure 8: Variation in spatial location of school type and student type

A. Tarapaca

B. Coquimbo

C. O'Higgins

D. Los Lagos

E. Magallanes
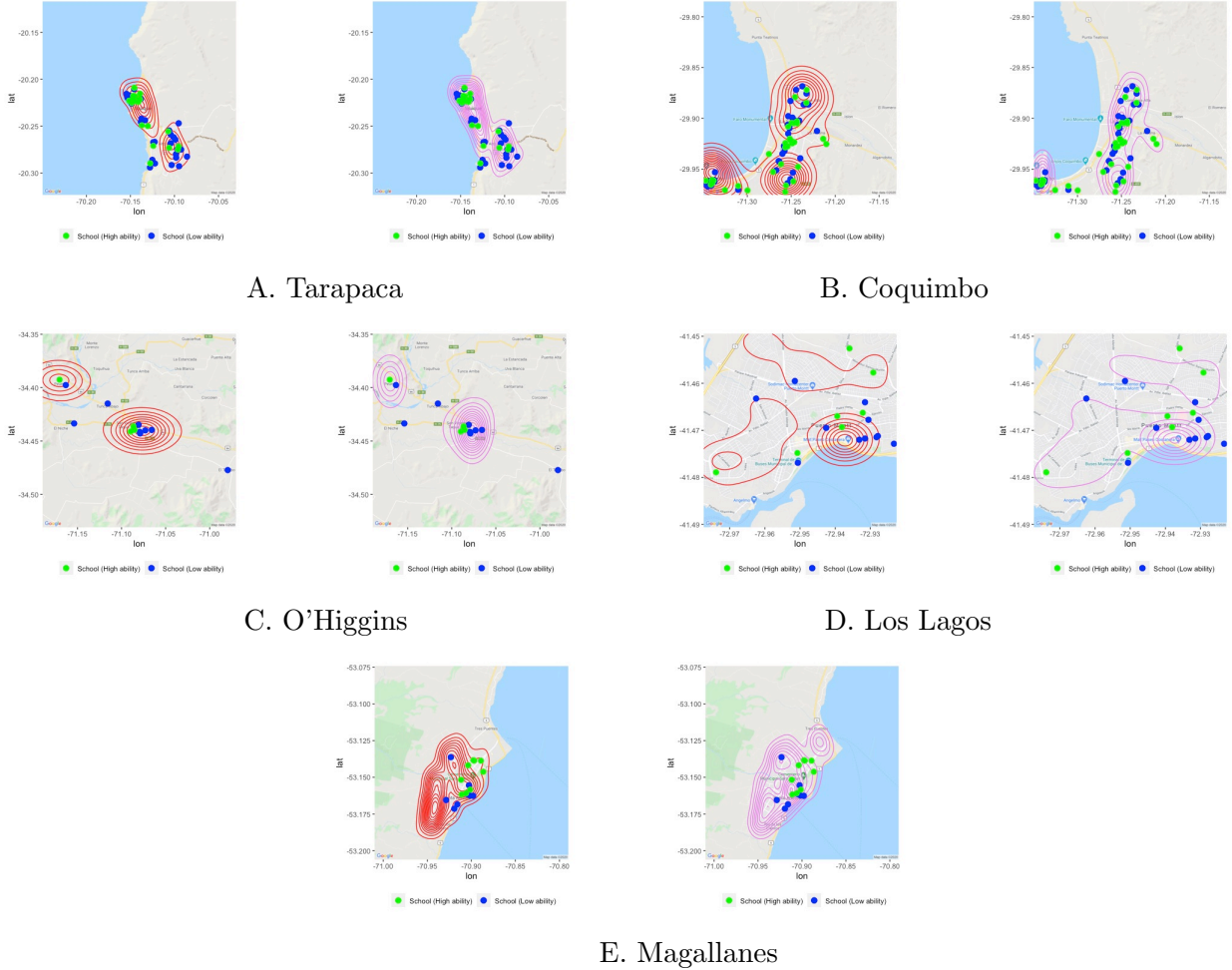
Notes: These graphs display the spatial density plots for high and low income students around high ability and low ability schools respectively. The plots with red density contours correspond to the low income students and the plots with violet density contours correspond to high income students.

4.3   Reduced form evidence

The government's key motivation behind the introduction of centralized assignment in Chile was to reduce the existing segregation levels based on socio-economic status (SES). The voucher system introduced in the 1990s and later modified in 2008 had led to an out-migration of middle-income category students to voucher schools. This resulted in overcrowding of low-income students in the free public schools that raised segregation in schools.

School segregation poses additional constraints if low SES students are under-represented in high quality schools. In figure B4, I divide the schools into two types, high and low academic quality and display the composition of student type across these schools. I do this analysis for the ninth grade cohort just before and after the introduction of DA for student assignment.

I divide the total enrolment in high and low quality schools by student income.[17] Figure B4 displays that low income students constitute a much lower fraction of total enrollment in high-quality schools. Such under-representation by school quality can have ramifications on student academic performance and the frequency of absenteeism suspension and drop out rates (Figlio, Karbownik, Roth, Wasserman, et al., 2016; Hanushek et al., 2008). In particular, lower school absenteeism is a precursor to achieving better student outcomes in the short and long term(Bergman & Chan, 2019; Gottfried & Kirksey, 2017; Jackson, 2018; Liu, Lee, & Gershenson, 2019). Moreover, it is often students at the margin who show up at the left tail of the attendance distribution. Given the existing segregation levels in the Chilean context, it might be relevant from the policy perspective to understand how school quality impacts attendance for those at the lower end of the distribution.

I display the differences in absenteeism by schools in the Chilean context. Panel A in Figure B5 shows the distribution of attendance for the ninth graders in 2018. I observe overall the distribution for low income students is shifted marginally to the left of the distribution for high income students. However, for both groups there are students at the left tail of the distribution. I intend to capture to what extent school quality impacts attendance. Panel B displays the coefficients on school dummies for the following reduced for model

$$\text{Attendance}_{is} = \sum_{s=1}^{M} \text{School dummy}_s + X_i\beta + \epsilon_{is}$$

The dependent variable in the above specification is the student $i's$ annual attendance rate in school s in 2018. I regress this on school dummies and a set of observed student-level characteristics. This suggests that there are differences in attendance rates by schools in Chile. Lastly, in panel C, I repeat the above analysis but using the two school types, high and low academic quality. The results of this estimation indicate there is a positive correlation between attendance and school quality.

5   Main Findings

I provide the main findings on the determinants of the underlying parental preferences in Chilean DA for the year 2016 and 2017. Chilean DA was implemented first in Magallanes in 2016 but it was expanded to Tarapaca, Coquimbo, O'Higgins and Los Lagos in 2017. I use province as the definition of schooling market in my analysis. Preliminary examination for the 2017 participants suggests that 91.6% of the students apply to schools within their province of residence. The choice set for student i residing in province p consists of the schools present in p^{th} provinces.[18]

---

[17]The percentage of low and high income students should add to 100 for each school type in figure B4.

[18]The students who applied for ninth-grade admission in 2017 switched to new schools in 2018. Similarly, the students who applied for admission in 2016, they were admitted to the new schools in 2017.

5.1   Results: 2016

In this section, I use the threshold rank model to estimate the determinants of student ROL in 2016. In 2016, the government introduced DA in Magallanes. This analysis focuses on the ninth-grade applications. Panel (A) in Table B4 illustrates the key characteristics of students who participated in the new system in 2016. I observe that the average number of listed schools in ROL is 4.1 with a standard deviation of 1.6. Additionally, I observe significant heterogeneity in their academic ability and background characteristics.

Table 5: School choice for ninth graders in Magallanes in 2016 (Parameter estimates)

Outcome variable:	Rank ordered list	
Variables	(1)	(2)
Travel distance (km)	-0.691*	-0.493*
	[-1.030,-0.250]	[-1.116,-0.045]
Fee dummy	-0.124*	-0.516*
	[-0.945,-0.036]	[-0.803,-0.127]
Student income	0.395	0.487
	[-0.047,0.789]	[-0.152,0.812]
Student score	0.358	0.335
	[-0.030,1.030]	[-0.038,0.757]
School score	0.589*	0.167
	[0.002,0.671]	[-0.087,0.650]
School score × Student score		0.418*
		[0.001,0.679]
Travel distance ×Student Income		0.043
		[-0.380,0.836]
Constant	0.498	-0.021
	[-0.435,0.597]	[-0.462 ,0.610]
Unobserved cost	✗	✗
J(Schools)	17	17
N(Students)	499	499

Notes: The 90% bootstrap confidence intervals are provided in the square brackets. The sample for these specifications include the eighth grade students who applied for ninth-grade admission in 2016.

In panel (B) in Table B4, I display the summary statistics for the top-choice school for the participants. As discussed in related work on school choice, parents do seem to have a strong preference for proximity. The average commuting distance by car to the top choice school is around $\sim$ 2.4 km, with a standard deviation of 1.5 km. As for the academic quality, the average math and language test score for the top choice school is marginally below the average for all public and voucher schools in Magallanes in 2016 (252 vs. 254: math and 242 vs. 243: language). Moreover, this average is significantly below the school

with the highest academic record in Magallanes (328: math and 299: language). This suggests that not all parents are necessarily aiming for the best academic school in their application.

Moreover, Figure B6 illustrates the spatial location of students and schools used for this analysis. Panel (A) depicts actual school assignment for ninth-grade admission for low-ability students in 2017. Panel (B) replicates the same information for high ability students. A comparison of the two graphs suggests that the enrollment of low ability students is lower in some of the high test score schools, particularly those schools in the northeast direction. I explain the determinants for such systemic variations in the actual assignment using the threshold rank order model.

I begin with the Threshold Rank Order Model with travel distance to school by car, student income, student test score, and the school average test score. The results for this model are displayed in column 1 of Table 5. Since it is a non-linear model, I need to evaluate the marginal effects for the impact of each of the covariates.

First, I compute the sensitivity to distance (model (1) in Table 5). The other covariates, such as school test scores, student test scores, and student income for this analysis, have been fixed at their average values. The school fee dummy is set at one suggesting that this school charges a fee. I compute the confidence intervals using the bootstrap samples used in Table 5. I observe a consistent drop in the likelihood of applying to this average school as the student gets further away. Second, I explore differences in average sensitivity to distance for high and low ability students. Panel (B) in Figure 9 suggests although high ability student has a higher probability of applying to the average school, the sensitivity to distance might vary by student ability.

Figure 9: Likelihood of applying to the average school

A. Travel distance by car

B. School ability and distance
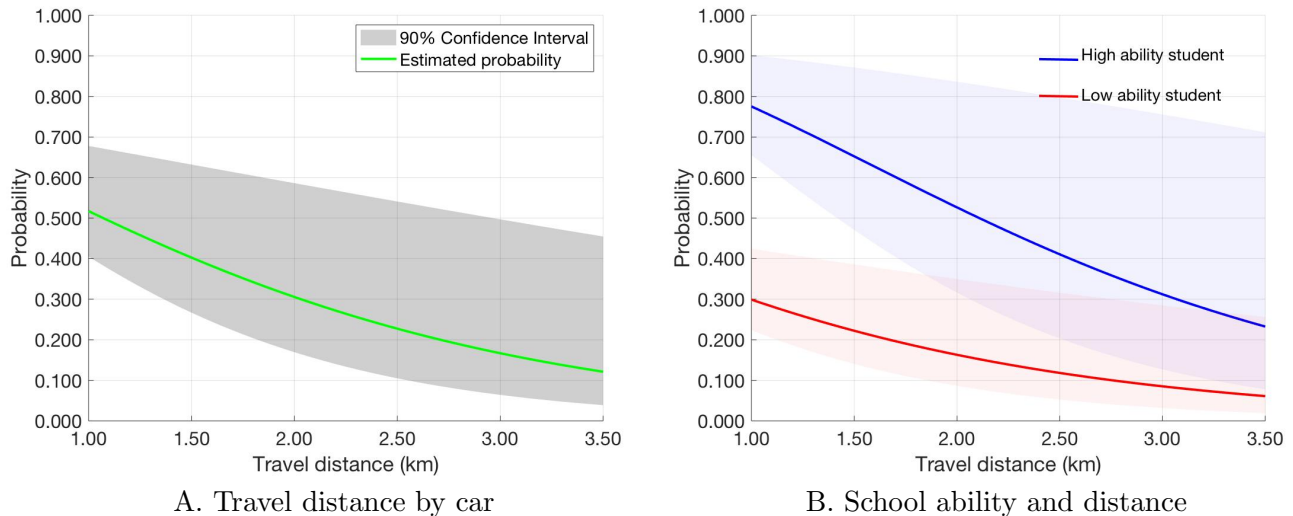
I expand the set of covariates to incorporate additional functional forms for travel distance and interactions between student test scores and school test scores (see Model (2) in 5). I also do a model selection test. The AIC for model (1) is 16653 with smaller set of covariates as compared to 14618 for

the model with additional covariates for additional functional forms for distance and the match between student and school ability. It suggests that the model fit is higher for the specification with additional covariates relative to model 1.

Figure 10: Likelihood of applying to the top school by student ability

| A. Low ability | B. High ability | C. Low income | D. High income |
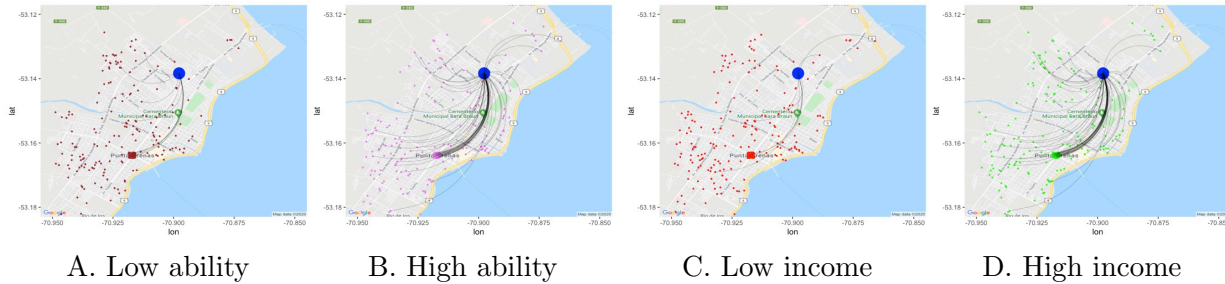
I examine whether parents prefer schools that are a better match for student ability. Such assessments are possible in model 2, where I allow for interaction between student test scores and school test scores. I display the students who have a high probability (≥ 0.5) of applying to the top school by ability (Figure 10). If the probability of application is high, it is indicated using the black curve. The absence of a curve joining the student and school suggests that the student had a low likelihood of applying to the top school by the model predictions. Panel (A) and (B) in Figure 10 show the presence of a higher density of curves for high ability as compared to low ability students. This illustrates that parents do have a preference to apply to schools closer to the student's ability.

I observe in all columns of Table 5 that distance is consistently negative. This supports the hypothesis that parents have a preference for school proximity. However, there might exist heterogeneity in this preference across income groups. Such differences will be captured by the interactions between travel distance and student income in the rank order model.

Panel (C) in Figure 10 shows the probability of applying to the best school. I use a curve to join the student location and school location if the probability of applying to the school is larger than 0.5. The density of such curves is much higher for high income students as compared to low income students.

Next, I allow for unobserved heterogeneity in the above specification. Individual specific school invariant characteristics such as student test score as well as background characteristics parental income and education can be allowed to be correlated with unobserved cost. I display these estimates in Table 6. Column (1) illustrates the results with a restricted set of covariates. Additional functional forms for travel distance and academic ability are introduced in column (2).

Table 6: School choice for ninth graders in Magallanes in 2016, allowing unobserved cost

Outcome variable:	Rank ordered list	
Variables	(1)	(2)
Travel distance (km)	-0.444*	-0.088*
	[-0.918,-0.038]	[-1.118,-0.030]
Fee dummy	-0.797*	-1.301
	[-1.945,-0.225]	[-1.538,0.306]
Student income	0.078	-0.137
	[-0.008,0.301]	[-0.646,1.436]
Student score	0.203	0.008
	[-0.060,0.280]	[-0.196,0.124]
School score	0.131	0.318
	[-0.113,0.564]	[-0.047,0.605]
School score × Student score		0.225*
		[0.024,0.670]
Travel distance ×Student Income		0.052
		[-1.151,0.275]
Constant	-0.423	-0.519
	[-0.916,0.101]	[-1.378,1.036]
Unobserved cost	✓	✓
J(Schools)	17	17
N(Students)	499	499

Notes: The 90% bootstrap confidence intervals are provided in the square brackets. The sample for these specifications include the eighth grade students who applied for ninth-grade admission in 2016.

The coefficient on travel distance is negative across the two specifications, in line with the previous estimates (Baseline model without unobserved cost). Since the school choice model is highly non-linear, I cannot interpret the parameter estimates directly. Therefore, I display the marginal effects of travel distance on ROL in Figure 11. Here, I display the probability of ranking a school with an academic rigor close to the median for the set of schools present in the choice set. I plot the travel time isochrones using the OSRM API. The x minute isochrone connects all the geo-coordinates around this school, which can be reached in x minutes using a car. I want to highlight, though, that the isochrones are likely a lower limit on the actual commuting time as not all students have access to a car to commute to school. I observe that more the travel time increases, the likelihood of ranking this median school decreases.

Figure 11: Marginal effect of travel distance on ranking the median school
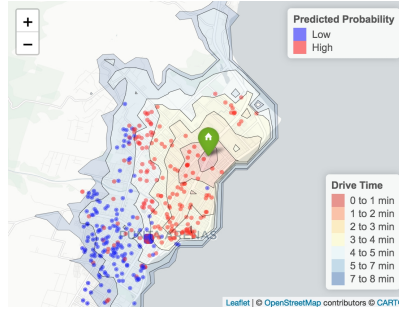
Notes: The probability of ranking this school with the median level of academic rigor is high (low) if the predicted probability is greater (lower) than the 75th percentile.

In addition to the median school, I show the predicted probability of ranking the best school and how it varies with the student's income. Figure 12 displays that the likelihood of applying to the best school in this region is low for low-income students. The fraction of students who have a high likelihood of applying to this school is significantly higher for high-income students. Additionally, high-income students are less responsive to travel time to this school.

Figure 12: Predicted probability of applying to top school by income
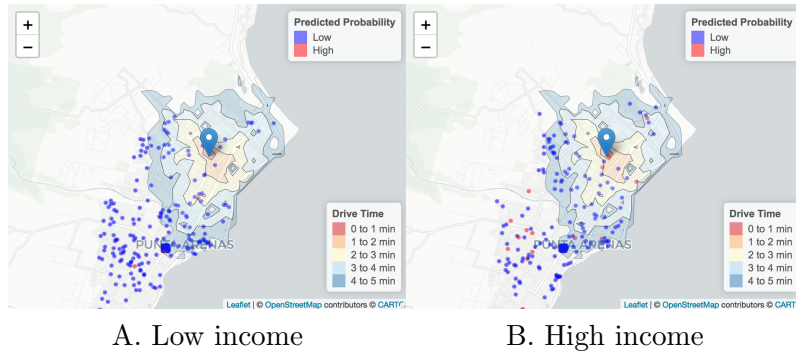
A. Low income   B. High income

Notes: The sample consists of students who participated in DA for ninth grade admissions in 2016 in Magallanes and therefore started ninth grade in the allocated school in 2017.

Lastly, I examine the extent to which the match between student and school ability impacts school choice. I calculate the predicted ROL and examine the intersection between the predicted ROL and the set of best schools in this region. I divide my sample into two groups based on ability (High>Mean ability, Low<Mean ability). Figure B8 shows the extent of overlaps for two samples between the predicted top five ROL and the set of best schools. High student ability increases the likelihood of listing the best schools high up in the ROL. In fact, the distribution is completely shifted to the right for high-ability students compared to the low ability students.

The school choice model for 2016 in Magallanes provides some meaningful predictions. However, since the new system was implemented only in one region in 2016, there are limitations in terms of the data variation that can be exploited to include the full set of covariates. I expand the covariates set for the

29

school choice model in 2017 as then DA was implemented in five regions.

5.2   Results: 2017

For this analysis, I am using the data on participants in 2017. The new system was implemented in 5 regions in 2017. Preliminary analysis for the 2017 participants suggests that 91.6% of the students apply to schools within their residence province. I use province as a school market, and the choice set for student i residing in province p consists of the schools present in p^{th} provinces. The students who applied for ninth-grade admission in 2017 switched to new schools in 2018. Table B5 provides summary statistics for the eighth grade participants in 2017.

The set of covariates in this model comprises of travel distance to school, school's academic score measured in 2017 (pre-DA), a dummy for fee as well as the socio-economic composition of the school, student pre-DA scores, student income, interactions between student income and school SES, student score and school scores as well as the interactions between distance and student score, distance and student income and triple interactions between student score, distance and student income. I account for the guaranteed school's academic score, and the school test score is used as a proxy variable for vacancies. Since a tiny fraction of students is eligible for a priority in the lotteries due to factors such as siblings enrolled in the same school, parents working in the same school or alumni, I include a priority indicator in the school choice model.

In table B6, I provide the parameter estimates of the school choice model for the regions that got reallocated in 2017. The covariates set across the provinces in the five regions remain the same except the Coquimbo school fee variable. I do not observe any school fee variation for the schools that participated in DA in 2017 in Choapa province in Coquimbo with all the required information.

In order to unmask the impact of different covariates on the predicted probability of either listing the school or ranking it higher up in the ROL, I provide the marginal effects. First, I illustrate the response to travel distance to the mean school in the schooling market. I plot the predicted probability conditional on the average value of other covariates except for student income. I calculate the predictive margins for all the regions and display distribution of these marginal effects.

Table 7: School choice for ninth graders in 2017 (Parameter estimates)

Outcome variable:	Rank ordered list				
Region	Tarapaca	Coquimbo	O'Higgins	Los Lagos	Magallanes
Province	Iquique	Choapa	Colchagua	Chiloe	Magallanes
Variables	(1)	(2)	(3)	(4)	(5)
Travel distance	-0.804*	-0.930*	-1.085*	-0.779*	-1.073*
	[-1.156,-0.509]	[-2.331,-0.702]	[-1.423,-0.678]	[-1.351,-0.470]	[-1.139,-0.558]
Students pre-DA score	0.306*	0.775*	-0.051	0.350*	0.261*
	[0.009,0.508]	[0.234,0.900]	[-0.234,0.034]	[0.241,0.473]	[0.019,0.318]
Student income	-0.217	-0.004	0.060	-1.776*	-0.102*
	[-0.479,0.005]	[-0.372,0.172]	[-0.053,0.108]	[-1.921,-1.664]	[-2.846,-0.070]
School's pre-DA score	0.408*	0.107	0.015	-0.136*	0.222
	[0.188,0.647]	[-0.292,0.292]	[-0.123,0.051]	[-0.263,-0.033]	[-0.039,0.278]
School SES	0.159	0.657*	0.021	-0.230*	-0.146
	[-0.109,0.387]	[0.397,0.794]	[-0.199,0.158]	[-0.353,-0.117]	[-0.645,0.327]
School fee	-1.024*		-0.029	-0.558*	-0.917*
	[-1.206,-0.870]		[-0.139,0.088]	[-0.685,-0.458]	[-1.005,-0.626]
Student score×School score	0.094	0.798*	0.028	0.421*	0.386*
	[-0.142,0.263]	[0.298,0.996]	[-0.098,0.154]	[0.284,0.558]	[0.277,0.486]
School SES×Student income	-0.350*	0.319*	0.146	0.885*	0.587*
	[-0.530,-0.185]	[0.121,0.487]	[-0.032,0.246]	[0.748,1.025]	[0.336,4.459]
Distance×Student score	-0.075	0.247	-0.854*	-0.285*	0.085
	[-0.329,0.181]	[-0.531,0.301]	[-1.264,-0.431]	[-0.606,-0.020]	[-0.273,0.165]
Distance×Student income	-0.743*	0.236	-0.281*	-0.229*	-0.215
	[-1.008,-0.526]	[-0.444,0.328]	[-0.570,-0.158]	[-0.438,-0.081]	[-0.449,0.012]
Student score×Student income	0.217	0.034	0.419*	0.058	-0.376*
	[-0.039,0.409]	[-0.206,0.242]	[0.268,0.526]	[-0.083,0.209]	[-0.433,-0.078]
Distance×Student score×Student income	0.801*	1.089*	0.504*	0.373*	0.449*
	[0.567,0.969]	[0.062,1.197]	[0.204,0.724]	[0.188,0.550]	[0.088,0.621]
Outside option value	-0.316*	-0.271	-0.230*	-0.738*	-0.816*
	[-0.667,0.149]	[-0.373,0.229]	[-0.354,-0.059]	[-0.864,-0.598]	[-0.659,-0.215]
Constant	-1.462*	-0.629*	-0.314*	0.373*	-1.178*
	[-1.857,-1.143]	[-0.927,-0.474]	[-0.496,-0.195]	[0.259,0.522]	[-1.400,-0.570]
Priority dummy	0.052	-0.030	-0.096	0.859*	2.021*
	[-0.136,0.254]	[-0.214,0.895]	[-0.213,0.008]	[0.763,1.003]	[1.219,2.253]

In panel A in figure 13, I display the responsiveness to the distance to average school for high-income students. There is a decline in the predicted probability of application by increasing distance to mean score school. However, compared to low-income students displayed in panel B, high-income students have lower likelihood of applying to the mean school than low-income students. In other words, probability of application drops more for high income students than low income students. In the Chilean context, there is variation in the value of the outside option. Student's whose pre-DA school offers ninth grade have a guaranteed seat if the algorithm fails to allocate the student to one of the listed schools in the ROL. If the pre-DA school does not offer ninth grade, then the student can be allocated to the nearest public school with vacancies. This nearest school is required to be within 17 km to the student's residence. Therefore, for students whose pre-DA school does not offer ninth grade, I extract set of public schools within 17 km using the travel distance and take the average school pre-DA test score. Figure B10 shows that the predicted probability of application to the mean school drops with an increase in the value of outside option.

Figure 13: Predictive probability: Distance to mean score school

A. Student high income   B. Student low income
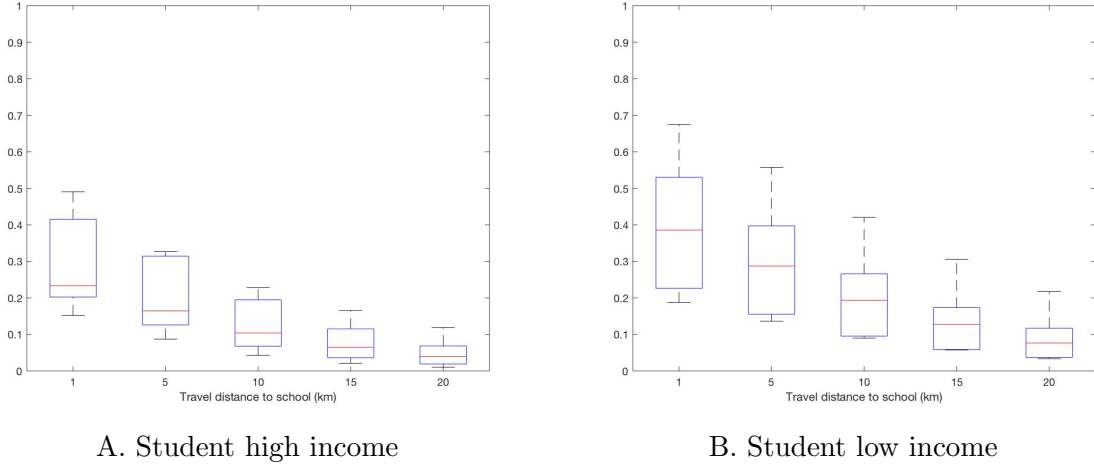
I illustrated in section 4.3 that low-income students are under-represented in high score schools in Chile. Reducing segregation through DA was high up in the government's agenda. However, a better representation will depend on whether parents from low-income families end up listing the good schools in ROL. Parents are likely to care about the ability match between the student's ability and school academic rigor. Consequently, it is interesting to disentangle the effect of travel distance conditioned on both income and ability. Mainly, I next illustrate that student ability is compensating for income and resource constraints in Chile. I condition on low-income students and illustrate the probability of listing a high score school. Figures in panels A and B in B10 showcase that high-ability low-income students are more likely to list a school with high test scores than low-ability low-income students. In panel C, I keep the distance constant at 1 km for low income category and calculate the probability of application to high score school by varying student ability. As student ability increases the likelihood of applying to good schools increase.

These charts display the predicted probability of application to a high score school for low income students by varying student ability and travel distance. The other covariates are kept constant at the average values. In figure 14, I plot the unconditional marginal effect of application to the best school in Coquimbo region by student income and ability. I explore the angle between ability and school quality further by computing a rich array of marginal effects. I do this by calculating the likelihood of ranking a high score school above a low score school for low-income students (Figure B11). The figure illustrates that the predicted probability of observing such a ranking behavior is higher for high-ability students than low-ability students.

Lastly, I estimate the school choice model for all the areas together instead of estimating separately for each area. The summary statistics for this analysis is provided in B6. I replicate the key predictions derived above using this joint analysis. Table 8 displays the predicted probability of ranking the mean school, low score and high score school by travel distance.

Figure 14: Marginal Effects: Distance, student score and income
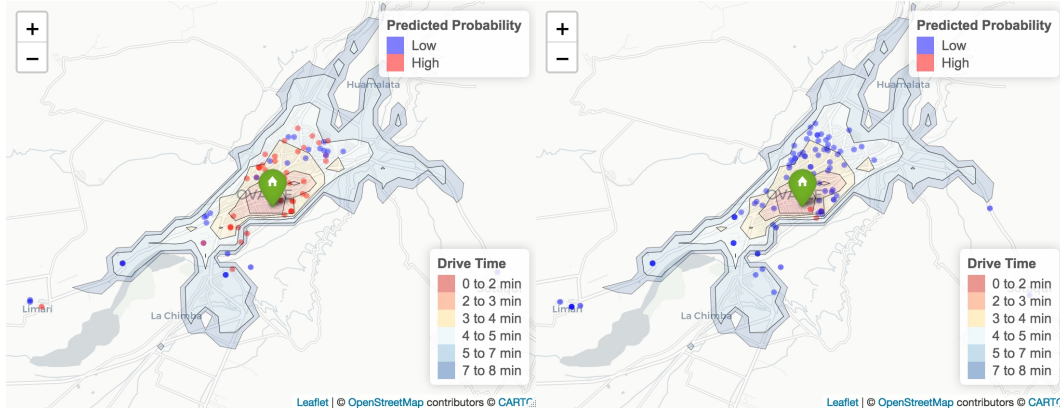
A. Low score high income

B. Low score low income
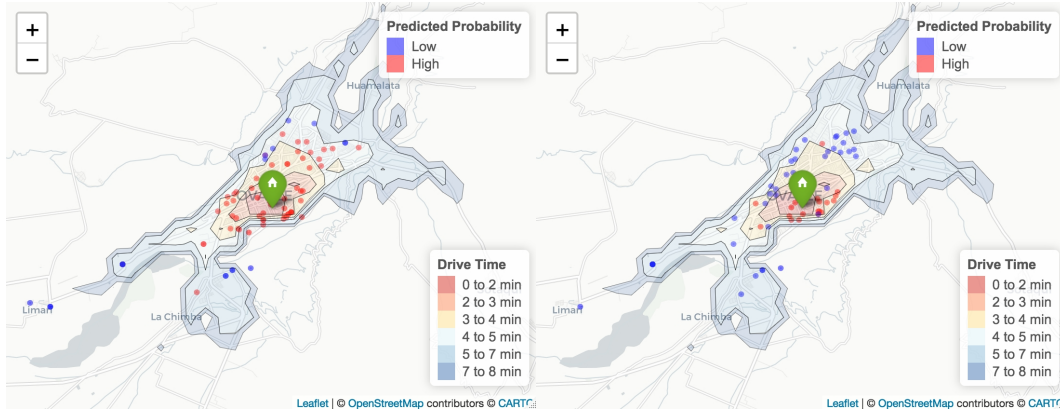
C. High score high income

D. High score low income

Notes: These graphs display the travel time isocrones from the best school (green marker) in the Coquimbo region in Chile. A x minute isocrone to the best school connects all points that are x minutes away (travel time by car) from this school. The red (blue) marker indicates a high (low) predicted probability of application to high score school. It displays conditioning on student score and income matters. Student score can compensate for travel costs for low income households.

Table 8: Predictive Probability: Joint Analysis

Outcome Variable	Predictive Margins				
A. Travel distance (km)	1	5	10	15	20
Probability of application					
High Score School	0.695	0.550	0.384	0.255	0.164
	[0.681,0.716]	[0.536,0.571]	[0.373,0.402]	[0.247,0.268]	[0.158,0.173]
Mean Score School	0.592	0.452	0.307	0.199	0.126
	[0.578, 0.578]	[0.440, 0.440]	[0.297, 0.297]	[0.193, 0.193]	[0.122,0.122]
Low Score School	0.492	0.365	0.241	0.155	0.097
	[0.479,0.512]	[0.355,0.382]	[0.234,0.254]	[0.149,0.163]	[0.093,0.102]

6 Discussion

A key motivation behind the new reform for the government was to reduce the existing school segregation based on Chile's socioeconomic status. However, Kutscher et al. (2020) has analyzed the implications of DA for entire Chile and found no evidence of an unambiguous positive impact on segregation due to the new policy. The school choice model illustrates that parents value travel costs and ability match while listing the ROL. Consequently, the combination of the school choice model and the government algorithm opens up the possibility of comparing the current policy's segregation outcomes with alternative counter-factual school networks.

I begin by examining the participants in 2016. I use the Duncan index as the primary measure of school segregation. The main advantage of using the Duncan index for this analysis is that it makes the results easily comparable to existing studies on school segregation as it is the most common index used to measure school segregation. I use household income to construct the Duncan index. Income is a common measure of socioeconomic status in related work on Chile (Alves et al., 2015; Valenzuela et al., 2014). The Duncan index for Magallanes is defined as follows

$$D = \frac{1}{2} \sum_{i=1}^{J} |\frac{N_{i,ses=l}}{N_{ses=l}} - \frac{N_{i,ses=h}}{N_{ses=h}}| \tag{5}$$

I compute the Duncan index for the public and voucher schools that participated in DA in 2016 in Magallanes for ninth-grade admissions. Figure B7 shows the changes in school segregation before and after DA. Overall, there has been a slight uptick in segregation levels post the reform.[19]

The school choice model illustrates that parents value the commuting distance to school. In my first simulation exercise, I use the Threshold Rank Order Model to generate alternative allocation under no travel cost. Such a simulation exercise can be interpreted as a thought experiment similar to busing. [20] The Duncan index for 2017 is 0.372 under the actual allocation. However, if the low income students are provided busing which eliminates their travel cost then this index drops to 0.351 for 2017. The drop is marginal as not all students participate in DA and there are factors other than distance that are critical for submitted ROL.

The school choice model for 2016 had limited set of covariates and this set has been expanded for 2017. The marginal effects suggest that although distance plays a critical role in school choice, student ability can compensate for such costs for low income students, specially the decision to list high score schools higher up in their ROL. Consequently, I address this interplay between costs and ability by conducting a simulation exercise which compensates for low ability and compute the probability of ranking high score

---

[19]I construct the Duncan index using student's family income and the income>200,000 Chilean pesos is used as the cut-off for low and high income students.

[20]I mimicked the government allocation rule for the applicants in Magallanes in 2016. Since, at the margin, there is a lottery, it is not possible to perfectly replicate the actual assignment. Overall, I could replicate the assignment for 82% of the sample correctly. Moreover, students who had some priority at any applied school have relatively less uncertainty in admission than students without a priority as they face a random lottery with a higher probability. I observe that the replicated algorithm accurately allocates schools to 87% of students who had indicated some priority in their application. The corresponding number for students without any priority is 79%.

schools under this alternative.

Figure B12 illustrates that students with mean or ability below mean often have a very low likelihood of ranking the high score school over a low score school. In other words, the match between student ability and school average test scores matters in school choice. Suppose a policy intervention can provide additional tutoring services to the students at or below the mean in the pre-DA test score distribution. In that case, that can go a long way in changing the DA ranking behavior of low-income low ability students.

7   Conclusion

Centralized algorithms are increasingly used across the globe to assign students to schools and universities. A key motivation for policymakers to use this mechanism is an improvement in equity. However, the expansion in school choice need not necessarily translate into lower levels of segregation. The overall impact hinges critically on parental preferences, which is a crucial input for these algorithms.

The urn model introduced by Plackett and Luce is the most commonly used model to estimate the determinants of the rank-ordered list (ROL). This model assumes that the decision to rank n schools can be split into n independent processes where each step can be thought as a multinomial choice by shrinking the choice set as one goes from ranking top to the bottom-most school. This assumption is unlikely to hold in ROL, where parents are expected to see the process of ranking n schools as a one-step process. Modeling the ROL is computationally intensive. Consequently, I develop a recursive algorithm to estimate the likelihood of this scenario efficiently. I also allow for heterogeneity in the rank cut-off across individuals. This heterogeneity is handled using the Expectation-Maximization algorithm in the maximum likelihood estimator.

I apply my new estimator to the centralized allocation in Chile. The Chilean government adopted a centralized system for student assignments starting in 2016. This new reform was sequentially implemented, and by 2019 all regions had a centralized assignment system for admissions to public and voucher schools in Chile. My analysis in Chile suggested that parents attach a high value to the school's distance in their ROL. Instead of crude distance proxies, I use the open street maps API and the entire road network of Chile to compute the actual travel distance to schools. I also find that the sensitivity to distance decreases with an improvement in student's socio-economic characteristics (SES) like income. Second, the results also suggest that parents care about the ability match between the student and the school while listing schools in ROL. Lastly, my simulation exercise shows policy interventions such as tutoring can help improve the representation of low income low ability students in high score schools.

A. Appendix

Proof of Theorem 1

Proof.

Suppose, the individual ranks K schools in their neighborhood where a rank of 1 corresponds to the top school and a rank of K corresponds to the lowest rank among the schools that were ranked. I do not putting any restriction on the number of non-ranked schools. The integral that I want to evaluate is

$$P(u_1 - \frac{c}{p_1^s} > u_2 - \frac{c}{p_2^s} > .... > u_K - \frac{c}{p_K^s} > u_0) = \left[\int_{-W_1}^{\infty} \int_{-W_2}^{W_1+\epsilon_1-W_2} ...... \int_{-W_K}^{W_{K-1}+\epsilon_{K-1}-W_K} f(\epsilon_K)d\epsilon_K.....f(\epsilon_1)d\epsilon_1 \right] \tag{6}$$

where $W_1 = u_1 - \frac{c}{p_1^s}$. First, I evaluate the innermost integral over ϵ_K

$$\int_{-W_K}^{W_{K-1}+\epsilon_{K-1}-W_K} f(\epsilon_K)d\epsilon_K = F(W_{K-1} + \epsilon_{K-1} - W_K) - F(-W_K)$$

Putting this part back into (5)

$$P(u_1 - \frac{c}{p_1^s} > u_2 - \frac{c}{p_2^s} > .... > u_K - \frac{c}{p_K^s} > u_0) =$$
$$\left[\int_{-W_1}^{\infty} \int_{-W_2}^{W_1+\epsilon_1-W_2} ...... \int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} f(\epsilon_{K-1})d\epsilon_{K-1}.....f(\epsilon_1)d\epsilon_1 \right] - \left[F(-W_K)\boldsymbol{I(K-1)} \right]$$

where $\boldsymbol{I(K-1)} = \int_{-W_1}^{\infty} \int_{-W_2}^{W_1+\epsilon_1-W_2} ...... \int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} f(\epsilon_{K-1})d\epsilon_{K-1}.....f(\epsilon_1)d\epsilon_1$. Next, I work with the innermost integral in for the first term in (2). I use integration by parts to solve the innermost integral in (2).

$$\int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} \underbrace{F(\epsilon_{K-1} + W_{K-1} - W_K)}_{g} \underbrace{f(\epsilon_{K-1})d\epsilon_{K-1}}_{\frac{dh}{dx}dx}$$

The integration by parts formula for definite integrals is given below.

$$\int_a^b g \frac{dh}{dx}dx = [gh]_a^b - \int_a^b h \frac{dg}{dx}dx$$

First, I compute $\frac{dg}{d\epsilon_{K-1}}$ and h.

$$\frac{dg}{d\epsilon_{K-1}} = f(\epsilon_{K-1} + W_{K-1} - W_K), h = F(\epsilon_{K-1})$$

Using all the components derived so far, term 1 in integration by parts is

$$[gh]_a^b = [F(\epsilon_{K-1} + W_{K-1} - W_K)F(\epsilon_{K-1})]_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}}$$

$$= F(W_{K-2} + \epsilon_{K-2} - W_K)F(W_{K-2} + \epsilon_{K-2} - W_{K-1}) - F(-W_K)F(-W_{K-1})$$

Next, using all the components derived so far the term two in integration by parts is

$$\int_a^b h\frac{dg}{dx}dx = \int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} F(\epsilon_{K-1})f(\epsilon_{K-1} + W_{K-1} - W_K)d\epsilon_{K-1}$$

I will apply result 2 derived below to term 2 of integration by parts.

$$F(x)f(x+c) = e^{-e^{-x}}e^{-(x+c)}e^{e^{-(x+c)}}$$

$$F(x+c)f(x) = e^{-e^{-(x+c)}}e^{-x}e^{e^{-x}}$$

$$F(x)f(x+c) = e^{-c}F(x+c)f(x)$$

Using result 2, term 2 is

$$\int_a^b h\frac{dg}{dx}dx = \int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} F(\epsilon_{K-1})f(\epsilon_{K-1} + W_{K-1} - W_K)d\epsilon_{K-1}$$

$$= e^{-(W_{K-1}-W_K)} \int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} \underbrace{F(\epsilon_{K-1} + W_{K-1} - W_K)}_{g} \underbrace{f(\epsilon_{K-1})d\epsilon_{K-1}}_{\frac{dh}{dx}dx}$$

<p style="text-align:center">LHS of integration by parts</p>

Bringing all the three terms together, I obtain

$$\int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} F(\epsilon_{K-1} + W_{K-1} - W_K)f(\epsilon_{K-1})d\epsilon_{K-1}$$

$$= \Big[F(W_{K-2} + \epsilon_{K-2} - W_K)F(W_{K-2} + \epsilon_{K-2} - W_{K-1}) - F(-W_K)F(-W_{K-1})$$

$$- e^{-(W_{K-1}-W_K)} \int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} F(\epsilon_{K-1} + W_{K-1} - W_K)f(\epsilon_{K-1})d\epsilon_{K-1} \Big]$$

$$(1 + e^{-(W_{K-1}-W_K)}) \int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} F(\epsilon_{K-1} + W_{K-1} - W_K)f(\epsilon_{K-1})d\epsilon_{K-1}$$

$$= F(W_{K-2} + \epsilon_{K-2} - W_K)F(W_{K-2} + \epsilon_{K-2} - W_{K-1}) - F(-W_K)F(-W_{K-1})$$

$$\int_{-W_{K-1}}^{W_{K-2}+\epsilon_{K-2}-W_{K-1}} F(\epsilon_{K-1} + W_{K-1} - W_K)f(\epsilon_{K-1})d\epsilon_{K-1}$$

$$= \kappa_{K-1}F(W_{K-2} + \epsilon_{K-2} - W_K)F(W_{K-2} + \epsilon_{K-2} - W_{K-1}) - \kappa_{K-1}F(-W_K)F(-W_{K-1})$$

where $\kappa_{K-1} = \frac{1}{1+e^{-(W_{K-1}-W_K)}}$ Next, I put this term back into the integral and obtain the following

$$P(u_1 - \frac{c}{p_1^s} > u_2 - \frac{c}{p_2^s} > .... > u_K - \frac{c}{p_K^s} > u_0) =$$

$$\left[\kappa_{K-1}\int_{-W_1}^{\infty}\cdots\int_{-W_{K-2}}^{W_{K-3}+\epsilon_{K-3}-W_{K-2}}F(W_{K-2}+\epsilon_{K-2}-W_K)F(W_{K-2}+\epsilon_{K-2}-W_{K-1})f(\epsilon_{K-2})d\epsilon_{K-2}.....f(\epsilon_1)d\epsilon_1\right]-$$

$$\left[\kappa_{K-1}F(-W_K)F(-W_{K-1})\boldsymbol{I(K-2)}\right]-\left[F(-W_K)\boldsymbol{I(K-1)}\right]$$

Again I repeat the same steps of integration by parts to solve for the innermost integral. I redefine $g=F(W_{K-2}+\epsilon_{K-2}-W_K)F(W_{K-2}+\epsilon_{K-2}-W_{K-1})$ and $\frac{dh}{dx}=f(\epsilon_{K-2})$. Repeating the similar scaling transformation in Integration by parts I obtain the following. Here,$g=F(W_{K-2}+\epsilon_{K-2}-W_K)F(W_{K-2}+\epsilon_{K-2}-W_{K-1})$ and $\frac{dh}{dx}=f(\epsilon_{K-2})$ which will result in the following two components

$$[gh]_a^b=F(W_{K-2}+\epsilon_{K-2}-W_K)F(W_{K-2}+\epsilon_{K-2}-W_{K-1})F(\epsilon_{K-2})|_{-W_{K-2}}^{W_{K-3}+\epsilon_{K-3}-W_{K-2}}$$

$$=F(W_{K-3}+\epsilon_{K-3}-W_K)F(W_{K-3}+\epsilon_{K-3}-W_{K-1})F(W_{K-3}+$$

$$\epsilon_{K-3}-W_{K-2})-F(-W_K)F(-W_{K-1})F(-W_{K-2})$$

The second component in integration by parts is

$$\int_a^b h\frac{dg}{dx}dx=\int_{W_{K-3}+\epsilon_{K-3}-W_{K-2}}^{-W_{K-2}}f(W_{K-2}+\epsilon_{K-2}-W_K)F(W_{K-2}+\epsilon_{K-2}-W_{K-1})F(\epsilon_{K-2})d\epsilon_{K-2}$$

$$+\int_{W_{K-3}+\epsilon_{K-3}-W_{K-2}}^{-W_{K-2}}f(W_{K-2}+\epsilon_{K-2}-W_{K-1})F(W_{K-2}+\epsilon_{K-2}-W_K)F(\epsilon_{K-2})d\epsilon_{K-2}$$

Applying the same transformation as before, I obtain result a and result b
Result a

$$F(x)F(x+c_2)f(x+c_1)=e^{e^{-x}}e^{-e^{-(x+c_2)}}e^{-e^{-(x+c_1)}}e^{e^{-(x+c_1)}}$$

$$=e^{-c_1}F(x+c_1)F(x+c_2)f(x)$$

Result b

$$F(x)F(x+c_1)f(x+c_2)=e^{e^{-x}}e^{-e^{-(x+c_1)}}e^{-e^{-(x+c_2)}}e^{e^{-(x+c_2)}}$$

$$=e^{-c_2}F(x+c_1)F(x+c_2)f(x)$$

In my setting $c_1=W_{K-2}-W_K$ and $c_2=W_{K-2}-W_{K-1}$. Therefore, I write term 2 of integration by parts as

$$\text{term 2}=e^{-c_1}LHS+e^{-c_2}LHS$$

$$\text{LHS}=\int_{-W_{K-2}}^{W_{K-3}+\epsilon_{K-3}-W_{K-2}}F(W_{K-2}+\epsilon_{K-2}-W_K)F(W_{K-2}+\epsilon_{K-2}-W_{K-1})f(\epsilon_{K-2})d\epsilon_{K-2}$$

$$P(u_1-\frac{c}{p_1^s}>u_2-\frac{c}{p_2^s}>....>u_K-\frac{c}{p_K^s}>u_0)=$$

$$\left[\kappa_{K-2}\kappa_{K-1}\int_{-W_1}^{\infty}..\int_{-W_{K-3}}^{W_{K-4}+\epsilon_{K-4}-W_{K-3}}\prod_{j=0,1,2}F(W_{K-3}+\epsilon_{K-3}-W_{K-j})f(\epsilon_{K-3})d\epsilon_{K-3}..f(\epsilon_1)d\epsilon_1\right]-$$

$$\left[\kappa_{K-2}\kappa_{K-1}F(-W_K)F(-W_{K-1})F(-W_{K-3})\boldsymbol{I(K-3)}\right]-\left[\kappa_{K-1}F(-W_K)F(-W_{K-1})\boldsymbol{I(K-2)}\right]-\left[F(-W_K)\boldsymbol{I(K-1)}\right]$$

where $\kappa_{K-2}=\dfrac{1}{1+e^{-(W_{K-2}-W_K)}+e^{-(W_{K-2}-W_{K-1})}}=\dfrac{1}{\sum_{j=n-2}^{K}e^{-(W_{K-2}-W_j)}}$. I keep on iterating backward to get the final solution stated as result (3). For completeness of the proof, I next derive the last step for a general m where $1<m<K$. First, the last result if I have iterated backward m times is

$$P(u_1-\frac{c}{p_1^s}>u_2-\frac{c}{p_2^s}>....>u_K-\frac{c}{p_K^s}>u_0)=\left[\kappa_{K-(m+1)}\right.$$

$$\left...\kappa_{K-1}\int_{-W_1}^{\infty}..\int_{-W_{K-m}}^{W_{K-(m+1)}+\epsilon_{K-(m+1)}-W_{K-m}}\prod_{j=0,1,..,m-1}F(W_{K-m}+\epsilon_{K-m}-W_{K-j})f(\epsilon_{K-m})d\epsilon_{K-m}..f(\epsilon_1)d\epsilon_1\right]$$

$$-\left[\kappa_{K-(m+1)}..\kappa_{K-1}F(-W_K)...F(-W_{K-m})\boldsymbol{I(K-m)}\right]-....-\left[\kappa_{K-1}F(-W_K)F(-W_{K-1})\boldsymbol{I(K-2)}\right]$$

$$-\left[F(-W_K)\boldsymbol{I(K-1)}\right]$$

Second, I prove that the result for having iterated m times is what I derived above. I begin with $m-1$ and then derive the solution for the mth iteration.

$$P(u_1-\frac{c}{p_1^s}>u_2-\frac{c}{p_2^s}>....>u_K-\frac{c}{p_K^s}>u_0)=$$

$$\left[\kappa_{K-m}..\kappa_{K-1}\int_{-W_1}^{\infty}..\int_{-W_{K-(m-1)}}^{W_{K-m}+\epsilon_{K-m}-W_{K-(m-1)}}\right.$$

$$\underbrace{\underbrace{\prod_{j=0,1,..,m-2}F(W_{K-(m-1)}+\epsilon_{K-(m-1)}-W_{K-j})}_{g}\underbrace{f(\epsilon_{K-(m-1)})d\epsilon_{K-(m-1)}}_{\frac{dh}{dx}}..f(\epsilon_1)d\epsilon_1}\Bigg]-$$

$$\left[\kappa_{K-m}..\kappa_{K-1}F(-W_K)F(-W_{K-1})F(-W_{K-3})\boldsymbol{I(K-(m-1))}\right]-..-\left[F(-W_K)\boldsymbol{I(K-1)}\right]$$

Using integration by parts for the innermost integral.

$$[gh]_a^b=[\prod_{j=0,1,..,m-2}F(W_{K-(m-1)}+\epsilon_{K-(m-1)}-W_{K-j})\times F(\epsilon_{K-(m-1)})]_{-W_{K-(m+1)}}^{W_{K-m}+\epsilon_{K-m}-W_{K-(m-1)}}$$

$$=[\prod_{j=0,1,..,m-1}F(W_{K-m}+\epsilon_{K-m}-W_{K-j})-F(-W_{K-(m-1)})..F(-W_K)]$$

Now for the second term in integration by parts I use the result $\prod_{j=2,...,b}F(x)F(x+c_j)f(x+c_1)=e^{-c_1}\prod_{j=1,...,b}F(x+c_j)f(x)$

$$\int_a^b h\frac{dg}{dx}dx=(\sum_j e^{-c_j})\int_{-W_{K-(m-1)}}^{W_{K-m}+\epsilon_{K-m}-W_{K-(m-1)}}$$

$$\prod_{j=0,1,..,m-2} F(W_{K-(m-1)} + \epsilon_{K-(m-1)} - W_{K-j}) f(\epsilon_{K-(m-1)}) d\epsilon_{K-(m-1)}$$

$$(1 + (\sum_j e^{-c_j})) \int_{-W_{K-(m-1)}}^{W_{K-m}+\epsilon_{K-m}-W_{K-(m-1)}} \prod_{j=0,1,..,m-2} F(W_{K-(m-1)} + \epsilon_{K-(m-1)} - W_{K-j}) f(\epsilon_{K-(m-1)}) d\epsilon_{K-(m-1)} =$$

$$\left[\prod_{j=0,1,..,m-1} F(W_{K-m} + \epsilon_{K-m} - W_{K-j}) - F(-W_{K-(m-1)})..F(-W_K) \right]$$

$$\int_{-W_{K-(m-1)}}^{W_{K-m}+\epsilon_{K-m}-W_{K-(m-1)}} \prod_{j=0,1,..,m-2} F(W_{K-(m-1)} + \epsilon_{K-(m-1)} - W_{K-j}) f(\epsilon_{K-(m-1)}) d\epsilon_{K-(m-1)} =$$

$$\kappa_{K-(m+1)} \left[\prod_{j=0,1,..,m-1} F(W_{K-m} + \epsilon_{K-m} - W_{K-j}) - F(-W_{K-(m-1)})..F(-W_K) \right]$$

Putting this term back into the integral, I obtain the following expression

$$P(u_1 - \frac{c}{p_1^s} > u_2 - \frac{c}{p_2^s} > .... > u_K - \frac{c}{p_K^s} > u_0) =$$

$$\left[\kappa_{K-(m+1)}..\kappa_{K-1} \int_{-W_1}^{\infty} .. \int_{-W_{K-m}}^{W_{K-(m+1)}+\epsilon_{K-(m+1)}-W_{K-m}} \prod_{j=0,1,..,m-1} F(W_{K-m} + \epsilon_{K-m} - W_{K-j}) f(\epsilon_{K-m}) d\epsilon_{K-m}..f(\epsilon_1) d\epsilon_1 \right]$$

$$- \left[\kappa_{K-(m+1)}..\kappa_{K-1} F(-W_K)...F(-W_{K-m}) \mathbf{I(K-m)} \right] - .... - \left[\kappa_{K-1} F(-W_K) F(-W_{K-1}) \mathbf{I(K-2)} \right]$$

$$- \left[F(-W_K) \mathbf{I(K-1)} \right]$$

The last component left is to solve for the topmost integral over the highest ranked school which is 1 in my setting.

$$P(u_1 - \frac{c}{p_1^s} > u_2 - \frac{c}{p_2^s} > .... > u_K - \frac{c}{p_K^s} > u_0) =$$

$$\kappa_2...\kappa_{K-1} \int_{-W_1}^{\infty} \prod_{j\neq 1} e^{-e^{-(\epsilon_1+W_1-W_j)}} e^{-\epsilon_1} e^{-e^{-\epsilon_1}} d\epsilon_1 - \left[\kappa_2..\kappa_{K-1} F(-W_K)..F(-W_2) \mathbf{I(1)} \right] -$$

$$\left[\kappa_3..\kappa_{K-1} F(-W_K) F(-W_{K-1})..F(-W_2) \mathbf{I(2)} \right] - ...... - \left[\kappa_{K-1} F(-W_K) F(-W_{K-1}) \mathbf{I(K-2)} \right] - \left[F(-W_K) \mathbf{I(K-1)} \right]$$

I solve for the first term in the above integral to complete the proof. The derivation of the closed form solution for $\kappa_2...\kappa_{K-1} \int_{-W_1}^{\infty} \prod_{j\neq 1} e^{-e^{-(\epsilon_1+W_1-W_j)}} e^{-\epsilon_1} e^{-e^{-\epsilon_1}} d\epsilon_1$ is given below

$$\kappa_2...\kappa_{K-1} \int_{-W_1}^{\infty} \prod_{j\neq 1} e^{-e^{-(\epsilon_1+W_1-W_j)}} e^{-\epsilon_1} e^{-e^{-\epsilon_1}} d\epsilon_1 = \kappa_2...\kappa_{K-1} \int_{-W_1}^{\infty} \prod_{j=1}^{K} e^{-e^{-(\epsilon_1+W_1-W_j)}} e^{-\epsilon_1} d\epsilon_1$$

$$= \kappa_2...\kappa_{K-1} \int_{-W_1}^{\infty} exp(-\sum_{j=1}^{K} -e^{-(\epsilon_1+W_1-W_j)}) e^{-\epsilon_1} d\epsilon_1$$

$$= \kappa_2...\kappa_{K-1} \int_{-W_1}^{\infty} exp(-e^{-\epsilon_1} \sum_{j=1}^{K} -e^{-(W_1-W_j)}) e^{-\epsilon_1} d\epsilon_1$$

I substitute $t = e^{-\epsilon_1}$ and obtain the following expression.

$$= \int_{e^{W_1}}^{0} exp(-t \sum_j e^{-(W_1 - W_j)}) - dt$$

$$= \int_{0}^{e^{W_1}} exp(-t \sum_j e^{-(W_1 - W_j)}) dt$$

$$= \frac{exp(-t \sum_j e^{-(W_1 - W_j)})}{\sum_j e^{-(W_1 - W_j)}} \Big|_{0}^{e_1^W}$$

$$= \frac{1}{\sum_j e^{-(W_1 - W_j)}} \Big[1 - F(-W_1) F(-W_2) .... F(-W_K) \Big]$$

Substituting for the closed form solution for term 1 as derived above, I obtain the final recursive solution for the likelihood.

$$P(u_1 - \frac{c}{p_1^s} > u_2 - \frac{c}{p_2^s} > .... > u_K - \frac{c}{p_K^s} > u_0) =$$

$$\kappa_1 \kappa_2 ... \kappa_{K-2} \kappa_{K-1} (1 - F(-W_1) ... F(-W_K)) - \Big[\kappa_2 .. \kappa_{K-1} F(-W_K) F(-W_{K-1}) .. F(-W_2) \boldsymbol{I(1)} \Big] \quad (7)$$

$$- ...... - \Big[\kappa_{K-1} F(-W_K) F(-W_{K-1}) \boldsymbol{I(K-2)} \Big] - \Big[F(-W_K) \boldsymbol{I(K-1)} \Big]$$

∎

Proof of Step 1, solving for $\hat{\sigma}^2$

Proof. The score function is $S_i(\sigma^2) = -\frac{1}{2\sigma^2} + \frac{(c_i - z_i\gamma)^2}{2(\sigma^2)^2}$. Solving for $\sum_{i=1}^{N} E_{h(c|R)} S_i(\sigma^2) = 0$

$$\sum_{i=1}^{N} E_{h(c|R)} S_i(\sigma^2) = 0$$

$$\sum_{i=1}^{N} \int [-\frac{1}{2\sigma^2} + \frac{(c_i - z_i\gamma)^2}{2(\sigma^2)^2}] h(c_i|R_i; \gamma, \sigma^2) dc = 0$$

$$-\frac{n}{2\sigma^2} + \sum_{i=1}^{N} \int [\frac{(c_i - z_i\gamma)^2}{2(\sigma^2)^2}] h(c_i|R_i; \gamma, \sigma^2) dc = 0$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{N} \int (c_i - z_i\gamma)^2 h(c_i|R_i; \gamma, \sigma^2) dc$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{N} \left[\frac{\sum_{t=1}^{T} (c_i - z_i\gamma)^2 f(R_i|c; \beta)}{\sum_{t=1}^{T} f(R_i|c; \beta)} \right]$$

∎

B. Appendix

Tables

Table B1: Student and school participation in DA by region

Regions	N	School % public	% voucher	N	Student % enroll
A. 2016					
Magallanes	24	45.83	54.17	1040	46.19
B. 2017					
Tarapacá	57	28.07	71.93	1757	32.78
Coquimbo	129	25.78	66.41	4824	43.14
O'Higgins	132	43.51	56.49	7872	59.47
Los Lagos	145	47.22	52.78	7439	57.86
Magallanes	25	48.00	52.00	1041	46.61
C. 2018					
Tarapacá	56	27.27	72.73	1804	35.33
Antofagasta	67	44.78	55.22	4459	51.79
Atacama	33	54.55	45.45	2856	67.41
Coquimbo	124	33.06	66.94	4517	41.67
Valparaíso	321	31.35	68.65	9909	42.29
O'Higgins	132	43.18	56.82	7517	58.22
Maule	153	44.08	55.92	8794	60.75
Biobío	309	47.23	52.77	15127	53.56
Araucanía	166	36.97	63.03	8502	60.20
Los Lagos	142	47.18	52.82	7290	58.89
Aysén	25	52.00	48.00	469	28.71
Magallanes	25	48.00	52.00	1012	46.36
Los Ríos	74	41.10	58.90	2990	56.49
Arica and Parinacota	31	35.48	64.52	1655	49.34

Notes: In this table I illustrate the distribution of school participation in DA by its type. Public and private voucher schools participated in DA. Private non-voucher schools did not participate in DA. The participation of students corresponds to those who applied for ninth-grade admissions.

Table B2: School variables for tenth grade

Variables	N	mean	sd	min	max
Tenth grade (2015)					
Language	2856	247.42	30.55	171	340
mathematics	2859	261.23	46.74	136	387
% public	2860	0.30	–	0	1
% private voucher	2860	0.57	–	0	1
% private nonvoucher	2860	0.13	–	0	1
% rural	2860	0.06	–	0	1
Tenth grade (2016)					
Language	2881	247.87	30.28	176	340
Mathematics	2881	264.04	46.72	138	391
% public	2884	0.30	–	0	1
% private voucher	2884	0.56	–	0	1
% private nonvoucher	2884	0.13	–	0	1
% rural	2884	0.06	–	0	1
Tenth grade (2017)					
Language	2901	251.89	28.82	169	338
mathematics	2900	264.71	45.13	149	393
% public	2901	0.31	–	0	1
% private voucher	2901	0.56	–	0	1
% private nonvoucher	2901	0.14	–	0	1
% rural	2901	0.06	–	0	1

Notes: I use tenth grade SIMCE for 2015, 2016 and 2017 and the enrolment files for these years to obtain the school variables for the school choice model.

Table B3: Distribution of school type and % charging copayment

Variables	N	% in total	% fee
	(1)	(2)	(3)
2015			
Public	877.0	30.0	7.3
Private Voucher	1671.0	57.1	77.7
Private Non-Voucher	380.0	13.0	100.0
Total	2928.0	100.0	–
2016			
Public	887.0	30.2	0.6
Private Voucher	1657.0	56.4	57.3
Private Non-Voucher	393.0	13.4	100.0
Total	2937.0	100.0	–
2017			
Public	902.0	30.5	0.4
Private Voucher	1659.0	56.0	53.3
Private Non-Voucher	401.0	13.5	100.0
Total	2962.0	100.0	–
2018			
Public	916.0	30.7	0.4
Private Voucher	1621.0	54.3	47.3
Private Non-Voucher	446.0	15.0	100.0
Total	2983.0	100.0	–

Notes: I use files on the details of schools in Chile to obtain the complete set of schools. Second, I construct the fee variable using the ficom data file that reports the fee structure of public and private non-voucher schools in Chile. The private non-voucher schools are allowed to charge any amount of fee. They are not required to give details of the fee to the Ministry of Education.

Table 1: Regular stage

Region	Participants (N)	Enrollment==Assignment (%)
2017		
Magallanes	1029	74.34
2018		
tarapaca	1664	76.62
Coquimbo	4655	75.32
Ohiggins	7670	74.81
loslagos	7182	76.65
Magallanes	1010	75.94
2019		
tarapaca	1827	71.81
antofagasta	4446	74.85
atacama	2830	75.09
coquimbo	4548	71.92
valparaiso	10008	75.70
ohiggins	7492	73.75
maule	8851	78.93
biobio	11597	79.34
araucania	8550	81.78
loslagos	7242	70.52
aysen	471	74.73
magallanes	996	75.40
rios	3055	79.18
ayp	1651	77.83

Notes: The program was sequentially implemented in Chile.

Table B4: Descriptive Statistics: Magallanes in 2016

Variables	N	Mean	Std. dev.	Min.	Max.
A. Student characteristics					
Math score	499	235.13	42.65	134	368
Language score	499	231.76	49.09	112	363
Mother's education	499	11.57	3.11	0	19
Father's education	499	11.41	3.03	3	19
Household income index	499	4.74	2.31	1	15
Number of ranked schools	499	4.10	1.54	2	11
B. Top choice school characteristics					
Math score	499	252.19	40.50	191	328
Language score	499	242.70	29.47	198	299
Household income index	499	5.66	1.27	4	9
Mother's education	499	11.89	1.40	10	15
Father's education	499	11.87	1.55	10	16
Travel distance (kms)	499	2.41	1.36	0	7

Notes: This table provides information on eighth-graders who applied for ninth-grade admission in 2016 in Magallanes. I obtain their background variables and test scores from sixth grade SIMCE. Travel distance has been computed using OSRM API.

Table B5: Descriptive Statistics by Region

	N	Mean	Std. dev.	Min	Max
Tarapaca					
Pre-reform Math (Standardized)	858	-0.09	1.05	-2.41	3
Pre-reform Language (Standardized)	858	-0.08	1.03	-2.58	3
Mother's education	858	11.30	3.31	0	22
Father's education	858	11.27	3.25	0	19
Income Index	858	4.87	2.59	1	15
Avg. academic score of outside option (Standardized)	858	-0.74	0.52	-1.76	1
Coquimbo					
Pre-reform Math (Standardized)	2693	-0.06	1.00	-2.44	3
Pre-reform Language (Standardized)	2693	-0.01	1.00	-2.63	3
Mother's education	2693	10.77	3.33	0	22
Father's education	2693	10.54	3.41	0	22
Income Index	2693	4.06	2.23	1	15
Avg. academic score of outside option (Standardized)	2693	-0.43	0.60	-1.60	2
O'Higgins					
Pre-reform Math (Standardized)	4765	0.03	0.99	-2.45	3
Pre-reform Language (Standardized)	4765	-0.01	1.01	-2.86	3
Mother's education	4765	10.55	3.34	0	19
Father's education	4765	10.28	3.44	0	22
Income Index	4765	4.08	2.03	1	15
Avg. academic score of outside option (Standardized)	4765	-0.59	0.43	-1.75	2
Los Lagos					
Pre-reform Math (Standardized)	4203	0.04	1.01	-2.48	3
Pre-reform Language (Standardized)	4203	0.03	0.99	-2.84	3
Mother's education	4203	10.15	3.59	0	22
Father's education	4203	9.95	3.41	0	22
Income Index	4203	3.92	2.09	1	15
Avg. academic score of outside option (Standardized)	4203	-0.58	0.49	-1.49	2
Magallanes					
Pre-reform Math (Standardized)	544	-0.03	0.97	-2.27	3
Pre-reform Language (Standardized)	544	0.01	0.97	-2.46	3
Mother's education	544	11.67	3.11	1	19
Father's education	544	11.52	3.18	0	17
Income Index	544	5.67	2.88	1	15
Avg. academic score of outside option (Standardized)	544	-0.66	0.40	-1.00	2

Notes: The program was sequentially implemented in Chile. In 2017 the program was introduced in a total of 5 regions.

Table B6: Joint Analysis: Summary Statistics

	N	Mean	Std. dev.	Min	Max
Pre-reform Math (Standardized)	7030	0.02	1.01	-2.44	3
Pre-reform Language (Standardized)	7030	0.03	0.99	-2.84	3
Mother's education	7030	10.64	3.42	0	22
Father's education	7030	10.39	3.40	0	22
income	7030	4.14	2.28	1	15
Avg. academic score of outside option school (Standardized)	7030	-0.49	0.53	-1.76	2
N	7030				
J	286				

Figures

Figure B1: Simulation results: Multiple covariates without individual heterogeneity
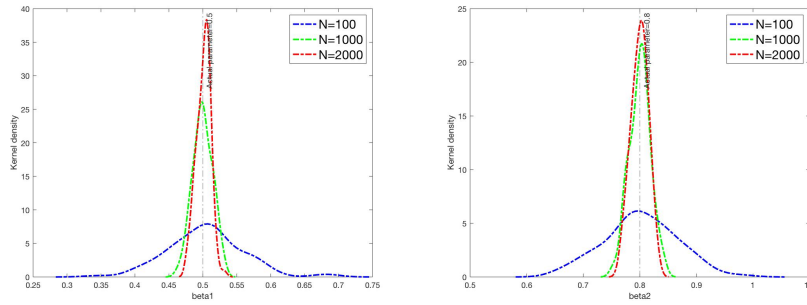
A. $\beta_4 = 0.5$

B. $\beta_2 = 0.8$
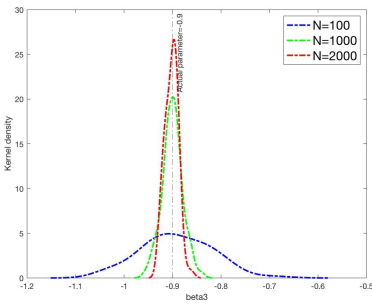
C. $\beta_3 = -0.9$
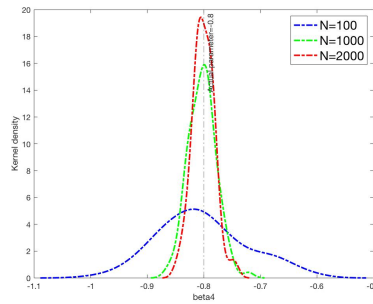
D. $\beta_4 = -0.8$
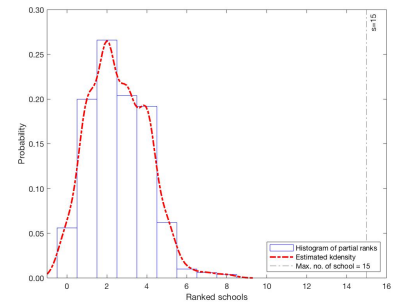
E. Rank

Notes: These graphs display the smoothed kernel density plots of the empirical distribution of the parameter estimates in repeated MC samples. The number of schools for this analysis is fixed at 15.

Figure B2: Simulation results: Multiple covariates with costs
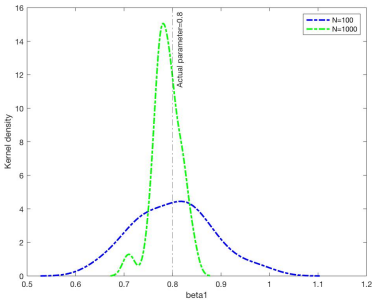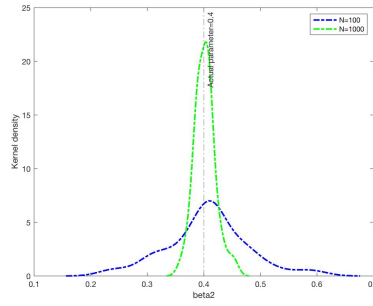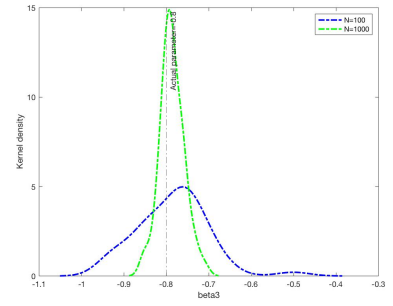
A. $\beta_4 = 0.8$ B. $\beta_2 = 0.4$ C. $\beta_3 = -0.8$
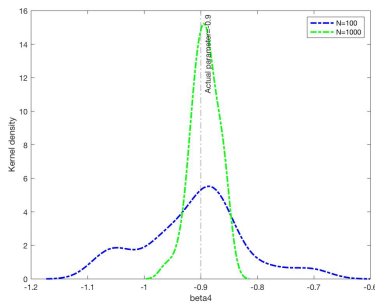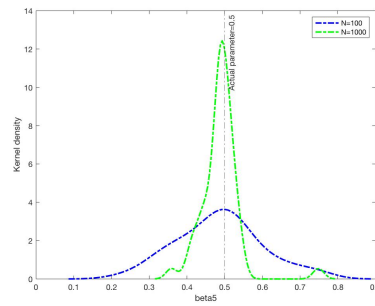
D. $\beta_4 = -0.9$ E. $\gamma = 0.5$

Notes: These graphs display the smoothed kernel density plots of the empirical distribution of the parameter estimates in repeated MC samples.

Figure B3: Comparison of threshold rank model with urn model
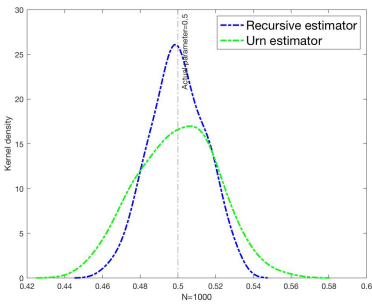
A. $\beta_1 = 0.5$
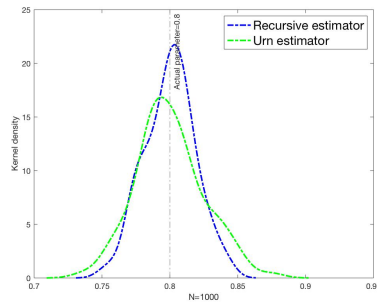
B. $\beta_2 = 0.8$
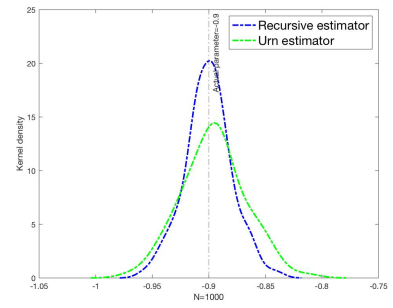
C. $\beta_3 = -0.9$
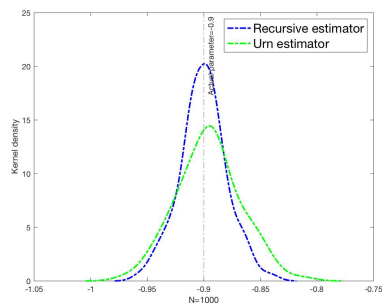
D. $\beta_4 = -0.8$

Notes: These graphs display the smoothed kernel density plots of the empirical distribution of the parameter estimates in repeated MC samples.

Figure B4: Segregation (Income) by school type
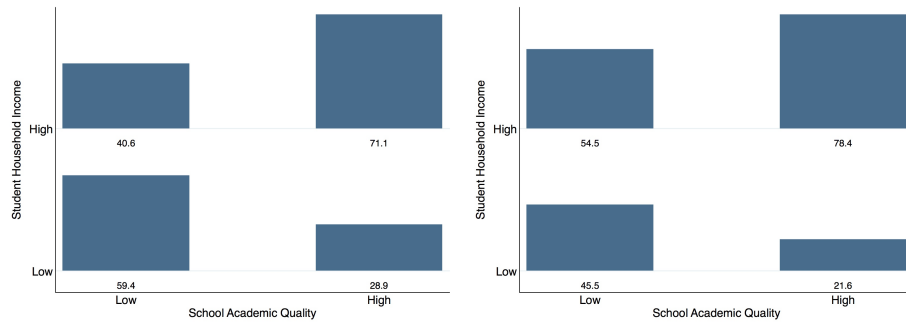
A. 2015 (Before)　　　　　D. 2018 (After)

Notes: These graphs display the composition of low income and high income students in high and low ability school types. For each school type I calculate the % of students who are low income and the % students who are high income. Students with low income have a monthly income of less than CLP 300,000. Low income students have a much lower representation in high score schools compared to their representation in low score schools.

Figure B5: Attendance and school dummies: OLS
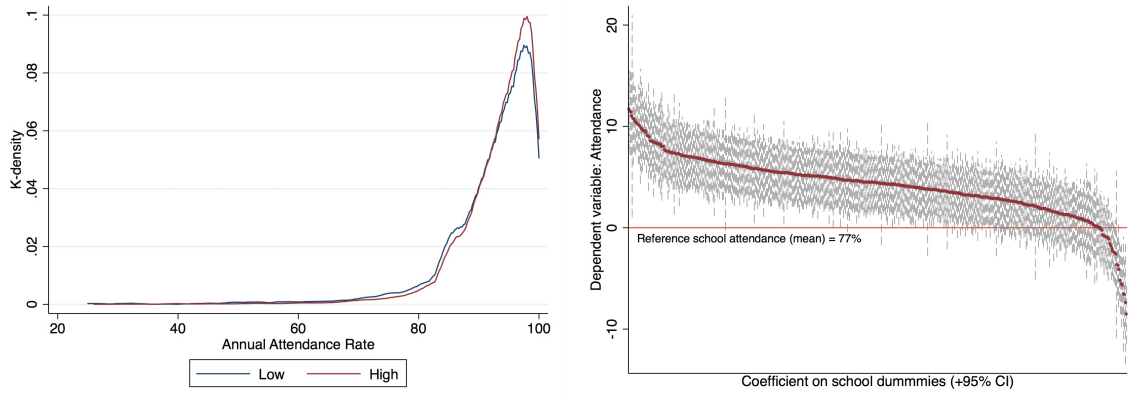
A. Density by student income

B. School dummies and attendance
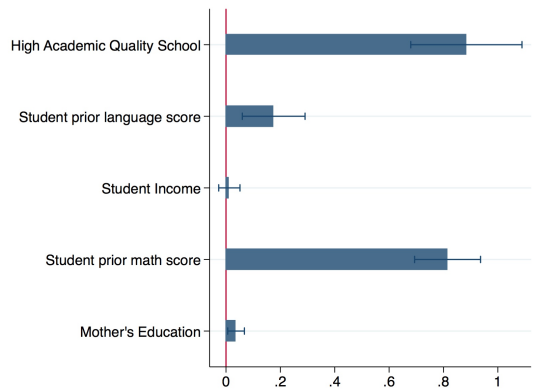
C. Attendance and school type

Notes: The graph in panel (A) illustrates the kernel density plots for annual attendance rates by student income. Students with low income have a monthly income of less than CLP 300,000. Panel (B) plots the coefficients of a regression of attendance on school dummies and student covariates measuring academic ability, income and socio-economic status. The graph in panel (B) displays the differences in attendance rate by school quality. I control for observed student characteristics in this regression as well.

Figure B6: School assignment of students that participated in DA, Magallanes in 2016
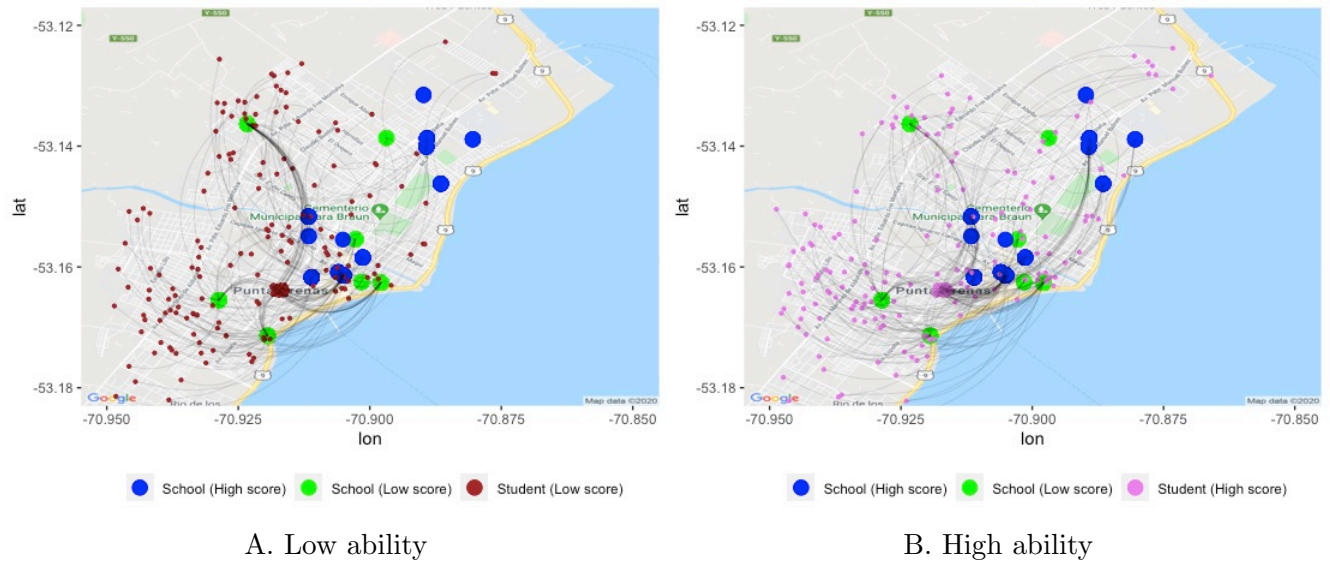
A. Low ability   B. High ability

Notes: The sample consists of students who participated in DA for ninth grade admissions in 2016 in Magallanes and therefore started ninth grade in the allocated school in 2017. The actual allocation data is taken from enrollment files for 2017.

Figure B7: Segregation in Magallanes
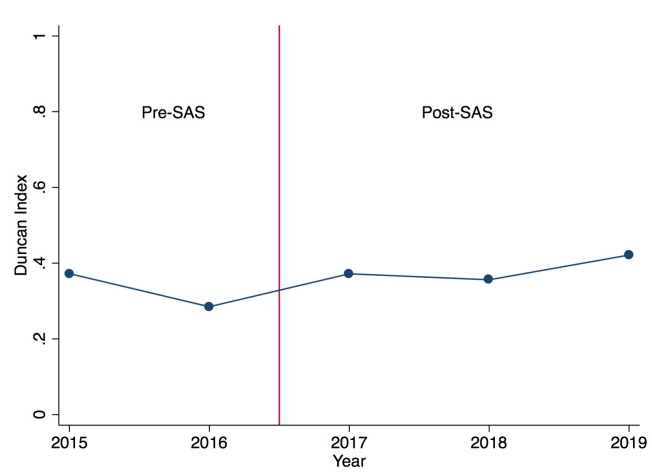
Notes: The sample consists of ninth-grade students and schools in Magallanes. The enrollment under DA will get reflected in 2017 enrollment files. I consider public and voucher schools for this analysis as these two types of schools participated in DA.

Figure B8: Overlap between predicted ROL and the best schools in Magallanes
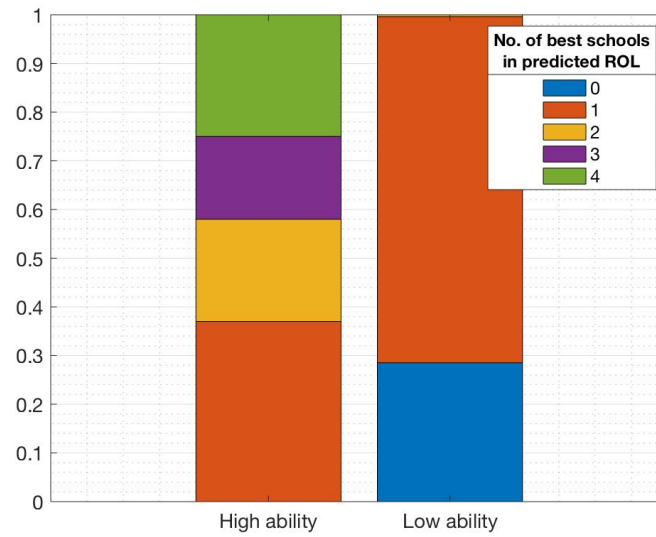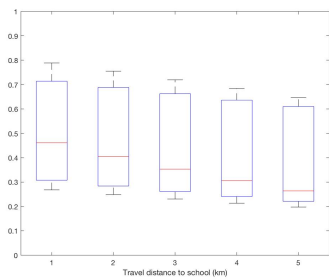

Notes: The predicted ROL has been generated using the estimates of model (2) in Table 6.
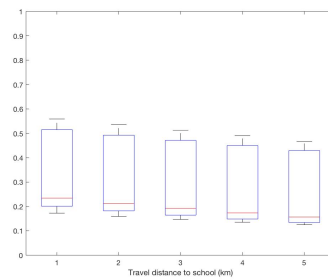
Figure B9: Predicted probability: Outside option
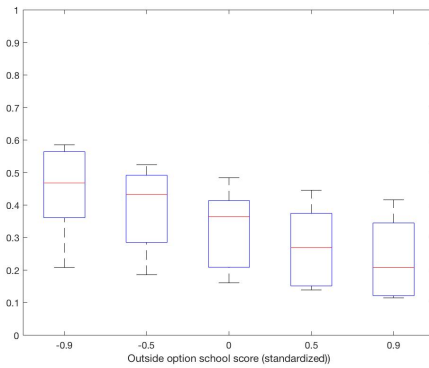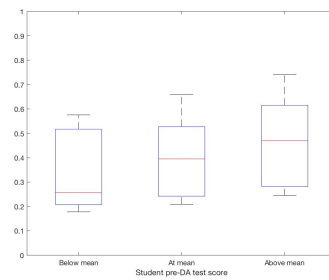

A. Mean school

Figure B10: Predictive margins for high score school

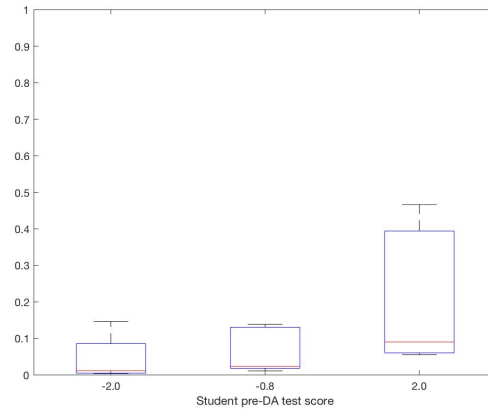

B. Low income high ability   C. Low income low ability   C. Low income ability

Figure B11: Predictive probability of ranking high score school over low score school
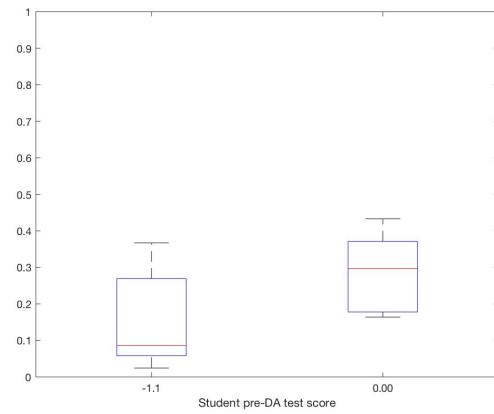

A. High score vs. low score

Notes: The predicted probability for this analysis is computed as $P(H > L > 0) = \frac{e^{V_H}}{e^{V_H}+e^{V_L}}(1 - F(-V_L) - \frac{e^{V_H}}{e^{V_H}+e^{V_L}}F(-V_L)(1 - F(-V_H))$, where H corresponds to the high score school and L corresponds to the low score school.

Figure B12: Simulated probability: Rank high score school over low score school
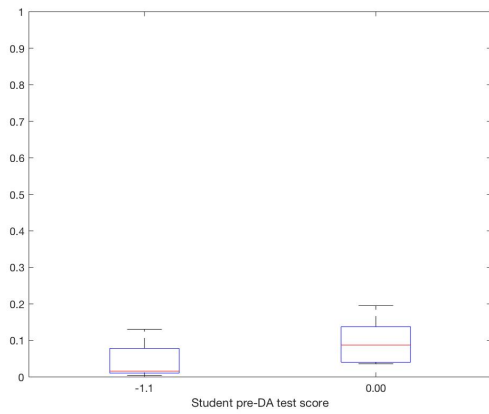
A. Actual predicted probability

B. Simulated predicted probability

References

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

Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., & Sönmez, T. (2005). The boston public school match. *American Economic Review*, *95*(2), 368–371.

Abdulkadiroğlu, A., & Sönmez, T. (2003). School choice: A mechanism design approach. *American economic review*, *93*(3), 729–747.

Ajayi, K. F. (2013). School choice and educational mobility: Lessons from secondary school applications in ghana. *Institute for Economic Development Working Paper*, *259*.

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

Allison, P. D., & Christakis, N. A. (1994). Logit models for sets of ranked items. *Sociological methodology*, 199–228.

Alves, F., Elacqua, G., Koslinki, M., Martinez, M., Santos, H., & Urbina, D. (2015). Winners and losers of school choice: Evidence from rio de janeiro, brazil and santiago, chile. *International Journal of Educational Development*, *41*, 25–34.

Bergman, P., & Chan, E. W. (2019). Leveraging parents through low-cost technology: The impact of high-frequency information on student achievement. *Journal of Human Resources*, 1118–9837R1.

Biró, P. (2012). University admission practices-hungary. *www. matching-in-practice. eu accessed July*, *21*, 2014.

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

Booth, J. G., & Hobert, J. P. (1999). Maximizing generalized linear mixed model likelihoods with an automated monte carlo em algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *61*(1), 265–285.

Burgess, S., & Briggs, A. (2010). School assignment, school choice and social mobility. *Economics of Education Review*, *29*(4), 639–649.

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

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

Croon, M. A., & Luijkx, R. (1993). Latent structure models for ranking data. In *Probability models and statistical analyses for ranking data* (pp. 53–74). Springer.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22.

Ergin, H., & Sönmez, T. (2006). Games of school choice under the boston mechanism. *Journal of public Economics*, *90*(1-2), 215–237.

Fack, G., Grenet, J., & He, Y. (2019). Beyond truth-telling: Preference estimation with centralized

school choice and college admissions. *American Economic Review*, *109*(4), 1486–1529.

Figlio, D., Karbownik, K., Roth, J., Wasserman, M., et al. (2016). School quality and the gender gap in educational achievement. *American Economic Review*, *106*(5), 289–95.

Francis, B., Dittrich, R., & Hatzinger, R. (2010). Modeling heterogeneity in ranked responses by nonparametric maximum likelihood: How do europeans get their scientific knowledge? *The annals of applied statistics*, 2181–2202.

Frenette, M. (2006). Too far to go on? distance to school and university participation. *Education Economics*, *14*(1), 31–58.

Fuller, W. C., Manski, C. F., & Wise, D. A. (1982). New evidence on the economic determinants of postsecondary schooling choices. *Journal of Human Resources*, 477–498.

Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, *69*(1), 9–15.

Gallego, F. A., & Hernando, A. (2010). School choice in chile: Looking at the demand side. *Pontificia Universidad Catolica de Chile Documento de Trabajo*(356).

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

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

Glewwe, P., & Jacoby, H. (1994). Student achievement and schooling choice in low-income countries: Evidence from ghana. *Journal of Human Resources*, 843–864.

Glick, P., & Sahn, D. E. (2006). The demand for primary schooling in madagascar: Price, quality, and the choice between public and private providers. *Journal of development economics*, *79*(1), 118–145.

Gottfried, M. A., & Kirksey, J. J. (2017). when students miss school: The role of timing of absenteeism on students test performance. *Educational Researcher*, *46*(3), 119–130.

Guiver, J., & Snelson, E. (2009). Bayesian inference for plackett-luce ranking models. In *proceedings of the 26th annual international conference on machine learning* (pp. 377–384).

Haeringer, G., & Klijn, F. (2009). Constrained school choice. *Journal of Economic theory*, *144*(5), 1921–1947.

Hanushek, E. A., Kain, J. F., Rivkin, S. G., & Branch, G. F. (2007). Charter school quality and parental decision making with school choice. *Journal of public economics*, *91*(5-6), 823–848.

Hanushek, E. A., Lavy, V., & Hitomi, K. (2008). Do students care about school quality? determinants of dropout behavior in developing countries. *Journal of Human Capital*, *2*(1), 69–105.

Hastings, J., Kane, T. J., & Staiger, D. O. (2005). *Parental preferences and school competition: Evidence from a public school choice program* (Tech. Rep.). National Bureau of Economic Research.

Hastings, J., Kane, T. J., & Staiger, D. O. (2009). Heterogeneous preferences and the efficacy of public school choice. *NBER Working Paper*, *2145*, 1–46.

Hastings, J., & Weinstein, J. M. (2007). *No child left behind: Estimating the impact on choices and student outcomes* (Tech. Rep.). National Bureau of Economic Research.

Hastings, J., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence

from two experiments. *The Quarterly journal of economics*, *123*(4), 1373–1414.

Hsieh, C.-T., & Urquiola, M. (2006). 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.

Ibrahim, J. G., Chen, M.-H., & Lipsitz, S. R. (1999). Monte carlo em for missing covariates in parametric regression models. *Biometrics*, *55*(2), 591–596.

Jackson, C. K. (2018). What do test scores miss? the importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, *126*(5), 2072–2107.

Kirabo Jackson, C. (2010). Do students benefit from attending better schools? evidence from rule-based student assignments in trinidad and tobago. *The Economic Journal*, *120*(549), 1399–1429.

Kuersteiner, G., Nath, S., & Urzua, S. (2020). Causal inference for deferred acceptance algorithms. *Working Paper*.

Kutscher, M., Nath, S., & Urzua, S. (2020). Centralized admission systems and school segregation: Evidence from a national reform. *IZA discussion paper No. 13305*.

Laverde, M. (2020). Unequal assignments to public schools and the limits of school choice.

Light, A., & Strayer, W. (2000). Determinants of college completion: School quality or student ability? *Journal of Human resources*, 299–332.

Liu, J., Lee, M., & Gershenson, S. (2019). The short-and long-run impacts of secondary school absences.

Luflade, M. (2017). The value of information in centralized school choice systems (job market paper).

Marden, J. I. (2014). *Analyzing and modeling rank data*. Chapman and Hall/CRC.

Pathak, P. A., & Sönmez, T. (2008). Leveling the playing field: Sincere and sophisticated players in the boston mechanism. *American Economic Review*, *98*(4), 1636–52.

Pathak, P. A., & Sönmez, T. (2013). School admissions reform in chicago and england: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review*, *103*(1), 80–106.

Reback, R. (2008). Demand (and supply) in an inter-district public school choice program. *Economics of Education Review*, *27*(4), 402–416.

Salonen, M., et al. (2014). Matching practices for secondary schools–finland. *Matching in Practice: European network for research on matching practices in education and related markets*.

Sammel, M. D., Ryan, L. M., & Legler, J. M. (1997). Latent variable models for mixed discrete and continuous outcomes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *59*(3), 667–678.

Skrondal, A., & Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, *68*(2), 267–287.

Torche, F. (2005). Privatization reform and inequality of educational opportunity: The case of chile. *Sociology of education*, *78*(4), 316–343.

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