Title
Essays on Centralized School Choice and Assignment Systems

Author
Krussig, Thomas

Publication Date
2023

Essays on Centralized School Choice and Assignment Systems

by

Thomas Krussig Vocke

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Benjamin Handel, Chair
Professor Matthew Backus
Professor Kei Kawai
Professor Jonathan Kolstad

Spring 2023

Essays on Centralized School Choice and Assignment Systems

Abstract

Essays on Centralized School Choice and Assignment Systems

by

Thomas Krussig Vocke

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Benjamin Handel, Chair

This dissertation studies how to improve the outcomes of Centralized Choice and Assignment Systems (CCAS) for schools, with a focus on both social objectives such as socioeconomic, cultural, or racial integration, as well as family welfare. CCAS systems are increasingly used worldwide for both student and teacher assignments, and their popularity is expected to rise further as digitalization continues to expand globally. Therefore, identifying cost-effective policies that can improve the functioning of CCAS systems can have significant implications, as they are likely to be applicable in many different contexts.

In the first chapter, co-authored with Isabel Jacas, we investigate an assignment rule policy: socioeconomic reserves (SR) for low-socioeconomic status (low-SES) families in the Chilean CCAS for PreK schools. Although the implementation of the CCAS system was expected to reduce school segregation by eliminating family background-based selection, similar levels of segregation persist. However, we find that optimizing the size of SR to local conditions can significantly improve outcomes. By setting SR at the schools' municipality low-SES applicants' share, educational segregation can be reduced by five times compared to the current flat 15% level when compared to a minimum segregation benchmark obtained by estimating family preferences and generating counterfactual applications that eliminate the drivers of differences in choice between socioeconomic groups. Additionally, leveraging demand-side policies as a complement to SR can further reduce segregation considerably, as SR effectively assigns additional low-SES applicants to congested schools where their share is underrepresented. In the second chapter, co-authored with Gregory Elacqua, Leidy Gómez, Luana Marotta, Carolina Méndez, and Christopher A. Neilson, we investigate an information demand-side policy that leverages the "smart" personalized feedback potential of digital application processes in CCAS. Specifically, we examine the causal impact of a personalized non-assignment risk warning, combined with a list of "achievable" teaching position recommendations, on teacher applications in the Ecuadorian "I Want to Become a Teacher" selection process. Our analysis reveals that treated teachers are significantly more likely to modify their application and secure an assignment, and we also provide evidence suggest-

ing that the intervention resulted in an increase in selection scores used by the Ecuadorian Ministry of Education to evaluate teacher performance.

The third chapter of this dissertation, co-authored with Gregory Elacqua, Isabel Jacas, Carolina Méndez, and Christopher A. Neilson, takes a step back from improving the outcomes of CCAS to examine the broader comparison between these systems and alternatives. Specifically, we compare the welfare effects on families of a CCAS implementation in the city of Manta, Ecuador in 2021, incorporating household preferences for the first time as an assignment criterion (using the deferred acceptance algorithm) and compare it to the alternative assignment mechanism previously used, which was also centrally coordinated but based on minimizing residence-to-school distances. Our findings reveal that considering applicant preferences leads to significant welfare gains, suggesting that CCAS can have a substantial impact on welfare in developing country contexts, even without complete optimization using policies such as those studied in the first two chapters.

Para Eva, Isidora y Mónica.

Contents

List of Figures

List of Tables

Acknowledgments

Coauthor information: Chapter 1 of this dissertation is coauthored with Isabel Jacas, currently at the United Nations Economic Commission for Latin America and the Caribbean. Chapters 2 and 3 are coauthored with researchers from the Inter-American Development Bank (IADB) and the NGO Consiliumbots (CB), who were involved in the implementation of the policies studied in those chapters and their subsequent analysis. Specifically, Gregory Elacqua and Carolina Méndez from the IADB are coauthors in both chapters and Luana Marotta is a coauthor in Chapter 2. From CB, Leidy Gómez, currently at TetherEducation, coauthors Chapter 2, Isabel Jacas coauthors Chapter 3, and Christopher A. Neilson, also Professor of Economics at Yale University, is a coauthor in both chapters.

I am deeply grateful to all of the coauthors from the IADB and CB teams for their dedicated and hard work, especially during the challenging implementation phases of the Covid-19 pandemic. Our collaborations have provided invaluable lessons that have enriched my academic and professional experience.

Acknowledgements: I am deeply grateful to Benjamin Handel for his unwavering support and guidance throughout my PhD journey. His kindness and encouragement were essential in helping me complete this dissertation. I am also extremely grateful to my Committee members, Matthew Backus, Kei Kawai, and Jonathan Kolstad, for their valuable time, exceptional advice, and feedback, as well as for their overall encouragement. I also want to thank Chris Walters for participating in my Qualifying Examination and offering extremely insightful comments and feedback. I would also like to express my heartfelt thanks to Janene Carol for her exceptional professionalism and support, which made the administrative process smoother, especially during my re-joining the program.

I am also grateful to Christopher A. Neilson for introducing me to the field of information interventions in education, particularly in relation to Centralized Choice and Assignment Systems, which are core issues studied in my dissertation. His encouragement and support were invaluable in these projects, and I deeply appreciate his generosity in sharing knowledge, ideas, and feedback. I would also like to acknowledge the contributions of all the collaborators at ConsiliumBots, including my coauthors and many other outstanding professionals, who made these projects possible. In particular, I want to thank Benjamin Madariaga for his collaboration in implementing assignment algorithms and simulations, which were critical to the success of this work.

I cherish many wonderful memories from my time at Berkeley, largely due to the exceptional individuals who enriched my experience. I am grateful to Juan Pablo Atal, Pierre Bachas, César Díaz, Ezequiel Garcia-Lembergman, Felipe González, Danae Varándano, Andrés González, Elena Carafí, Pablo Muñoz, Manuela Urrutia, Román Zárate, Darío Tortarolo, Malena Arcidiácono, Mathieu Pedemonte, Daphné Skandalis, and others, for their friendship and support. I am thankful for Nano Barahona and Sebastián Otero's hospitality in hosting me while I took a class at Stanford, for their friendship and support. I also want to thank Sebastián for his valuable feedback on socioeconomic reserves and affirmative action policies more broadly.

Last, but certainly not least, my deepest appreciation goes to my family, especially my mother Monica for her unwavering support, and my wife Isidora for her love, patience, and encouragement. I could not have completed this journey without them.

Chapter 1

The Effects of Adjusting Socioeconomic Reserves to Local Conditions in Chile's Centralized PreK Choice

1.1 Introduction

Affirmative action is a controversial issue, particularly in the context of admission to selective high school and college programs. Part of the controversy stems from the displacement of non-targeted candidates who scored higher on the selection metric in use, necessitating a balance between social objectives and potential inefficiencies that may arise from enrolling potentially less qualified students (Arcidiacono and Lovenheim, 2016). In extreme cases, targeted students may be worse off by being placed in programs for which they are ill-prepared.[1] In the case of PreK admissions, the debate differs. Although disparities in children's readiness for school may exist, selecting students based on these disparities is much less sensible. Affirmative action policies can aim to counter schools that do select on them or that select on family background, as well as aim to counter systematic differences in family choice resulting from residential sorting or heterogeneous preferences between families in targeted and non-targeted groups. By prioritizing targeted applicants, the objective is to

[1] On the other end to this "mismatch theory" risk (see Sander (2004)), affirmative action policies can not only benefit targeted students but also be efficient, which is likelier when the outside option of non-targeted students is comparatively better (e.g. more expensive private institutions). For example, Otero, Barahona, and Dobbin (2021) study an affirmative action policy in the context of Brazilian college enrollment and estimate that, in terms of their predicted average salaries, average gains by targeted individuals are slightly higher than the average losses of non-targeted ones, which is notable considering the size of reserves being 50% of seats (target marginalized students represent 82% of the overall student population).

reduce segregation and assignment gaps to higher-quality programs to the extent possible.[2] The main challenge at the PreK level lies in effectively achieving this objective as attainable results are constrained by context factors such as residential sorting.[3]

In Chile, the implementation of a new centralized school choice and assignment system (CCAS) aimed to regulate admissions to public and private voucher (subsidized) schools was expected to reduce segregation by eliminating one of its drivers: the possibility of schools discriminating against low-socioeconomic status (low-SeS) applicants. However, the segregation reduction effect of the new system has been limited, even with the incorporation of socioeconomic reserves (SR) as an affirmative action rule that prioritizes 15% of the seats at all schools. The SR give priority to low-SeS applicants who choose popular schools that would otherwise not meet the reserve's seat target while filling all their available seats.

For CCAS like the Chilean, targeted reserves are identified in the literature as the recommended instrument,[4] but evidence on two key elements is lacking. First, the extent to which the effects of SR may be constrained by residential sorting and heterogeneity in family preferences between targeted and un-targeted groups. Second, to maximize their effectiveness, the size of SR at each school, as well as any complementary policies, need to respond to these constraints and to variations in the share of the targeted population across different geographic units, even if residential sorting within each unit is small. Implementing a flat reserve size, as is done in Chile, is sub-optimal.

In this chapter, we estimate family preferences over PreK programs in the Chilean CCAS, focusing on applications in the city of Santiago between 2019 and 2021, and contribute to filling the lack of evidence in three dimensions.

First, we decompose the contribution of each segregation driver and assess its interaction with SR. We find that residential sorting accounts for 40% of the segregation in the absence of SR, while differences in preferences between socioeconomic groups account for the remaining portion. Implementing SR can only close a fraction of the segregation produced by both factors. However, we further show that targeted reserves complement reductions in residential segregation more effectively. When comparing the gap in segregation between implementing only SR and implementing SR along with the elimination of both drivers of dif-

[2] Moreover, school segregation may cause a widening in achievement gaps between groups (Card and Rothstein, 2007; Billings, Deming, and Rockoff, 2013).

[3] From a normative standpoint, the main contention is treating students differently based on their socioeconomic, race, or ethnic status. We disregard this debate and assume less segregation and a smaller achievement gap are strictly preferred.

[4] This instrument is ideal because it acts as a "soft-bound" (as opposed to quotas), in the sense that reserved seats are allocated to other applicants when there are not enough in the target group to fill them. On the contrary, when the target quota in a school is not satisfied, congestion in other schools increases, pushing unassignment levels upwards and reducing welfare (Hafalir, Yenmez, and Yildirim, 2013; Ehlers, Hafalir, Yenmez, and Yildirim, 2014; Echeñique and Yenmez, 2015). More specifically, Ehlers et al. (2014) demonstrate that the "student-proposing deferred acceptance algorithm produces an assignment that Pareto dominates all other fair assignments while eliciting true preferences" (here, fair assignment is the same as the absence of justified envy, which using the priorities determined by the intended affirmative action policies implies that the matching is strictly better for minority students). As Chile uses the student-proposing deferred acceptance mechanism, reserves are the optimal policy.

ferences in choice across socioeconomic groups (residential sorting and preferences), we find that SR and only eliminating residential sorting reduces segregation by around 60% of that gap, compared to around 30% when paired with the elimination of preference heterogeneity between groups.

Our second contribution is to examine the importance of tailoring SR to local conditions. We define the achievable segregation gap as the difference between a no-reserve benchmark and a minimum segregation scenario that accounts for the elimination of both drivers of differences in family choice between socioeconomic groups, and find that, relative to this gap, current SR reduce average segregation at the Municipality level by 3% to 9.2%. However, the impact of SR varies significantly depending on the level of reserves in each school. We estimate that the effect of SR is multiplied by around five if their size equals the share of low-SES applicants in each school's Municipality. Setting SR at such a level is not necessarily optimal, as computing the optimal level would require precisely defining the target segregation minimization objective, which, as discussed in Section 1.4, is not straightforward. The key message nevertheless is that adapting the reserve size to local conditions is important for their effectiveness. To further strengthen this point, we estimate the impact of setting SR at 40% in each school, equivalent to the share of low-SES applicants in the whole city, and find that the average impact is significantly smaller compared to SR set at the Municipality level. Furthermore, in Municipalities with a smaller share of low-SES applicants, segregation can worsen considerably.

Our third contribution is to highlight the strong complementarities between SR and demand-side policies, as SR make the most of each additional low-SES applicant to a congested program under-representing their share. We simulate arbitrarily demand-side policies that, when implemented with reserves, reduce the segregation gap by an additional 20-30% on top of the effect of SR alone. In contrast, implementing them in isolation only achieves a reduction of 4-10%, emphasizing the potential of considering both policies jointly.

Understanding the effects of affirmative action in CCAS is relevant because insights gained from one setting can be applied to others more directly, and the implementation of CCAS is on the rise, as documented by Neilson (2019). While relatively fewer of these systems encompass PreK admissions, the trend towards digitalization will likely increase their prevalence in the coming years. Moreover, promoting integration at the PreK level can have positive externalities for integration in later grades. Additionally, these systems provide an ideal environment to design, implement, and monitor affirmative action policies.

Most of the literature on CCAS has focused on alternative assignment mechanisms and their impact on the welfare of individual families (Pathak and Sönmez, 2008; Abdulkadiroğlu, Che, and Yasuda, 2011; Pathak, 2017; Abdulkadiroğlu et al., 2017; Kapor, Neilson, and Zimmerman, 2020; Elacqua, Jacas, Krussig, Marotta, Méndez, and Neilson, 2022b). More recently, a growing body of research has attempted to disentangle the underlying determinants of segregation and the achievement gap, such as the role of family preferences and residential sorting (Kessel and Olme, 2018; Oosterbeek, Sóvágó, and van der Klaauw, 2021; Laverde, 2021; Son, 2020; Idoux, 2022). Some studies have also examined changes to assignment rules, but to the best of our knowledge, none have focused on targeted reserves.[5] This is partly due to some jurisdictions not allowing such direct affirmative action (e.g., the US in the case of race (Ellison and Pathak, 2021)). However, countries such as Brazil, Chile, Colombia, Ecuador, Peru, The Netherlands, and Sweden have used or continue to use targeted reserves as part of their affirmative action policies.

Similar to other studies, we employ a measurement strategy based on constructing counterfactual assignments (Kessel and Olme, 2018; Son, 2020; Laverde, 2021; Oosterbeek et al., 2021; Idoux, 2022), estimating family preferences to create counterfactual applications under three alternative demand estimation models drawn from the related literature that we adapt to our context and integrate into one unique framework. The reason for using different models is that each has its own advantages and limitations in addressing the "short list challenge", that is, the fact common to other school choice contexts of observing short application lists despite applicants having many different alternatives that they can include in their submitted preference list free of cost. In our context, rank-order lists (ROLs) have an average length of 3.11 programs, despite families having an average of 21 alternatives within a distance equal to or smaller than the farthest school ranked. Moreover, from the on average, 8.13% of applicants that resulted unassigned during the three years studied, 61.6% of them ended up enrolling in a public or voucher alternative available at the time of portfolio choice, highlighting that other options are also acceptable to them (30.2% did not enroll in any school, as PreK is voluntary, and only 8.2% enrolled in a private non-subsidized alternative).[6]

The models implemented differ only in how families form their ROLs. The first assumes that applicants consider all available alternatives and rank all those preferred to the outside option (similar to (Abdulkadiroğlu et al., 2017)). The second assumes that applicants consider all alternatives but face a ranking cost of including them in their ROL (as in Idoux (2022)). The third assumes that applicants face a cost of considering alternatives and thus optimally decide to consider a subset of available options, based on the information they have available, ranking at the end of the process only those considered and preferred to the outside option (similar to Son (2020)).

We show that results obtained with the different models are broadly consistent, and we mainly focus the more detailed comparisons on the last alternative for two reasons. First, in our view it is more realistic in assuming a portfolio formation process where information acquisition and processing is difficult and thus limited. Second, the model also allows us to

[5] Escobar and Huerta (2021) are an exception studying socioeconomic reserves also in Chile, but uniquely changing their size and computing segregation results, without considering geographic heterogeneity nor the interaction of reserves with residential sorting and family preference heterogeneity. Their analysis revolves mainly around comparing theoretically derived predictions of the effect of reserves on the properties of the assignment mechanism with actual results, assuming that preferences would remain unchanged.

[6] The percentages for each year are as follows: 11.59%, 6.37%, and 5.91% applicants were unassigned in 2019, 2020, and 2021 respectively. 61.34%, 61.74%, and 61.74% of those applicants enroll in a public or voucher alternative. 32.24%, 29.78%, and 28.65% did not enroll, and 6.42%, 8.48%, and 9.61% enrolled in a non-subsidized private alternative. Note that non-enrollment levels, as well as school desertion levels, rose in Chile during the COVID-19 pandemic.

simulate policies that reduce information acquisition costs, for example, through information interventions (Hastings and Weinstein, 2008; Andrabi, Das, and Khwaja, 2017; Allende, Gallego, and Neilson, 2019; Arteaga, Elacqua, Krussig, Méndez, and Neilson, 2022). It is worth noting that the results with the alternative models would nevertheless be very similar given their consistency observed in the main counterfactuals. As a result, key takeaways of the additional counterfactuals would be the same and focusing on the costly consideration alternative is not significant for the analysis.

One limitation of our study to consider is that we focus only on the Main assignment round and do not examine the aftermarket dynamics that may shape some of our conclusions. The reason is due to data restrictions and assumptions required to predict counterfactual behavior beyond the main round. To address the potential bias resulting from this limitation, we report an alternative segregation measure that includes unassigned applicants in each Municipality.[7] Future research should investigate the effects of these dynamics on our findings in more detail.[8]

As for the rest of this chapter, Section 1.2 discusses the Chilean educational context, focusing on the Inclusion Law that regulates the CCAS. Here we also introduce the segregation measures used in this chapter. Section 1.3 details our available data, while, given the relevance of challenges posed by the geographic context, Section 1.4 is dedicated to explaining the methodology used to determine the geographic units used for preference estimation and the criteria used to select applicants included in counterfactuals, as well as the rationale for using Municipalities as our geographic unit to measure segregation. Section 1.5 lays out the models and estimation results, Section 1.6 covers the different counterfactuals implemented and presents our main results, and Section 1.7 concludes.

For readers unfamiliar with SR or interested in better understanding how they operate, Appendix A.9 describes the deferred acceptance algorithm used in Chile (see Correa, Epstein, Escobar, Rios, Bahamondes, Bonet, Epstein, Aramayo, Castillo, Cristi, and Epstein (2019) for further detail).

1.2 The Chilean Educational Context

In 1981, a significant reform introduced competition between education providers, establishing the structure of the current educational system. The reform decentralized the administration of public schools, transferring responsibility from the state to Municipalities, which are the smallest administrative units of the country. The reform also introduced a

[7] Thus, for example, if setting socioeconomic reserves closer to the share of low-SeS applicants in the Municipality increases the unevenness of unassignment between socioeconomic groups, the segregation measure is reduced by less. Moreover, including unassigned applicants, we keep the pool of applicants in the segregation computation of the different counterfactuals roughly constant (some variation is due to changes in the Municipality of the school to which an applicant is assigned). That increases the comparability of results.

[8] The impact of this limitation on other outcomes considered in this chapter, related to the efficiency of the assignment, is relatively small because levels of unassignment are not substantial, especially for the 2020 and 2021 years affected by the COVID-19 pandemic, having fewer overall applicants.

voucher-based system linked to schools' enrollment, providing funds for public and a fraction of private schools. There are thus three types of schools in this system: public, subsidized, and private (non-subsidized). Private subsidized or private voucher schools were allowed to charge unrestricted top-off fees over the government subsidies.[9]

After 1990, the Chilean education system underwent two significant reforms. The first was the enactment of the Preferential School Voucher (SEP) in 2008. Prior to this policy, the voucher linked to enrollment was the same for all students. The SEP reform introduced a targeted voucher for more vulnerable students, which was added to the flat subsidy (see Neilson (2021)).[10] This change increased schools' incentives to enroll low-SES students, whose categorization is determined by measures linked to social protection programs (in the CCAS, low-SES status determines priority for SR).[11]

The Inclusion Law Reform

In 2015, the Inclusion Law was enacted with three primary goals: to eliminate for-profit institutions from subsidized education, gradually eliminate top-off fees, and regulate the application and enrollment processes of public and subsidized schools through a centralized system. To achieve the first objective, for-profit private voucher schools were required to switch to non-profit entities to access public resources. The second objective was accomplished by freezing top-off fees and implementing a schedule for their gradual reduction, which was supported by public funding. The third objective involved the implementation of the School Admission System (SAE, for its initials in Spanish), the CCAS studied in this chapter, which was gradually rolled out starting with entry grades in the Magallanes region

[9] The Chilean educational system comprises three levels: preschool, primary, and secondary education, with schools able to offer any combination of grades. Preschool education includes nurseries, pre-kindergarten (PreK), and kindergarten, none of which are mandatory. There was a law project aiming to make kindergarten compulsory since 2013, but it was rejected on September 2021 by the Congress' Chamber of Deputies. However, according to 2017 data from the population census and the enrollment data, only 7% of children do not attend PreK, and only 1% do not attend kindergarten. These numbers rose, however, during the COVID-19 pandemic. It is important to note that families must make strategic enrollment decisions, as finding a desirable school during the PreK application process is easier due to the greater number of available seats relative to kindergarten admissions the following year. Primary education spans from first to eighth grade, with secondary education consisting of four grades covering the academic material in the college entry exam (PAES). While students can change schools at any grade, the main entry points to schools in the system are PreK and the first grade of secondary education.

[10] The SEP Law defines low-SeS or "priority" students as "students with a family's socioeconomic situation that hampers the possibilities of facing the educational process." Concretely, this represents around 40% of students in the city of Santiago.

[11] The SEP reform also enhanced accountability and control over the use of public resources. To receive additional funds for their low-SeS students, schools must sign an agreement with the Ministry of Education. The agreement requires schools to eliminate top-off payments for these students and use the extra resources to implement educational improvement plans specifically targeting them. This agreement is renewable and lasts for four years, subject to specific academic goals.

in 2016 and all the region's grades in 2017.[12]

By regulating the school admission process and eradicating the possibility of schools selecting students based on their family background, the system was expected to contribute significantly to reducing segregation.[13] However, evidence suggests that the new system's success in reducing school segregation has been, at best, moderate. Kutscher, Nath, and Urzua (2020) use a difference in differences strategy to measure the SAE's impact on segregation, leveraging the progressive policy roll-out in different regions of the country and find that there is "(no) overall improvement in the evenness of the student background distribution". Their findings also suggest that Municipalities with higher residential segregation or a larger share of private non-subsidized schools experienced an uptick in school segregation.

However, caution should be exercised when interpreting Kutscher et al. (2020) findings. They only study segregation in 9th-grade enrollment and schools offering 9th grade are fewer and enroll more students than those offering PreK. Moreover, families are more likely to enroll in schools further from home when students are older. Furthermore, recent difference-in-differences literature highlights some challenges in using that methodology to study the effect of the Inclusion law reform in Chile that may not be adequately incorporated in Kutscher et al. (2020) empirical approach (see Roth, Sant'Anna, Bilinski, and Poe (2022) for a summary of that literature).[14] In any case, substantial levels of segregation in Chilean school enrollment remain.

The SAE uses the Deferred Acceptance (DA) algorithm by Gale and Shapley (1962), starting the assignment with applicants to the highest grade (4th year of secondary education) and finishing assigning PreK applicants. Total available seats in a program are distributed in reserve groups, with the priority orderings shown in Table 1.1.[15]

The Ministry of Education's Decree Law 196 from 2006 required private voucher institutions to have a student body with at least 15% low-SeS (or "priority") students to receive

[12] The entry levels of four more regions (Tarapacá, Coquimbo, O'Higgins, and Los Lagos) were added in 2017 (and all their grades in 2018), ten more regions in 2018 and all their grades in 2019 (Antofagasta, Atacama, Valparaíso, El Maule, Bío Bío, La Araucanía, Aysén, Los Ríos, Arica and Parinacota and in 2020 the Region of El Ñuble was created, whose Comunas are in this group as part of the Bío Bío region). The Metropolitan Region entry levels were added in 2019, covering all grades in public and private voucher schools, for the first time in 2020.

[13] A quote from Adriana del Piano, the Minister of Education that was in office during the Congress' discussion of the Inclusion Law project is illustrative (own translation): "the project aims to improve the quality of education children receive, independently from the school they attend, and to eliminate structural segregation factors, like the ones that existed before this law. Some families have paid to segregate their children (...). The country has a system that generates segregation. Our aim with this law is to implement an integrated system, like the ones implemented in most countries around the world, that guarantees peer effects, which are very important for learning, knowledge, and coexistence within schools. The idea is to give students from different socioeconomic backgrounds the possibility of attending the same school, (...)."

[14] Specifically, the year of treatment differs by region, and it is reasonable to expect geographic and time heterogeneity in treatment effects. Moreover, it is unlikely that the "parallel trends" assumption holds, and therefore, its validity needs to be thoroughly examined.

[15] In the 2022 application process reserves for applicants with a disability were eliminated.

public funding.[16] Compliance with the Decree was loosely monitored, but when implementing the SAE, this percentage was adopted as the reserve group size to comply with the spirit of existing legislation. Since the low-SeS population share is significantly higher in many cities of the country, and given that reserves operate as a soft bound that is only active when low-SeS applicants are interested in a given program, revising their size to better achieve greater equity in school admissions is a clear opportunity.

In regards to the other reserve groups, the Academic Achievement group has no direct relevance for PreK applicants and has a minor impact due to its limited implementation in select schools. Meanwhile, the disability group is comprised of only two seats per academic program class and is applicable to a relatively small population. As a result, our focus will solely be on groups three and four of Table 1.1.

Enrollment priorities are determined in the following order: secured enrollment priority is given to students currently enrolled in a school that offers their next corresponding grade (not applicable for PreK, except for students repeating the grade).[17] Sibling priority comes next and is only applicable to students who have the same mother or father. This priority is distinguished by whether the sibling is already enrolled in the school (static priority) or is assigned to a higher grade during the assignment process (dynamic priority). In reserve group three, priority is then given to all low-SeS applicants. The last two priorities in all reserve groups are for applicants with at least one parent working at the school, followed by the priority for those who were previously enrolled in the school during their education (not relevant for PreK). Lotteries are used to break ties in priority groups, as is standard in such systems.[18]

To understand how the reserve groups function, it is helpful to consider each program and reserve group combination as a distinct "new program". The algorithm processes applicants in sequence for these program-group combinations using the applicant's rank-order list (ROL), and group three followed by group four as the default processing order. However, if an applicant has sibling priority (static or dynamic) but not low-SeS priority, they will first be considered for the Regular group and then for group three, in order to avoid using low-SeS reserves for non-low-SeS applicants with higher priority.[19]

[16] The Decree is accessible at this url: https://bcn.cl/3cfy9.

[17] That ensures that if the student does not get a seat in a more preferred school, she does not lose her automatic enrollment for next year.

[18] The system employs a sibling multiple tie-breaking rule at the school level. This means that a different lottery number is given to each group (of one or more) of siblings that share a school in their application (the lottery number is the same for all the school's programs). In the case of siblings applying to the same grade, a different lottery is used to determine which sibling has a better lottery number.

[19] Students with secured enrollment are assigned to the group where they generate the seat first, depending on their characteristics such as having a disability, being high academic achievers, or being low-SeS. Additionally, families with two or more siblings applying to different grades may choose a "family application". In this case, the ROLs of students applying to lower grades are reordered to prioritize schools where a sibling has already been assigned by the algorithm above others while preserving the original order if it applies to more than one program. For a more comprehensive explanation of the deferred acceptance algorithm utilized in Chile, please refer to Correa et al. (2019), as well as the algorithm description provided in Appendix A.9.

The SAE has two application rounds: the Main and the Complementary. Students who did not participate in the Main round or were not assigned to any program during this round can apply for the remaining seats during the Complementary process. During both rounds, students can apply to as many programs as they wish, with a minimum of two at entry levels. If an applicant is not assigned to any program in their ROL during the Complementary round, a distance-to-school minimization algorithm is used to offer an alternative program. If students want to enroll in a different program, or are not assigned to any program after the Complementary round, they can register at a public registrar at each school. They will then be processed manually based on their order of registration when seats become available in their desired program.

Segregation Measures

Depending on the data and objectives, there are several ways to measure segregation. For socioeconomic segregation, the most commonly used is the dissimilarity index proposed by Duncan and Duncan (1955). In our context, it measures the unevenness of low-SeS students across schools in a given geographic unit, which, in our case, will be Municipalities.

Duncan & Duncan Dissimilarity Index (DD index)

For schools s within a specific Municipality m, the index is computed as:

$$SEG^{DD}_{m,t} = \frac{1}{2}\sum_{s=1}^{S}\left|\frac{\textit{Low-SeS}_{s,t}}{\textit{Low-SeS}_{m,t}} - \frac{\textit{No-SeS}_{s,t}}{\textit{No-SeS}_{m,t}}\right|$$

We define $\textit{Low-SeS}_{g,t}$ as the number of assigned students with low-SeS status in a given school or Municipality unit g during period t, and $\textit{No-SeS}_{g,t}$ as the number of assigned students without low-SeS status priority in the same unit during the same period.

The dissimilarity index measures the percentage of the minority population that needs to be reallocated to achieve an even distribution across schools. However, one drawback of this measure is that it does not account for the scale of the necessary redistribution in relation to the overall population. As noted by Meister and Niebuhr (2021), in certain cases, the index's value may decrease mechanically as the proportion of the minority group increases, i.e., as both groups approach 50%. To illustrate this point, consider the following example: School A has two low-SeS and 100 no-SeS assigned students, while School B has 100 assigned students, none of whom have low-SeS. The value of the dissimilarity index in this example is $^1/_2$. By moving one of the two low-SeS students from School A to School B, we can achieve an equal fraction of $^1/_{100}$ in each school, which requires reallocating around 0.5% of the overall student population. In contrast, suppose there are 20 low-SeS students, with 19 assigned to School A and an equal number of 100 no-SeS assigned in each school. In this scenario, the value of the index drops to 0.45, a 10% reduction. To achieve an even distribution, we need to reallocate 9 of the 20 low-SeS students from School A to School B. However, the magnitude of the segregation problem is much higher, requiring the reallocation of approximately 5% of the student population to achieve an even distribution.

The example above also illustrates that the DD index does not consider school capacity constraints (whether School B has the physical capacity to accommodate nine additional students).

Minority-size Adjusted Dissimilarity Index (MA-DD index)

To address these shortcomings of the DD index, van Mourik, Poot, and Siegers (1989) proposes the minority-adjusted dissimilarity index. The index calculates the minimum proportion of students in both socioeconomic groups that would have to relocate to achieve an even distribution across schools, subject to the constraint that the number of students in each school remains unchanged (Nijkamp and Poot, 2015). As van Mourik et al. (1989) demonstrate, the index is a function of the dissimilarity index:

$$
\begin{aligned}
SEG^{MA\text{-}DD}_{m,t} &= 2 f^{SeS}_{m,t}(1 - f^{SeS}_{m,t}) SEG^{DD}_{m,t} \\
f^{SeS}_{g,t} &\equiv \frac{Low\text{-}SeS_{g,t}}{Low\text{-}SeS_{g,t} + No\text{-}SeS_{g,t}}
\end{aligned}
$$

We can intuitively see that the correction factor $f^{SeS}_{m,t}(1 - f^{SeS}_{m,t})$ is largest when both groups have an equal number of individuals, which is when a given value of the dissimilarity index -expressing the share of the minority population to be reallocated- is applied to the largest share of individuals. Following our example above, the index increases from 0.0098 to 0.0744 to reflect the magnitude of the reshuffling multiplying by almost eight.

We prefer using the MA-DD index to interpret our results as it provides a more precise measurement of the segregation issue across Municipalities in terms of the proportion of the population that would need to relocate to achieve an even distribution of students. However, the dissimilarity index also offers valuable insights as it captures the concentration of low-SeS students, and therefore we present both indices in our results.

MAPPD and MA-DD equivalence

The Mean absolute percentage point difference (MAPPD) is another alternative used in the literature to measure school segregation. For instance, Margolis and Hashim (2020) used this index, and it was slightly adapted by Idoux (2022) to account for the difference between assigned and enrolled students. The index measures the absolute difference between the share of applicants from a particular socioeconomic group at a school and the share of the group in the geographic unit of interest.[20] However, when there are only two socioeconomic groups, this index is just a reformulation of the MA-DD index, which can be demonstrated by rearranging the terms in each formulation.

Finally, note that in this chapter we examine school-level segregation instead of program-level segregation. When students and their families are in the same school and grade, they are highly likely to interact and form strong connections during PreK or in later grades. While this likelihood is slightly lower than if they were in the same program within a school, it is still significant and much higher than if they were assigned to different schools. Educational segregation dynamics in Chile are briefly discussed in Appendix A.11.

1.3 Data

Administrative Data

Schools and Programs

Each year, the Ministry of Education publishes a directory of certified schools containing each school's name and identification code and some general characteristics like geolocation, grades offered, rurality, enrollment, and payment ranges for tuition and monthly fees. In the geographic area used in our analysis, there were 3,713 schools in 2019, 1,517 of which offered PreK. In 2020, there were 3,781 schools and 1,496 that had PreK, and in 2021 there were 3,827 schools and 1,487 offering PreK.

The Ministry of Education publishes the SAE's list of available programs to choose from during the application process. The unique combination of school, location (campus), grade, specialty, gender, and school day (full-day, morning, or evening) defines a program. For our analysis, we keep only public and voucher schools that offer PreK on regular educational programs (omitting programs for students with special needs). In 2019, the system had 1,010 public and private voucher schools offering 1,237 PreK programs. In 2020, there were 1,012 schools with 1,247 PreK programs, and in 2021, 1,015 schools with 1,249 programs.

For each school -and all its programs- we gather information about their SEP policy agreement, average fees charged[21] and school performance -the achievement category-, which we use as a proxy for school quality. The SEP agreement and average fees are public

[20] Mathematically, it is expressed as:

$$SEG_{m,t}^{MAPPD} = \sum_{s=1}^{S} \pi_{s,t} \left| f_{s,t}^{Low\text{-}SeS} - f_{m,t}^{Low\text{-}SeS} \right|$$

$$\pi_{s,t} \equiv \frac{Low\text{-}SeS_{s,t} + No\text{-}SeS_{s,t}}{Low\text{-}SeS_{m,t} + No\text{-}SeS_{m,t}}$$

[21] Fees contained missing values, so we imputed them using i) rates of increase or decrease within the years available if the school has 2 out of 3 years of information, and ii) rates of increase or decrease at the Municipality level using the information of other schools that charge fees, for schools that have only one year of data.

information published annually by the Ministry of Education, as is the achievement category issued by the Quality Agency of Education. The latter, however, hasn't been updated since 2019, so we use 2019 values for all years. Table 1.2 summarizes the main programs' characteristics used in the demand estimation.

The achievement categories are high, medium, medium-low, and low. We group the latter two in our analysis. Most programs are during the morning, a similar share during the afternoon, with a remainder of around 18% of full-day programs. Only a decreasing fraction of schools charge top-off fees over public subsidies (20% in 2019, 19% in 2020, and 16% in 2021). The percentage of these schools inscribed in the SEP policy -in which low-SeS applicants are free of tuition- was 44% in 2019, 42% in 2020, and 45% in 2021.

Applicants

The centralized assignment system data includes information for all applicants and their applications. We observe low-SeS status as determined by the SEP policy, geolocation with its quality (accuracy),[22] gender, special needs, and priorities. The Municipalities in our sample include 29,524 applicants for PreK in 2019, 25,235 in 2020, and 24,057 in 2021. Of these applicants, we use a fraction in preference estimation and counterfactuals. The criteria to determine these two groups of applicants and their characteristics will be discussed in Section 1.4, after characterizing the geographical context.

Expected Assignment Probabilities

As described in Section 1.5, the expected assignment probabilities of applicants for different programs are important in preference estimation and counterfactual analysis for both the costly consideration and costly ranking models. These probabilities are calculated from observed applications and available seats in each program. However, accurately estimating these probabilities is challenging, as discussed in detail in Appendix A.10. Specifically, the interaction between the heterogeneity in priorities and the fact that lotteries of applicants unassigned to a quota are, on average, worse when considered for another makes it difficult to identify probabilities without using lottery cutoff values. Unfortunately, lottery cutoff values tend to underestimate assignment probabilities, and correcting this bias is not straightforward.

Without considering the challenges anticipated in the previous paragraph, estimating assignment probabilities consists of re-drawing the applicant pool and applicant lotteries one thousand times,[23] obtaining empirical assignment probabilities in each of these simulations and then averaging them out.

[22] Geolocation public data is not perfectly accurate, as each coordinate has a random noise to avoid the identification of applicant addresses. Nevertheless, the noise is relatively small, and we interpret it as part of the measurement error in the relevant distance to schools that we approximate using the linear distance between family and school coordinates.

[23] And one-hundred times in each counterfactual loop to obtain equilibrium probabilities and applications in the simulated counterfactuals of the costly consideration and costly ranking models

1.4 Geographical Context

Santiago is a large city with diverse neighborhoods that present different local realities. Although the CCAS does not include any consideration for applicant residences, choices and segregation dynamics are strongly influenced by city characteristics. This poses a challenge for determining the appropriate geographic levels for our study.

When studying the supply-side dynamics of education in Santiago, optimal decisions of schools regarding the quality of education offered or prices charged are an equilibrium result that is influenced by the interactions of all schools in the city. Thus, studies such as Neilson (2021), focusing on understanding supply responses, in his case to the SEP policy, must consider all contiguous urban areas when defining geographic units of interest. However, our study has different objectives. We aim to model individual choice counterfactuals under fixed supply conditions and understand segregation in schools (and other outcomes) that are not meaningfully comparable across schools in distant areas of the city. To achieve this objective, we propose constraining the set of alternatives that applicants consider when forming their portfolio choices to those within a "reasonable distance" from their residence.

There are different ways to restrict the geographic space of school choice and to partition a large city into different geographic units to study segregation. The appropriate approach depends on the context and normative considerations. In Santiago, we suggest leveraging the administrative units that the city is divided into (Municipalities), for two reasons (Abdulkadiroğlu, Pathak, Schellenberg, and Walters (2020) use a similar approach in estimating preferences at the Borough level for New York City). First, local governments are elected at this level, meaning that neighborhoods located on either side of a Municipality border are impacted. Additionally, Municipalities are responsible for administrating public schools within their territory.[24] Second, families are aware of these administrative units and their borders because elections are held at this level, and they are often used to describe various social outcomes across the city, such as crime, average income, and public goods provision. Moreover, choosing the Municipality of interest is one of the search criteria in the application interface. However, using Municipalities is not a perfect solution, and we need to consider its impact on our results. To discuss this, Figure 1.1 shows Santiago's map divided into different Municipalities. Panel A's points indicate the locations of schools that offer PreK programs, and Panel B shows the share of low-SeS applicants in each Municipality.

The Municipalities in Santiago vary in size and in their share of low-SES applicants. This means that, especially in smaller Municipalities that are adjacent to others with different low-SES population shares, interactions with schools and applicants in neighboring areas can strongly influence segregation measures and other outcomes. Conversely, larger Municipalities that are closer to the city limits tend to be more insulated from the impact of the rest of the city.

[24] There is an ongoing project to change the administration of public schools within geographic units not always identical to their Municipality to a specialized public entity. It has no significant impact on the schools during our period.

To gain a better understanding of how these dynamics affect our results, throughout the chapter we will exemplify using two Municipalities: Maipú and La Cisterna (indicated in Figure 1.1). Although these two Municipalities have a similar proportion of low-SES applicants (32.22% in Maipú and 35.37% in La Cisterna), they differ in size and exposure to neighboring Municipalities. Specifically, we will observe that results in Maipú are consistent with average results for the different Municipalities (and with what we expect from our counterfactuals), while results in La Cisterna differ as they are heavily influenced by the interactions with other geographic units.

Definition of Municipality + Buffer

In the Metropolitan region, 92% of PreK urban applicants list a school in their Municipality as their first choice. However, this percentage varies across different Municipalities. To determine the choice sets, we assume that all applicants in a given Municipality have the same set of schools, including all PreKs offered in that Municipality and some others likely to be relevant in neighboring ones.[25] We define a buffer radius around the Municipality set such that 90% of schools in the first preference lie within the Municipality plus buffer geographical unit. A more detailed description of this process is outlined in Appendix A.3. In the case of the Municipality of Maipú, no buffer is required to reach the target, while the buffer for the case of La Cisterna in presented in Figure 1.2. The area of urban Maipú is approximately $58km^2$, while in La Cisterna, the area within the Municipality limits is approximately $10km^2$, and that encompassed by its buffer is approximately $69km^2$. The figure also shows schools considered in the different exercises. In the case of La Cisterna, only 16.5% of schools (33 schools) are located within the administrative limits of the Municipality, while 83.5% (167 schools) are located in 7 adjacent Municipalities.[26]

Applicant Sample for Demand Estimation and Counterfactuals

To estimate preferences and counterfactuals, we restrict the set of applicants within each Municipality to those who i) have a high-accuracy georeference,[27] and ii) apply only to schools within the Municipality plus buffer. The reason is that it is unfeasible to accurately study the choices of applicants not meeting either one of the criteria (at least without additional information). Table A.2.4 in Appendix A.2 shows the share of applicants included in each Municipality and the share of low-SES applicants in the whole population of each Municipality and in the restricted sample (for all years, 2019-2021). Additionally, Table 1.3 compares the sample population with those not included in the estimation and counterfactuals for the whole city. The share of applicants in the estimation and counterfactual sample

[25] This assumption simplifies estimation without dramatically affecting our results as the added programs are still not too far from applicant residences.

[26] The buffer zone even contains schools from the Municipalities of San Bernardo and La Pintana, which are not adjacent, but are very close.

[27] A more detailed description of the different georeference quality levels is explained in Appendix A.3.

varies across Municipalities, as does the share of low-SeS applicants and the characteristics of applications.[28]

1.5 Preference Estimation

We first describe the specification of the three models implemented in this chapter and then move to presenting the estimation results.

Model Specifications

Let v_{ijt} to denote the utility that applicant i would derive from being assigned to program j in application year t. We separate this utility in three additive parts. First, the utility that applicants with the same characteristics as i derive, on average, from an assignment to j, denoted with δ_{ijt}. Second, the individual idiosyncratic preferences that the applicant has for the program ϵ_{ijt}. And third, a dis-utility portion proportional to the distance between the applicant's residence and the school offering the program denoted by d_{ijt}. Utilities can be thus expressed as:

$$\begin{aligned} v_{ijt} &= \underbrace{\delta_{ijt} + \epsilon_{ijt}}_{u_{ijt}} - \tau_{it} d_{ijt} \\ v_{i0t} &= 0 \end{aligned}$$

We impose a location normalization over indirect utilities by assuming that applicants have an outside option with zero utility ($v_{i0t} = 0$). As a result, only programs with $v_{ijt} \geq 0$ are acceptable to be included in i's reported portfolio.

Under this utility structure, additively separable in u_{ijt} and $\tau_{it} d_{ijt}$, using g_i to denote a group of applicants with the same observable characteristics as applicant i, if $\tau_{it} = \tau_{g_i t}$ and g_i is sufficiently large, variation in distances to different programs and relative rankings across applicants allows the non-parametric identification of utilities under the assumption that idiosyncratic tastes and distance to schools are independent, conditional on the rest of the variables affecting utility (Agarwal and Somaini, 2018):

$$\epsilon_{ijt} \perp d_{ijt} | \delta_{ijt}$$

[28] The applicants used in demand estimation have shorter distances to their preferred and assigned schools, more often receive assignments to the Municipality in which they reside, and have higher exposure to both general and high achievement schools. They also have a higher proportion of low-SeS applicants and submit slightly shorter applications. In contrast, those not included in the demand estimation are less likely to be low-SeS, apply to more schools, and are less likely to be assigned within their Municipality. These differences may be due to income-related factors (those willing to travel to more distant schools are likely wealthier), such as car ownership and transportation costs.

This assumption, that we upheld across all three models specifications described below, is violated if family residences are influenced by idiosyncratic tastes for programs conditional on δ_{ijt}. Essentially, distance serves as a "special regressor" (as defined by Agarwal and Somaini (2018)) since small variations in its value for different programs, along with the relative ranking of these alternatives, allow the mapping of the share of utilities taking specific values.

For our empirical application we divide applicants into two groups denoted by g_i (low-SeS or not), and specify utilities as presented in equation 1.1:

$$v_{ijt} = \underbrace{\underbrace{Family_{ijt}\lambda + x_{jt}\beta^{g_i} + \xi_j^{g_i}}_{\delta_{ijt}} + \epsilon_{ijt}}_{u_{ijt}} - \underbrace{\left(1 + \Delta_{d,C19}\mathbf{1}_{t>2019}\right)}_{\tau_t} d_{ijt} \tag{1.1}$$

$Family_{ijt}$ represents the family-related priorities of applicant i in program j at time t, which includes sibling and parent-worker priorities. ξ_j^g is a program-specific vertical coefficient, unobservable to the econometrician. Observable attributes of a program in year t are represented by x_{jt}.

As explained below, we assume that ϵ_{ijt} follows a normal distribution, estimating its variance along with other parameters. Thus, unlike logit models, we do not impose a scale normalization over the distribution of ϵ_{ijt} and instead impose a scale normalization over the distance to school parameter,[29] which is measured in kilometers of linear distance between the applicant's residence and the program's school. As can be observed in equation 1.1, we assume that, in 2019, each additional kilometer reduced utility by a factor of -1, while, for 2020 and 2021, we introduce the parameter $\Delta_{d,C19}$ to account for the possible impact of the COVID-19 pandemic on distance dis-utility. Note that, there is no relevant impact of assuming a common distance dis-utility parameters for both socioeconomic groups ($\tau_{g_i t} = \tau_t$), as it only normalizes the values of parameters in δ_{ijt} that do depend on the applicant's group.

Portfolio formation

The three models considered in this chapter differ in the portfolio formation process. The "full consideration" model, based on Abdulkadiroğlu et al. (2017), assumes that families have complete information about the v_{ijt} values for all programs and form their ROLs without any frictions. Unlike their model, we do not impose an exogenously fixed length of ROLs and instead assume that all alternatives preferred to the outside option are ranked. As a result, unranked programs have negative utilities rather than utilities below that of the program ranked last.

[29]This normalization is standard in the school choice literature implementing the same Markov-Chain Monte-Carlo estimation technique we use in this chapter (e.g. Abdulkadiroğlu et al. (2017); Idoux (2022)).

In the costly ranking case,[30] families are assumed to perfectly know their utility if assigned to any program, but face a cognitive or information processing cost when adding programs to their ranking. We specifically assume that the total ranking cost increases linearly at a rate of c_i with the number of programs ranked. This cost implies that some programs with positive utility might not be included in the ROL if they do not sufficiently increase the portfolio's expected utility, taking assignment probabilities into account.

Regarding the model of costly consideration, rather than assuming that consideration is determined by a latent variable as in Son (2020), we assume that applicants need to incur the c_i cost to learn the value of $\epsilon_{ijt} + \xi_{jt} \equiv \theta_{ijt}$.[31] If, after learning θ_{ijt}, the utility is above the outside option, the program is included in the applicant's ranking free of cost. The rationale for this specific assumption is that the information observed by the econometrician perfectly coincides with the information displayed on the Government applications website and that this information is easy to process for applicants. Learning θ_{ijt} requires effort to better understand the program's unique characteristics and the quality of the family-program match. Only after learning θ_{ijt}, families can decide whether to include the program in their ROL based on whether it provides higher utility than their outside option (we assume that the cost of ranking programs without bearing the cost is infinite). We also assume that $\epsilon_{ijt} \perp \xi_{jt}$ and that, for simplicity, the distribution used by applicants to compute expectations is the same that we estimate for these two parameters under the normality assumption used in estimation ($\sigma^2_\theta = \sigma^2_\epsilon + \sigma^2_{\xi^g}$) detailed below.

In both the costly ranking and costly consideration alternatives, we model c_i using the same structure, following Idoux (2022), presented in Equation 1.2:[32]

$$\begin{aligned} c_i =& c^{g_i} + \zeta_i \\ \zeta_i \sim& TN(0, \sigma^2_\zeta, -c^{g_i}, UB_{\zeta_i}) \end{aligned} \tag{1.2}$$

To complete the costly ranking and costly consideration models, we need to determine how applicants evaluate and decide which alternative to rank or consider next.

In the costly consideration case, since evaluating each alternative incurs an incremental cost and provides additional information, the optimal evaluation process is sequential. Applicants learn about the unknown program characteristics one by one and decide which program (if any) to consider next. We will assume that applicants are risk neutral and that they are endowed with the information of programs in which they have a family-related pri-

[30] This model has been introduced in other papers such as Fack, Grenet, and He (2019); Idoux (2022), following the rationale in Chade and Smith (2006). We implement a specification very close to that of Idoux (2022).

[31] Re-arranging the utility equation for expositional clarity of this assumption:

$$v_{ijt} = \underbrace{Family_{ijt}\lambda + x_{jt}\beta^{g_i} - (1 + \Delta_{d,C19}\mathbf{1}_{t>2019})d_{ijt}}_{\equiv v_{ijt} - \theta_{ijt},\ known\ before\ consideration} + \underbrace{\xi^{g_i}_j + \epsilon_{ijt}}_{\equiv \theta_{ijt}}$$

[32] The function $TN(\mu, \sigma^2, LB, UB)$ represents a truncated normal distribution with mean μ, variance σ^2, and truncated between a lower bound LB and an upper bound UB.

ority, and as a result use the decision-making rule in Definition 1 to evaluate which program to consider next. The applicant's partial ranking corresponds to a subset of the submitted ROL, formed by the programs in it already considered at a given point during the portfolio formation process.

Definition 1. ***Optimal decision-making rule in costly consideration model***
Given a (possibly empty) partial ranking of the alternatives considered so far with utility above that of the outside option, the next program to be considered is the one that increases the expected utility of the applicant's partial portfolio the most, provided that the increase is above the consideration cost.

In the case of the costly ranking model, applicants have complete information regarding their utilities for all alternatives. The process is thus not sequential as they need to find the subset of programs that maximizes their expected utility. However, the computation burden of comparing all possible subsets is extremely large as the set of ROL permutations grows exponentially with the size of the choice set, making it an NP-hard problem. To address this issue, Idoux (2022) proposes a sub-optimal sequential decision-making heuristic, which we adopt in this chapter.[33]

Definition 2. ***Decision-making heuristic in costly ranking model***

1. *Starting with an empty partial ranking, applicants choose for the first position the alternative with the highest utility among those that have an expected utility (assignment probability times utility if assigned) above the ranking cost.*

2. *To determine the next ranking positions, programs not included in the partial ranking where the multiplication of the probability of non-assignment to any program in the partial ranking and the expected utility is greater than the ranking cost are candidates, and the program with the highest utility among this set is selected. If the set is empty, the process ends.*

As a result of this heuristic, applicants will potentially neglect programs with a relatively lower utility but a higher assignment probability than others included in the ranking, even though those programs would be part of the optimal portfolio.

Discussion

The strengths of the costly ranking and costly consideration models are that they introduce frictions in the portfolio formation process that help rationalize short ROLs without having to either assume that ranking length is exogenously fixed or that all programs not ranked are less preferred than the outside option. And in the case of the costly consideration model,

[33] Alternatively, other studies using similar models, such as Larroucau and Rios (2020), use context-specific criteria to limit the space of ROLs considered.

the partial information structure it assumes is, in our view, the more realistic. However, a weakness that our implementation shares for both models is that we assume that families accurately know their assignment probabilities to different programs. However, recent evidence suggests that families are, on average, optimistic about their assignment chances.[34]

Additionally, in the case of the costly consideration model, the introduction of incomplete information in the portfolio formation process generates a cost. As explained in Appendix A.5, the model is not identified non-parametrically, resulting in estimated parameters depending on functional forms and on included covariates that are assumed to be observed when deciding whether to pay the consideration cost. Further differences between the models are discussed in Appendix A.6.

Model to data mappings

We can use the models and portfolio formation processes described earlier to determine how observed applications impose constraints on utilities, as well as on ranking and consideration costs. For the full consideration model, the relationship between observed ROLs and utilities is straightforward. Programs not included in a submitted portfolio have negative utility, while programs included have a positive utility that increases with the rank position. To express this restriction, we denote the program ranked in position r as k_r:[35]

$$\begin{aligned} j \notin ROL_i &\Leftrightarrow v_{ijt} < 0 \\ \infty > v_{ik_1t} &> \ldots > v_{ik_{|ROL_i|}t} \geq 0 \end{aligned}$$

To account for portfolio formation frictions in the other models, two additional elements must be incorporated into the mapping implied by observed applications. First, some or all programs not ranked may have positive utility. Second, the mapping must also consider the value of c_i, along with restrictions on utilities. We first introduce these elements in the costly ranking case and then explain the differences in the costly consideration alternative, focusing on that model's mapping and its intuition. A more detailed and formal derivation is presented in Appendix A.4.[36]

However, before moving to the other models, it is necessary to introduce q_{ijt} to represent applicant i's subjective probability of being assigned to program j if processed by the assignment algorithm to that program, i.e., if not assigned to something more preferred, or

[34] For instance, Arteaga, Kapor, Neilson, and Zimmerman (2021) show this for the case of Chile, and Corcoran, Jennings, Cohodes, and Sattin-Bajaj (2018) and Kapor et al. (2020) for the US cities of New York and New Haven, respectively. Other studies use survey evidence, as in Kapor et al. (2020), or additional identification assumptions and data variation, as in Son (2020), to simultaneously estimate preferences and biases in subjective assignment probabilities. We lack data for such an approach and thus leave the incorporation of subjective preference estimation for future research, using data similar to that of Arteaga et al. (2021) in the case of Chile, for example.

[35] $|ROL_i|$ represents the ranking's length and thus $k_{|ROL_i|}$ the program ranked last.

[36] And in the case of the costly ranking model, further details can be found in Idoux (2022).

the probability of assignment if j is ranked in the first position. We assume that applicants have the best possible rational guess of q_{ijt}, which is computed as described in Appendix A.10. This implies that all applicants with the same priorities in the different reserve seat groups of a program have the same q_{ijt}.

Costly ranking

In the costly ranking model, three restrictions are added to the utility inequalities for ranked programs:

Result 1. ***Costly ranking model to data mapping***

1. *Utilities of ranked programs are above that of the outside option and ordered increasing in ranking position:*

$$\infty > v_{ik_1t} > ... > v_{ik_{|ROL_i|}t} \geq 0$$

2. *The program ranked in position k increased the expected utility of the portfolio more than the ranking cost:*[37]

$$\prod_{n=1}^{r-1} \{(1 - q_{ik_nt})\} \, q_{ik_rt} v_{ik_rt} \geq c_i$$

3. *If any program j not included in the portfolio submitted by applicant i had an expected utility above the ranking cost at one or more steps during portfolio formation before the last program was added to ROL_i, then its utility has to be below the program added in that step:*

$$\forall N \leq |ROL_i|, j \notin ROL_i : \prod_{n=1}^{N-1} \{(1 - q_{ik_nt})\} \, q_{ijt} v_{ijt} \geq c_i \Rightarrow v_{ijt} < v_{ik_Nt}$$

4. *For all programs j not included in the portfolio submitted by applicant i, their utility is consistent with an expected utility below the ranking cost, given a partial ranking that is equal to the submitted ROL_i:*

$$\prod_{n=1}^{|ROL_i|} \{(1 - q_{ik_nt})\} \, q_{ijt} v_{ijt} < c_i$$

These restrictions are intuitive. The second restriction simply reiterates the model's basic assumption that a program must offer enough additional expected utility to be included in the portfolio. In the case of the third restriction: if multiple programs offer more extra expected utility than the ranking cost, the program included in the portfolio must have the

[37] Here and throughout the chapter we use the convention $\prod_{n=1}^{0} x_n \equiv 1$

highest utility, and we can identify it based on the applicant's ROL. As for the fourth, if it were not true, then the applicant would have included at least one more program in the ROL.

Costly consideration

The additional restrictions in the case of the costly consideration model are similar to the ones of the costly ranking model, but the main difference is that applicants are assumed to have incomplete information about their utilities. As a result, we need to compute the expected increase in the expected portfolio utility if a program were to be considered and compare that value with the consideration cost. To do that, we start by defining, in any given step of the portfolio formation process and for any program j not yet considered, the consideration value (CV) given the partial ranking $\mathcal{R}_i$ at that step of the consideration process, as the expected increase in expected portfolio utility if i learns θ_{ijt}. Dispensing of the time subscript to simplify notation, we denote the CV of program j, given the applicant's partial ranking $\mathcal{R}_i$ as:

$$CV_{j|\mathcal{R}_i} = CV\left(v_{ij} - \theta_{ij}, q_{ij} | v_{ik}, q_{ik} \forall k \in \mathcal{R}_i; \sigma_\theta^2\right)$$

The full expression for $CV_{j|\mathcal{R}_i}$ and its derivation can be found in Appendix A.4. Using this notation, we can express the restrictions that ROL_i imposes on utilities and consideration cost in Result 2.

Result 2. ***Mapping between observed ROLs, utilities, and consideration costs***

1. *j was considered by i if and only if*
   a) *i has a family-priority in j;* ***or***
   b) *$j \in ROL_i$;* ***or***
   c) *$j \notin ROL_i$, and there exists at least one partial ranking $\mathcal{R}_i \subseteq ROL_i$ (generated during the portfolio formation) that satisfies:*[38]

$$\forall m \in ROL_i \setminus \mathcal{R}_i : CV_{j|\mathcal{R}_i} \geq CV_{m|\mathcal{R}_i}$$

2. *If j is in the set of considered alternatives:*
   a) *j ranked in position r of $ROL_i \Rightarrow v_{ij} \in (v_{ik_{r-1}}, v_{ik_{r+1}})$, where $v_{ik_{r=0}} = \infty$ and $v_{ik_{r=|ROL_i|+1}} = 0$*
   b) *$j \notin ROL_i \Rightarrow v_{ij} < 0$*

[38]Partial rankings depend on the step during portfolio formation when each program had the highest consideration value (excluding those considered due to family priorities). The details of how these rankings are computed for the empirical application can be found in Appendix A.6.

3. *If j is **not** in the set of considered alternatives its utility is unbounded:*

$$v_{ij} \in \mathbb{R}$$

4. *Defining $k_{cr=|ROL_i|}$ as the program in ROL_i that was considered last, the consideration cost can be bounded by:*

$$CV_{k_{cr=|ROL_i|}|ROL_i \setminus \{k_{cr=|ROL_i|}\}} \geq c_i \geq 0$$

These restrictions are similar in intuition to those in the costly ranking case, but there are two key differences. Firstly, even if an applicant bears the cost to consider a program, it may not be included in their ROL if the unobserved portion θ_{ijt} is too low. Secondly, when analyzing only the programs in ROL_i, the order in which they are considered may not correspond to their ranking order, as the former depends on the program's observed portion $v_{ijt} - \theta_{ijt}$ and its assignment probability q_{ijt}, while the latter only on the value of the utility v_{ijt}.

Parts (a) and (b) of the first restriction reiterate the model's assumptions that applicants with family priority in j have their θ_{ij} observed without cost and that all ranked programs have been considered. Part (c) states that for a program not included in the applicant's portfolio to have been considered, it must have had the highest consideration value at some point during the formation process. As in the costly ranking case, that includes the last step, where the partial ranking corresponds to the ROL. The second constraint has two parts. Part (a) is the same as in the full consideration and costly ranking cases, requiring that ranked alternatives have ordered utilities above zero according to their ranking order. Part (b) specifies that a program not included in ROL_i must provide less utility than the outside option if it were considered. The third is not a restriction and states that utilities of programs not considered are unbounded. Similar to the costly ranking case, that implies some programs with positive utility are not included in ROLs. Finally, restriction four requires that the consideration cost must be lower than or equal to the smallest consideration value among the programs included in the applicant's portfolio. This means that it must be lower than the consideration value of the last program included in the ROL, as this program has the smallest consideration value among all the programs in the portfolio. We can identify that program by mapping the list of partial rankings, starting with the initial one that includes only programs ranked with family priorities, which may be an empty set. In the estimation, we draw the value of the consideration cost constrained to this bound to identify alternatives that were likely considered but not ranked.[39]

[39] Note also that we do not consider the requirement that applicants must submit ROLs with at least two programs to avoid further complicating our model. This implies that for some applicants, their second-ranked program may not conform with the consideration process detailed so far. This is particularly relevant given that 48% of applicants in our estimation sample rank two programs. However, our results are relative to the different counterfactuals computed, and therefore, we can confidently make valid comparisons between them as long as all of our estimations and counterfactuals are internally consistent.

The comparison between the costly ranking and costly consideration mappings reveals an important difference regarding the rationalization of a scenario where an applicant includes a program with a 100% assignment probability above the last program in the observed ROL. In the costly ranking case, the ranking cost is necessarily equal to zero, implying that all programs not ranked must have a utility below zero, as in the full consideration model. On the other hand, in the costly consideration case, the consideration cost is potentially above zero in this situation because there is always some probability that an alternative not considered has a utility above all (or some) of those in a partial ranking, after learning the value θ_{ijt}.[40]

Estimation Methodology

We estimate model parameters using the Gibbs sampling technique, adapted from McCulloch and Rossi (1994), which asymptotically approximates maximum likelihood estimates (Vaart, 1998). This technique is computationally efficient and simpler to specify when demand-side data includes ranked order lists of preferences and applicant-specific covariates such as the distance between residences and schools. Additionally, it avoids biases that can affect simulated maximum likelihood estimates in datasets with a large number of choices, as noted by Abdulkadiroğlu et al. (2017) (see Train (2009)).[41]

The Gibbs sampling technique is a Bayesian estimation method from the Markov Chain Monte Carlo family that iteratively estimates the different parameters in the model over many iterations. A detailed explanation of how we implement the Gibbs sampler is provided in Appendix A.6. We now present the central elements.

Prior distributions

For Bayesian estimation, we need to start defining prior values for the distribution of the different parameters in the models. These distributions are outlined below. The specific prior values used are described in the Appendix, but a key point to highlight is that they need to strike a balance between having diffused distributions, so that they have a relatively small effect on posterior parameters values, but also being "proper" in the sense of being centered at a value reasonable for were posteriors can be expected to lie, and not having a too diffuse variance, so that initial draws are not too slow to converge to the Markov Chain of posteriors that we are interested in finding.

The distribution of fixed parameters is assumed to be:

[40] To have a zero consideration cost, one of the ranked programs needs to have $q_{ijt} = 0$. If that is the case, the ranking cost is also zero in the costly ranking specification.

[41] Other methods used in similar studies include maximum likelihood, such as generalized method of moments (Son, 2020), direct likelihood maximization (Laverde, 2021), and rank-ordered mixed logit models (Kessel and Olme, 2018; Oosterbeek et al., 2021).

$$\begin{aligned}
\lambda &\sim N(\mu_\lambda, \Sigma_\lambda)\\
\beta^g &\sim N(\mu_{\beta^g}, \Sigma_{\beta^g})\\
\Delta_{d,C19} &\sim N(\mu_{\Delta_{d,C19}}, \sigma^2_{\Delta_{d,C19}})\\
c^g &\sim TN(\mu_{c^g}, \sigma^2_{c^g}, 0, \infty)
\end{aligned}$$

In the case of random coefficients, these are assumed to be centered at zero and the object of interest is their variance, which requires additionally assuming a prior distribution to draw their values. Both distributions complement each other and are labeled as a conjugate-prior structure in the literatue. Their values are given by:[42]

$$\begin{aligned}
\xi^g_j &\sim N(0, \sigma^2_{\xi^g}); & \sigma^2_{\xi^g} &\sim IW(\tau_{\xi^g}, df_{\xi^g})\\
\epsilon_{ij} &\sim N(0, \sigma^2_\epsilon); & \sigma^2_\epsilon &\sim IW(\tau_\epsilon, df_\epsilon)\\
\zeta_i &\sim TN(0, \sigma^2_\zeta, -c^{g_i}, \infty); & \sigma^2_\zeta &\sim IW(\tau_\zeta, df_\zeta)
\end{aligned}$$

Gibbs Sampling Iteration

To initiate the process, parameters from the prior distributions mentioned above are drawn, and initial utility and ranking or consideration costs consistent with observed choices are determined. Subsequently, each Gibbs Sampling iteration comprises the following steps:

1. Sample λ_1 given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, $\beta^{g_i}_0$, $\Delta_{d,C19,0}$ and $\xi_{j,0}$.
2. Sample β^g_1 given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, $\Delta_{d,C19,0}$ and $\xi_{j,0}$.
3. Sample $\xi^g_{j,1}$ given $\sigma^2_{\xi^g,0}$, $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β^g_1, $\Delta_{d,C19,0}$.
4. Sample $\Delta_{d,C19,1}$ given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β^g_1 and $\xi_{j,1}$.
5. Sample covariances for $\sigma^2_{\xi^g,1}$ and $\sigma^2_{\epsilon,1}$
6. Then proceed with utilities and ranking or consideration costs following the data to model mapping results described in Section 1.5 (see Appendix A.6 for details).

The process involves repeating these iterations many times to obtain parameter values that fluctuate within a bound of possible values -the Markov Chain-. Using the values of the different iterations, we can compute the mean and variances of interest. To remove the influence of initial draws and chosen starting utilities and ranking or consideration costs,

[42]The inverse-wishart distribution (IW) is used in our model as a conjugate prior for the normal distribution. This is a common choice in similar applications.

a number of initial iterations are discarded as a burn period necessary for the sampler to converge.

Identification and Estimation Results

Identification

One important point to note before discussing parameter identification is that our data only includes changes in program prices over time, and only if they were above zero in 2019. Other program attributes are unchanged during the period. Thus, considering that we include a school fixed-effect $\xi_j^{g_i}$ for each socioeconomic group in our estimation, β^{g_i} parameters for attributes that are constant across time cannot be separately identified, and including them in the estimation could therefore be regarded as unnecessary. However, these attributes are included in the model for three reasons. First, in the costly consideration model, we assume that applicants use program attributes to predict their group's mean utility and their idiosyncratic taste values. Second, including these attributes (and a constant) in x_{jt} adds structure to the model that allows us to interpret $\xi_j^{g_i}$ values as deviations from the linear prediction of their value based on observable information, which in turn makes the structure imposed over their distribution more reasonable. Third, including these parameters in the full consideration and costly ranking models allows us to compare the results for the three alternative specifications.

Regarding the identification argument, let's first consider the simpler case of the full consideration model. As previously explained, under the additively separable utility structure and the assumption $\epsilon_{ijt} \perp d_{ijt}|\delta_{ijt}$, we can map the density of the utilities for different programs in different years, denoted as $\mathbf{u}_{jt}$. Using this mapping, variation in the utilities of applicants with and without family priorities allows us to identify the parameter λ. Next, $x_{jt}\beta^{g_i} + \xi_j^{g_i}$ corresponds to the average program utility in year t, after subtracting for the effect of family priorities on $\mathbf{u}_{jt}$. The parameters $\beta_{Price}^{g_i}$ (or more generally the parameters over attributes that vary across years), are identified by variation in average program utilities and prices over the different years in our sample. Finally, variation in ROLs over the different years under the assumption that the outside option's value is fixed at zero allows for the identification of $\Delta_{d,C19}$.

In the case of the costly ranking model, the argument follows a similar logic but including the restrictions outlined in Result 1. Importantly, we can identify ranking costs with the variation in assignment probabilities, which is largely driven by differences in program congestion across different years. This assumes that applicants use rationally expected assignment probabilities when forming their portfolios, and that they do so as described in the portfolio formation process.

In the costly consideration model, applicants are assumed to make decisions with partial information on which alternatives to consider. However, it is not possible to separate θ_{ijt} from the rest of the utility using assignment probabilities, since consideration depends on both the observed variables for programs not yet considered and θ_{ijt} for those included in a partial ranking. And moreover, including a program in the ranking depends on both values. To separate these parts, data variation correlating only with the consideration part of the process is needed. For instance, Son (2020) argues that the order in which schools are presented in New York's school brochure correlates only with consideration but not with utilities. Unfortunately, such data variation is, at least to our knowledge, not available in our context. Moreover, the model is currently not designed to accommodate it. As a result, the model parameters are not identified, and the parameters obtained come from the functional structure assumed. We provide an intuitive explanation of how the model's structure affects the parameters in Appendix A.5 and, as we show in the next section, estimated parameters are not significantly different from those obtained with the other models.

Estimation Results

Due to the large size of some Municipalities, for computational efficiency we include a sample share of applicants in our estimation, whose size is presented in Table 1.4.

We implement two Gibbs sampler chains consisting of 25,000 iterations each for every Municipality. The first 10,000 iterations are discarded as the burn period. To assess convergence and mixing, we examine potential scale reduction factor (PSRF) values (Gelman, Rubin, et al., 1992) and the trace plot of ϵ_{ijt}'s variance. All PSRF values are close to one for all parameters and Municipalities, and the trace plots indicate that σ_{ϵ}^2 is bounded (see Appendix A.7).

Table 1.5 summarizes the key focus of our analysis: the average mean utilities across programs for different achievement categories and socioeconomic groups. Our aim is to measure differences in preferences between these groups, which, combined with differences in residential locations, can help explain differences in program choices.

The mean utility estimates reveal three main findings. Firstly, the estimates are consistent across all models. Secondly, non low-SeS applicants are estimated to have higher utilities for all program types on average. Thirdly, the differences are mainly concentrated in higher mean utilities for programs with higher achievement categories, especially for those in the highest category. Specifically, low-SeS applicants are willing to travel approximately 750m further to attend a program in the highest category than to one in the lowest, whereas non low-SeS applicants are willing to travel around 1 km further. For programs in the high and medium achievement categories, low-SeS applicants are willing to travel approximately 325m more to attend a high-category program than a medium-category one, while non low-SeS applicants are willing to travel around 475m further.

To provide a better understanding of the economic significance of the estimated differences in mean utilities, we compare them with two reference values. First, we note that the average distance to the first preference for applicants in our sample is 1.17km (see Table 1.3). This means that, compared to the average distance to the first preference, no-SeS applicants are willing to travel an additional 21% to attend a program in the highest achievement category relative to one in the lowest, and an additional 13% to attend a high category program relative to a medium category one. Second, we compare the differences in mean utilities with the standard deviation of idiosyncratic utility shock estimates presented in Table 1.6. Using this measure as a reference, no-SeS applicants are willing to travel an additional 0.28 and 0.17 standard deviations to attend a high category program relative to low and medium category programs, respectively. Considering that these are differences in differences in relative utilities for programs of different achievement categories, these values are significant, consistent with the estimate in Section 1.6 that approximately 60% of educational segregation absent SR is explained by differences in preferences over programs between socioeconomic groups.

Furthermore, Table 1.5 reveals that in the costly consideration model, the linear prediction for the unobserved portion of mean program utilities using observed program attributes (represented by $\delta^{g_i} - \xi^{g_i}$) almost perfectly matches the mean utility of the group. This suggests that the linear model, which separately adds the effect of differences in each program attribute, is very similar to a strategy of implementing cells for each attribute combination, or at least cells based solely on program achievement. Additionally, the table presents statistics on the estimation sample, showing that, on average, there are 10 choices for each program in the Municipality plus Buffer geographic unit.[43]

With respect to differences in mean utility estimates for different Municipalities, Figure 1.3 presents the values obtained in the case of the costly consideration model. We can note three things from the figure. First, there is a significant variance in mean utility differences between achievement categories across different Municipalities. Second, in the high vs. medium and high vs. low achievement category comparisons, the mean utility for high achievement is higher in the majority of cases. Third, the difference in mean utilities between no-SeS and low-SeS applicants is in favor of the high achievement category in more cases and to a larger extent when compared to the low achievement category programs than to medium category ones.

The results for our two example Municipalities in the costly consideration case are as follows. For Maipú, the average mean utility of high achievement programs is almost the same for both groups, with a value of 0.819 for low-SeS applicants and 0.801 for the rest. In terms of the difference with respect to low achievement category programs, the values for the low-SeS and no-SeS groups are 0.630 and 1.408, respectively, while for the medium achievement case, the values are 0.203 and 0.758, respectively. As for La Cisterna, the mean utility of the high achievement categories are 0.909 and 1.373, respectively. Differences between mean utilities of high and low achievement categories are 0.970 and 0.716, respectively, while differences between mean utilities of high and medium achievement categories are 0.387 and 0.385, respectively. Therefore, we can expect a larger share of low-SeS applicants to rank high-achievement programs in La Cisterna in the absence of residential sorting (and to rank them higher in their ROL).

Table 1.6 presents the results of the main model parameters of interest, while the values of the parameters related to program attributes can be found in Appendix A.7.[44] As seen

[43] Although the sample share is restricted, larger Municipalities have a larger ratio of choices over programs due to their buffers including fewer programs from other Municipalities.

[44] Additionally, figures that display the values of these parameters in each municipality are included. These

in the table, the parameters for the COVID-19 pandemic distance dis-utility are estimated to be very small. Surprisingly, the consideration and ranking costs are slightly larger for no-SeS applicants. This could be partly compensating for the higher mean utility values observed in this group, particularly in the costly consideration model where cost parameters are not identified. However, this is likely also the case in the costly ranking model, as identifying ranking costs with assignment probabilities requires strong assumptions. Finally, the parameters related to family priorities and variances are reasonable and consistent across different specifications.

Using these estimates and the steps of the portfolio formation process, we now move to presenting our counterfactual results.

1.6 Results

The goal of this section is to investigate the impact of SR on educational segregation and other outcome variables, as well as to assess the role of residential sorting in existing segregation and the interaction of SR and demand-side policies. However, the results can be sensitive to methodological choices, such as the geographic level used to measure segregation and whether unassigned applicants are included in the analysis. To clarify the latter point, we focus initially on measuring educational segregation while keeping all applications constant, and vary only SR.

Preliminary analysis: The effect of changes in global SR conditional on submitted ROLs

Figure 1.4 illustrates the difference in measuring segregation at the Municipality versus city level, and the impact of including unassigned applicants. The figure shows the relationship between the share of reserves and educational segregation, using the minority-adjusted dissimilarity index while keeping applications constant. Without unassigned applicants, segregation at the city level is minimized at around 40% of reserves (the share of low-SeS applicants in Santiago), which replicates results in Escobar and Huerta (2021) albeit at a different scale due to the minority adjustment.

When unassigned applicants are included in the city-level computation, grouped together as if they were assigned to one alternative school in the city, the effect of reserves on segregation is significantly reduced. Then, segregation diminishes only slightly up to a point before the 40% level where it begins to increase. This effect is particularly salient in 2019 due to the higher level of unassignment.[45]

figures reveal significant variation in parameter estimates across Municipalities, with larger Municipalities tending to have more consistent estimates.

[45] The minimization of segregation at lower SR levels in this case, when unassigned applicants are included, is partly mechanical and driven by the increasing size of the group of unassigned applicants. The relevant point is that the inclusion of unassigned applicants in the measurement leads to a smaller reduction in segregation

If we measure segregation at the Municipality level, using a weighted average by the total number of applicants in each Municipality to compare results with those obtained at the city level, we observe that segregation is smaller, as segregation within Municipalities is typically lower than for the whole city combined. Additionally, segregation is minimized at a smaller SR value, as seen in the figure, because, as reserves are further pushed to the 40% city value, increasingly more Municipalities have a level of SR above their share of low-SeS applicants, which increases their local segregation level. Furthermore, when including unassigned applicants grouped at their Municipality of residence, we observe that reserves have a negligible effect, and segregation starts increasing at a lower value (as in the case of the city measure).[46]

To minimize segregation when measured at the Municipality level, reserves at each school should be adjusted based on the proportion of low-SeS applicants in its Municipality and neighboring areas. We approximate the optimal level of reserves by setting them at the share of low-SeS applicants in the Municipality where the school is located and observe that it leads to a reduction in segregation. To see that comparison, we include obtained values in Figure 1.4, with and without including unassigned applicants in the measurement, as a proxy benchmark for the achievable reduction adapting SR to local conditions (the horizontal dashed and solid lines respectively). The magnitudes of these reductions are discussed below with the introduction of application counterfactuals.

We believe that measuring segregation at the Municipality level is superior to doing so at the city level because it provides a better approximation of the level of integration of schools in the same neighborhoods. However, partitioning the city into different geographic units generates interactions between them, which can be particularly relevant for measurements in contiguous units with different levels of low-SES populations. Schools near the border may appear segregated even if all applicants in each Municipality are randomly distributed and have the same choice patterns. But it is desirable for these schools to have a share of low-SES applicants between the two geographic units to reflect their local reality. Additionally, implementing differentiated SR in each Municipality changes the share of applicants assigned to their Municipality in different socioeconomic groups, as reserve shares affect the likelihood of applicants on either side of the border being placed, which can affect measurements. As a result, while our average results at the Municipality level indicate the importance of adapting SR to local conditions, specific Municipalities may exhibit counter-intuitive segregation measurements, as exemplified by La Cisterna (see discussion below).

The inclusion of unassigned applicants is intended to capture the aftermarket dynamics that affect segregation in enrollment across different programs, but including them can lead to an excessive reduction in measured segregation in some counterfactual scenarios. This is because when an additional no-SES applicant is unassigned, the unassigned group, which over-represents their share, becomes less evenly distributed. Nevertheless, some of these

as some no-SeS applicants are re-shuffled from programs where they are over-represented to the unassigned group where they are over-represented as well.

[46]This is due to the same mechanical effect discussed in that case.

unassigned applicants may eventually enroll in schools where low-SES students are over-represented, resulting in a reduction in segregation instead of an increase.[47]

While precisely estimating counterfactual aftermarket dynamics would be ideal, it exceeds the scope of this chapter and is left for future research. Nonetheless, it is important to note that interpreting the results as bounds is a reasonable approach given these limitations. In particular, we can interpret results including unassigned applicants as a lower bound due to the effect of unevenness in the unassigned group produced by SR, while results without including them are an upper bound as the unevenness is disregarded. We present the results obtained with various measurements but additionally, and more importantly, we emphasize the policy implications and relative magnitudes of our findings, which remain remarkably consistent across different measurement alternatives, rather than their specific values.

Analysis including ROL counterfactuals

Leveraging estimated preferences, we can enrich the analysis on three main dimensions. First, we can provide benchmarks of achievable reductions in segregation by simulating applications that eliminate differences in choice behavior across socioeconomic groups. Second, we can account for potential changes in application behavior that may result from variations in assignment probabilities, which can arise due to changes in SR or the application behavior of other applicants. And third, we can conduct simulated demand-side policy exercises to study their interaction with SR. Next, we describe how we construct counterfactual scenarios based on the estimated preferences and then present the results obtained.

Methodology

The process of obtaining each counterfactual application in the different models is detailed in Appendix A.8. Two key points to highlight are: (i) In some counterfactual exercises, we randomize applicant locations, using the same randomization across these exercises and across the different models to ensure comparability of results. (ii) Prior to any counterfactual simulations in each model, we obtain random coefficients that rationalize the observed ROLs with current SR, and then use these coefficients throughout the counterfactuals. The only exception is the minimum segregation benchmark counterfactual, where we generate new random coefficients that are independent of reported ROLs. The reason is that, in that case, we want equal choice behavior between socioeconomic groups and thus removing any influence from observed ROLs or model selection.[48]

[47] Furthermore, including unassigned applicants in measuring segregation does not address the biases in our measurements of justice and efficiency, which only consider assigned or ideally enrolled students.

[48] A relatively minor additional point to consider is that the assignments in the baseline scenario (where no assignment rule or choice behavior is changed) are very similar but not equivalent to the official assignments. This is because to increase consistency across counterfactuals and model comparisons, we use the same lotteries throughout all exercises, and official lotteries are published only for programs included in an applicant's ROL and only relative to other applicants to the program.

The first step in our counterfactual analysis is to identify benchmark levels of educational segregation. We establish our maximum benchmark by considering the scenario of only eliminating SR. Our minimum benchmark is determined by eliminating the contributions of residential sorting and differences in average preferences over programs between socioeconomic groups while also eliminating SR. To eliminate the effect of residential sorting, we randomize residence locations.[49] To eliminate systematic differences in preferences between socioeconomic groups, we average the estimated model parameters that differ between the groups and equalize top-off fees for both groups at half their value in the case of schools subscribed to the SEP policy. The relative contribution of residential sorting is obtained by implementing a counterfactual in between where only applicant residences are randomized.

A final methodological consideration when estimating segregation measures using the costly ranking and costly consideration models is that it is important to note that counterfactual applications and assignment probabilities are interdependent. Therefore, to obtain the final segregation measures reported in the following section, we iteratively estimate one and then the other for each counterfactual. We start the process with assignment probabilities obtained using observed applications with the corresponding SR level. As presented in Figures A.1.1 and A.1.2 in Appendix A.1, changes in segregation are relatively minor.

Counterfactual Results

To begin our discussion of the counterfactual results, we present the main estimates for the three alternative models in Table 1.7. We then conduct further analysis for additional counterfactuals, focusing solely on the costly consideration case. We average the results over the three years in our data and focus only on applicants included in the estimation sample, as the applications of the other third of applicants remain unchanged. In Appendix A.2, we provide additional findings obtained for each year separately and considering all applicants in the costly consideration model's case.[50] The main takeaways of our analysis remain consistent whether focusing on all applicants and/or a specific year separately, although specific values vary.

Panel A presents the main outcome of interest, which is the reduction achieved in the segregation gap under different scenarios for our two segregation measures, as well as for the

[49] This follows the approach taken by Laverde (2021). However, in that paper, new locations and assignments are computed one at a time, which is the ideal ceteris paribus exercise to avoid changing school composition or assignment probabilities. This approach is too computationally demanding in our case. Nonetheless, our model incorporates optimal responses to changes in assignment probabilities. Moreover, we do not include school composition in our covariates, as this information is not presented to families, as opposed to racial statistics in the context studied by Laverde (2021).

[50] Including all applicants overstates the impact of SR, as it implies eliminating choice differences between socioeconomic groups only for two-thirds of the applicants leading to a higher minimum segregation benchmark. Furthermore, the low-SeS share used to calculate SR is based on the overall population, which differs from the sample population in some Municipalities, potentially affecting results (the latter effect is likely to be much less relevant than the former).

minority-adjusted Duncan index measure that includes unassigned applicants. It also reports the level of unassignment of applicants in both socioeconomic groups. Three counterfactuals are considered in addition to the current case with observed applications and the current SR level, presented in the first row of each model. First, the baseline or maximum segregation level that would exist if SR were completely eliminated (second row of each model). Second, the reduction in the segregation gap obtained by eliminating residential sorting (third row of each model). And third, the reduction in the segregation gap obtained by implementing SR at the program's Municipality low-SeS share level (fourth row of each model). Panel B presents the segregation values obtained in the different models under the MA-DD segregation measurement, including and excluding unassigned applicants, for the two benchmark counterfactuals used to estimate the segregation gap.

A key result from the table is the consistency in results across different specifications. The main common findings are highlighted first, followed by a brief discussion of the differences. The main takeaway is that SR's effect in closing the segregation gap, when set at their current 15% level, ranges from 3% (using the MA-DD index including unassigned applicants in the full consideration model case) to 9.2% (when unassigned applicants are not included under the costly consideration model). In all models and measurements, the effect of SR is approximately multiplied by five when their level is set at the low-SeS share in the Municipality. Another important takeaway is that residential sorting contributes to educational segregation by around 40%. Lastly, Panel B indicates that estimated values of segregation are very similar, particularly when focusing on weighted averages.

Turning to differences, each panel has one difference worth noting. In Panel A, the unassignment weighted average flattens for both socioeconomic groups in all three models when residential sorting is eliminated. However, in the costly consideration model, the overall level of unassignment increases slightly, while in the other two models, it decreases by about 1.5%. This difference is driven by the fact that in the costly consideration model, applicants choose which schools to consider based only on observed program attributes (and priorities and distances), which generates more congestion in counterfactual scenarios, not only between groups but also within applicants in the same group. In this model, dispersion in program unobservables (among those with the same attributes) and in idiosyncratic tastes impact only which considered alternatives are ranked and in what order.

Moving to Panel B, we can observe that the segregation gap is larger in the full consideration model, and this is due to a smaller level of segregation in the minimum benchmark. This difference highlights that while the results are robust to the different specifications, they are affected by the different assumptions made regarding portfolio formation. Further research to better understand the process and obtain more precise estimates is therefore also valuable for papers studying similar settings and questions.

Figure 1.5 provides a graphical representation of these results, highlighting their consistency overall, and that the main difference is the estimated minimum segregation benchmark.

The horizontal lines in the figure represent the educational segregation gap, standardized to zero with the minimum segregation benchmark (MSB) of the costly consideration model and to one hundred with the maximum segregation level using that model. The current level of segregation is marked by a vertical dotted line (naturally, the same for all models). The level of segregation achieved by setting SR at the low-SeS share of the Municipality is also shown, with its range across different years in parentheses. The estimate of the segregation value with SR is similar for the two models with portfolio formation frictions but larger for the full consideration model, which partly explains its lower contribution to closing the segregation gap in that model.[51] The contribution of residential sorting is represented by a shaded region with its variability at the bottom of the figure. Notably, the level of segregation is nearly identical across all three models, with the difference in the gap reduction estimate resulting from the denominator, the different segregation gap sizes.

Table 1.8 reports two additional outcomes of interest. The first is the percentage of applicants from the low- and no-SES groups who are assigned to high-achievement or medium-achievement programs. This measure only includes assigned applicants and is a weighted average of the different Municipalities to control for differences in the applicant's low-SES share and the availability of higher-achievement programs across Municipalities. The second measure, labeled as the percentage use of ideal occupancy, shows the share of seats used relative to a benchmark where high-achieving schools fill first, followed by medium-achievement schools.[52]

The two measures presented exhibit little variation across different counterfactuals, with one notable exception. Setting SR at the Municipality level reduces the achievement gap by about one-third. In contrast, setting SR at the current level has little effect. However, SR at the Municipality level also results in a slight decrease in the share of ideal occupancy of high-achievement programs.

As the takeaways from the different models are essentially the same, we will focus on the costly consideration model in the last section of our counterfactual analysis. This section will examine the interaction between demand-side policies and SR, as well as the complementarity between SR and the two drivers of differences in choice between the socioeconomic groups. Additionally, using the costly consideration model will allow us to study the impact of reducing consideration costs, simulating a policy that reduces information acquisition frictions. But before presenting those results, it is insightful to compare the values of segregation gaps, benchmarks, and the different counterfactuals in our two example Municipalities to highlight some of the nuances in our geographic methodological approach. These are presented in Table 1.9.

Both segregation measures align in terms of the relative comparison of the different counterfactuals and their magnitudes. Maipú shows relatively little variation in the segregation value across counterfactuals, while La Cisterna exhibits more significant variation due to its sensitivity to interaction with other Municipalities. In Maipú, results are consistent with expectations and city-wide averages, but in La Cisterna, segregation is smaller under the maximum rather than the minimum benchmark in 2019 and 2020. Moreover, in that Mu-

[51] Note that the estimated effect of SR in reducing the segregation gap when set at the Municipality level and at the current level move up or down together across different years, roughly maintaining their proportionality.

[52] To calculate this measure, we need to include all applicants, not only those with counterfactual applications.

nicipality, eliminating residential sorting slightly reduces segregation in 2019 relative to both benchmarks, but increases it above them in 2020. A final point to highlight from the table is that SR set at the Municipality level consistently reduces segregation for both Municipalities relative to the maximum benchmark.[53]

These results demonstrate that the interaction between geographic units can strongly affect segregation measures. While our methodology of presenting measures averaged for the different Municipalities to discuss counterfactuals provides a comprehensive picture in our view, as interactions can drive measures both upwards and downwards, future research should explore the best geographic methodology strategy depending on policy objectives.

Table 1.10 summarizes the results obtained with the costly consideration model, averaged over the three-year period, for the next set of counterfactuals. The results for each year separately and including measures with all applicants, inorporating also previously discussed counterfactuals, are presented in Tables A.2.1, A.2.2, and A.2.3 in Appendix A.2. Panel A of Table 1.10 shows the primary outcomes of segregation and unassignment, while Panel B presents the other outcomes of interest introduced above.

The first additional counterfactual involves implementing SR at the city low-SeS share. The next four counterfactuals correspond to implementing demand-side policies together with SR at the Municipality level, either individually or all together. In the first simulated policy, high-achievement programs' indirect utilities are increased by 583 meters for all applicants, which is half of the average distance to the first preference of applicants included in the estimations. This intervention aims to simulate a successful policy that promotes the highest quality schools. In the second counterfactual scenario, individual consideration costs are reduced by 90%, nearly eliminating them. This intervention simulates a policy such as a digital platform where each school presents all its information in detail and interacts with applicants, enabling families to acquire relevant information with significantly less effort. The third scenario cuts the dis-utility of distance-to-school by 20% for all applicants, intending to simulate a considerable transportation cost reduction.

Importantly, note that none of these policies are targeted or have heterogeneous effects in different socioeconomic groups. However, they do have large distributional effects due to the simultaneous implementation of targeted reserves. To see this, the table includes a final scenario that combines the effects of all three demand-side policies without SR.

Implementing SR at the city level is found to be inferior to implementing it at the Municipality level across all dimensions. Moreover, the primary differentiation is observed in terms of segregation, which even worsens in the measure that includes unassigned applicants. This is because the reserves primarily operate over congested programs, leading to no-SeS applicants being unassigned from schools where their share is underrepresented relative to their Municipality and being moved to the unassigned group, where they are the majority. Moreover, even when comparing results with the measure without unassigned applicants, SR at the city level only achieves around half of the reduction in segregation achieved by SR at

[53] In Appendix A.3, we visually illustrate the level of exposure of La Cisterna to the diverse socioeconomic composition of neighboring areas.

the Municipality level.

With regard to demand-side policy interventions, we find that both increasing preference for high-achievement programs and reducing transportation costs significantly reduced segregation by 10-20% in addition to SR alone. However, they differ in other aspects. As expected, increasing preference for high-achievement programs produces a significant increase in the overall assignment level to these programs and a greater reduction in the gap in assignment between socioeconomic groups. However, it also results in a slight increase in unassignment by about 0.5%. On the other hand, reducing transportation costs generates a significant 3% reduction in overall unassignment (a decrease of almost 40%), particularly benefiting the population of no-SeS applicants, but has limited impact on assignment to higher-quality programs as results in this case are driven by an increase in the average ROL length. The near-elimination of consideration costs has a minor impact on all dimensions, with only a slight decrease in overall unassignment from 7.89% to 7.76% and an additional 2% reduction in segregation on top of SR alone. These findings suggest that information acquisition frictions have a relatively small role in observed ROLs. However, this could also be due to limitations in the model's identification of such frictions, as discussed in Section 1.5.

If all demand-side interventions are implemented together with SR, educational segregation can be reduced by 20-30%, while also achieving the gains on the other dimensions obtained by implementing the policies individually. However, if these policies are implemented alone without reserves, while their effect on unassignment and ideal occupancy is almost identical, their impact on segregation is considerably smaller, resulting in values similar to what is obtained with SR at their current level. Moreover, only implementing the demand-side policies, the achievement gap is increased, albeit with more applicants in each group assigned to high-achievement category programs. These results highlight the complementarity between SR and demand-side policies in reducing segregation and the achievement gap. We interpret this as SR playing the role of targeting the effects of un-targeted demand-side policies, potentially facilitating their design and implementation significantly. A well-designed policy mix can attain significant improvements in all outcomes of interest cost-effectively.

Figure 1.6 provides a graphical representation of the results under the MA-DD index with and without unassigned applicants. There are two additions to note relative to Figure 1.5. First, we include a MSB that implements SR at the program's Municipality low-SeS share on top of eliminating differences in choice between the socioeconomic groups. Intuitively, implementing SR should lead to an additional reduction in segregation because of the randomness in residences and preferences. Second, we compare the complementarity between SR implemented at the Municipality level together with the elimination of residential sorting or together with the elimination of systematic differences in preferences between the socioeconomic groups.

Regarding the addition of the MSB including SR, Figure 1.6a shows that segregation is further reduced by 6% in terms of the segregation gap when unassigned applicants are included. This reduction results from SR eliminating some of the segregation produced by the randomness in choice discussed earlier. However, if unassigned applicants are excluded from the measure, the MSB with reserves is measured at around 30% less. The reason is simple: in many cases, SR make the schools' socioeconomic composition more even while making the unassigned group more uneven. Both forces roughly balance out in Figure 1.6a, but the latter is not incorporated in Figure 1.6b.

With regard to the complementarity between SR and either eliminating residential sorting or differences in preferences between socioeconomic groups, our findings indicate consistent results for both segregation measures. The impact of SR on reducing segregation is greater when combined with eliminating residential sorting, resulting in a reduction of 53.1% to 61.8% compared to using only SR at the Municipality level and the MSB with SR. The impact of SR and eliminating differences in preferences between socioeconomic groups on reducing segregation is only about half. This is due to the greater operational range of SR in less residentially segregated neighborhoods. When residential sorting is maintained and differences in preferences are eliminated, SR only operate over a small fraction of applicants in neighborhoods where there are few low-SES applicants. In contrast, in less segregated neighborhoods or where low-SES applicants are the majority, additional low-SES applicants to a program with enough to achieve an even assignment can only make the end result more uneven by over-representing their share, depending on lotteries.

1.7 Discussion

The use of Centralized Choice and Assignment Systems is rapidly increasing as digitalization becomes more widespread, and policymakers recognize their benefits in terms of efficiency and transparency. To further reduce segregation in schools based on socioeconomic, racial, or ethnic backgrounds, targeted reserves should be integrated into these systems. This can also help close gaps in assignment to higher-quality programs, which is a common goal for policymakers in the context of school choice, especially at the entry levels of the educational system. This chapter highlights that for targeted reserves to be effective, they need to take local conditions into account and be implemented alongside demand-side policies, such as those that reduce transportation costs or increase demand for higher quality schools.[54]

In the case of PreK assignment in Chile's CCAS, our analysis reveals that the current flat level of SR at 15% leads to only a modest reduction of 3-9.2% in the average educational segregation gap in Santiago's Municipalities. Simply increasing SR uniformly across all schools may provide some improvement, but it can also have unintended consequences such as intensively exacerbating segregation in Municipalities with fewer low-SES families. Setting SR at the level of low-SES applicants in each Municipality can be a more effective and amplify the impact of SR by five times relative to the current level. Moreover, setting this level is at least twice as effective as setting reserves homogeneously at the city level.

[54]Moreover, reducing information frictions leveraging digital systems for cost-effective personalized feedback and information delivery is a promising approach (Arteaga et al., 2021, 2022; Elacqua, Gómez, Krussig, Marotta, Méndez, and Neilson, 2022a; Elacqua et al., 2022b).

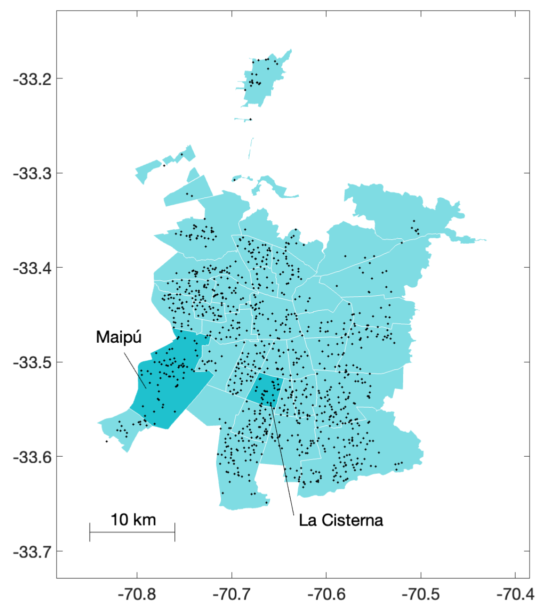
Additionally, implementing targeted reserves together with demand-side policies can significantly reduce segregation further. Our simulations show that with SR they can reduce educational segregation in an additional 20-30% of the segregation gap, compared to only 4-10% without SR, highlighting the complementarity of these policies.

Analyzing targeted reserves poses challenges due to two main factors: the geographic component and aftermarket dynamics affecting final enrollment. Deciding on the appropriate level to measure segregation requires normative decisions regarding geographic units of interest. However, even with a clear definition of geographic units, families may choose schooling alternatives outside of their residential unit, making it difficult to accurately measure changes in segregation. In addition, accounting for aftermarket dynamics, such as administrative processes that allow families to change their assigned program or obtain an assignment if initially unassigned, and private schools with separate admission processes, is important for a complete analysis.

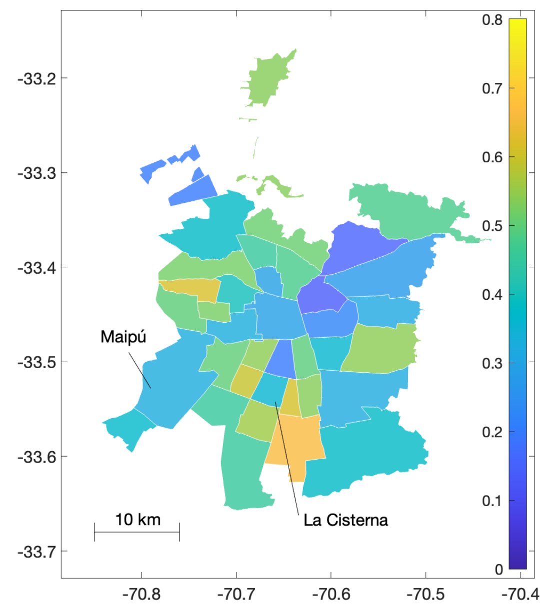
Further research is necessary to better understand the optimal way to set reserves at each school and to design complementary policies appropriately. However, our findings underscore the importance of prompt policy action in Chile and other similar contexts. First, our results call for targeted reserves to be set at each educational program based on local conditions. In Chile, for instance, being conservative one could set SR at a fixed percentage below the Municipality's low-SeS applicants share across the country to achieve better results than the current flat 15% level. Additionally, if SR are implemented, policies promoting higher quality options or more broadly options were low-SeS applicants are under-represented, and policies increasing the number of alternatives considered and acceptable (and thus ranked) by families should be deployed, even if targeting their effects is unfeasible and segregation is a relevant policy concern, as SR can target their effects.

Figures

Figure 1.1: Santiago's Municipalities



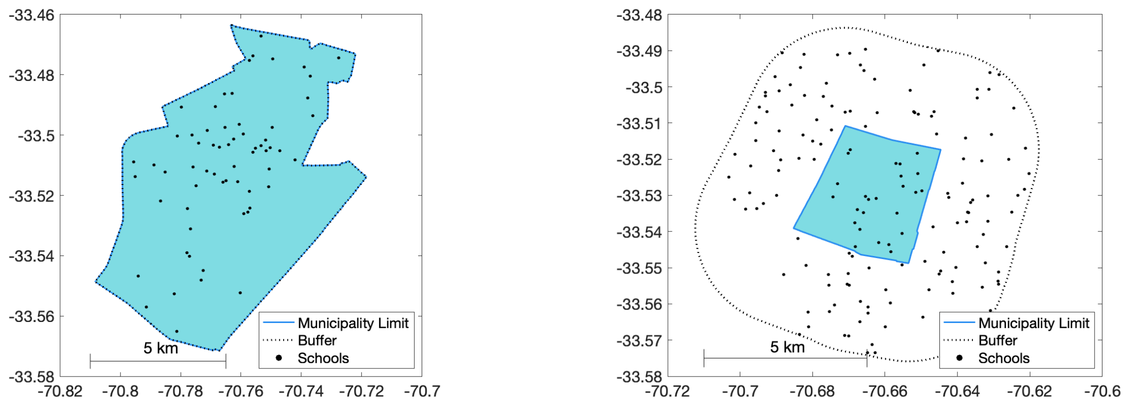
(a) Schools with PreK programs in Santiago City



(b) Share of low-SeS applicants in each Municipality

Note: This figure shows the urban limits of Santiago's Municipalities. We do not consider the Municipality of Pirque because it only has 14 applicants accurately geolocated applying to schools in or close to it. Panel A shows schools considered in our exercises, and they can be outside urban limits because of the buffers drawn around each Municipality. This procedure is explained with more detail in Subsection 1.4. Panel B shows the share of low-SeS Pre-K applicants in each Municipality, with a minimum value of 13.5% in the Municipality of Providencia and a maximum value of 66.21% in the Municipality of La Pintana.

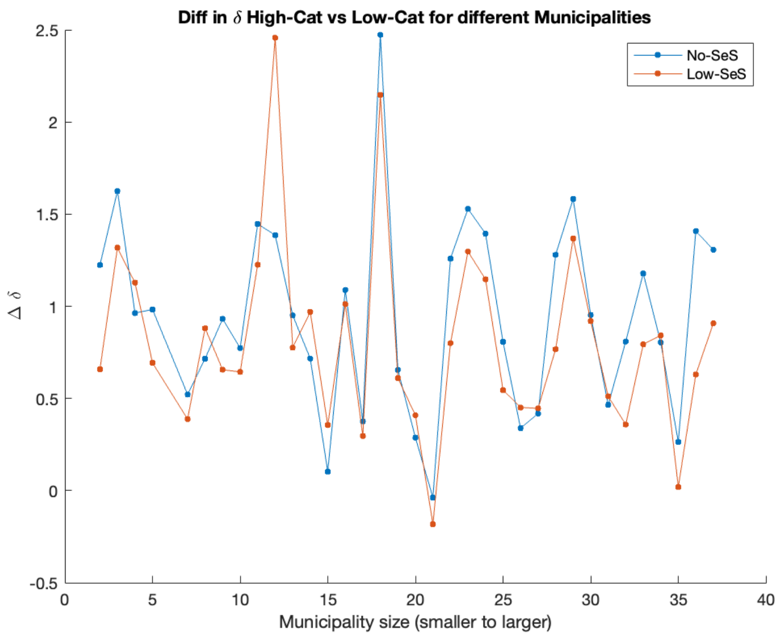
Figure 1.2: Municipality + Buffer Area in Maipú and La Cisterna



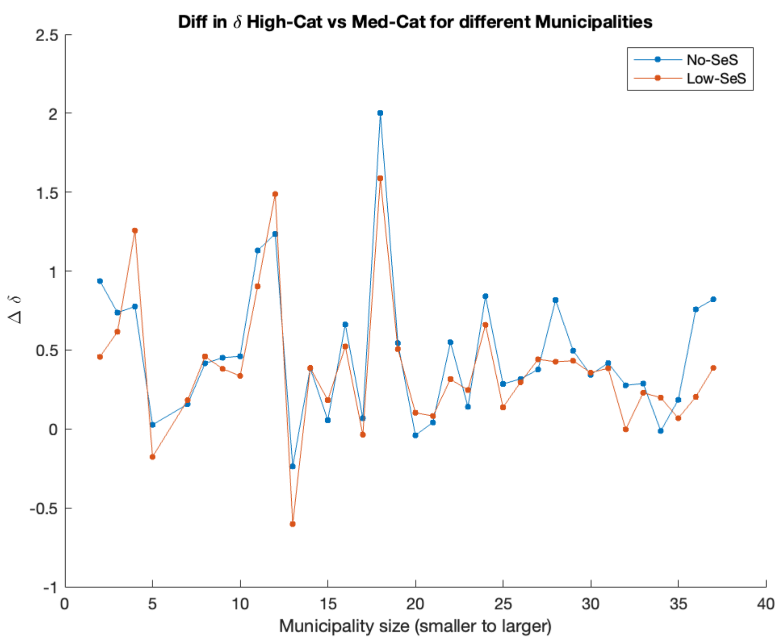
(a) Maipú (b) La Cisterna

Note: This figure shows the Muncipilities of Maipú and La Cisterna, differentiating the urban limit with the buffer drawn around it. In the case of Maipú, there is no buffer around, so the limits match.

Figure 1.3: Differences in Mean Utilities for Different Municipalities in Costly Consideration Model (kms)

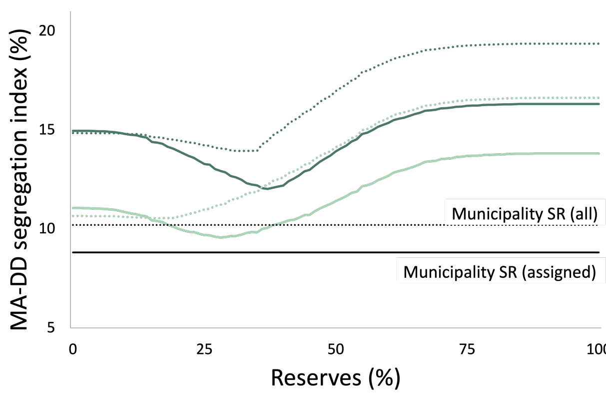


(a) High vs. Low Achievement

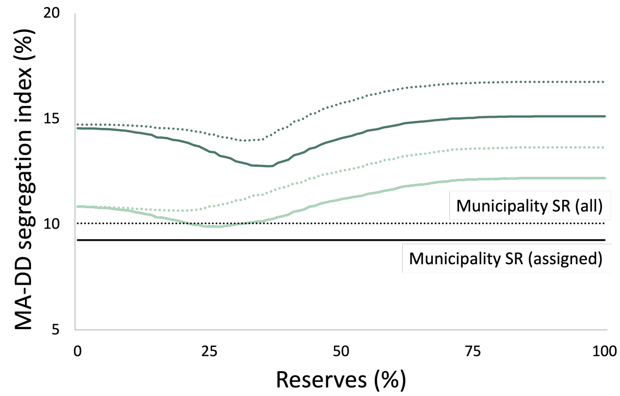


(b) High vs. Med. Achievement

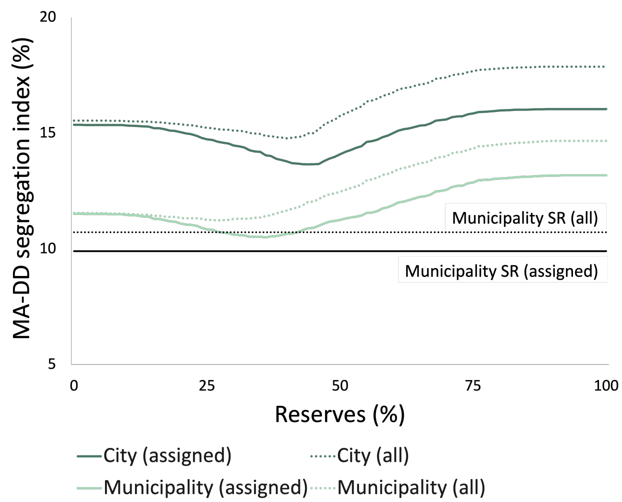
Figure 1.4: The effect of reserves on segregation using observed applications



(a) 2019



(b) 2020



(c) 2021

Figure 1.5: Educational Segregation Decomposition in All Models: Only Counterfactual Applicant's Sample Using $MA\text{-}DD^{WUN}$ Index

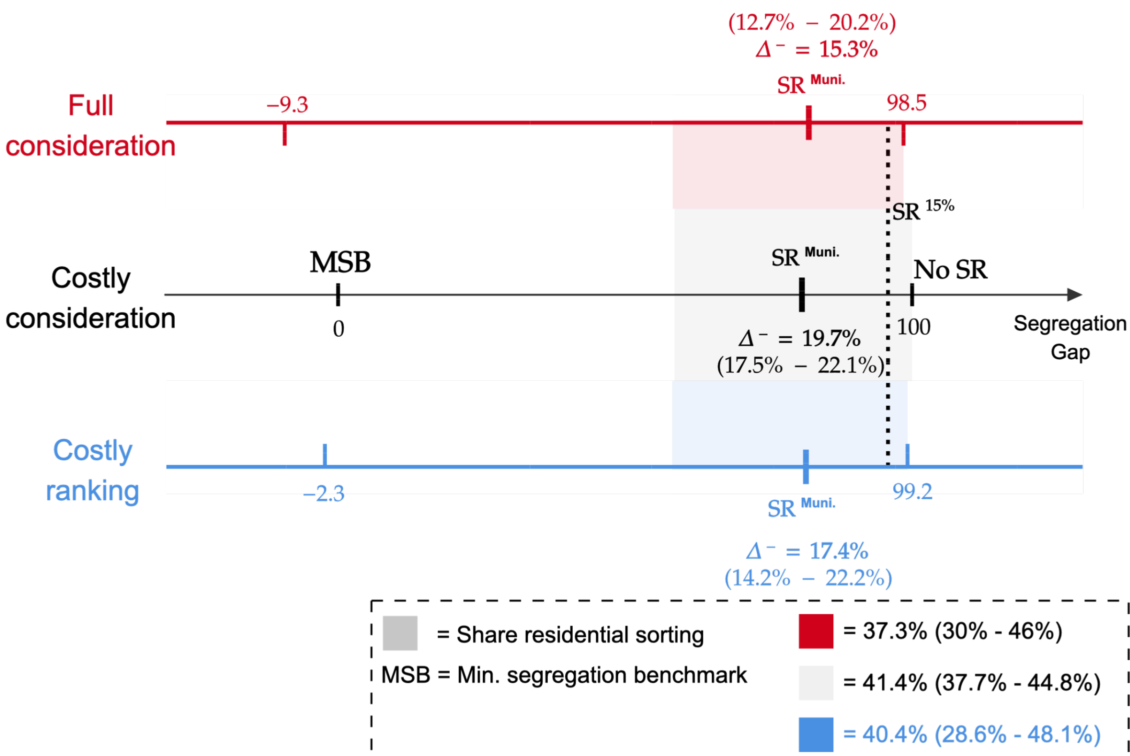
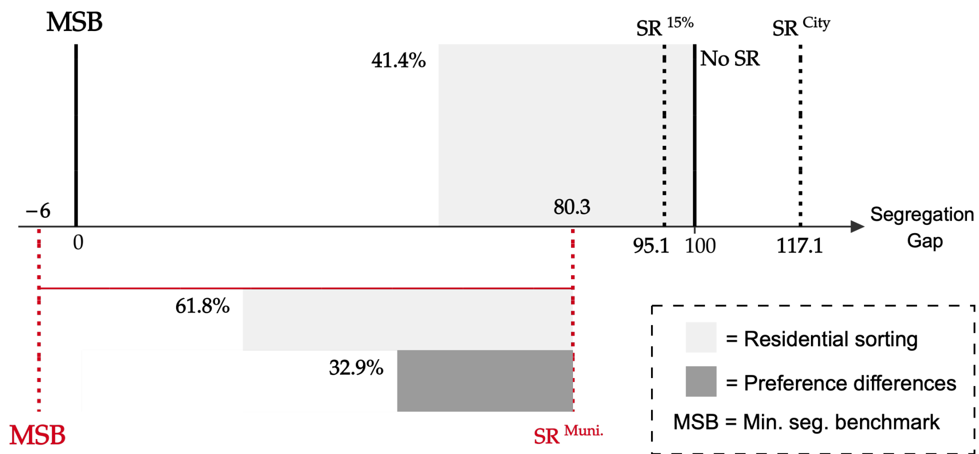
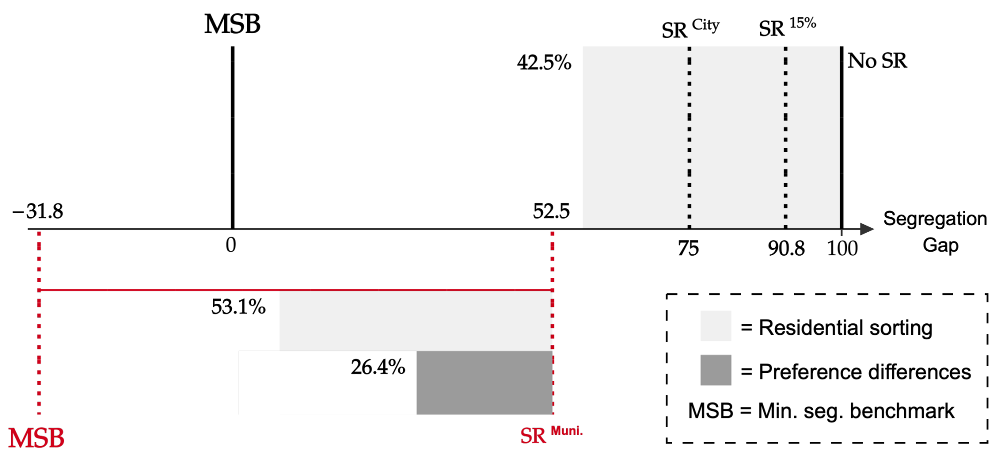


Figure 1.6: Educational Segregation Decomposition in Costly Consideration Model: Three-Year Average Including Only Counterfactual Applicant's Sample



(a) Assigned and unassigned applicants



(b) Only assigned applicants

Tables

Table 1.1: Priority orderings in each reserve group

Priority	(1) Disability	(2) Academic Ach.	(3) low-SeS	(4) Regular
6 (highest)	Secured Enrollment			-
5	Disability	Academic Achiev.	Sibling (S)	Secured E.
4	Sibling (S)	Sibling (S)	Sibling (D)	Sibling (S)
3	Sibling (D)	Sibling (D)	low-SeS	Sibling (D)
2	One parent of the applicant works at the school			
1	The applicant is a former student of the school			
0	No priority			

Note: Sibling (S) applies when the sibling of the applicant is enrolled in a program at the same school of the program to which the applicant is applying. Sibling (D), where the "D" stands for dynamic as opposed to a static priority, applies when the sibling is assigned to a program at the same school by the algorithm.

Table 1.2: Program Characteristics

2019 (1,237 Programs)	N obs	Mean	Min	Max
Subsidized School (=1)	1237	0.65	0	1
High Achievement Category (=1)	1217	0.11	0	1
Medium Achievement Category (=1)	1217	0.48	0	1
Full School Day (=1)	1237	0.17	0	1
Morning School Day (=1)	1237	0.43	0	1
Positive Price (=1)	1237	0.23	0	1
Price (UF^1)	290	1.169	0.023	5.095
2020 (1,247 Programs)	N obs	Mean	Min	Max
Subsidized School (=1)	1247	0.65	0	1
High Achievement Category (=1)	1224	0.11	0	1
Medium Achievement Category (=1)	1224	0.48	0	1
Full School Day (=1)	1247	0.18	0	1
Morning School Day (=1)	1247	0.42	0	1
Positive Price (=1)	1247	0.22	0	1
Price (UF^1)	273	0.096	0.002	3.258
2021 (1,249 Programs)	N obs	Mean	Min	Max
Subsidized School (=1)	1249	0.65	0	1
High Achievement Category (=1)	1225	0.11	0	1
Medium Achievement Category (=1)	1225	0.48	0	1
Full School Day (=1)	1249	0.18	0	1
Morning School Day (=1)	1249	0.42	0	1
Positive Price (=1)	1249	0.19	0	1
Price (UF^1)	233	1.211	0.021	3.147

Note: UF means *Unidad de Fomento*, which is a unit of account used in Chile to adjust the Chilean peso for inflation (it is a non-circulating currency). On December 31, 2019, the exchange rate was 0.03 UF/USD. Same date in 2020, it was 0.0258; and in 2021, 0.0309.

Table 1.3: Applicants' characteristics: Considered and Not Considered in Demand Estimation

	Sample Population			Not included in sample		
	Mean	Min	Max	Mean	Min	Max
High-Accuracy Georef.	1	1	1	0.75	0	1
Low-SeS	0.42	0	1	0.34	0	1
Application Lenght[1]	3.11	1	44	3.78	1	119
Distance[2] First Preference (km)	1.17	0.003	13.74	3.45	0.002	88.69
Assigned	0.92	0	1	0.90	0	1
Distance[2] Assigned (km)	1.19	0.003	13.74	3.51	0.007	88.69
Assigned in their Municipality	0.93	0	1	0.55	0	1
N Schools in a 2km radius	21.40	0	48	20.13	0	48
High Achiev. Schools in a 2km radius	1.91	0	7	1.84	0	7
N Applicants	78,816			34,121		

Note: Even when there is a minimum length of application at entry levels (2 programs), we can observe some cases of 1 program applications due to rurality (for applicants not in demand estimation), being located in urban areas with no enough PreK programs nearby (usually in peripheral Municipalities near the limits of urbanity), or having applied to more than one program but then face schools or programs closures. The percentage of applicants with one program in their application in the Municipalities of Santiago city was 0.15% in 2019, 0.22% in 2020, and only 0.02% in 2021.

[2] Distances to the first preference and the school assigned only consider schools in the Metropolitan region, to not account for students moving to other regions, which leads to maximum distances of more than 2,000 kilometers.

Table 1.4: Share of applicants used to estimate preferences

Apps. in counterfactual	Share
> 5,000	15%
$[5,000, 4,000)$	20%
$[4,000, 3,000)$	25%
$[3,000, 2,500)$	30%
$[2,500, 2,000)$	35%
$[2,000, 1,500)$	42.5%
$[1,500, 1,000)$	50%
< 1,000	100%

Table 1.5: Estimated Mean Utilities

Panel A: Mean utilities (kms)

	Costly consideration		Costly ranking		Full consideration	
	wav	range	wav	range	wav	range
$\overline{\delta}^{Low-SeS}_{HA\ Cat}$	.664	-.227 - 2.491	.772	-.302 - 2.602	.591	-.362 - 2.399
HA vs MA Cat	.320	-.603 - 1.588	.350	-.610 - 2.351	.311	-.483 - 1.639
HA vs LA Cat	.735	-.183 - 2.457	.793	-.329 - 3.819	.719	-.240 - 2.621
$\overline{\delta}^{No-SeS}_{HA\ Cat}$	.902	-.238 - 2.670	1.054	-.094 - 3.011	.827	-.343 - 2.597
HA vs MA Cat	.470	-.238 - 2.001	.504	-.290 - 2.214	.459	-.302 - 2.023
HA vs LA Cat	.952	-.038 - 2.473	1.024	-.143 - 2.740	.936	-.071 - 2.465
$\overline{\delta}^{LS}_{HA} - \xi^{LS}_{HA}$	.702	-.294 - 3.089				
HA vs MA Cat	.358	-.904 - 2.564				
HA vs LA Cat	.758	-.407 - 3.606				
$\overline{\delta}^{NS}_{HA} - \xi^{NS}_{HA}$	.933	-.320 - 3.809				
HA vs MA Cat	.511	-.665 - 2.990				
HA vs LA Cat	.938	-.318 - 3.080				

Panel B: Estimation sample

	All Models	
	wav	range
$N_{Applicants}$	841	140 - 1,293
$N_{Programs}$	85	4 - 200
Applicants/Programs	12.4	3.3 - 76.4
Choices/Programs	39.1	10.2 - 202.6

Note: $Low-SeS$ refers to applicants with low socioeconomic status, and this term is abbreviated as LS in the third group of rows in the table. Likewise, $No-SeS$ refers to applicants who do not have low socioeconomic status (non Low-SeS or no-SeS), and it is abbreviated as NS in the fourth group of rows.

Table 1.6: Main Estimated Parameters

	Costly consideration		Costly ranking		Full consideration	
	wav	range	wav	range	wav	range
λ_{SIB}	2.772	1.187 - 8.034	2.869	1.233 - 8.489	2.804	1.209 - 7.483
	(29.7)	(7.85 - 42.47)	(29.9)	(8.73 - 42.36)	(30.5)	(8.84 - 44.38)
λ_{PW}	4.094	-.038 - 12.278	4.494	.023 - 16.124	4.329	.026 - 14.89
	(6.5)	(-.01 - 13.7)	(6.5)	(.01 - 13.81)	(6.5)	(.01 - 13.76)
$\Delta_{d,C19}$	.3%	0% - 3.3%	.2%	-.4% - 2.7%	.3%	0% - 2.8%
	(.3)	(.04 - .81)	(.2)	(-.09 - .75)	(.3)	(.02 - .79)
$\sigma^2_{\xi^{NSeS}}$	.527	.134 - 4.262	.526	.137 - 5.249	.449	.126 - 3.568
	(4.9)	(.78 - 6.99)	(4.9)	(.51 - 6.81)	(5)	(.77 - 6.7)
$\sigma^2_{\xi^{SeS}}$	.491	.134 - 4.429	.496	.118 - 4.255	.433	.115 - 3.99
	(4.8)	(.71 - 6.7)	(4.7)	(.68 - 6.53)	(4.8)	(.66 - 6.36)
σ^2_{ϵ}	.781	.047 - 7.433	.829	.046 - 7.8	.761	.048 - 6.148
	(45.1)	(8.85 - 77.75)	(45)	(9.51 - 80.29)	(46.9)	(9.76 - 77.8)
c^{NSeS}	.043	.001 - .381	.023	.004 - .068		
	(6.2)	(1.3 - 8.31)	(7.3)	(.99 - 12.87)		
c^{SeS}	.037	.003 - .391	.021	0 - .096		
	(5.3)	(1.44 - 6.99)	(5.3)	(.03 - 8.13)		
σ^2_{zeta}	.564	.01 - 4.922	.307	.026 - 1.73		
	(4.8)	(1.87 - 6.52)	(6.4)	(1.39 - 9.64)		

Note: weighted av. **t-stats in parenthesis** (if opposite sign from parameter set to zero)

Table 1.7: Segregation Reduction Outcomes in All Models: Three-Year Average Including Only Counterfactual Applicant's Sample

Countefactual Scenario			**Main Outcomes**					
Panel A: Segregation gap			**Integration: Reduction in educational seg. gap (Δ^{-}%)**			**Unassignment (% of group)**		
Model	SR share (%)	Randomized residences	*MA-DD*^{WUN}	*MA-DD*	*DD*	L-SeS	N-SeS	All
	15	No	4.9	9.2	10.6	5.98	8.8	7.9
Costly	0	No	-	-	-	6.54	8.62	7.9
cons.	0	Yes	41.4	42.5	40.9	8.16	8.02	8.01
	Muni.	No	19.7	47.5	51.3	4.26	9.9	7.89
	15	No	3	7.1	9	5.98	8.8	7.9
Full	0	No	-	-	-	6.66	8.6	7.92
cons.	0	Yes	37.3	40.1	38.7	6.76	6.69	6.67
	Muni.	No	15.3	38.9	41.6	4.23	9.91	7.88
	15	No	3.9	8.4	10.1	5.98	8.8	7.9
Costly	0	No	-	-	-	6.67	8.72	8
rank.	0	Yes	40.4	43.7	41.6	6.52	6.29	6.34
	Muni.	No	17.4	43.2	46.1	4.33	9.95	7.94

Panel B: Segregation values		**Segregation Measure**					
		MA-DD^{WUN}			*MA-DD*		
Model	Benchmark	min/max	wav	gap	min/max	wav	gap
Costly cons.	Achievable min.	.036/.136	.079	.035	.027/.135	.083	.034
	Baseline (max)	.04/.189	.114		.037/.19	.117	
Full cons.	Achievable min.	.041/.117	.076	.038	.044/.115	.077	.039
	Baseline (max)	.039/.189	.114		.037/.192	.117	
Costly rank.	Achievable min.	.036/.126	.079	.035	.045/.124	.081	.036
	Baseline (max)	.039/.191	.114		.037/.192	.117	

Table 1.8: Other Outcomes in All Models: Three-Year Average Including Only Counterfactual Applicant's Sample

Countefactual Scenario			**Other Outcomes**					
			Achievement Gap (% of)				**% of ideal occup.**	
	SR	Randomized	**HA category**		**MA category**		**All Applicants**	
Model	share (%)	Residences	L-SeS	N-SeS	L-SeS	N-SeS	HA cat.	MA cat.
Costly cons.	15	No	16.4	20.4	52.3	53.9	84.5	82.9
	0	No	16.2	20.4	52.1	53.9	84.4	82.9
	0	Yes	16.6	20	52.1	53.4	84.9	82.5
	Muni.	No	17.3	19.7	52.1	54	84.5	82.9
Full cons.	15	No	16.4	20.4	52.3	53.9	84.5	82.9
	0	No	16.2	20.4	52.1	53.9	84.4	82.9
	0	Yes	16.3	19.8	52.2	53.2	85.2	83.2
	Muni.	No	17.2	19.8	52.2	54	84.5	82.9
Costly rank.	15	No	16.4	20.4	52.3	53.9	84.5	82.9
	0	No	16.3	20.4	52.1	54	84.4	82.9
	0	Yes	16.5	19.8	52.2	53.5	85.3	83.5
	Muni.	No	17.3	19.7	52.1	54.1	84.5	82.8

Table 1.9: Counterfactual Segregation Values in Example Municipalities: Costly Consideration Model and Counterfactual Applicant's Sample Results

Maipú	**2019**		**2020**		**2021**	
Counterfactual	$MA\text{-}DD^{WUN}$	$MA\text{-}DD$	$MA\text{-}DD^{WUN}$	$MA\text{-}DD$	$MA\text{-}DD^{WUN}$	$MA\text{-}DD$
Max. Benchmark	.144	.153	.151	.152	.150	.148
Min. Benchmark	.074	.083	.069	.069	.080	.082
No Res. Sorting	.124	.134	.132	.134	.132	.135
SR Municipality	.135	.130	.147	.143	.134	.128
La Cisterna	**2019**		**2020**		**2021**	
Counterfactual	$MA\text{-}DD^{WUN}$	$MA\text{-}DD$	$MA\text{-}DD^{WUN}$	$MA\text{-}DD$	$MA\text{-}DD^{WUN}$	$MA\text{-}DD$
Max. Benchmark	.106	.116	.088	.088	.156	.154
Min. Benchmark	.114	.116	.106	.106	.115	.114
No Res. Sorting	.104	.109	.114	.111	.130	.130
SR Municipality	.100	.094	.082	.076	.127	.121

Table 1.10: Additional Counterfactuals in Costly Consideration Model: Three-Year Average Including Only Counterfactual Applicant's Sample

Panel A: Main Outcomes

Demand-side simulation	SR share (%)	**Integration: Reduction in educational seg. gap (Δ^{-}%)** *MA-DD*^{WUN}	*MA-DD*	*DD*	**Unassignment (% of group)** L-SeS	N-SeS	All
None	Muni.	19.7	47.5	51.3	4.26	9.9	7.89
None	City	-17.1	25	32.8	2.28	10.4	7.95
$\Delta^{+}\delta_{HA}$	Muni.	33	58.8	62.4	4.41	9.66	7.8
$\Delta^{-}c_i$	Muni.	21.6	49.7	53.5	4.15	9.76	7.76
$\Delta^{-}\tau_t$	Muni.	38.8	57.9	60.3	2.9	6	4.95
All	Muni.	50.1	69.5	71.8	2.72	5.55	4.61
All	0	3.9	10.3	9.1	4.46	4.66	4.57

Panel B: Other Outcomes

Demand-side simulations	SR share (%)	**Achievement Gap (% of group)** **HA category** L-SeS	N-SeS	**MA category** L-SeS	N-SeS	**% of ideal occup. All Applicants** HA cat.	MA cat.
None	Muni.	17.3	19.7	52.1	54	84.5	82.9
None	City	17.6	19.8	52.3	53.9	84.4	82.8
$\Delta^{+}\delta_{HA}$	Muni.	21.6	23	49.4	52	90.9	81.1
$\Delta^{-}c_i$	Muni.	17.3	19.6	52.1	53.8	84.4	82.7
$\Delta^{-}\tau_t$	Muni.	17.4	19.4	52.2	53.8	85.2	84
All	Muni.	21.3	22.2	49.6	51.8	90.9	82.2
All	0	19.1	23.7	50.1	51.4	90.8	82.3

Chapter 2

The Potential of Smart Matching Platforms in Teacher Assignment: The Case of Ecuador

2.1 Introduction

Making a "good" or optimal choice is a difficult task, particularly when faced with information frictions. Providing agents with personalized information can facilitate the decision-making process. Such informational interventions are potentially beneficial not only at the individual level (by bettering people's outcomes) but also at the system level (by improving efficiency). The effects of informational interventions have been studied in the context of school selection (Arteaga et al., 2021; Cohodes, Corcoran, Jennings, and Sattin-Bajaj, 2022; Weixler, Valant, Bassok, Doromal, and Gerry, 2020; Andrabi et al., 2017), financial choices (Saez, 2009; Duflo and Saez, 2003), health care (Kling, Mullainathan, Shafir, Vermeulen, and Wrobel, 2012), and consumer behavior (Allcott and Rogers, 2014; Jin and Leslie, 2003), with researchers widely concluding that they can have a low-cost, positive impact on the decision-making process.

In Chapter 1, we discussed how information frictions play a crucial role in the formation of an application portfolio and simulated a policy to reduce the cost of acquiring information. However, we did not delve into specific policies to implement this. This chapter explores the role of reducing information frictions by providing timely, personalized feedback in teacher job markets. Given that applicants for teaching positions have already undergone the necessary examinations and requirements to apply, they can be expected to invest considerable effort into the portfolio formation process. However, the scope of choice is vast, as teachers can potentially move to different cities in the country to begin or continue their teaching career. Moreover, information about demand for different teaching positions might be unavailable or unrepresentative of the current environment. On average, teachers tend to prefer working close to where they grew up or live, in urban areas, or in schools with specific features such as higher enrollment, better infrastructure, or a higher percentage of socioeconomically advantaged students (Bertoni, Elacqua, Hincapié, Méndez, and Paredes, 2019; Boyd, Lankford, Loeb, and Wyckoff, 2005; Reininger, 2012). These preferences can result in inefficiencies in the job market, where many candidates are unable to get a vacancy in more desirable schools that are in high demand, while positions in other schools, often vulnerable and remote, go unfilled. In fact, such positions may remain vacant despite the existence of willing candidates who might have applied if they had known that doing so would increase their chances of securing a job.

We tested a low-cost intervention that provides teachers with information aimed at increasing their chances of securing a position and improving system outcomes (i.e., increasing the score of assigned teachers and the number of filled positions). The intervention was implemented in Ecuador as part of the "I Want to Be a Teacher" (*Quiero Ser Maestro*) program, which assigns teachers to schools through a centralized choice and assignment system that uses a deferred acceptance algorithm (Gale and Shapley, 1962).[1] Specifically, to better inform teacher candidates who participated in the Ecuador's 2021 selection process, the latter received a personalized report via WhatsApp and email containing a summary of their application.[2] For candidates whose estimated risk of not being assigned was "high" (above a defined cutoff level), the report also included a non-assignment risk warning and a list of recommended schools where they had higher chances of securing a position.[3] We evaluate the impact of the intervention on teachers' submitted ranked ordered lists (ROLs) and assess the equilibrium effect on their probability of assignment.

To this end, we use a regression discontinuity design that allows us to estimate the causal effect of providing teachers with information about their non-assignment risk and possible schools to which they could apply. Similar to Arteaga et al. (2021), the running variable is defined as estimated non-assignment risk and the cutoff is set to 30%. Additionally, after the end of the application period but before results were distributed, we conducted a survey aimed at measuring applicants' opinions on different dimensions of the process, as

[1] Centralized choice and application systems (CCAS) refers to algorithmic assignment processes that take applicants' preferences and priorities into account when allocating available vacancies (see www.ccas-project.org). Elacqua, Olsen, and Velez-Ferro (2020) identify several advantages of this kind of system for teacher assignment: (i) a potentially sharp reduction in search costs, (ii) increased transparency in assignment criteria, thanks to the use of scoring systems that facilitate the prioritization of teachers with higher potential, and (iii) efficiency gains in relation to teacher preferences, due to assignment algorithms suitable for improving school-teacher matching, potentially impacting teacher satisfaction and retention rates. The authors report that CCAS have been successfully implemented in recent years in a number of countries including France, Germany, Turkey, Peru, Portugal and Ecuador. For additional evidence on the benefits of CCAS for teacher assignment see, for example, Pereyra (2013); Terrier (2014); Cechlárová, Fleiner, Manlove, McBride, and Potpinková (2015); Dur and Kesten (2019); Combe, Tercieux, and Terrier (2022).

[2] The personalized report was prepared using a personalized url and a responsive front-end design that was adapted to mobile devices.

[3] In other studies, such as Arteaga et al. (2021), the non-assignment risk intervention focuses on helping applicants to find and add more alternatives (and potentially re-order their portfolio). In this system, teacher candidates can only apply to a maximum of 5 schools. As a result, some needed to change their original application to improve their chances of obtaining a position.

well as assignment beliefs and their knowledge of available alternatives within their area of specialization.

We find that receiving the warning and school vacancy recommendations increased the probability of changing their ROL by 52%.[4] The effect on the equilibrium chances of being assigned to a school is an increase of 37% at the discontinuity. As explained in Section 2.6, this result is an equilibrium effect in the sense that it is affected by changes in the applications of all the participants, both close to and far from the discontinuity. Additionally, the descriptive results presented in Section 2.6 suggest that the overall program results improved after the intervention because additional positions were filled, or because the relatively high-performing candidates who received the personalized report and changed their application displaced lower-score applicants.[5]

Our study adds to the literature on informational frictions by showing the positive impact of a low-cost informational intervention on teacher preference and assignment. The intervention also seems to have generated system-level efficiency gains. This is of particular importance, given that teachers are the most expensive schooling input and the greatest influential educational factor for student outcomes. Several papers have shown the effects of providing agents with information. For instance, in a similarly configured intervention, Arteaga et al. (2021) use "real-time" feedback on applicants' admissions probabilities in the context of student school choice in the Chilean CCAS to study the effect of non-assignment risk warning pop-ups and SMS/WhatsApp messages on submitted ROLs and assignment probability. The authors find that this real-time feedback led families to add more schools and increased their likelihood of assignment to a more preferred school,[6] both on the order of a 20-25% increase relative to applicants with a similar non-assignment risk that did not receive the warning.

In a similar vein, other studies have shown that information about the characteristics of available choices can guide individuals to make better decisions. For example, Hastings and Weinstein (2008) and Allende et al. (2019) demonstrate that when lower-income families have access to information about school quality, they are more likely to choose high-performing schools.

Finally, our findings have important policy implications when it comes to reducing inefficiencies in teacher assignment and improving educational effectiveness. Indeed, teacher recruitment and assignment processes can be lengthy and costly (Allen, 2005). Yet, given that teachers have a strong and long-lasting impact on student outcomes (Rivkin, Hanushek, and Kain, 2005; Kane and Staiger, 2008; Chetty, Friedman, and Rockoff, 2014a,b), school vacancies should ideally be filled on time and with the best possible candidates. Otherwise, there is a risk of assigning vacancies to less-qualified teachers through temporary contracts

[4] More precisely, these are the estimated effects of the RDD described in Section 2.5 and presented in Section 2.6; namely, the local average treatment effect at the 30% non-assignment risk threshold.

[5] Because the algorithm is based on candidates' selection assessment score and preferences, these high-performing treated candidates were ranked higher by schools.

[6] The paper also shows that treated applicants ended up assigned to better-quality schools at the end of the assignment process.

(Bertoni, Elacqua, Marotta, Martínez, Méndez, Montalva, Olsen, Santos, and Soares, 2020), which can have a negative impact on student achievement (Marotta, 2019). Reducing inefficiencies in teacher assignment can ultimately improve education quality if high-performing candidates who are unable to obtain a position due to congestion at their preferred schools are encouraged to apply to less demanded schools with unfilled vacancies or with slots that are instead filled by candidates with lower scores.

To improve equity and efficiency in teacher assignment, some policies use monetary incentives to influence teacher preferences, though this has been found to have a small or non-significant effect (Clotfelter, Glennie, Ladd, and Vigdor, 2008; Falch, 2011; Glazerman, Protik, Teh, Bruch, and Seftor, 2012; Springer, Swain, and Rodriguez, 2016; Rosa, 2017; Bueno and Sass, 2018; Feng and Sass, 2018; Elacqua, Hincapie, Hincapié, and Montalva, 2022c). More recently, studies have examined the impact of low-cost non-monetary interventions on teacher preferences. For instance, Ajzenman, Bertoni, Elacqua, Marotta, and Vargas (2020) evaluate an intervention aimed at attracting teacher candidates to rural and more vulnerable schools in Peru using behavioral nudges that cultivated their extrinsic and intrinsic motives for pursuing these alternatives. The nudges led to a 3.4% increase in the probability that a candidate included a vulnerable school in their choice set, and a 6% increase in the probability that the applicant would be assigned to one of these schools. Similarly, Ajzenman, Elacqua, Marotta, and Olsen (2021) assess an intervention in Ecuador that highlighted teaching vacancies in vulnerable schools and displayed them in first place within an application platform. The intervention increased the share of applicants that included these schools in their portfolio by almost 9% and raised their probability of assignment by 4%. We build on this literature by testing the effectiveness of a low-cost intervention that provides non-assignment risk information and direct recommendations to teacher candidates. Our results suggest that information that reduces search frictions can have a significant effect on teachers' preferences and, thus, may complement other policies aimed at providing extrinsic or intrinsic incentives.

The chapter is organized as follows. Section 2.2 describes the institutional context of the Ecuadorian teacher assignment system. Section 2.3 presents the design and implementation of the intervention. Section 2.4 provides descriptive statistics and Section 2.5 introduces the empirical strategy employed in the analysis. Section 2.6 discusses our results and Section 2.7 concludes.

2.2 The Ecuadorian Teacher Assignment System

Since 2013, the Ecuadorian Ministry of Education has implemented a centralized teacher selection and assignment program known as *Quiero Ser Maestro* (QSM). Here, we focus on the seventh annual intake to the QSM program (QSM7), which took place in 2021.

The QSM includes three phases: i) the eligibility phase, ii) the "merits and public examination" (*méritos y oposición*) phase, and iii) the application phase. In the eligibility phase, teacher candidates must pass a psychometric test comprised of personality and reasoning questions, and a knowledge test that is specific to the specialty area for which candidates are applying (e.g., general primary education, secondary school math, etc.). To proceed to the next phase, candidates must have passed the psychometric test and obtained a minimum score of 70 percent in the knowledge exam.

In the "merits and public examination" phase, candidates are evaluated according to their academic and professional credentials (the merits portion). Candidates who pass the merits portion move on to the "public examination" portion. In the latter, candidates are scored based on their performance in the specific knowledge test taken in the first phase[7] and a mock class.[8] Candidates must obtain a minimum score of 70 percent in this mock class in order to apply for job offers.

The total score for the "merits and public examination" phase is weighted 35% for the merits portion and 65% for the public examination portion, as described in Table B.2.1 in Appendix B.2. Additionally, candidates can receive up to ten "bonus" points for meeting certain criteria, such as living in an indigenous community, having a disability, or residing in the same educational circuit where their preferred school is located.[9]

In the last phase, eligible candidates have 10 days to apply for up to 5 open positions in schools located in any region of the country by submitting a ranked ordered list (ROL) on an online platform. Once candidates submit their application, they cannot change it during the initial 10-day period. However, after this application period is closed, candidates are allowed to go back to the platform and modify their preferences during a two-day validation period. In this validation phase, they have a single opportunity to add, delete, and change the order of their submitted choices.

ROLs and school rankings based on the candidates' final score are then processed using a deferred acceptance algorithm (Gale and Shapley, 1962). The candidates' final scores take into account the results obtained in each of the components listed in Table B.2.1 in Appendix B.2 and the bonus points computed according to their choices.[10] A more in-depth description of the QSM selection process is provided by Drouet and Westh (2020).

[7] Scores on the disciplinary knowledge test are also admissible.

[8] Its consists of a 40-minute teaching assessment in which the teacher demonstrates his/her teaching ability on a topic in his/her specialty.

[9] Appendix B.3 provides details on the bonus score.

[10] It is important to note that the rationale behind the bonus scores is not made clear to applicants when they are applying, since they are assigned after the application period. Applicants also have little insight of the points awarded to other candidates. This has two implications. First, since applicants do not know their exact final score, it is more difficult for them to assess their assignment probabilities for each vacancy and, therefore, to act upon these probabilities and change their choices. Moreover, not knowing their bonus scores makes it harder to obtain personalized feedback on assignment probabilities, which this intervention aims to address. The fact that candidates can only apply to a maximum of 5 schools (as opposed to an unlimited number of positions) makes the preference selection exercise even more challenging. This restriction clearly influences teacher behavior, as shown by the share of applicants applying to 5 schools (Figure B.1.1 in Appendix B.1), and by the survey result presented in Figure B.5.2 of Appendix B.5, which shows that 92% of the teachers that responded the survey would have liked to apply to more schools. However, we leave the analysis of the effect of the 5-school restriction to future studies.

Ecuador's teacher selection and assignment process has significantly improved over time. In 2019, for example, the country's Ministry of Education changed the QSM to allow candidates to apply directly to schools rather than school districts, which reduced the margin of discretion (Drouet and Westh, 2020). However, there remain inefficiencies in the selection process. For example, some vacancies are congested, while others (vulnerable and remote) have few applicants (Bertoni et al., 2020). In fact, Elacqua, Westh Olsen, and Velez-Ferro (2021) show that 27% of vacancies went unfilled in the 2019 QSM program, mainly in schools of low socioeconomic status. Our intervention aims to further diminish these inefficiencies and improve market outcomes by reducing informational frictions.

2.3 Intervention

As discussed in the previous section, the QSM7 application period consisted of two stages: the application stage, in which candidates could submit a single ranked list of their choices in a 10-day period, and a two-day validation stage during which they were able to modify their application. We implemented our intervention between these two stages. Specifically, a day before the application stage ended, we processed the applications of teachers that had participated up to that point (or the "pre-validation applicants"), and used this information to provide them with a personalized report. Before the end of the validation period, 20.3% of contacted teachers opened the report.

A template of the personalized report can be found in Figure B.4.1 in Appendix B.4. Applicants with no risk of not being assigned received an introductory message, an invitation to visit the application interface (panels B.4.1a and B.4.1e respectively), and a summary of their application, including the following information about each of the selected schools: location, distance from the candidate's home, type of financing, number of enrolled students and number of vacancies (panel B.4.1b). Applicants with a high risk of non-assignment received the same information and, additionally, a warning of non-assignment, along with a list of recommended schools (panels B.4.1c and B.4.1d). Thus, the two main groups of the intervention are formed by the candidates who received the sections with the warning and recommendations (treatment group) and those who only got a summary of their application (control group).

The personalized reports were sent via WhatsApp to applicants that provided a valid phone number during the process.[11] Additionally, some applicants received their personalized report via email. For the emailing, priority was given to applicants without a telephone number, while a sample of the remaining applicants was randomly selected.[12]

We will now discuss how we defined the risk of non-assignment, how we constructed the lists of recommendations and how treatment and control groups were formed.

[11] Some of these numbers were validated through complementary communications inviting registered teachers to apply.

[12] The government had a restriction on the number of emails that could be sent each day and emails had to be sent from the official government account.

Non-Assignment Risk

To determine which applicants were at risk of non-assignment, we estimated this risk by simulating the partial assignment using the following procedure:

- **Applicants and simulations:** One day prior to the closing date of the application stage (before the validation stage started), we generated 200 assignment simulations[13] with 19,190 pre-validation applicants.[14] We also sampled 40%[15] of the 2,527 potential applicants who had not participated in the process at the time of the calculation. Of the latter, the information used was mainly their score and location.[16]

- **ROLs:** For pre-validation applicants, we considered their reported preferences in each simulation. Since we did not have information on the preferences of sampled applicants, we followed Arteaga et al. (2021) to match each sampled applicant with an existing applicant (pre-validation applicant) to impute their preferences. To do this, we searched for the "closest" pre-validation applicant for each sampled applicant as follows:

    1. All applicants within the same geographic unit and specialty were considered. The geographical scales, in increasing order of size, were the circuit (*circuito*), canton, and province.[17]

    2. Among the applicants drawn from the same geographic unit, we selected those of the same specialization and tercile score. If there were no applicants within the same tercile, we used the closest one(s).

    3. Where there was more than one applicant in the same geographic unit and tercile (or the closest tercile when applicable), we selected the match randomly.

[13] Given the number of pre-validation applicants and the fact that, in nearly all iterations, a considerable fraction would either be assigned or remain unassigned based on their application score, the goal was to generate enough dispersion in the estimated risk to be able to implement a regression discontinuity design. To that end, as shown in Figure B.1.2 in Appendix B.1, we ended up with around 5 applicants in each 0.5% risk bin around the discontinuity (200 simulations implies that we estimated risks in 0.5% intervals).

[14] A total of 22,015 eligible teachers participated in the program. However, 2,527 candidates had not yet submitted their applications by August 4, when we processed the partial pre-validation data. Candidates whose applications were submitted after our deadline were not considered as pre-validation applicants. Additionally, we considered only candidates from specialties with at least 20 applicants prior to validation and 80 after the validation period, leaving us with a total of 19,190 pre-validation applicants.

[15] This percentage was defined based on the guidance of policy-makers who, at the time, estimated that approximately 93% of all applicants would participate based on previous QSM programs. The participation rate ended up being significantly higher, implying that our risk estimation was somewhat conservative.

[16] Location has been extensively documented in the literature as a key determinant of teachers' preferences for schools (Bertoni et al., 2019; Boyd et al., 2005; Reininger, 2012; Rosa, 2017).

[17] The Ecuadorian educational system is split into two regimes, one for the interior (Sierra) and another for the coast (Costa). In addition to these regimes, the territory is divided into nine administrative zones, which are further divided into educational districts and educational circuits.

- **Scores:** After generating the preferences in each simulation, we calculated the final score of each applicant, which corresponds to the sum of the "merits and public examination" score and the bonus. Although we had data for the first component, unfortunately neither we nor the applicants knew the bonus score that applicants would receive at each school. However, it was possible to anticipate part of the bonus score using the bonus criteria described in Appendix B.3. In cases where bonus criteria could be identified in the registration data (e.g., when the applicant resides in the province where the school is located), we assigned bonus points. We also generated a random uniform bonus between 0 and 10 points to represent bonuses that we could not identify in the available data, truncating final bonus scores at 10, in line with the rules of the process.[18]

- **Algorithm:** Following Gale and Shapley (1962), we ran a deferred acceptance assignment algorithm for each simulation.

  In summary, what varies from one simulation to another are the sampled applicants, the random bonus and the imputation of preferences for sampled applicants when more than one applicant met the matching criteria.

For each pre-validation applicant, we used the 200 simulated assignments to compute the proportion of simulations in which they were not assigned to a position, generating a running variable for the risk of non-assignment. This allowed us to implement a regression discontinuity approach to study the impact of the informational intervention, following Arteaga et al. (2021). We defined risky pre-validation applicants as those who were unassigned in 30% or more of the simulations, using the same cutoff value as in Arteaga et al. (2021). In this sense, we have a sharp discontinuity scenario given that treatment compliance is divided precisely at the 30% cutoff. Figure B.1.2 in Appendix B.1 shows the density of estimated risk for applicants who opened the personalized report, excluding sizable groups of applicants whose risk was evaluated as 0 and 100%.[19] The figure shows that the density of the running variable is similar on both sides of the cutoff.

Recommendations

The objective of the recommendations was to assist applicants with a low estimated probability of assignment in their applications by showing them alternatives where they would have a better chance of obtaining a position. Specifically, risky applicants were pointed to

[18] We implemented this approximation because we did not have data from previous QSM programs to simulate more precisely the potential bonuses. This implies that our risk calculations were less accurate than we would have liked. That said, we could still compare applicants with this imperfect measure and identify participants who were more likely to be at risk of non-assignment.

[19] The inclusion of these groups would make it difficult to visualize the distribution for the 30% cutoff of interest.

vacancies where the scores of other applicants likely to be assigned (if any) were lower than theirs.[20]

When generating the list of recommended schools for treated applicants, we did not consider the general equilibrium effect that some schools might end up very congested if too many applicants were recommended. However, to reduce the concern of generating excessive congestion, and to learn about applicant preferences -particularly, whether they would be willing to apply to schools in other provinces- we did create recommendation lists that varied in the number of recommended schools in provinces other than that where the applicant resided. Additionally, we randomly selected some of the recommendations in all list and implemented a fully randomized recommendation list for some applicants, thus providing additional random variation to the intervention.

The process was as follows:

1. We selected all vacancies in the applicant's area of specialization for which the cutoff score was lower than the applicant's score. The cutoff score was derived from a partial assignment that included only pre-validation applicants and their bonus scores only included identifiable points.[21]

2. If there were more than 10 options, we selected 10 using the following criteria[22]:

   a) For a quarter of the applicants (randomly selected), we included up to 4 recommendations within their province and the remainder were randomly selected among the applicant's feasible options (from the rest of the provinces).

   b) For the rest of the applicants, we included stratified recommendations using the following criteria:

      i. **Geographic criteria**
         - Schools in the same province
         - Schools in provinces included in the application
         - Schools in other provinces

      ii. **Rural**
         - Rural schools

[20] The rationale for not including recommendations for applicants below the 30% risk cutoff was that, given the maximum of 5 choices imposed by the system, recommendations could mistakenly lead to a higher non-assignment risk if they eliminated a lower-risk option from their portfolio. To include recommendations for the latter group, it would be necessary to implement a strategy that accurately communicates the non-assignment risk of each choice. As explained above, this risk could not be precisely estimated, so it was preferred not to recommend low-risk applicants.

[21] The cutoff score of each program was calculated as the score of the last applicant admitted to the program in an assignment in which only applications from teachers who had applied before we implemented this procedure were considered.

[22] 1,000 applicants (10.7% of treated applicants) were mistakenly processed twice in the recommendation algorithm, and some of them received more than 10 recommendations (but no duplicates). Specifically, 832 applicants (8.9% of treated applicants) received more than 10 recommendations.

- Urban schools

These alternatives were sampled using the following procedure:

i. Two rural and urban alternatives in the applicant's province of residence (maximum of 4).
ii. Two rural and urban alternatives in other provinces included in the application (maximum of 8 in total).
iii. One rural and one urban alternative in provinces other than those included in the application or the applicant's province of residence.

If fewer than 10 alternatives met the criteria, the remaining alternatives were randomly sampled from the applicant's viable options. Some applicants had fewer than 10 feasible alternatives.[23]

Treatment and Control groups

The treatment and control groups were selected from the 19,428 teachers who had passed the "merits and public examination" phase and submitted an application by August 4 (i.e., pre-validation applicants). From this group, we omitted teachers of specialties with fewer than 80 registered or 20 pre-validation applicants, which left us with 19,190 teachers.
For our treatment, teachers must meet two conditions: (i) have a high non-assignment risk (above 30%) and (ii) have a high enough score to obtain at least one vacancy in their specialty (i.e., those who have at least one alternative to be recommended, as detailed in Section 2.3). Treatment and control groups were formed using the following procedure:

1. From the group of 19,190 teachers described above, we ruled out applicants who did not meet the second condition. This left us with 14,810 teachers, who we refer to as the analysis group.

2. Of the analysis group, 9,334 were at high risk of not being assigned, so they were placed in the treatment group, while the remaining 5,476 were placed in the control group.

A total of 3,653 (24.7%) teachers in the analysis group opened their personalized report, who we refer to as the "compliers". To estimate the effect of our treatment, we focus exclusively on compliers because we want to study the impact of the informational intervention among comparable teachers who did or did not receive the treatment. Of the compliers, 65.5% belonged to the treated group and 35.5% to the control group (who we call "treatment compliers" and "control compliers", respectively). This means that 25.6% of teachers in the treatment group and 23.7% of those in the control group opened the personalized report.

Finally, by the end of the validation period, which coincides with the end of the QSM, 3 pre-validation applicants dropped their application and 2,388 new eligible teachers submitted an application. Hence, we define post-validation applicants as the group consisting of pre-validation applicants who continued until the end of the process, plus new applicants.

[23]Specifically, 33% of the treated applicants were recommended less than 10 schools.

2.4 Data

We use administrative data from the registration and application process of the 2021 QSM7 program collected by the Ecuadorian Ministry of Education. The data includes individual records of teachers' registrations and choices as well as school-level data with vacancy information.

Individual-Level Data

The dataset contains information on candidates' socio-demographic characteristics (gender, marital status, date and place of birth, ethnicity), residential address, area of specialization, score on the "merits and public examination" phase by category, and ranked school preferences.

Table 2.1 presents descriptive statistics of the registered applicants. Column (1) shows the statistics for all eligible applicants (i.e., all the teachers who passed the examination phase): 72% of applicants are female, 9% belong to an ethnic minority (non-mestizo), 55% are married, 7% hold a master's degree, and 43% have more than 5 years of work experience. The most common specialization for which candidates applied was basic general education from second to seventh grade; this accounts for 22% of all eligible applicants. The province of Guayas, where the large city of Guayaquil is located, is the most common region and comprises 14% of total applicants. On average, teachers are 39 years old and scored about 65 points in the "merits and public examination" phase. We see in column (2) that the statistics are similar for pre-validation applicants.

Columns (3) to (5) present the same statistics for compliers (see Section 2.3). For total compliers (Column (3)), all their indicators are slightly higher compared to those in columns (1) and (2). The largest differences are observed in the percentage of teachers with more than 5 years of experience, the percentage who applied to the most common specialization and the average score. Finally, columns (4) and (5) show the statistics of compliers in the treatment and control groups, respectively. Control compliers are, on average, more educated and more experienced, and have higher scores (which correlates with lower risk) than treated compliers. Additionally, the share of females is higher in the treated group. The differences between both groups are not surprising, and, as explained below, these differences are minor around the discontinuity threshold (30% probability of non-assignment), which is the relevant test for our identification strategy.

School-Level Data

School data provides information on the location, specialties offered and available vacancies for each school. The dataset contains a total of 3,345 schools, 33 specialties and 8,057 vacancies. To generate recommendations, we only consider specializations with at least 80 registered and 20 pre-validation applicants, leaving us with 24 specializations and 8,009 vacancies in our main sample. Table B.2.2 in Appendix B.2 shows the shares of pre-validation applicants by specialty.

Outcomes

We are mainly interested in estimating the effect of the informational intervention described in Section 2.3 on candidates' choices during the validation period (probability of changing the application and/or adding new schools). Additionally, we assess how the intervention affected the equilibrium assignment at the end of the QSM.

Table 2.2 shows that treated compliers change their application twice as often as those in the control group. Specifically, 65% of treated compliers added schools and 35% added positions from the list of recommendations. In contrast, 29% of control compliers added a school to their application and only 6% would have added a position from the list of recommendations if they had received it.[24] The proportion of treated compliers who were assigned increased from 14% before the validation period to 20% after the validation period, while these proportions for control compliers were 99.6% and 91%, respectively. Note that these shifts are due to both changes in the compliers' applications as well as the entry of new applicants who sent their application after August 4.[25] Interestingly, among the treated compliers who were assigned at the end of the process, 61% were assigned to one of the recommended alternatives. This may indicate the importance of such recommendations, a topic we return to in Section 2.6.

2.5 Empirical Strategy

To explore the causal effect of providing information to teachers at risk of non-assignment, we rely on a regression discontinuity strategy. The underlying assumption of this strategy is that observations near either side of the threshold are, on average, similar in all their characteristics except for treatment status.

Formally, treatment is assigned as shown in equation 2.1, where z_i represents the risk of non-assignment of applicant i, and c represents the non-assignment risk threshold of 30%.

$$T_i = \mathbb{1}\{z_i \geq c\} \tag{2.1}$$

[24] Using the procedure described in Section 2.3, we can also identify the list of recommendations that would have been included if candidates with non-risky applications had received them. It is important to note that this is possible because we did not include general equilibrium effects into the recommendation-generating procedure, meaning that including these applicants would not have affected recommendations for treated applicants (and vice versa).

[25] As explained above, candidates initially had 10 days to apply for vacancies. We generated the simulations and produced personalized reports for candidates who sent their applications no later than one day before the end of the 10-day period (so-called pre-validation applicants). The new applicants are those who completed their applications at the very end of the application period and, therefore, did not receive a personalized report.

Figure B.1.3 in Appendix B.1 confirms that the probability of treatment rises sharply at the discontinuity. Consequently, as shown by Imbens and Lemieux (2008), for a given outcome of interest Y, the estimated impact of the treatment at the discontinuity point is given by:

$$\tau = lim_{z\downarrow c}\mathbb{E}[Y|z = c] - lim_{z\uparrow c}\mathbb{E}[Y|z = c] \tag{2.2}$$

In this setting, an appropriate econometric model to estimate the impact of the intervention is the following:

$$Y_i = \beta_0 + \beta_1 T_i + h(z_i) + \varepsilon_i \tag{2.3}$$

Where Y_i represents an outcome of interest, β_1 is the estimator of the treatment effect of the informational intervention, and h is a continuous function of z_i. We specify h as linear or quadratic following Gelman and Imbens (2019).

The main drawback of the regression discontinuity design is that we can only identify the treatment effect at the discontinuity, known as the local average treatment effect (LATE), which implies that we cannot simply extrapolate estimates to the entire population of interest.

2.6 Results

Baseline Results

We are primarily interested in assessing the extent to which the treatment induces changes in reported preferences during the validation period. Although we also explore the effects on assignment, these depend on the equilibrium (see Section 2.3). The latter, in turn, depends on applicants both close to and far from the threshold, new applicants and even on those who did not open their personalized reports but did alter their ROL. Figure 2.1a shows the probability of changing the application as a function of the risk level of the compliers. We observe a clear discontinuity at the 30% non-assignment risk, with a large increase at the threshold in the probability of changing the application. Figure 2.1b confirms the statistical significance of the jump at the threshold by showing the same plot but now with confidence intervals and using the optimal non-assignment risk bandwidth (using the one common MSE-optimal bandwidth selector following Imbens and Kalyanaraman (2012), which minimizes the mean squared error).

Figures B.1.4a and B.1.4b in Appendix B.1 replicates the plots for the probability of obtaining an assignment. We observe that compliers close to the threshold had a large and statistically significant increase in their probability of obtaining an assignment. Figure B.1.5 in Appendix B.1 compares applications before and after the validation period (which we call "partially assigned" and "finally assigned", respectively), graphically demonstrating that this difference is an equilibrium result. Specifically, we see that the difference in assignment probability at the discontinuity is a result of a drop in assignment probability for applicants on the left of the threshold (both close and far), and an increase among applicants on the right of the threshold, specially among those with a higher estimated risk of non-assignment. This suggests that the treatment, on average, induced changes in candidates' applications that helped them obtain an assignment at equilibrium.

Next, we confirm the graphical evidence by formally estimating the effects of the intervention using alternative RDD specifications. Table 2.3 shows the main results for the outcomes of interest. We report 5 models with different specifications and optimal bandwidths. Model (1) is estimated using a parametric approach with a linear interaction and the bandwidth is calculated using the one common MSE-optimal method following Imbens and Kalyanaraman (2012), which minimizes the mean squared error. Model (2) is the same as model (1), except that it calculates two different bandwidths (one for above and one for below the cutoff) instead of one common bandwidth. Model (3) also relies on a parametric regression with linear interaction, but the bandwidth is calculated using the one common CE-optimal method following Calonico, Cattaneo, and Farrell (2020), which minimizes the coverage error of the interval estimator. Model (4) is the same as model (3), but estimates two different bandwidths for either side of the threshold. Finally, model (5) is estimated using a parametric approach with a quadratic interaction and the one common MSE-optimal bandwidth selector. All models use robust standard errors and the total number of observations in the optimal bandwidth are reported. Though the results are relatively consistent among the different models, we will focus on the estimates of model (1) as it is the most standard in the literature.

We observe that choice behavior changed due to the warning and recommendations. Conditional on opening the personalized report, receiving the treatment increased the likelihood that applicants would change their application by 52%. Specifically, the probability of adding a preference to the application increased by about 59%. Treated compliers added, on average, 0.4 schools after the validation period, while control compliers added an average of 0.1 schools. Moreover, conditional on having added a preference, the probability of adding any of the schools recommended in the personalized report increased by 43% (compared to the recommendations that would have been given to control applicants using the same process to generate recommendations). This is consistent with the fact that most teachers were not entirely sure about all the schools to which they wanted to apply, as shown in Figure B.5.1 in Appendix B.5 on the survey results.[26] That is, recommendations seemed to help treated teachers learn about new schools they had not considered before.

The final assignment at the discontinuity shows that treatment was helpful even after non-compliers edited their application during the validation period. Treated compliers were 37% more likely to obtain an assignment than control compliers. Additionally, those who obtained a position were 35% more likely to be assigned to one of the recommended schools. The survey results are consistent with the fact that people evaluated the information received in the personalized reports and acted accordingly. Specifically, 82% of respondents said they

[26] As mentioned in the introduction, the survey was implemented after the application period, but before the results of the QSM were published. See Appendix B.5 for more information about the survey.

wanted more information about their assignment chances. In addition, teachers rated the information received in the personalized report at 8.22 on average, on a scale of 1 to 10 (see Table B.5.3 of Appendix B.5).

In general, the estimates are robust to the different specifications used in Table 2.3. As an additional robustness check, we test the sensitivity of our main specification (model 1) to different arbitrary bandwidths. Table B.2.3 in Appendix B.2 shows that if we vary the bandwidth between 0.1 and 0.3 our estimates lead to the same conclusions. In all cases, the probability of modifying preferences (and ultimately obtaining an assignment) is significantly larger for treated applicants close to the threshold.

To assess the validity of the regression discontinuity design, we test the balance on covariates on either side of the threshold. Table B.2.4 in Appendix B.2 replicates the estimates of Table 2.3 using other outcomes. They are consistently not significant with the exception of the marital status variable in model (5). This implies that observable characteristics are, in general, similar in the neighborhood of the threshold, suggesting that the identifying assumptions are met. Graphic evidence is reported in Figure B.1.6 in Appendix B.1.

Additionally, we further assess the validity of the estimates by introducing a placebo test. To check whether there is any significant effect when we know that there should not be, we use arbitrary fake cutoffs at the 0.5 and 0.2 non-assignment risk levels. Figure B.1.7 in Appendix B.1 shows that there are no unexpected discontinuities at these cutoffs. These results, combined with the covariates test, suggest that the positive effects we find are caused by the informational intervention.

We do not study the content of the recommendations in this chapter, leaving such analysis for future research. Our main goal in introducing variation in the recommended alternatives was to study teachers' preferences and, more specifically, to be able to identify variation in consideration sets (i.e., schools known and of interest to a particular applicant). Though we do not extensively analyze these results here, Table B.2.5 in Appendix B.2 presents evidence suggesting that, as expected, recommendations were more likely to be included when the institution was either in the province where the candidate resided, or in the same province as one of the preferences given in the pre-validation ranking. Additionally, rural institutions are marginally less preferred and male teachers are marginally more likely to add a recommendation (both coefficients show a 1% change in the probability of adding a recommendation).

Heterogeneous Effects

In this section, we explore whether certain factors related to applicant characteristics can explain or amplify our results. To this end, we estimate our RDD model allowing for heterogeneous effects of teachers' gender, marital status, skill level, and experience. We then estimate a specification based on equation (3) in which treatment is interacted with these characteristics. Panel A of Table 2.4 shows the results of the same 5 models described in Section 2.6 for the probability of changing the application. Similarly, Panel B of Table 2.4 presents the findings for the probability of equilibrium assignment.

Our results suggest that males were more affected by the treatment in terms of their likelihood of modifying their application, but the difference in equilibrium assignment is smaller (with a non-statistically significant coefficient). As expected, married people seem to be less affected by the intervention, likely because they are more restricted by location (e.g., they prefer places where their spouse can find better work opportunities) and may therefore be less willing to change their original choices.[27] However, the coefficients of the interaction between treatment and marital status are not significant.

To explore the potential role of skills, we interact the treatment variable with a dummy variable that identifies whether a teacher has a score above the median on the public examination portion (which evaluates specific skills). The results suggest that skilled teachers are no more likely than others to change their original application after treatment. However, the treatment effect on assignment probability is statistically larger for highly skilled teachers (model (1)), an unsurprising result given that these teachers had more potential recommendations thanks to their higher scores. Nevertheless, it is important to note that the non-interacted effect remains economically large (around 20% or more), and that it is also statistically significant in specifications (2) and (4), implying that treated "unskilled" teachers also had better assignment chances than unskilled teachers in the control group.

Similarly, we look at the role of experience by interacting the treatment variable with a dummy variable that indicates whether an applicant has worked for more than 5 years as a teacher. As shown in Table B.2.4 and Panel D of Figure B.1.6, experience shows no change at the threshold and, as we can see in Panel A of Table 2.4, its interaction with treatment has no significant effect on the probability of changing the application. However, we do find that the treatment effect on assignment probability is statistically smaller for experienced teachers (models (1), (2) and (5)), as shown in Panel B of Table 2.4, that could be explained by the large negative correlation between experience and the public examination score (-0.29).

We do not have enough statistical power to analyze heterogeneous effects on the probability of adding a recommended school or on the probability of being assigned to one.[28]

System-Level Outcomes

We now descriptively explore the effects of the intervention on system-level outcomes such as the number of filled vacancies and the general quality of assigned teachers. It should be noted that, due to the large number of applicants relative to offered vacancies, most positions were filled even when only pre-validation applications were considered. Thus, with regard to the overall effect of the intervention, we should expect a relatively small impact in equilibrium.

[27] The direction of the interactions change across bandwidth specifications, so we focus again on our preferred model from column (1).

[28] As shown in Table 2.3, the number of observations in the regressions of Panel C and E are less than the others because the first is conditional on having added something to the application and the second is conditional on having been assigned. This reduces the statistical power of both regressions and increases the standard deviation.

As mentioned in Section 2.3, when generating recommendations we did not consider the general equilibrium effect that some schools might end up very congested if they were recommended to many applicants. That said, we design the intervention to recommend many diverse alternatives, in order to reduce the risk of generating excessive congestion for highly demanded vacancies. Our aim was also, as explained above, to better understand consideration sets and preferences. The negative spillover effects of the recommendations could, in theory, have increased the number of unassigned applicants, as well as potentially reduced the scores of assigned teachers.[29] However, descriptive evidence suggests that, although there was some congestion and a few teachers remained unassigned despite having added recommended schools to their applications, a much larger percentage of teachers who followed the recommendations were assigned. We interpret this as a positive result despite the fact that the intervention was not designed with the general equilibrium in mind. As shown in Table B.2.6 of Appendix B.2, the total number of vacancies filled increased slightly after the intervention.[30] Column (1) considers only the initial applications submitted by pre-validation applicants before changes during the validation period plus the actual final application of teachers who applied only after the validation period. This column therefore represents how the assignment would have ended up if no one had made any changes in the validation period. Column (2) shows the actual scenario in which pre-validation applicants changed their preferences and some teachers only applied during the validation period.[31]

To explore the overall quality of the assigned teachers, we first analyze the scores of those who changed their application. Figure B.1.8 of Appendix B.1 presents the score distribution for teachers moving from partially non-assigned to assigned, as well as those moving from partially assigned to non-assigned. The mean scores of teachers that were assigned to a vacancy is 68.02, which is 1.77 points above those who did not receive an assignment. This provides preliminary evidence that the intervention may have increase overall assignment scores and, thus, the general quality of assigned teachers.[32]

Focusing on the assigned vacancies before and after the validation period, we observe in Figure B.1.9 in Appendix B.1 that the distribution shifts to the right. This shift becomes more pronounced when looking solely at the vacancies that were assigned to different appli-

[29] Because if high scoring teachers compete for the same vacancies, we may end up with a bi-modal distribution of assigned scores for vacancies with high and low demand.

[30] It is important to note that this does not necessarily mean that the intervention's potential to affect the total number of assignments is small. Rather, the magnitude of the effect depends on the overall congestion of the available alternatives. Congestion in the context of the QSM is considerable, with an average of over three applicants for each available vacancy and over 86% of positions filled. Moreover, some vacancies in specific specialties and schools may be structurally unappealing, making reaching a goal of 100% unrealistic. The effect of the intervention on the aggregate also depends on the uptake of the intervention, which in this case was around 20%. This rate could be improved by introducing similar interventions directly within the application interface.

[31] We would ideally want to have a counterfactual scenario without treatment to ascertain the causal effect on the general equilibrium. However, we do not have an identification strategy to estimate who would have changed their application in the counterfactual scenario and how.

[32] Note that 1.77 points is a significant difference, representing an increase of 0.23 standard deviations in the evaluation scores of assigned teachers, and this from an extremely low-cost informational intervention.

cants before and after the validation period (Panel B). Together with the overall increase in assigned vacancies, we interpret these results as positive, although not causal, evidence that the intervention had a positive impact on QSM7 equilibrium outcomes.

2.7 Discussion

This chapter evaluates a low-cost informational intervention in the context of Ecuador's centralized teacher assignment system. We show that teachers in the treatment group, who received and opened a non-assignment risk warning and a list of recommended schools, were much more likely to change their choices and add new schools to their applications. Ultimately, this translated into a significant difference in the equilibrium assignment of teachers close to the treatment threshold. Our results are robust to different specifications, suggesting that these changes were caused by the intervention.

Moreover, the findings point to a positive general equilibrium effect by improving both the average scores of teachers who obtained an assignment and the number of assigned vacancies, even though we did not design the intervention to maximize spillover effects. Similar interventions that incorporate general equilibrium effects in their design might be the subject of future work.

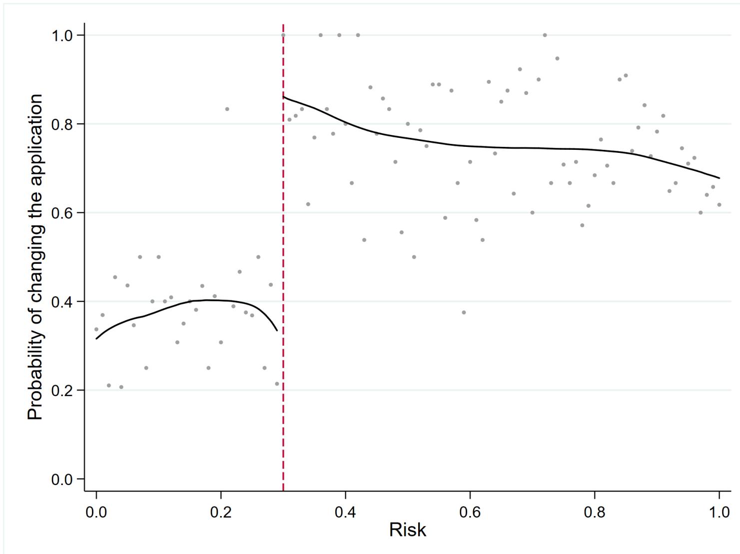
It is important to note that our strategy identifies a local average treatment effect on compliers (teachers who opened the personalized feedback report). This implies that our estimates do not extend directly to the whole population, or to compliers with a non-assignment risk level far from the 30% non-assignment risk cutoff. More research is needed to understand how these results would have changed if we had increased the compliance rate (e.g., with a more salient intervention), or if we had implemented a different threshold (the impact of which could be explored, for example, with an RCT design).

The low-cost intervention studied in this chapter has important policy implications, in the sense that teachers are the most expensive and valuable educational input, and significantly impact student outcomes in the short and long term (Chetty et al., 2014a,b). Improving teacher assignment is likely to have a positive effect on resource allocation and learner success. Moreover, centralized choice and assignment systems have been gaining popularity around the world as a tool for organizing student and teacher application and assignment processes. Informational interventions will potentially play an important role in optimizing the results obtained through these systems. In this chapter, we demonstrate the capacity of such interventions to affect teacher behavior, while other works show their impact on student behavior in the context of school choice (see, for example, Arteaga et al. (2021)).

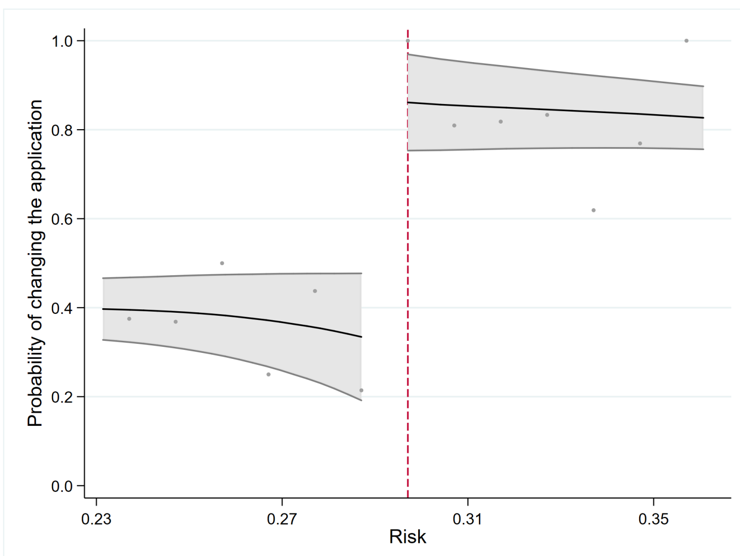
Future studies should consider changes to assignment rule, since these can lead to significant outcome improvements, as for example discussed for the case of targeted reserves in Chapter 1. In the case of the QSM program, a relevant rule to study would be the effect of expanding or eliminating the restriction on portfolio sizes. While such restrictions are often implemented to force applicants to limit their applications to a small number of relevant alternatives (i.e., schools where they would actually be willing to work), it introduces strategic considerations into the submitted preferences, particularly in contexts with high-congestion, such as the QSM program. Moreover, Figure B.1.1 in Appendix B.1, illustrates that approximately 90% of applicants end up ranking the maximum number of alternatives allowed. In other settings, it has been shown that applicants face significant difficulties in formulating optimal application strategies (see, for example, Kapor et al. (2020)), which can also make informational interventions more challenging. When the assignment system is strategy-proof, communication efforts (and informational interventions) can focus on expanding searches and recommending that candidates apply to all the schools where they would be willing to work, in the true order of their preferences.

Figures

Figure 2.1: RDD results on application changes during the validation period



(a) Probability of changing application



(b) Probability of changing application with CI

Note: Figure (a) plots the probability of an applicant changing their application using linear polynomials. Figure (b) plots the same but within the optimal bandwidth and with confidence intervals. Total observations: 3,653. The size of the bins is 0.01. The bin that contains 0 consists of 727 observations. The bin that contains 1 consists of 1,078 observations. The remaining bins consist, on average, of 18.7 observations.

Tables

Table 2.1: Summary statistics

	(1) All eligible applicants	(2) Pre-validation applicants	(3) Compliers	(4) Treated compliers	(5) Control compliers
Total	22,015	19,190	3,653	2,392	1,261
Share female (%)	72	73	74	77	0.69
Share non-mestizo (%)	9	9	10	9	11
Share married (%)	55	55	57	58	55
Share with master degree (%)	7	7	9	7	12
Share with more than 5 years of experience (%)	43	44	52	48	60
Share in the most common specialty (%)	22	22	29	29	29
Share in the most common province (%)	14	15	16	14	19
Mean age	38.57	38.69	38.70	38.69	38.70
Mean score	64.62	64.69	67.15	64.64	71.91

Note: Eligible applicants are the teachers who passed the "merits and public examination" phase. Pre-validation applicants are the teachers who had a personalized report available, that is the ones who applied before the validation period and applied to specialties with at least 80 registered and 20 partial applicants. Compliers are the ones who opened the personalized report. The most common specialty is basic general education from second to seventh grade. The most common province is Guayas.

Table 2.2: Summary statistics of outcomes within the analysis group

	(1) Treated compliers	(2) Control compliers	(3) Mean difference
Total	2,392	1,261	
Any modification (%)	68.02	35.37	32.65
			(0.02)
Add any (%)	65.34	28.71	36.64
			(0.02)
Add any from recommendations (%)	35.58	6.03	29.55
			(0.01)
Partially assigned (%)	13.92	99.68	-85.76
			(0.01)
Finally assigned (%)	20.32	90.96	-70.64
			(0.01)
Assigned in recommendation (%)	12.46	0.79	11.67
			(0.01)

Note: Standard errors in parentheses. Compliers are the ones who opened the personalized report. Treated refers to teachers that received the warning and the list of recommended schools. Control refers to teachers that received only the summary of the applications. Column (3) shows a mean difference test.

Table 2.3: RDD Results

	(1)	(2)	(3)	(4)	(5)
Panel A. Any modification					
RDD estimate	0.519	0.521	0.751	0.648	0.860
	(0.130)	(0.124)	(0.168)	(0.136)	(0.180)
Left BW	0.069	0.065	0.046	0.043	0.069
Right BW	0.069	0.236	0.046	0.157	0.069
Total observations in BW	170	357	111	229	170
Panel B. Add any					
RDD estimate	0.591	0.515	0.796	0.602	0.780
	(0.114)	(0.106)	(0.140)	(0.113)	(0.159)
Left BW	0.090	0.084	0.060	0.056	0.090
Right BW	0.090	0.285	0.060	0.189	0.090
Total observations in BW	216	440	137	283	216
Panel C. Add any from recommendations					
RDD estimate	0.427	0.562	0.697	-0.017	0.754
	(0.248)	(0.315)	(0.261)	(0.459)	(0.581)
Left BW	0.064	0.051	0.044	0.035	0.064
Right BW	0.064	0.234	0.044	0.160	0.064
Total observations in BW	75	216	36	149	75
Panel D. Assigned					
RDD estimate	0.371	0.352	0.365	0.470	0.324
	(0.124)	(0.119)	(0.157)	(0.153)	(0.188)
Left BW	0.101	0.089	0.067	0.059	0.101
Right BW	0.101	0.263	0.067	0.174	0.101
Total observations in BW	248	416	170	264	248
Panel E. Assigned in recommendation					
RDD estimate	0.347	0.343	0.356	0.331	0.310
	(0.084)	(0.109)	(0.119)	(0.146)	(0.127)
Left BW	0.148	0.091	0.102	0.063	0.148
Right BW	0.148	0.144	0.102	0.099	0.148
Total observations in BW	241	173	162	125	241

Note: Robust standard errors in parentheses. This table reports parametric estimates using different strategies to calculate the optimal bandwidth (BW) and different types of polynomials. (1) is estimated using a linear polynomial and the BW is calculated using the "one common MSE-optimal". (2) is estimated using a linear polynomial and the BW is calculated using the "two different MSE-optimal" that calculates two different BW below and above the cutoff. (3) is estimated using a linear polynomial and the BW is calculated using the "one common CER-optimal" bandwidth selector. (4) is estimated using a linear polynomial and the BW is calculated using the "two different CER-optimal" that calculates two different BW below and above the cutoff. (5) is estimated using a quadratic polynomial and the BW is calculated using the "one common MSE-optimal" method. All estimates control for specialty, sex, marital status, and region. Panel C is conditional on having added something to the application. Panel E is conditional on having been assigned.

Note: Robust standard errors in parentheses. This table reports parametric estimates using different strategies to calculate the optimal bandwidth and different types of polynomials. (1) is estimated using a linear polynomial and the BW is calculated using the "one common MSE-optimal". (2) is estimated using a linear polynomial and the BW is calculated using the "two different MSE-optimal" that calculates two different BW below and above the cutoff. (3) is estimated using a linear polynomial and the BW is calculated using the "one common CER-optimal" bandwidth selector. (4) is estimated using a linear polynomial and the BW is calculated using the "two different CER-optimal" that calculates two different BW below and above the cutoff. (5) is estimated using a quadratic polynomial and the BW is calculated using the "one common MSE-optimal" method. All estimates control for specialty, sex, marital status, and region.

Table 2.4: Heterogeneous effects

Panel A. Any modification					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Male					
Treated	0.496	0.456	1.014	0.650	0.824
	(0.144)	(0.137)	(0.191)	(0.148)	(0.193)
Male	-0.149	-0.149	-0.037	-0.142	-0.116
	(0.142)	(0.120)	(0.155)	(0.148)	(0.146)
Treated × Male	0.032	0.046	-0.151	-0.054	0.019
	(0.175)	(0.137)	(0.180)	(0.167)	(0.173)
Left BW	0.068	0.065	0.045	0.043	0.068
Right BW	0.068	0.241	0.045	0.160	0.068
Total observations in BW	170	353	94	232	170
Married					
Treated	0.534	0.453	0.898	0.543	0.828
	(0.181)	(0.159)	(0.241)	(0.180)	(0.217)
Married	0.032	-0.038	0.056	-0.137	0.030
	(0.141)	(0.129)	(0.158)	(0.159)	(0.140)
Treated × Married	-0.043	-0.013	0.094	0.139	0.006
	(0.175)	(0.137)	(0.197)	(0.171)	(0.176)
Left BW	0.067	0.065	0.044	0.043	0.067
Right BW	0.067	0.254	0.044	0.169	0.067
Total observations in BW	170	371	94	239	170
High skilled					
Treated	0.566	0.459	0.923	0.577	0.903
	(0.170)	(0.156)	(0.212)	(0.170)	(0.224)
High skilled	0.038	0.018	0.157	-0.056	0.054
	(0.151)	(0.128)	(0.186)	(0.162)	(0.158)
Treated × High skilled	-0.095	0.002	-0.245	0.088	-0.101
	(0.163)	(0.130)	(0.193)	(0.157)	(0.166)
Left BW	0.068	0.064	0.045	0.043	0.068

Table 2.4 (continued)

Right BW	0.068	0.247	0.045	0.164	0.068
Total observations in BW	170	358	111	232	170
More than 5 years of experience					
Treated	0.371	0.441	0.567	0.645	0.707
	(0.167)	(0.145)	(0.207)	(0.154)	(0.202)
Experienced	0.014	-0.143	0.087	-0.062	-0.005
	(0.133)	(0.126)	(0.149)	(0.153)	(0.132)
Treated × Experienced	0.214	0.080	0.236	0.009	0.236
	(0.179)	(0.134)	(0.201)	(0.165)	(0.178)
Left BW	0.068	0.064	0.045	0.043	0.068
Right BW	0.068	0.247	0.045	0.164	0.068
Total observations in BW	170	358	111	232	170

Panel B. Assignment	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Male					
Treated	0.412	0.418	0.413	0.508	0.383
	(0.130)	(0.116)	(0.163)	(0.149)	(0.191)
Male	0.016	-0.007	0.094	0.095	0.017
	(0.120)	(0.119)	(0.158)	(0.167)	(0.123)
Treated × Male	-0.114	-0.037	-0.107	-0.139	-0.115
	(0.147)	(0.130)	(0.187)	(0.188)	(0.148)
Left BW	0.102	0.090	0.068	0.060	0.102
Right BW	0.102	0.263	0.068	0.175	0.102
Total observations in BW	248	422	170	264	248
Married					
Treated	0.421	0.387	0.408	0.473	0.391
	(0.154)	(0.146)	(0.211)	(0.182)	(0.212)
Married	0.066	0.072	0.055	0.054	0.066
	(0.099)	(0.106)	(0.141)	(0.158)	(0.100)
Treated × Married	-0.054	-0.039	-0.033	0.004	-0.056
	(0.134)	(0.119)	(0.184)	(0.175)	(0.134)
Left BW	0.101	0.089	0.067	0.059	0.101
Right BW	0.101	0.263	0.067	0.174	0.101
Total observations in BW	248	416	170	264	248
High skilled					
Treated	0.196	0.254	0.254	0.390	0.179
	(0.157)	(0.138)	(0.204)	(0.184)	(0.211)
High skilled	-0.065	-0.022	-0.111	0.026	-0.069
	(0.115)	(0.113)	(0.149)	(0.147)	(0.116)

Table 2.4 (continued)

Treated × High skilled	0.281	0.179	0.203	0.153	0.283
	(0.133)	(0.117)	(0.172)	(0.151)	(0.133)
Left BW	0.101	0.089	0.067	0.059	0.101
Right BW	0.101	0.263	0.067	0.174	0.101
Total observations in BW	248	416	170	264	248
More than 5 years of experience					
Treated	0.563	0.506	0.398	0.588	0.479
	(0.157)	(0.129)	(0.181)	(0.162)	(0.201)
Experienced	0.182	0.248	0.070	0.151	0.194
	(0.098)	(0.101)	(0.131)	(0.140)	(0.104)
Treated × Experienced	-0.281	-0.294	-0.044	-0.262	-0.295
	(0.130)	(0.115)	(0.165)	(0.156)	(0.135)
Left BW	0.101	0.089	0.067	0.059	0.101
Right BW	0.101	0.263	0.067	0.174	0.101
Total observations in BW	248	416	170	264	248

Chapter 3

The Welfare Effects of including Household Preferences in School Assignment Systems: Evidence from Ecuador

3.1 Introduction

In this chapter, we study the welfare effects of a significant policy change to the coordinated school assignment system in Manta, Ecuador. The assignment system in place before the policy change aimed to minimize the (linear) travel distance between homes and schools and its implementation proved challenging due to the considerable effort required to geo-reference all students and ensure that the assignment process results were consistent with actual transportation options as well as the existence of hills, rivers, etc. Given the costs and difficulties of reviewing linear distance-based assignments to correct for geographic features, the Ecuadorian Ministry of Education piloted an alternative assignment system in partnership with the Inter-American Development Bank (IADB) and the NGO ConsiliumBots in which applicants' preferences were the main driver. This was motivated by findings in the school choice literature on the benefits of coordinated assignment systems that take family preferences into account (Abdulkadiroğlu et al. (2017)). The new system followed standard best practices, including the use of the deferred acceptance algorithm (Gale and Shapley, 1962), unlimited ranked ordered lists, and information provision systems (Abdulkadiroğlu and Sönmez, 2003; Pathak, 2011, 2017; Arteaga et al., 2021). If an applicant could not be assigned to one of their reported preferences due to excess demand, the system maintained the linear distance-based assignment criteria, but only considered seats left by applicants who were assigned to a submitted preference.

Families in Ecuador were not accustomed to a system that allowed them to express their preferences over schooling alternatives. Added to the limited information available about schools likely resulted in the new system's welfare potential not being fully realized. An indication of it is the significant number of applicants including only one alternative in their portfolios. While optimizing assignment rules and implementing complementary policies, as for example discussed in chapters 1 and 2, and increased familiarity of applicants with the new system in subsequent implementations may lead to improved outcomes, the contribution of this chapter is to show that even a sub-optimal but reasonably well-designed initial CCAS system can produce superior results compared to alternative systems. This highlights the value of CCAS adoption in other contexts, particularly in developing countries like Ecuador.

To compare the assignment alternatives, we take advantage of the fact that the system implemented in Manta elicited the true preferences and locations of all participating applicants. This allows us to compare the assignments made by the new centralized choice and assignment system (CCAS) with the simulated assignments of the prior alternative, whose rules we can replicate, as well as with assignments resulting from alternative algorithms under a CCAS option. We use a counterfactual strategy to simulate assignments, replicating the rules of the previous process.

Our main finding is that implementing a coordinated mechanism that incorporates the preferences of applicants has large welfare benefits. When compared to the previous system, using the deferred acceptance (DA) algorithm increases the portion of applicants assigned to any of their chosen schools from 49.96% to 78.44%, while the percentage of applicants assigned to their first choice increases from 42.42% to 69.76%. The main trade-off of implementing the DA alternative is that average linear distance to the school increases by 0.29km. To measure the importance of preferences and distance-to-school, we estimate applicant utilities over the different alternatives and find that the use of the DA algorithm results in significant utility increases.

We measure utilities in kms of linear distance between the applicant's residence and the assigned school and find that comparing the DA algorithm to the prior mechanism, average utility for Pre-School 1 applicants is 0.655 km higher, while that for Pre-School 2 applicants is 0.315 km greater, and for Primary 1, it is 0.021 km smaller. In the first year of primary school, the difference in utility is much smaller, and in favor of the previous system, due to greater congestion, as many seats are already taken by applicants enrolled during pre-school. As a result, many primary school applicants are not assigned to any of their preferred schools under either alternative and are instead assigned to another school using the distance-based criteria.

When we restrict the sample to applicants who receive a different assignment under the distance and DA mechanisms (i.e., applicants who improve or worsen their utility when the mechanism is changed), the differences for each level increase to 1.411km for Pre-School 1, 0.487km for Pre-School 2, and -0.027km for Primary 1.

We focus our analysis on estimated welfare and on the share of applicants being assigned to more or less preferred alternatives based on their reported preferences, without consideration to school quality for two reasons: i) we do not aim to study whether families in Ecuador prefer higher-quality schools, but rather to assess the welfare consequences of the assignment system as it relates to applicants' valuation of different schools, and ii) we cannot (at least directly) observe school quality.[1]

These results contribute to a better understanding of the advantages of coordinated school choice and assignment systems. While several studies demonstrate the welfare benefits of using one mechanism compared to others, few have directly estimated the welfare benefits of a coordinated assignment system that takes household preferences into account. The most closely related paper is Abdulkadiroğlu et al. (2017), which examines the welfare effects of switching the previously uncoordinated New York City assignment system to a coordinated alternative that incorporates family preferences. The authors find that most of the welfare gains are obtained from the coordination using the standard deferred acceptance algorithm, with only marginal gains when implementing alternatives. We find similar results in the context of Ecuador, though here we are comparing the DA algorithm to a coordinated one that centers on the distance of the household to the school.

This chapter is structured as follows. In Section 3.2, we provide an overview of the Ecuadorian school system, including the previous and DA assignment mechanisms. Section 3.3 describes the available data. In Section 3.4, we present our model for estimating preferences and the methodology used to compare the different assignment mechanism alternatives. Section 3.5 presents our main results comparing the different assignment alternatives, followed by a brief discussion of the aftermarket dynamics. Finally, in Section 3.6, we conclude.

3.2 Context and Algorithm Descriptions

We study school assignment in the coastal region of Manta, Ecuador. Specifically, we concentrate on the urban areas within and around the city of Manta,[2] including the geographic units ("*cantones*") of Manta, Montecristi, and Jaramijó.

Manta was selected through a process aimed at identifying a small yet representative city in the coastal region of Ecuador. The objective was to test and evaluate the school assignment policy in this context before scaling it up to other cities, and particularly to the largest city in the country, Guayaquil.[3] The selection process took into account students in the urban area, school coverage, distribution of school types (mainly public and private), as well as city size. Ultimately, Manta was chosen for its relative similarity to the alternatives. Table C.2.1 of Appendix C.2 compares the main characteristics of Manta and Guayaquil,

[1] The Ecuadorian government does not currently apply census-based student learning assessments in primary grades. We also do not study the impact of the system on other measures of interest, such as educational segregation, as we lack socioeconomic data for participating applicants.

[2] The educational system in Ecuador is split into two educational regimes nationally, one for the coastal region and another for the country's interior. The academic year in the coastal regime (where Manta is located) begins in May and ends in January, while it begins in September and ends in June in the interior.

[3] There was a change in government in Ecuador in 2021, and the new administration recently decided to scale up the system in coastal districts beginning in 2023. Because of the COVID pandemic and the fact that the distance-centric alternative required several in-person interactions during the process, the Ministry is currently using a First-Come, First-Serve digital system.

another coastal city and the country's largest, using data from the 2010 Census and school transfer requests in the 2019-2020 school year.

The Ecuadorian educational system is organized into three levels: Pre-school (Educación Inicial), Primary School (Educación General Básica) and Secondary School (Bachillerato). In this chapter, we focus on school assignments at the "entry level", a designation that encompasses enrollment in Pre-school 1, Pre-school 2, and Primary 1.[4][5]

There are three types of schools in Ecuador: public ("*fiscal*" and municipal), "*fiscomisional*," and private. Public institutions are funded by the government, "*fiscomisionales*" receive mixed funding from the state and families, and private schools are fully funded by families. Nationwide, free public schools account for 73.8% of enrollment, split between the "*fiscales*" and municipal schools (the latter of which represent only 0.8% of total enrollment). *Fiscomisional* schools receive 6% of all enrollments, while private schools take the remaining 20%. In Manta, free public schools account for 66% of enrollment at the entry-level grades, while *fiscomisional* and private alternatives account for 4% and 30% respectively.

This pilot only included free public schools, meaning that private schools represent an outside option for families that is not explicitly incorporated into our model. This is important for our welfare comparisons, which might be overestimated considering that families can switch from a less desirable free public school to a private alternative. As such, our welfare comparisons should be interpreted as the difference between the utility offered to families by the free public system. It is nonetheless also relevant to note that, at least for the applicants that participated in the pilot, there seems to be only limited overlap between free public schools and private alternatives. This is illustrated by a survey conducted after the application period but before results were distributed.[6] Only 1% of respondents stated that their reason for not including more alternatives was because of the alternative of enrolling their child in a school outside the public system. This is a critical observation because, as we explain in Section 3.3, many applicants submitted only one or two preferences.[7] While the short lists may have been due to applicants' preferences for outside options, the survey results suggest that this was not a frequent consideration for participating families.

[4] Pre-school is divided into two grades, called "*Inicial*" one and two. Primary school is divided into four levels. The first, which we call Primary 1, is for five-year-old children, while the other three levels cover children from six to eight years old, nine to eleven years old, and twelve to fourteen years old.

[5] The education system also offers different tracks: the regular track accounts for 98% of enrollment, with the remaining 2% distributed between schooling for students with special educational needs, artistic education, and adult schooling.

[6] The objectives of this survey were to gather information about parents' overall satisfaction level with the system, information sources they used to apply for schools, awareness of the school supply, among other aspects. The survey was completed by 1,484 parents.

[7] All applicants were ultimately assigned, though some to a school outside their reported list of preferences. In such cases, the assigned school was the closest possible alternative, as explained in Subsection 3.2.

Distance-Centric Algorithm

Prior to the COVID-19 pandemic, the school assignment system in Ecuador was primarily based on the applicant's location, which was reported through their electricity account code (CUEN). The assignment system was part of a broader enrollment process comprising six phases, as described in Appendix C.4. The Assignment Phase was the third of these phases, and also operated in stages. In the first stage of the Assignment Phase, different types of enrollment in the system were identified, depending on whether students preferred to attend a regular program or a rural, bilingual, or special education program. Also, where possible, applicants with siblings already in the system were assigned to the same school. Students registering for non-regular programs (e.g., bilingual or special education) and those with already-schooled siblings were processed before the other applicants.

Once these groups were assigned via a process that was carried out directly at the district headquarters, the rest of the students were assigned using what we will label as the distance-centric (DC) algorithm, which the Ministry called the "mathematical model". Legal guardians could complete an individual registration (of a single applicant) or one for a "group of siblings". While this latter option suggests that the system prioritized assigning groups of siblings to the same school over distance-based considerations, this was not confirmed by the Ministry experts with whom we interacted, and it was not formally included in the DC algorithm. Therefore, we did not take it into account in the DC assignment.

However, the process gave priority to applicants with siblings already enrolled in a public school, processing them first to be assigned to the same school as their sibling. Otherwise, these applicants were assigned using the distance criterion. In summary, the processing of regular assignments under the DC algorithm was as follows:

- Each applicant was assigned a random number at the beginning of the assignment process. The random numbers were used to determine the order in which applicants were processed.

- Applicants with siblings already enrolled in a public school were given priority and processed first, using their assigned random numbers to determine the order. The system attempted to assign them to the same school as their sibling. If this was not possible, these applicants were considered along with the rest of the students.

- All remaining applicants were processed next, with the order determined by their random numbers. These applicants were assigned to the closest school with vacancies based on the linear distance criterion.

When applicants' stated preferences are used to determine assignment, the mechanism of sequential processing is commonly known as a "random serial dictatorship" in the school assignment literature (Abdulkadiroğlu and Sönmez, 1998). In this case, the assignment process can be regarded as a random serial dictatorship with inputted preferences that are based on the linear distance criterion.

As explained above, an applicant's home address was based on the legal guardians' electricity bill. Using the latter to identify family location has proven highly effective, but may also incentivize families to procure (and even buy) electricity bills closer to their schools of interest. Moreover, there are still areas where households do not have electricity meters. These facts were reported in a series of interviews carried out by the IADB in Quito and Guayaquil, where families and officials recounted different factors affecting the registration processes.[8] Given that we do not have precise estimates of location misreporting rates, we conduct a sensitivity analysis in Section 3.4 and simulate assignments under different levels of misreporting.

The Ecuadorian government's concern with minimizing the distance to school arises from public policy considerations, and not because this aspect affects other dimensions such as, for example, public expenditure on free busing to schools. The latter consideration is nevertheless relevant in other contexts (e.g., many US cities), meaning that analyses comparing assignment mechanisms in similar cases should consider inclusion of these budget factors.

Deferred Acceptance (DA) Mechanism

The pilot used the deferred acceptance mechanism (Gale and Shapley, 1962), following the best practices in school choice mechanism design (Pathak, 2011; Correa et al., 2019). The specification of the assignment algorithm included static and dynamic sibling priorities, family linking, and a multiple tie-breaking rule.[9]

The static and dynamic sibling priorities indicate that an applicant will be prioritized for assignment to a school/program if their sibling is already assigned to the school (static). If the applicant is applying at the same time with another sibling, and one of them is assigned to a school,[10] the applicant that has not been assigned yet will receive priority for

[8] For example, district officials commented: (1) *"In District 24, Durán, Guayaquil, families lend their electricity bills to each other so they can all have access to the education system. We estimate that more than 60% of families in this district do not use their own electricity bill, so they do not register their real geolocation."* (2) *"In District 8, Monte Sinaí, Guayaquil, families maintain that there are "illegal invasions" of other families in areas where popular schools are located, using electricity bills from that area to get a seat in these schools."*

[9] The deferred acceptance algorithm was selected because it is both non-strategic and stable, and because it allows policy makers to implement desirable features such as dynamic sibling priority, family linked applications, and different priority-quota combinations. The only relevant drawback of the algorithm is that it is not Pareto efficient (i.e., it might be possible to assign an applicant to a higher priority without negatively affecting another one). The Stable Improvement Cycle (SIC) algorithm and the Top Trading Cycles (TTC) algorithms are more efficient, but at the cost of losing the strategy-proofness property in the case of the SIC mechanism, and the stability property in the case of the TTC mechanism.

Moreover, as shown in Section 3.4 of this chapter, and in Abdulkadiroğlu et al. (2017) for the case of New York, the efficiency gains obtained from using these alternatives are marginal when compared to the gains due to a transition from an uncoordinated system or a centralized one that cannot be classified as a CCAS, which is the case in this analysis.

[10] This can happen if one sibling is older than the other and will depend on the order in which the algorithm is run. If it is descending, the older sibling will give dynamic priority to the younger sibling. If it is ascending,

being assigned to that same school (dynamic). The dynamic sibling priority is lower than the static sibling priority because the latter is already defined (the sibling is attending the school), while the former will depend on the answer from the applicant after the assignment.

The family linking feature consists of trying to assign all siblings applying together to the same schools. Following a descending order, where older applicants are assigned first, if an older sibling is assigned to school A, the applications of the younger siblings will be modified to put school A as the first-ranked school to improve the probability of being assigned together. Finally, a multiple tie-breaking rule gives each applicant a different lottery number for each school to which they apply. Lottery numbers are used to break ties within priority groups (siblings and non-siblings) when a school receives more applications than spaces available.

3.3 Data

The data used in this chapter mainly comes from the centralized choice and assignment system (CCAS) pilot web page created in 2021 in the region of Manta, Ecuador.[11] The first data set comprises the supply of vacancies for all schools and programs offered in the pilot, where an educational program consists of a combination of grade and school. The pilot was implemented for all students entering Pre-School 1, Pre-School 2, and the first year of primary school (i.e., ages 3 to 5) for the first time. Vacancies are presented in Panel A of Table 3.1. Pre-School 1 has the most vacancies and is the least congested grade while Primary 1 is the most congested.[12]

The second data set consists of student and legal guardian information, including geo-location, applicants' sibling relationships, special educational needs, and nationality.[13] For each applicant, we have their ranked ordered list (ROL) of reported preferences, which had no length limit, and the lotteries drawn up for each program. To assign applicants to a program close to their residence, if they were not assigned to one of their reported preferences, the applicants' initial ranked order lists (containing their preferred programs) were appended with all other alternatives sorted by distance. The assigned lottery numbers were different for each of the programs listed by the student, but the same lottery number was drawn for all of the programs appended to the list.

it will be the other way around.

[11] All PII data was eliminated for that purpose.

[12] This is most likely explained by a combination of factors: i) students prefer schools closer to their homes (ceteris paribus) and establishments in more crowded areas have been filled by the previously implemented distance-based algorithm; ii) students that are not satisfied with their assigned school can ask for a transfer; iii) applicants strategically reported addresses close to the more preferred schools under the previous location-based assignment system.

[13] The preferences of applicants who have a sibling already enrolled at their school of interest are specified in Panel A of Table C.3.5 in Appendix C.2. We do not have information on cases in which an applicant's siblings are enrolled in schools not included in their reported preferences.

The distribution of applicants by geographic unit and grade is presented in Panel B of Table 3.1. Notably, at least in the case of the geographic unit of Manta (Cantón), the number of applicants in Pre-School 1 and Pre-School 2 is roughly equivalent. Although this poses a challenge from a public policy standpoint in that it is desirable to enroll students earlier, it is also an interesting dynamic for the application system since families' decision to postpone the enrollment of their child(ren) puts them at a strategic disadvantage. This is because there are fewer available seats in Pre-School 2, given that currently enrolled Pre-School 1 students move automatically to the next level.

Figure C.1.1 of Appendix C.1 provides an overview of applicant priorities and the lengths of ranked ordered lists. Note that most applicants declared only a single preference despite there being no limits placed on the length of the preference list. This may be a legacy of the previous system in which applicants did not choose a portfolio of schools and in which it was implied that applicants were largely assigned to a school based on distance (walking or driving, as obtained from Google Maps) rather than their preferences.

A complementary explanation for the large number of short application lists is that, as observed in Figure C.1.2 of Appendix C.1, applicants who replied to the survey distributed after the application period had ended indicated high expectations that they would be assigned to their top choice. These responses were obtained before the results were published to avoid bias. In the same survey, when asked why they did not add more programs to their portfolio, 56% of respondents replied that they had no information on alternatives close to their home, 33% said that they were sure that they would be assigned to their reported preference, 6% said that it was difficult to find more schools, 4% declared that they preferred receiving no assignment to adding more alternatives, and 1% declared that their preferred option was a non-public school (i.e., an outside option).

In any case, the fact that the CCAS was new to families in Manta likely also resulted in them not fully adapting their behavior to the new system and rules, meaning that they may not have taken full advantage of the introduction of parental choice and preference reporting. If this is the case, our findings on the welfare gains obtained with the introduction of the CCAS system are probably downward biased when compared with the longer-term results that will eventually be obtained once families are fully accustomed to the new system.

To analyze aftermarket dynamics, we also requested official data on school change requests and actual enrollment. This additional data is discussed in Section 3.5, where we show that applicants assigned to a school based on distance criteria rather than one of their preferred schools are over twice as likely to request a school change or be enrolled in a different school.

3.4 Welfare Estimation and Mechanism Comparison Methodology

Utility Estimation

To estimate welfare, we first need to estimate the parameters determining the utility that families would receive from an assignment to a particular school. Our estimation is based on applicants' reported preference orderings being an accurate representation of true family preferences on two dimensions. First, that the schools included in the ROL are preferred to those not included. Second, that alternatives ranked higher in the ROL are preferred to those ranked lower.

Regarding the first dimension, the assumption is that families have complete knowledge or full consideration of the schools in their choice set. To make this assumption more plausible, we limit the schools included in choice sets geographically. Specifically, we focus on a relatively small set of alternatives in the urban area of Manta that are less geographically spread than other similar studies, while still encompassing most of the schools and applicants involved in the pilot. This is depicted in Figure C.1.5 of Appendix C.1. We do not consider portfolio formation dynamics in our analysis since our main interest is to estimate welfare, rather than producing alternative counterfactual applications, as was the case in Chapter 1. Additionally, we lack sufficient data to address this issue. The second dimension is supported by the non-strategic nature of the DA mechanism, which was furthermore emphasized in the pilot program's communication strategy.

Our utility model is presented in Equation 3.1. As explained in Chapter 1, under the model's additively separable structure in u_{ij} and $\tau_i d_{ij}$, variation in residence-to-schools distances and relative rankings of schools allows for non-parametric identification of utilities, given the assumption that utility shocks and distances are independent of observable and unobservable school and applicant characteristics, and assuming that τ_i is the same for a sufficiently large set of individuals (Agarwal and Somaini, 2018).

$$
\begin{aligned}
v_{ij} &= \underbrace{S_{ij}\lambda + \delta_j + \epsilon_{ij}}_{u_{ij}} - \tau_i d_{ij} \\
\delta_j &= \bar{\delta} + \xi_j \\
\tau_i &= \tau + \gamma_i
\end{aligned}
\tag{3.1}
$$

In this equation, S_{ij} is a binary variable that takes the value of 1 if applicant i has a sibling in school j. d_{ij} denotes the distance between the residence of applicant i and school j. Finally, ξ_j represents a common school-specific preference or mean-utility, that is unobserved by the econometrician, but taken into account by families when making decisions about which schools to apply to and how to rank them.[14]

[14] We have a limited set of observable school characteristics and thus, to keep the model simple, we choose not to

To account for differences in how applicants trade-off distance and average school utility, we introduce a random coefficient over the utility parameter for distance-to-school in our model. This allows for applicants who only consider schools close to their residence to place more weight on distance than those who apply to schools further away. To identify utilities, we additionally assume that distances are independent of the random coefficient over its value, given school unobservables and sibling priorities. The identification assumption is thus given by:

$$(\gamma_i, \epsilon_{ij}) \perp d_{ij} | \xi_j, S_{ij}$$

In essence, this assumption implies that applicants do not systematically choose their residence based on the distance to a specific school they prefer or to be in proximity to different schooling alternatives. The previous system, in which families could borrow or buy an electricity bill near their preferred school instead of actually changing their residence, aligns with this conditional independence assumption. Nevertheless, to test the robustness of our results, In Appendix C.3 we estimate a model without the random coefficient over distances, which yields similar outcomes.

λ is identified by the variation in choices between applicants with and without sibling priority in schools. This, however, does not take into account that the reason for the sibling being enrolled in that school may be because the family liked the school when the sibling enrolled (or transferred) in the first place. This implies that ϵ_{ij} is likely to be positive for such cases, introducing bias in its parameter estimation. In Appendix C.3 we therefore present our preference parameter estimates and welfare calculations with no sibling-related considerations.[15] Findings and conclusions remain the same.

Finally, we also introduce a scale normalization for the utility by setting $\tau \equiv 1$, and a mean-utility location normalization by setting $\overline{\delta} \equiv 0$, following Abdulkadiroğlu et al. (2017). Random components are assumed to have the following distributions:

$$\begin{aligned} \gamma_i &\sim \mathcal{N}(0, \sigma_\gamma) \\ \xi_j &\sim \mathcal{N}(0, \sigma_\xi) \\ \epsilon_{i,j} &\sim \mathcal{N}(0, \sigma_\epsilon) \end{aligned}$$

To estimate parameters, we implement the Gibbs sampler technique, which, as explained in Chapter 1, is a Markov Chain Monte Carlo estimation procedure that approximates parameters estimated by maximum likelihood but is better suited for school choice contexts (see Rossi, McCulloch, and Allenby (1996) for a more detailed description and Kapor et al.

include a conditional mean component based on these attributes into δ_j. Specifically, we may have included school infrastructure dummies as covariates, as families had access to a list of the school infrastructure in the application interface.

[15] In Chapter 1, the objective is to create counterfactual applications, and therefore the estimation of λ with bias is not crucial. The focus is on identifying whether the alternative is more valuable to the applicant and should be included in counterfactual rankings. However, for measuring welfare, it is more important to have accurate measurements.

(2020); Idoux (2022) for other recent implementations of this method in a school choice context).[16] The model's full specification to implement this technique, including conjugate priors over to obtain the standard deviations of random coefficients and the prior values used, is presented in Appendix C.5. The Gibbs sampling process starts with draws from the prior distributions and starting utility values consistent with observed choices and then iterates over the following steps:

1. Utilities are drawn using the estimated parameters of the previous iteration and using reported preference rankings to restrict possible values. Note that here, as in Abdulkadiroğlu et al. (2017), we do not explicitly include an outside option, and instead assume that rankings have exogenously fixed size. Alternatives not ranked thus have a utility below that of the last school ranked instead of the outside option.

   Specifically, assuming that i's ranking is of size R (1 being the most preferred alternative and R the least preferred), utilities are drawn iteratively using a truncated normal distribution so that:

   $$u_{i,j(r=1)} > u_{i,j(r=2)} > \dots > u_{i,j(r=R)} > u_{i,j(r=\hat{r})}, \ \forall \hat{r} > R$$

   To do this iterative sampling, $u_{i,j(r=1)}$ is drawn from $(u_{i,j(r=2)}, \infty)$ if $R > 1$ (using $u_{i,j(r=2)}$ from the previous iteration), and from $(-\infty, \infty)$ when $R = 1$. Alternatives not included in i's ROL are sampled from $(-\infty, u_{i,j(r=R)})$.

2. Using the updated utilities, estimates for λ, ξ_j and γ_i are obtained sequentially.

3. The parameters of the distribution of random coefficients are updated, and a new iteration is started. More details on the Gibbs sampling process can be found in Appendix A.6 of Chapter 1.

Utility Parameters Estimation Results

We implemented two 150,000 iteration chains of the Gibbs sampler, discarding the initial 50,000 iterations as a burn-in period to eliminate the effect of starting values. We then used the following 100,000 iterations of both chains to compute the mean parameters and standard deviations. Appendix C.2 presents the estimates from Equation 3.1 and the potential scale reduction factors (Gelman et al., 1992) to assess mixing and convergence of the Gibbs sampling procedure (values close to one imply convergence). In addition, we provide trace plots of the estimated σ_ϵ in one chain of the Gibbs sampling in Figure C.1.6 of Appendix C.1.

The potential scale reduction factors are close to one, indicating convergence, and the trace plots demonstrate that the values of σ_ϵ are bounded. We also provide estimates for the

[16] In the case of Kapor et al. (2020), to compare the Immediate Acceptance and the DA algorithms, and in the case of Idoux (2022), to weight the effect of policy changes, family preferences, and residential sorting.

model without random coefficients (Table C.3.1 in Appendix C.3) and without siblings (Table C.3.2 in the same appendix section). Across all specifications, the estimated parameters and welfare estimates are similar.

To estimate welfare, we compute the expected utility of each individual i for each school j based on the estimated parameters and reported preference rankings, following the approach of Abdulkadiroğlu et al. (2017):

$$\mathbb{E}\left[u_{i,j}|ROL_i, \xi, \lambda, \sigma_\epsilon, \sigma_\xi, \Sigma_\gamma, d_i\right]$$

Specifically, that implies that we obtain welfare estimates by averaging the expected utilities over the iterations of the Gibbs sampler, thus directly conditioning on reported preference rankings.

Counterfactual Assignment Strategy

We compare the distance-centric (DC) algorithm described in Subsection 3.2 with the DA algorithm using the Stable Improvements Cycles (SIC) (Erdil and Ergin, 2008) and the Top Trading Cycles (TTC) algorithms (see Abdulkadiroğlu and Sönmez (2003)) as benchmarks. The TTC algorithm is our welfare benchmark since it delivers a student-optimal assignment, and thus (weakly) higher welfare than the SIC algorithm, given that the latter delivers a stable student-optimal assignment (and the stability restriction reduces attainable welfare). The TTC algorithm also results in a higher welfare than the DA algorithm, which delivers a stable but not necessarily student-optimal assignment. Given the potential multiplicity of the DA, SIC and TTC assignments, we follow the procedure described in Abdulkadiroğlu et al. (2017) to ensure a monotonic welfare comparison across all simulations. This implies that we first run the SIC algorithm over each DA assignment (as described in Erdil and Ergin (2008)), and then the TTC algorithm over the obtained DA-SIC result.[17] Following Abdulkadiroğlu et al. (2017), we run 100 lottery simulations of the DA and DC algorithms to get our welfare calculations.

With regard to the DC algorithm, one relevant point is that parents could strategically report a different address, using someone else's electricity bill (CUEN) in order to be placed at a preferred school. To include this possibility in the analysis, we run counterfactual assignments in which a random proportion of the applicants strategically choose an address close to their most preferred program. We use different random proportions as we do not have a good estimate of CUEN misreporting under the previous system.

To compare mechanisms, we first re-run the DA algorithm used in the pilot. We use the same inputs, except that we do not include students with special needs in order to make the assignment comparable to that of the DC algorithm.[18] In the implemented DA, the re-

[17] In this case, the the SIC and TTC algorithms obtain the same assignment. In other words, the are no attainable Pareto efficiency improvements from relaxing stability constraints in this particular context.

[18] We eliminate both students and vacancies related to special needs, which account for only 0.23% of applicants. This decision was made because students with special needs had a special assignment round before the regular

ported preference rankings were appended to all non-ranked programs using a linear distance sorting criterion. Applicants received a lower priority in the distance-imputed preferences to maximize assignment to the reported preferences.[19] We define assignments to imputed preference as non-preference assignments to distinguish them from the overall assignment obtained with the DC algorithm.

To replicate Ecuador's previous system (described in Subsection 3.2), we consider all available programs and rank them using linear distance sorting. Students with siblings in the system were assigned (if possible) to their sibling's school before the main process was initiated. To this end, we create a priority group for these students that only applies at the schools in which their siblings are enrolled. Finally, given that applicants were processed sequentially, we run a single tie-breaking lottery to break ties.

3.5 Results

In this section, we begin by describing the differences between systems in terms of assignment to preferences and linear distance to home. We then present our welfare comparison using the utility model introduced above, and finish with an analysis of aftermarket dynamics observed after the CCAS system's operation.

Assignment Comparison

Notably, the reported preferences of 55.5% of the applicants (2,206 applicants) match the ranking used to simulate the distance assignment mechanism, reflecting the importance of distance between home and school to families. Specifically, this means that the first preference of over half of the applicants was the closest school to their homes (or the closest one where a sibling is enrolled). Table 3.2 compares the results of a single simulation of the DA and DC algorithms. In the distance-centric alternative, a significantly lower percentage of applicants are assigned to their preferred school. However, the percentages are quite similar for applicants with a strong preference for distance, as can be observed in rows 1 and 2 of Panel B. This outcome highlights that using a coordinated alternative that considers preferences does not harm (at least on average) applicants worried mainly about distance to school. In terms of the average linear distance of the assignment, we see that the DA algorithm assigned students to schools an average of 0.29 km farther away than the DC algorithm.

Figure C.1.3 in Appendix C.1 shows the assignment to different declared preference rankings for both systems. As we can see, the DA algorithm assigns more students to their first preference than the DC algorithm (70% to 42%) and much fewer students to an alternative outside of their reported preference list (22% vs 50%). Tables C.2.3 to C.2.5 in Appendix

one.

[19] When referring to applicant preferences, we intend reported preferences without the distance-imputed preferences.

C.2 and Figure C.1.4 in Appendix C.1 display these results for the different grades. Greater congestion leads to smaller differences between the two mechanisms in terms of applicants being assigned to their preferred options. However, there are two forces at play. On the one hand, more congestion implies that fewer applicants are assigned to a reported preference when using the DA alternative. On the other hand, under the distance-based alternative, more congestion increases the probability that one applicant who is placed in a closer school displaces another who would have ranked that school at the top of their list (particularly in the cases where the latter applicant's first preference and closest school coincide).

To evaluate the effect of location misreporting, Table 3.3 shows the results of computing counterfactual assignments in which a random sample of applicants report the location of their most preferred school as their address instead of their true residence. The exercise simulates cases in which families submit another household's electricity bill to maximize the likelihood of being assigned to their most preferred school. We compute assignments with misreporting levels of 10%, 30%, 50%, 70%, and 90%.

The results of this exercise show that, as the percentage of applicants who change their location increases, the percentage of applicants assigned to one of their preferences rises as well (from 50% to 59%). Nevertheless, the rates of assignment to a preferred option does not reach the level of the DA algorithm, since applicants who misreport their location can only signal a preference for a single alternative. If, however, they are not assigned to that alternative, they may end up in a school farther from home and less preferred to other options (some of which may be closer to their true location). When misreporting is greater, the true average distance to school (i.e., using the real and not the reported location) increases as well. At 90% misreporting, the distance to home reaches the same level as in the DA alternative. This implies that misreporting can close the gap to the DA mechanism only partially and at the cost of rapidly increasing the (true) distance between applicants' home and school.

Welfare Comparison

Table 3.4 presents the estimated differences in mean utilities (both in km), with respect to the benchmark student-optimal (TTC) assignment. Additionally, we include a standardization of utility differences in units of the utility standard deviation using a TTC assignment.[20] In Panel A, we can see that differences between the DA and TTC algorithms are small in terms of welfare (less than 70 meters) at the pre-school levels, and significantly smaller than the welfare loss under the DC (distance) alternative (0.655km and 0.315km on average, respectively). The difference is naturally larger when we consider only applicants assigned to different schools under the different algorithms, as shown in Panel B.

In Primary 1, the difference between the DA and DC algorithms is smaller due to increased congestion, which leads to more applicants being assigned to a non-preferred alternative, as shown in Figure C.1.4 of Appendix C.1. Furthermore, Table C.2.6 in Appendix C.2 shows that the share of applicants assigned to the same non-preferred school in both

[20] SIC and TTC actually have the exact same assignment in all 100 simulations, as explained in Appendix C.6.

algorithms increases significantly in Primary 1 (56.7% of applicants assigned to the same school, compared to 2.69% and 21.35% in Pre-School 1 and 2, respectively). Conditional on having a different assignment in the DC and DA algorithms, the share of applicants who move from a non-preferred to a preferred assignment under the DA algorithm is 40.85% in Primary 1, compared to 77% in Pre-School 1 and 53% in Pre-School 2, indicating a smaller share of applicants with improved outcomes in later years. As the DC algorithm finds schools that are on average closer to home, which is a result of not prioritizing reported preferences, and given that utility is, all else equal greater for applicants with a lower home-to-school distance, there is actually a reversal in the average utility difference between mechanisms in Primary 1.

The distribution of estimated welfare overall and in each grade is presented in Figure C.1.7 of Appendix C.1. Here, we observe that the phenomenon described in the above paragraph occurs in all grades, with two peaks in utility in each figure: one among applicants assigned to a preferred option and another for those assigned to a non-preferred option that is close to home. The DA and TTC (and SIC) algorithms have very similar distributions. However, the TTC algorithm does improve the assignment relative to the DA algorithm in Primary 1, which is explained by the fact that, with higher congestion, stability constraints imposed by tie-breaking lotteries are more restrictive. By eliminating them, TTC (and SIC) achieve a significant improvement (0.302 km overall over the DA assignment and 0.396 km if restricted to applicants with different assignments), as shown in Table 3.4.

Finally, Figure 3.1 compares welfare gains under different mechanisms, similar to Figure 5 in Abdulkadiroğlu et al. (2017) (included in Appendix C.1). The main takeaway is that, while the differences in magnitudes between results in this chapter and in Abdulkadiroğlu et al. (2017) are large, the proportions are actually very similar. Thus, the coordinated mechanisms that include applicant preferences and the alternatives considered in each case are roughly the same. Meanwhile, the differences in magnitudes are explained by the different settings of New York and Manta. In addition, Abdulkadiroğlu et al. (2017) study assignment to secondary school (where applicants are willing to travel more), while we assess assignment to pre-school and early primary school (where families place a greater value on distance from home).

With regard to improvements over the DA algorithm, the potential is context-dependent, as shown by the differences observed in the various grades. The margin, therefore, is not irrelevant, but likely to be more important in more congested grades (i.e., post-entry-level grades). It would arguably be best to focus on implementing a CCAS first, then use the reported preferences to study the potential of the SIC and TTC algorithms (and possibly others), before weighing the trade-offs between improving Pareto efficiency and losing stability or the strategy-proof properties.

Aftermarket Dynamics

In closing this section, we consider what happens after the assignment process using data on family requests to transfer to a different school, and data on the actual enrollment of students in different schools, which we can compare to the assignment.[21] This analysis is beyond the chapter's scope of comparing the assignment mechanisms but offers valuable insights into their possible effects on aftermarket dynamics. Moreover, transfer requests impose costs on the Ministry of Education that Ecuadorian officials explained were a concern. We will show that while the number of overall transfers and changes in assignment is significant, it is much more likely for applicants assigned by the distance to school criterion, highlighting that the coordinated approach is likely to reduce aftermarket movements. The main challenge is providing information and incentives for families to learn about more schools and report more preferences to have a larger share assigned to one alternative that is known and acceptable.

Table C.2.7 in Appendix C.2 presents descriptive data of aftermarket dynamics. The table shows that around 15% of applicants enroll in a different school than their assigned school, and only about two-thirds of these differences are explained by families formally requesting transfers (most transfer requests are accepted, as shown in the table), while the rest are a result of different administrative processes. A large fraction of both differences in enrollment and transfer requests come from applicants to Primary 1, already indicating that these dynamics are driven by non-assignment to a preferred school. Specifically, around 40% of applicants end up enrolling in a school different from their assigned school in Primary 1, compared to around 5% in Pre-School 1 and around 8% in Pre-School 2.

To assess the relationship between assignment to a reported preference and aftermarket movements, we leverage the use of lotteries to determine assignment at congested schools and compare applicants that are and are not assigned in such cases. We consider two groups of applicants: those applying to a congested first preference who are either assigned to it or to a different option due to their lottery number value, and those who are processed for their least preferred school (i.e. not assigned to one more preferred) and are either assigned to it or to a school using the distance criterion due to their lottery number value. Note that, as many applicants rank only one school, there is significant overlap between the two groups. For each group of applicants, we implement the following linear probability model regression:

$$Y_i = \beta_0 + \beta_1 NA_{i,s} + \Delta\beta_{0|s(i),g} + \epsilon_i \tag{3.2}$$

Y_i represents the aftermarket movement considered, which could be a request for transfer, a different assignment, or either of them. The coefficient β_0 represents the average probability of an applicant who is assigned to a congested school (either as a first or last preference) having the aftermarket movement considered. The coefficient β_1 represents the additional probability of an applicant who is not assigned to a congested school having the aftermarket movement considered. Additionally, $\Delta\beta_{0|s(i),g}$ is a fixed effect for congested schools and grades, which controls for any differences across different school-grade combinations.

[21] Note that all applicants participating in the CCAS pilot are included and not only those in the urban area of Manta included in the welfare comparison.

The results from regressions 3.2 are presented in Table C.2.8 of Appendix C.2. The table shows that the probability of requesting a transfer or having a different assignment is much higher for applicants not assigned to their congested first or last preference. For instance, for applicants processed at a congested last preference, the probability of either requesting a transfer or having a different enrollment than the assigned school is 9.61% if assigned to the least preferred option, compared to 22.38% if assigned by the distance criterion.

Overall, a significant portion of applicants are interested in changing their assignment even if it is their reported preference. This could be due, in part, to the impact of Covid-19 on family preferences during the period. Nonetheless, the share of applicants interested in changing their assignment is much greater for those assigned by the distance criterion.

3.6 Discussion

In this chapter, we document and study the effects of the Centralized Choice and Assignment System (CCAS) pilot implemented in early 2021 in the Ecuadorian city of Manta, where the previously existing system assigned students exclusively by the linear distance between their home and school. We contrast these systems using a sudden change in policy and the data generated by the new non-strategic deferred acceptance (DA) algorithm (Gale and Shapley, 1962). We estimate preferences closely following Abdulkadiroğlu et al. (2017) and use these to quantitatively study the welfare effects of the policy change.

Our main result is that implementing a coordinated mechanism that incorporates applicants' preferences has relatively large welfare benefits. This finding echoes that observed for New York City in Abdulkadiroğlu et al. (2017), though here, our setting is a developing country. The extent of the differences depends on different factors such as the congestion of schools, the grades considered, the characteristics of the city (e.g., residential segregation), and family preferences.

Specifically, we document that when compared to the previous distance mechanism used by the government, the DA algorithm increases the percentage of applicants assigned to a preferred school from 49.96% to 78.44%, while the percentage of applicants assigned to their first preference rises from 42.42% to 69.76%. The main trade-off of implementing the DA alternative is that the average linear distance between applicants' home and school increases by 0.29km. To assess the overall effect of the policy, we turn to our estimated welfare comparisons and show that the welfare gains are between 0.655km and 0.315km higher when the DA algorithm is used in Pre-School grades. Meanwhile, if the analysis is restricted only to applicants who are assigned to a different school under each alternative mechanism, these gaps more or less double in magnitude.

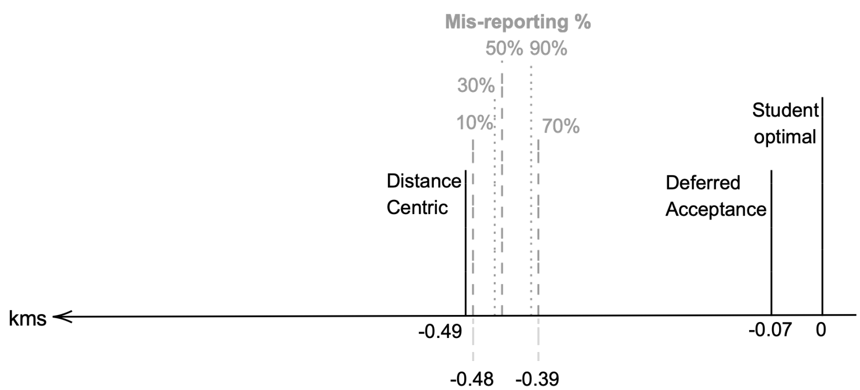
In the first year of primary school, the difference in welfare gains between the preference-based and distance-based mechanisms is much smaller and is actually reversed to a value of -0.021km. This is due to the high level of congestion in that grade arising from many seats being already taken by applicants enrolled in prior grades, which makes it difficult for many applicants to be assigned to their preferred schools under either mechanism. In this situation, the distance-based alternative seems to provide a very slightly better distribution of distances to schools and assignment to preferred options. This finding supports the intuitive expectation that incorporating preferences is particularly beneficial in less-congested entry grades of the schooling system.

The results showing that coordinated school choice systems are beneficial for families in developing country contexts is important in that more and more countries are adopting similar systems worldwide.[22] These findings are also timely since the COVID-19 pandemic has accelerated the adoption of online application and enrollment systems, facilitated by the fact that an increasing number of countries now have the pre-conditions to realize them. While various aspects of these policies beg further study, the growing body of evidence that preferences and coordination are a central driver of welfare gains is one important step in better understanding how to implement this type of market design in practice. The next step is to optimize coordinated school choice systems by studying and altering their rules, as well as implementing complementary policies that enhance their operation and yield better outcomes. As discussed in the preceding two chapters of this dissertation, these measures can significantly augment the benefits of such systems for families and help achieve important social objectives, such as promoting school integration and enhancing teacher selection.

[22]See Neilson (2021) and www.ccas-project.org.

Figures

Figure 3.1: Welfare Differences Between Algorithms



(a) All Applicants

(b) Only Applicants With Different Assignments

Tables

Table 3.1: Vacancies and Applicants by Geographic Unit (*Cantón*) and Grade

Panel A: Vacancies			
Cantón	Pre-School 1	Pre-School 2	Primary 1
Manta	1,830	1,394	425
Montecristi	905	668	654
Jaramijó	110	47	37
Total Grade	**2,845**	**2,109**	**1,116**
Total Global	**6,070**		
Panel B: Applicants			
Cantón	Pre-School 1	Pre-School 2	Primary 1
Manta	1,101	1,143	338
Montecristi	481	437	124
Jaramijó	125	125	107
Other	2	0	1
Total Grade	**1,709**	**1,705**	**570**
Total Global	**3,984**		

Table 3.2: Mechanism Comparison - Results

Panel A: All applicants		
Assigned in:	DA	Distance
Any preference	3,118	1,986
	(78.44%)	(49.96%)
First preference	2,773	1,686
	(69.76%)	(42.42%)
Average assignment distance	1.30km	1.01km
Panel B: Applicants maintaining 1st preference (2,206)		
Assigned in:	DA	Distance
Any preference	1,779	1,537
	(81.16%)	(70.12%)
First preference	1,644	1,472
	(75.00%)	(67.15%)

Table 3.3: Mechanism Comparison with Location Misreporting

	Applicants assigned to any preference	*Average Distance*
Distance Mech without misreporting	1,986 (49.96%)	1.01km
Distance Mech + 10% misreporting	2,055 (51.70%)	1.04km
Distance Mech + 30% misreporting	2,142 (53.89%)	1.09km
Distance Mech + 50% misreporting	2,230 (56.10%)	1.18km
Distance Mech + 70% misreporting	2,332 (58.67%)	1.21km
Distance Mech + 90% misreporting	2,403 (60.45%)	1.29km
DA Algorithm	3,118 (78.44%)	1.30km

Table 3.4: Differences in Welfare: Student-Optimal vs. DC and DA Algorithms

Panel A: All simulated applicants

	Pre-School 1		Pre-School 2		Primary 1	
Measure	DC	DA	DC	DA	DC	DA
Δ Mean utility (km)	-0.657	-0.002	-0.385	-0.070	-0.281	-0.302
$({}^{\Delta Mean\ utility}/_{\sigma_{Ut.\ FB}})$	-0.714	-0.003	-0.156	-0.029	-0.084	-0.090

Panel B: Applicants with different assignments across algorithms

	Pre-School 1		Pre-School 2		Primary 1	
Measure	DC	DA	DC	DA	DC	DA
Δ Mean utility (km)	-1.416	-0.005	-0.596	-0.109	-0.369	-0.396
$({}^{\Delta Mean\ utility}/_{\sigma_{Ut.\ FB}})$	-1.517	-0.006	-0.264	-0.048	-0.114	-0.122

Note: Δ Mean utility (km) is measured computing $u_{i,j(\mu)} - u_{i,j(TTC)}$, where $j(\mu)$ represents the school to which individual i is assigned under mechanism μ. We then compute average utilities for each algorithm and simulation and finally compute the average for each algorithm across simulations. $\frac{\Delta Mean\ utility}{\sigma_{Ut.\ FB}}$ simply uses the utility variance under the TTC mechanism to scale this difference in each simulation. This is done to facilitate extrapolations to other contexts. This same table is presented in Appendix C.3 for the specification without siblings.

Bibliography

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

Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, June 2003. doi: 10.1257/000282803322157061. URL http://www.aeaweb.org/articles?id=10.1257/000282803322157061.

Atila Abdulkadiroğlu, Yeon-Koo Che, and Yosuke Yasuda. Resolving conflicting preferences in school choice: The "boston mechanism" reconsidered. *American Economic Review*, 101(1):399–410, February 2011. doi: 10.1257/aer.101.1.399. URL https://www.aeaweb.org/articles?id=10.1257/aer.101.1.399.

Atila Abdulkadiroğlu, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. Do parents value school effectiveness? *American Economic Review*, 110(5):1502–39, May 2020. doi: 10.1257/aer.20172040. URL https://www.aeaweb.org/articles?id=10.1257/aer.20172040.

Atila Abdulkadiroğlu and Tayfun Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2998580.

Nikhil Agarwal and Paulo Somaini. Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica*, 86(2):391–444, 2018. doi: 10.3982/ECTA13615. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA13615.

Nicolás Ajzenman, Eleonora Bertoni, Gregory Elacqua, Luana Marotta, and Carolina Méndez Vargas. Altruism or money?: Reducing teacher sorting using behavioral strategies in peru. Working Paper 1143, Inter-American Development Bank, August 2020. URL https://doi.org/10.18235/0002625.

Nicolás Ajzenman, Gregory Elacqua, Luana Marotta, and Anne Olsen. Order effects and employment decisions: Experimental evidence from a nationwide program. Working Paper 1266, Inter-American Development Bank, 2021.

Hunt Allcott and Todd Rogers. The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104 (10):3003–37, 2014.

Michael B. Allen. Eight questions on teacher recruitment and retention: What does the research say? *Education Commission of the States (NJ3)*, 2005.

Claudia Allende, Francisco Gallego, and Christopher A. Neilson. Approximating the Equilibrium Effects of Informed School Choice. Working Papers 628, Princeton University, Department of Economics, Industrial Relations Section., July 2019. URL https://ideas.repec.org/p/pri/indrel/628.html.

Tahir Andrabi, Jishnu Das, and Asim Ijaz Khwaja. Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review*, 107(6): 1535–1563, 2017.

Peter Arcidiacono and Michael Lovenheim. Affirmative action and the quality-fit trade-off. *Journal of Economic Literature*, 54(1):3–51, March 2016. doi: 10.1257/jel.54.1.3. URL https://www.aeaweb.org/articles?id=10.1257/jel.54.1.3.

Felipe Arteaga, Adam J Kapor, Christopher A Neilson, and Seth D Zimmerman. Smart matching platforms and heterogeneous beliefs in centralized school choice. Working Paper 28946, National Bureau of Economic Research, June 2021. URL http://www.nber.org/papers/w28946.

Felipe Arteaga, Gregory Elacqua, Thomas Krussig, Carolina Méndez, and Christopher Neilson. Can information on school attributes and placement probabilities direct search and choice? evidence from choice platforms in ecuador and peru. 2022.

Cristián Bellei, Juan Valenzuela, and Danae De los Rios. *Segregación Escolar en Chile*, pages 209–229. 01 2010.

Eleonora Bertoni, Gregory Elacqua, Diana Hincapié, Carolina Méndez, and Diana Paredes. Teachers' preferences for proximity and the implications for staffing schools: Evidence from peru. *Education Finance and Policy*, pages 1–56, 2019.

Eleonora Bertoni, Gregory Elacqua, Luana Marotta, Matías Martínez, Carolina Méndez, Verónica Montalva, Anne Sofie Olsen, Humberto Santos, and Sammara Soares. El Problema de la Escasez de Docentes en Latinoamérica y las Políticas para Enfrentarlo. Technical Note 1883, Inter-American Development Bank, 2020.

Stephen B. Billings, David J. Deming, and Jonah Rockoff. School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg *. *The Quarterly Journal of Economics*, 129(1):435–476, 09 2013. ISSN 0033-5533. doi: 10.1093/qje/qjt026. URL https://doi.org/10.1093/qje/qjt026.

Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff. Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review*, 95(2):166–171, 2005.

Carycruz Bueno and Tim R. Sass. The effects of differential pay on teacher recruitment and retention. Research Paper Series 18-07, Andrew Young School of Policy Studies, 2018.

Sebastian Calonico, Matias D. Cattaneo, and Max H. Farrell. Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192–210, 2020.

David Card and Jesse Rothstein. Racial segregation and the black–white test score gap. *Journal of Public Economics*, 91(11):2158–2184, 2007. ISSN 0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2007.03.006. URL https://www.sciencedirect.com/science/article/pii/S0047272707000503.

Alejandro Carrasco, Alejandra Mizala, Dante Contreras, Humberto Santos, Gregory Elacqua, Florencia Torche, Carolina Flores, and Juan Valenzuela. Hacia un sistema escolar más inclusivo: Cómo reducir la segregación escolar en chile. *Informe de Políticas Públicas - Espacio Público*, pages 1–31, 05 2014.

Alejandro Carrasco, Francisca Bogolasky, Carolina Flores, Ernesto San Martin, and Gabriel Gutierrez. 'selección escolar' y desigualdad educacional en chile: ¿qué tan coactiva es la regulación que la prohíbe? *Estudios de Política Educativa - MINEDUC*, I:107–159, 01 2015.

Katarína Cechlárová, Tamás Fleiner, David F. Manlove, Iain McBride, and Eva Potpinková. Modelling practical placement of trainee teachers to schools. *Central European Journal of Operations Research*, 23(3):547–562, 2015.

Hector Chade and Lones Smith. Simultaneous search. *Econometrica*, 74(5):1293–1307, 2006. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/3805926.

Raj Chetty, John N. Friedman, and Jonah E. Rockoff. Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104 (9):2593–2632, September 2014a. doi: 10.1257/aer.104.9.2593. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2593.

Raj Chetty, John N. Friedman, and Jonah E. Rockoff. Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9):2633–79, 2014b. doi: 10.1257/aer.104.9.2633.

Charles Clotfelter, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor. Would higher salaries keep teachers in high-poverty schools? evidence from a policy intervention in north carolina. *Journal of Public Economics*, 92(5-6):1352–1370, 2008.

Sarah Cohodes, Sean Corcoran, Jennifer Jennings, and Carolyn Sattin-Bajaj. When do informational interventions work? experimental evidence from new york city high school choice. Working Paper 29690, National Bureau of Economic Research, 2022.

Julien Combe, Olivier Tercieux, and Camille Terrier. The design of teacher assignment: Theory and evidence. *The Review of Economic Studies*, 89(6):3154–3222, 2022. doi: 10.1093/restud/rdac002.

Sean P Corcoran, Jennifer L Jennings, Sarah R Cohodes, and Carolyn Sattin-Bajaj. Leveling the playing field for high school choice: Results from a field experiment of informational interventions. Working Paper 24471, National Bureau of Economic Research, March 2018. URL http://www.nber.org/papers/w24471.

Jose Correa, Rafael Epstein, Juan Escobar, Ignacio Rios, Bastian Bahamondes, Carlos Bonet, Natalie Epstein, Nicolas Aramayo, Martin Castillo, Andres Cristi, and Boris Epstein. School choice in chile. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, EC '19, page 325–343, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450367929. doi: 10.1145/3328526.3329580. URL https://doi.org/10.1145/3328526.3329580.

Arias Drouet and Olsen Westh. "quiero ser maestro": Mejoras del concurso docente del ministerio de educación de ecuador. Technical Note 1880, Inter-American Development Bank, 2020.

Esther Duflo and Emmanuel Saez. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, 118(3):815–842, 2003.

Otis Dudley Duncan and Beverly Duncan. A methodological analysis of segregation indexes. *American Sociological Review*, 20(2):210–217, 1955. ISSN 00031224. URL http://www.jstor.org/stable/2088328.

Umut Dur and Onur Kesten. Sequential versus simultaneous assignment systems and two applications. *Economic Theory*, 68(2):251–283, 2019.

Federico Echeñique and M. Bumin Yenmez. How to control controlled school choice. *American Economic Review*, 105(8):2679–94, August 2015. doi: 10.1257/aer.20130929. URL https://www.aeaweb.org/articles?id=10.1257/aer.20130929.

Lars Ehlers, Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory*, 153(C):648–683, 2014. doi: 10.1016/j.jet.2014.03.004. URL https://ideas.repec.org/a/eee/jetheo/v153y2014icp648-683.html.

G. Elacqua. The impact of school choice and public policy on segregation: Evidence from chile. Technical report, Documento de Trabajo N10, Centro de Políticas Comparadas de Educación, Universidad Diego Portales, Santiago, Chile, 2009.

Greg Elacqua, leidy Gómez, Thomas Krussig, Luana Marotta, Carolina Méndez, and Christopher A. Neilson. The potential of smart matching platforms in teacher assignment: The case of ecuador. Technical report, Inter-American Development Bank, 2022a.

Greg Elacqua, Isabel Jacas, Thomas Krussig, Luana Marotta, Carolina Méndez, and Christopher A. Neilson. The welfare effects of including household preferences in school assignment systems: Evidence from ecuador. Technical report, Inter-American Development Bank, 2022b.

Gregory Elacqua, Anne Sofie Olsen, and Santiago Velez-Ferro. Centralized Teacher Assignment. 2020.

Gregory Elacqua, Anne Sofie Westh Olsen, and Santiago Velez-Ferro. The market design approach to teacher assignment: Evidence from ecuador. Working Paper 11854, Inter-American Development Bank, 2021.

Gregory Elacqua, Diana Hincapie, Isabel Hincapié, and Verónica Montalva. Can financial incentives help disadvantaged schools to attract and retain high-performing teachers? evidence from chile. *Journal of Policy Analysis and Management*, 41(2):603–631, 2022c.

Glenn Ellison and Parag A. Pathak. The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from chicago's exam schools. *American Economic Review*, 111(3):943–75, March 2021. doi: 10.1257/aer.20161290. URL https://www.aeaweb.org/articles?id=10.1257/aer.20161290.

Aytek Erdil and Haluk Ergin. What's the matter with tie-breaking? improving efficiency in school choice. *American Economic Review*, 98(3):669–89, June 2008. doi: 10.1257/aer.98.3.669. URL http://www.aeaweb.org/articles?id=10.1257/aer.98.3.669.

Juan F Escobar and Leonel Huerta. School choice and segregation. *Working Paper, Department of Industrial Engineering, University of Chile*, 2021.

Gabrielle Fack, Julien Grenet, and Yinghua He. Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review*, 109(4):1486–1529, April 2019. doi: 10.1257/aer.20151422. URL https://www.aeaweb.org/articles?id=10.1257/aer.20151422.

Torberg Falch. Teacher mobility responses to wage changes: Evidence from a quasi-natural experiment. *American Economic Review*, 101(3):460–65, 2011.

Li Feng and Tim R. Sass. The impact of incentives to recruit and retain teachers in "hard-to-dtaff" subjects. *Journal of Policy Analysis and Management*, 37(1):112–135, 2018.

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

Andrew Gelman and Guido Imbens. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456, 2019.

Andrew Gelman, Donald B Rubin, et al. Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472, 1992.

Steven Glazerman, Ali Protik, Bing-ru Teh, Julie Bruch, and Neil Seftor. Moving high-performing teachers implementation of transfer incentives in seven districts. Report, Mathematica Policy Research, 2012.

Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. Effective affirmative action in school choice. *Theoretical Economics*, 8(2):325–363, 2013. doi: https://doi.org/10.3982/TE1135. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/TE1135.

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

Clemence Idoux. Integrating new york city schools: The role of admission criteria and family preferences. *Job Market Paper, MIT Economics Department*, 2022.

Guido Imbens and Karthik Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3):933–959, 2012.

Guido Imbens and Thomas Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635, 2008. URL https://EconPapers.repec.org/RePEc:eee:econom:v:142:y:2008:i:2:p:615-635.

Ginger Zhe Jin and Phillip Leslie. The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451, 2003.

Thomas J. Kane and Douglas O. Staiger. Estimating teacher impacts on student achievement: An experimental evaluation. Working Paper 14607, National Bureau of Economic Research, 2008.

Adam J. Kapor, Christopher A. Neilson, and Seth D. Zimmerman. Heterogeneous beliefs and school choice mechanisms. *American Economic Review*, 110(5):1274–1315, May 2020. doi: 10.1257/aer.20170129. URL https://www.aeaweb.org/articles?id=10.1257/aer.20170129.

Dany Kessel and Elisabet Olme. School choice priority structures and school segregation. 2018.

Jeffrey R. Kling, Sendhil Mullainathan, Eldar Shafir, Lee C. Vermeulen, and Marian V. Wrobel. Comparison friction: Experimental evidence from medicare drug plans. *The Quarterly Journal of Economics*, 127(1):199–235, 2012.

Macarena Kutscher, Shanjukta Nath, and Sergio Urzua. Centralized Admission Systems and School Segregation: Evidence from a National Reform. IZA Discussion Papers 13305, Institute of Labor Economics (IZA), May 2020. URL https://ideas.repec.org/p/iza/izadps/dp13305.html.

T. Larroucau and I. Rios. Do "short-list" students report truthfully? strategic behavior in the chilean college admissions problem. Technical report, Working Paper, 2020.

Mariana Laverde. Unequal Assignments to Public Schools and the Limits of School Choice. Job market paper, Economics Department, University of Chicago, 2021. URL https://drive.google.com/file/d/19HJFGmaf2HA56k3rNBgB_at4aSkG3Bpz/view.

J. Margolis and S. Hashim. The impact of middle school integration efforts on segregation in two new york city districts. Technical report, 2020.

Luana Marotta. Teachers' contractual ties and student achievement: The effect of temporary and multiple-school teachers in brazil. *Comparative Education Review*, 63(3), 2019. doi: 10.1086/703981.

Robert McCulloch and Peter E Rossi. An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, 64(1-2):207–240, 1994.

Moritz Meister and Annekatrin Niebuhr. Comparing ethnic segregation across cities—measurement issues matter. *Review of Regional Research*, 41(1):33–54, 2021.

Christopher A. Neilson. The rise of centralized mechanisms in education markets around the world. Technical report, August 2019. URL https://christopherneilson.github.io/work/CSAS.html.

Christopher A. Neilson. Targeted vouchers, competition among schools, and the academic achievement of poor students. *RR Econometrica*, 2021.

Peter Nijkamp and Jacques Poot. Cultural Diversity - A Matter of Measurement. CReAM Discussion Paper Series 1502, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London, January 2015. URL https://ideas.repec.org/p/crm/wpaper/1502.html.

Hessel Oosterbeek, Sándor Sóvágó, and Bas van der Klaauw. Preference heterogeneity and school segregation. *Journal of Public Economics*, 197:104400, 2021. ISSN 0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2021.104400. URL https://www.sciencedirect.com/science/article/pii/S0047272721000360.

Sebastián Otero, Nano Barahona, and Caue Dobbin. Affirmative action in centralized college admission systems: Evidence from brazil. *Job Market Paper*, 2021. URL https://sebotero.github.io/papers/cotas_JMP.pdf.

Parag A. Pathak. The mechanism design approach to student assignment. *Annual Review of Economics*, 3(1):513–536, 2011. doi: 10.1146/annurev-economics-061109-080213. URL https://doi.org/10.1146/annurev-economics-061109-080213.

Parag A. Pathak. *What Really Matters in Designing School Choice Mechanisms*, volume 1 of *Econometric Society Monographs*, page 176–214. Cambridge University Press, 2017. doi: 10.1017/9781108227162.006.

Parag A. Pathak and Tayfun Sönmez. Leveling the playing field: Sincere and sophisticated players in the boston mechanism. *American Economic Review*, 98(4):1636–52, September 2008. doi: 10.1257/aer.98.4.1636. URL https://www.aeaweb.org/articles?id=10.1257/aer.98.4.1636.

Juan Sebastián Pereyra. A dynamic school choice model. *Games and Economic Behavior*, 80:100–114, 2013.

Michelle Reininger. Hometown disadvantage? it depends on where you're from: Teachers' location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis*, 34(2):127–145, 2012.

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

Leonardo Rosa. Teacher preferences in developing countries: Evidences from the city of sao paulo, 2017.

Peter E Rossi, Robert E McCulloch, and Greg M Allenby. The value of purchase history data in target marketing. *Marketing Science*, 15(4):321–340, 1996.

Jonathan Roth, Pedro Sant'Anna, Alyssa Bilinski, and John Poe. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. Technical report, Working Paper, 01 2022.

Emmanuel Saez. Details matter: The impact of presentation and information on the take-up of financial incentives for retirement saving. *American Economic Journal: Economic Policy*, 1(1):204–28, 2009.

Richard H. Sander. A systemic analysis of affirmative action in american law schools. *Stanford Law Review*, 57(2):367–483, 2004. ISSN 00389765. URL http://www.jstor.org/stable/40040209.

Suk Joon Son. Distributional impacts of centralized school choice. Job market paper, Economics Department, Yale University, 2020. URL https://drive.google.com/file/d/1T33dkDPdVILUz_wUghfLeR0pe9CPJobN/view.

Matthew G. Springer, Walker A. Swain, and Luis A. Rodriguez. Effective teacher retention bonuses: Evidence from tennessee. *Educational Evaluation and Policy Analysis*, 38(2): 199–221, 2016.

Camille Terrier. Matching practices for secondary public school teachers – france. *Matching in Practice*, 2014.

Kenneth E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2 edition, 2009.

A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

J.P Valenzuela, C. Bellei, and D. de los Ríos. Evolución de la segregación socioeconómica de los estudiantes chilenos y su relación con el financiamiento compartido. Technical report, Evidencias para Políticas Públicas en Educación, Fonide, Ministerio de Educación, 2009.

Aad van Mourik, Jacques Poot, and Jacques J. Siegers. Trends in occupational segregation of women and men in new zealand: Some new evidence. *New Zealand Economic Papers*, 23(1):29–50, 1989. doi: 10.1080/00779958909544145. URL https://doi.org/10.1080/00779958909544145.

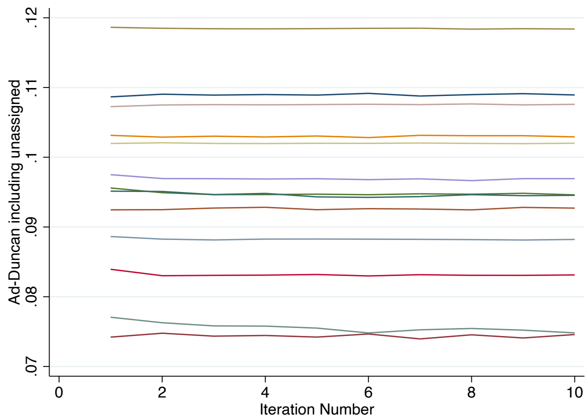
Lindsay Weixler, Jon Valant, Daphna Bassok, Justin B. Doromal, and Alica Gerry. Helping parents navigate the early childhood education enrollment process: Experimental evidence from new orleans. *Educational Evaluation and Policy Analysis*, 42(3):307–330, 2020.

Appendix A

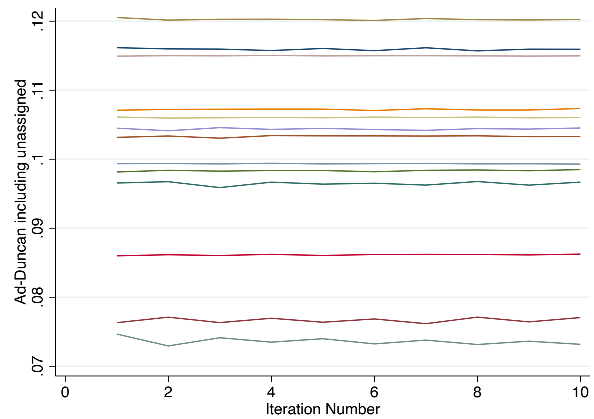
Appendix: The Effects of Adjusting Socioeconomic Reserves to Local Conditions in Chile's Centralized PreK Choice

A.1 Additional Figures

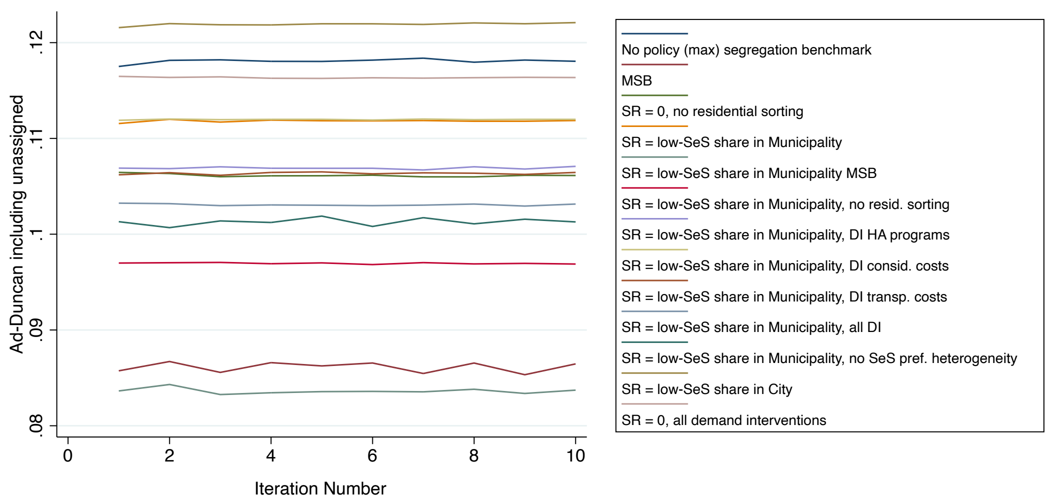
Figure A.1.1: Counterfactual loops: Costly consideration



(a) 2019

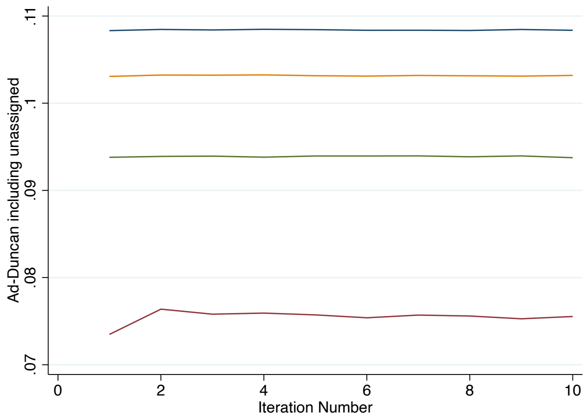


(b) 2020

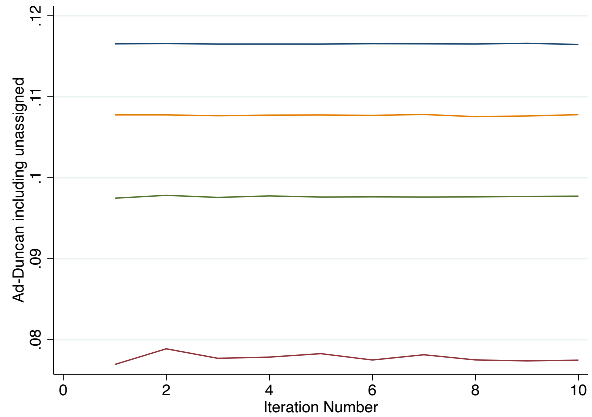


(c) 2021

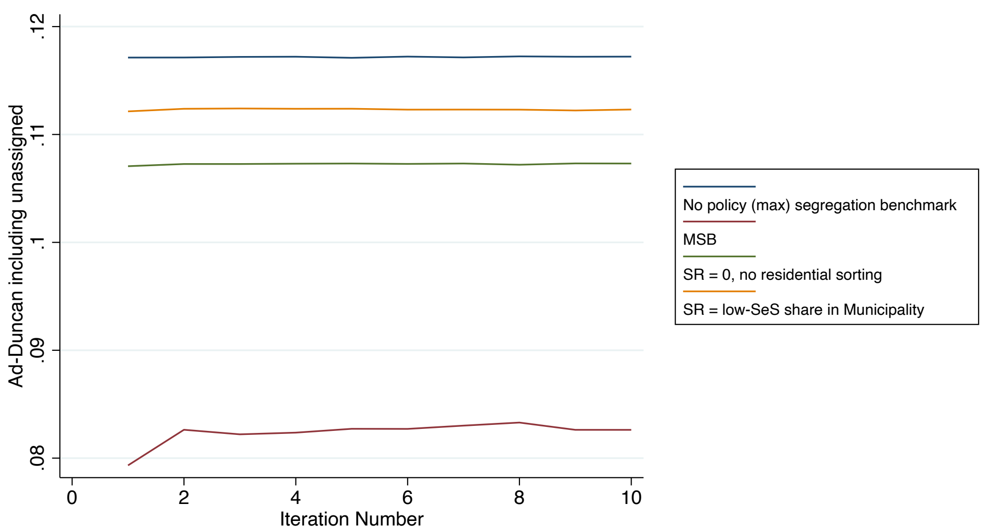
Figure A.1.2: Counterfactual loops: Costly ranking



(a) 2019



(b) 2020



(c) 2021

A.2 Additional Tables

Table A.2.1: Estimated impact of counterfactual social-welfare results in consideration cost model 2019

Countefactual			**Outcomes**											
Panel A: Only counterfactual sample			**Integration: Reduction in educational segregation gap ($\Delta^{-}\%$)**			**Achievement Gap (% of group))**				**Efficiency**				
						HA cat.		**MA cat.**		**% of ideal occupancy**		**Unassignment (% of group)**		
Demand-side simulation	SR share (%)	Randomized residences	*MA-DD*^{WUN}	*MA-DD*	*DD*	L-SeS	N-SeS	L-SeS	N-SeS	HA cat.	MA cat.	L-SeS	N-SeS	All
None	15	No	7.1	15	18.9	15.4	19.4	52.9	53.4	-	-	9.97	13.32	12.3
None	0	No	-	-	-	15.2	19.5	52.6	53.3	-	-	10.74	13.03	12.28
None	0	Yes	41.7	41.8	40.1	15.4	18.8	52	52.9	-	-	12.38	11.9	11.98
None	Muni.	No	17.5	63.1	71.9	16.5	18.8	53.1	53.4	-	-	7.13	14.93	12.25
None	City	No	-27.5	41.1	55.6	16.5	19.1	53.4	53.2	-	-	4.18	15.44	12.24
$\Delta^{+}\delta_{HA}$	Muni.	No	34.9	77.1	85.8	20.2	20.7	50.8	52.2	-	-	7.21	14.59	12.11
$\Delta^{-}c_i$	Muni.	No	20.2	64.6	73.9	16.5	18.7	53.1	53.2	-	-	7.05	14.68	12.07
$\Delta^{-}\tau_t$	Muni.	No	47.2	80.9	88.1	16.7	18.2	53.1	52.9	-	-	5.24	9.86	8.37
All	Muni.	No	60.3	95.7	103.4	20	19.9	50.9	51.5	-	-	4.91	9.24	7.88
All	City	No	3.9	15.8	13.9	17.1	21.8	51.3	51.2	-	-	7.93	7.83	7.83
Panel B: All applicants			**Integration: Reduction in educational segregation gap ($\Delta^{-}\%$)**			**Achievement Gap (% of group))**				**Efficiency**				
						HA cat.		**MA cat.**		**% of ideal occupancy**		**Unassignment (% of group)**		
Demand-side simulation	SR share (%)	Randomized residences	*MA-DD*^{WUN}	*MA-DD*	*DD*	L-SeS	N-SeS	L-SeS	N-SeS	HA cat.	MA cat.	L-SeS	N-SeS	All
None	15	No	7.2	18.7	26.5	16.3	21.1	52.8	53.8	89.5	86.7	9.82	14.5	13.03
None	0	No	-	-	-	16.1	21.1	52.8	53.8	89.4	86.7	10.66	14.1	12.97
None	0	Yes	49.5	47.2	44.7	16.3	20.8	52.4	53.4	89.7	86.5	11.6	13.18	12.59
None	Muni.	No	23.9	89.6	101.6	17.4	20.2	52.9	53.9	89.3	86.7	6.88	16.38	13.15
None	City	No	-41.8	48.7	65.4	17.5	20.5	53	53.8	89.1	86.7	4.27	16.89	13.21
$\Delta^{+}\delta_{HA}$	Muni.	No	37.5	105.1	117.7	19.3	20.2	52	54.1	92.4	86	7.25	16.77	13.54
$\Delta^{-}c_i$	Muni.	No	26.3	92.3	104.7	17.4	20.2	52.9	53.8	89.2	86.5	6.93	16.36	13.16
$\Delta^{-}\tau_t$	Muni.	No	52.7	115.7	126.9	17.5	19.6	52.8	53.5	89.8	88	5.92	14.02	11.3
All	Muni.	No	63.4	130	142.2	19.2	19.6	51.8	53.4	92.5	87	6.14	14.38	11.59
All	City	No	11.3	17.8	17.1	16.6	21.3	52.2	53.1	92.5	87.3	9.55	12.42	11.37

Table A.2.2: Estimated impact of counterfactual social-welfare results in consideration cost model 2020

Countefactual			**Outcomes**											
Panel A: Only counterfactual sample			**Integration: Reduction in educational segregation gap (Δ^{-}%)**			**Achievement Gap (% of group))**				**Efficiency**				
						HA cat.		**MA cat.**		**% of ideal occupancy**		**Unassignment (% of group)**		
Demand-side simulation	SR share (%)	Randomized residences	*MA-DD*^{WUN}	*MA-DD*	*DD*	L-SeS	N-SeS	L-SeS	N-SeS	HA cat.	MA cat.	L-SeS	N-SeS	All
None	15	No	1.9	6.8	6.8	16.4	20.7	51.3	53.8	-	-	4.04	7.41	6.4
None	0	No	-	-	-	16.1	20.6	51	53.8	-	-	5.02	7.25	6.48
None	0	Yes	44.8	45.8	42.2	16.7	20.3	51.3	53.5	-	-	6.93	6.69	6.67
None	Muni.	No	22.1	39.1	39.9	17.1	20	51.1	53.8	-	-	3.32	8.17	6.54
None	City	No	-11.1	17	22.4	17.4	20.1	51.3	53.8	-	-	1.44	8.65	6.63
$\Delta^{+}\delta_{HA}$	Muni.	No	29.3	44.8	45.5	21.5	23.4	48.1	51.8	-	-	3.51	7.85	6.38
$\Delta^{-}c_i$	Muni.	No	25.5	42.7	42.9	17.2	20.1	50.7	53.8	-	-	3.23	8.08	6.43
$\Delta^{-}\tau_t$	Muni.	No	32.5	41.5	40.9	17.3	19.7	50.9	53.6	-	-	2.22	4.63	3.83
All	Muni.	No	42.8	51.5	49.7	21.1	22.9	48.1	51.7	-	-	2.17	4.23	3.54
All	City	No	2.5	5.6	4.4	19.2	24	48.5	51.5	-	-	3.51	3.56	3.49
Panel B: All applicants			**Integration: Reduction in educational segregation gap (Δ^{-}%)**			**Achievement Gap (% of group))**				**Efficiency**				
						HA cat.		**MA cat.**		**% of ideal occupancy**		**Unassignment (% of group)**		
Demand-side simulation	SR share (%)	Randomized residences	*MA-DD*^{WUN}	*MA-DD*	*DD*	L-SeS	N-SeS	L-SeS	N-SeS	HA cat.	MA cat.	L-SeS	N-SeS	All
None	15	No	3.5	13.1	19.1	17.6	22	52.5	53.8	83.3	81.3	4.25	8.07	6.9
None	0	No	-	-	-	17.3	22	52.6	53.8	83	81.3	5.17	7.87	6.97
None	0	Yes	45.9	48.6	45.8	17.8	21.8	52.5	53.8	83.8	80.9	6.44	7.32	6.94
None	Muni.	No	31.1	59	64.1	18.4	21.5	52.6	53.9	83.3	81.3	3.28	8.85	7.04
None	City	No	-23.3	21.8	31.5	18.9	21.5	52.3	53.9	83.3	81.2	1.4	9.38	7.11
$\Delta^{+}\delta_{HA}$	Muni.	No	36.7	63.8	68.6	21.1	22.7	50.9	53.3	90.8	79.3	3.59	9.07	7.25
$\Delta^{-}c_i$	Muni.	No	33.6	62.5	67.6	18.4	21.5	52.3	53.8	83.3	80.9	3.22	9	7.09
$\Delta^{-}\tau_t$	Muni.	No	41.6	65.7	69.6	18.4	21.3	52.4	53.6	84	82	2.54	6.97	5.54
All	Muni.	No	50.9	76.8	81	20.7	22.3	50.7	53.1	90.7	80.1	2.76	7.3	5.81
All	City	No	3.1	6.1	5.1	19	23.4	51	52.8	90.6	80.2	4.65	6.35	5.73

Table A.2.3: Estimated impact of counterfactual social-welfare results in consideration cost model 2021

Countefactual			**Outcomes**											
Panel A: Only counterfactual sample			**Integration: Reduction in educational segregation gap (Δ^{-}%)**			**Achievement Gap (% of group))**				**Efficiency**				
						HA cat.		**MA cat.**		**% of ideal occupancy**		**Unassignment (% of group)**		
Demand-side simulation	SR share (%)	Randomized residences	*MA-DD*^{WUN}	*MA-DD*	*DD*	L-SeS	N-SeS	L-SeS	N-SeS	HA cat.	MA cat.	L-SeS	N-SeS	All
None	15	No	5.6	5.6	6.1	17.4	21.1	52.8	54.5	-	-	3.94	5.66	5.01
None	0	No	-	-	-	17.4	21	52.7	54.4	-	-	3.87	5.58	4.94
None	0	Yes	37.7	40	40.2	17.7	21	52.9	53.8	-	-	5.16	5.46	5.38
None	Muni.	No	19.6	40.3	42	18.4	20.3	52.3	54.8	-	-	2.33	6.6	4.88
None	City	No	-12.8	17	20.6	18.8	20.3	52.1	54.7	-	-	1.23	7.1	4.97
$\Delta^{+}\delta_{HA}$	Muni.	No	34.7	54.4	55.8	23.1	25	49.2	52	-	-	2.51	6.52	4.92
$\Delta^{-}c_i$	Muni.	No	19.2	41.8	43.6	18.2	20.2	52.5	54.5	-	-	2.18	6.53	4.78
$\Delta^{-}\tau_t$	Muni.	No	36.7	51.2	52	18.2	20.2	52.5	54.9	-	-	1.23	3.51	2.64
All	Muni.	No	47.2	61.3	62.3	22.7	23.8	49.7	52.2	-	-	1.08	3.18	2.41
All	City	No	5.4	9.5	9	21.1	25.4	50.5	51.6	-	-	1.94	2.6	2.39
Panel B: All applicants			**Integration: Reduction in educational segregation gap (Δ^{-}%)**			**Achievement Gap (% of group))**				**Efficiency**				
						HA cat.		**MA cat.**		**% of ideal occupancy**		**Unassignment (% of group)**		
Demand-side simulation	SR share (%)	Randomized residences	*MA-DD*^{WUN}	*MA-DD*	*DD*	L-SeS	N-SeS	L-SeS	N-SeS	HA cat.	MA cat.	L-SeS	N-SeS	All
None	15	No	5.1	5.6	6.5	18.4	23	52.7	54.3	80.8	80.8	4.01	6.21	5.31
None	0	No	-	-	-	18.5	22.9	52.7	54.3	80.7	80.8	3.96	6.15	5.26
None	0	Yes	48	46.2	45.8	18.9	23.2	52.8	53.7	81.3	80.1	4.95	5.95	5.49
None	Muni.	No	34	63.6	65.6	19.4	22.3	52.4	54.6	80.8	80.7	2.47	7.39	5.35
None	City	No	-13.8	25.9	29.2	19.9	22.1	52.2	54.6	80.6	80.6	1.32	7.96	5.45
$\Delta^{+}\delta_{HA}$	Muni.	No	45.8	79	81	22.5	24.3	50.5	53.6	89.6	78	2.79	7.85	5.7
$\Delta^{-}c_i$	Muni.	No	33.7	65.3	67.5	19.3	22.2	52.5	54.4	80.6	80.6	2.35	7.4	5.31
$\Delta^{-}\tau_t$	Muni.	No	51.7	79.8	81.3	19.3	22.1	52.6	54.7	81.9	82.1	1.65	5.58	3.96
All	Muni.	No	56.8	87.5	89.5	22.2	23.5	50.8	53.6	89.4	79.4	1.77	6.12	4.28
All	City	No	5.6	9	8.5	20.8	24.8	51.3	53.1	89.4	79.5	2.99	5.15	4.21

Table A.2.4: All Applicants and Applicants in Preference Estimation by Municipality

	All applicants		In pref. estimation	
Municipality	*N*	low-SeS (%)	*N* (%)	low-SeS (%)
Puente Alto	8,623	36	75	37
Maipu	6,526	32	78	32
Santiago	5,198	26	74	29
San Bernardo	4,307	44	72	45
Quilicura	3,974	36	71	38
La Florida	3,790	32	76	32
Pudahuel	3,312	45	69	50
La Pintana	2,871	64	73	66
Renca	2,446	48	73	52
Peñalolen	2,402	49	67	54
Cerro Navia	2,211	58	76	62
Estacion Central	2,177	30	72	33
El Bosque	2,047	52	76	56
Recoleta	1,898	46	70	48
Colina	1,874	51	64	54
San Miguel	1,830	19	63	20
Independencia	1,758	28	74	29
Conchali	1,748	43	74	45
Quinta Normal	1,729	37	71	40
Ñuñoa	1,459	18	60	22
La Granja	1,381	50	70	53
Lo Espejo	1,357	57	71	60
Lo Prado	1,342	49	78	52
La Cisterna	1,326	30	71	35
Huechuraba	1,212	47	73	51
Las Condes	1,169	22	62	27
Macul	1,168	32	63	36
Pedro Aguirre Cerda	1150	47	64	54
Cerrillos	1,054	44	75	50
San Joaquin	1,051	44	73	50
San Ramon	1,006	56	69	61
Padre Hurtado	1,003	37	61	37
Lo Barnechea	611	41	61	48
Providencia	600	13	64	14
La Reina	599	26	69	31
Lampa	467	43	14	20
Vitacura	140	15	50	15

A.3 Definition of Municipalities' Buffers, Schools and Applicants

In Santiago, most municipalities are not separated by natural geographical features and are instead interconnected, although they may be separated by highways, which can complicate transportation. Given this context of interconnectedness, there is a need to expand the set of programs included in the choice sets of applicants, particularly in smaller municipalities that are highly interconnected. To do that, we implemented the following approach.

1. We identified all Municipalities within the urban limits of Santiago city, as illustrated in Panel A of Figure A.3.1.

2. We locate all PreK applicants in each Municipality as shown in Panel B, limiting our analysis to those applicants applying to schools in the Metropolitan Region and with accurate geolocation data. Specifically, the data includes five levels of accuracy based on the responses obtained from the Google Geocoding API or user information: (1) a unique response (rooftop or range_interpolated quality), (2) a unique response (geometric_center quality), (3) multiple coherent responses, where the centroid of the responses is used, (4) the user's shared location, and (5) imputed coordinates of the Municipality centroid. For the analysis, we considered only the applicants with a quality score ranging from (1) to (4). The Municipality that is painted in red in Panel B of Figure A.3.1 is Pirque, excluded from the analysis due to its limited urban area and to only having 14 applicants located in it.

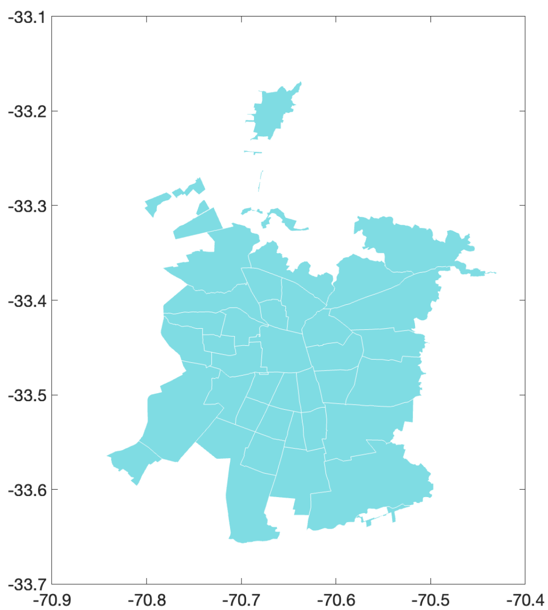
3. After locating all applicants, we identify all programs that were listed as first preferences in their ranked order lists and locate them.[1] For each Municipality, we determined the schools that were first listed by the applicants located in that particular Municipality. We drew a buffer radius around the Municipality such that 90% of these schools lay within the Municipality + Buffer area. The buffer zones around all Municipalities are shown in Panel C of Figure A.3.1. Given these geographical units, for each Municipality, we do not consider applicants that have schools listed on their applications outside the defined limits.

To summarize, this process defines (i) the geographical units, (ii) the applicants, and (iii) the schools we used in the analysis.

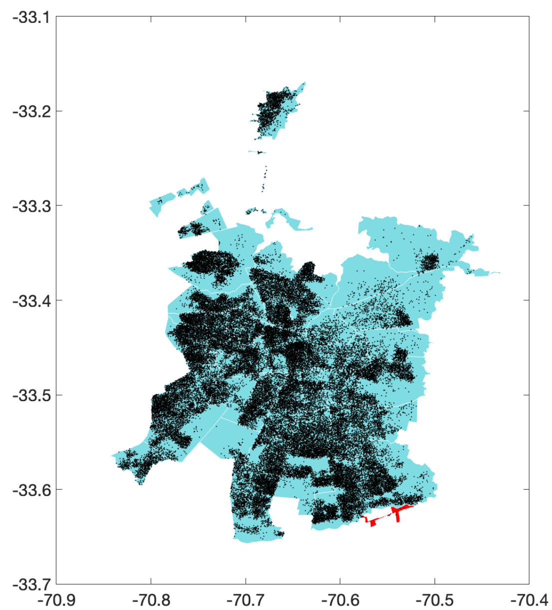
One of the main conclusions of the chapter is that the design of SR should take into account local conditions, but it is important to consider the interaction between different geographic units used to measure segregation. This interaction can affect how segregation is measured, as shown in the examples of Maipú and La Cisterna in Section 2.6. To provide further context, Figure A.3.2 presents the socioeconomic composition of neighborhoods in

[1] In this step of the procedure, we consider first preferences from all applicants across all years pooled together.

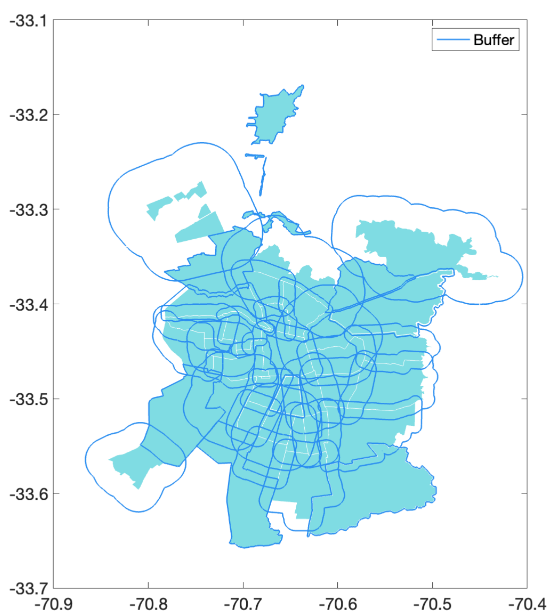
Figure A.3.1: Geographic units definition



(a) Municipalities in Santiago City



(b) Applicants distribution in Santiago City



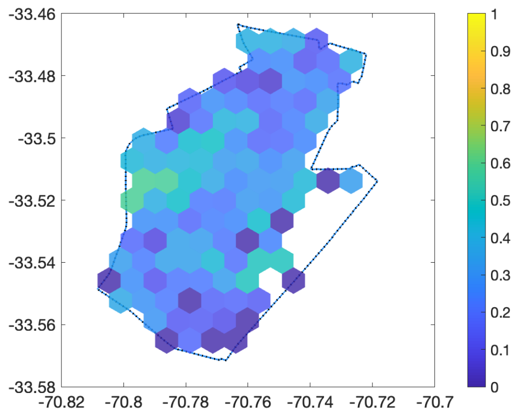
(c) Municipalities Buffers

these Municipalities, and in the case of La Cisterna, in neighboring areas, by dividing the area within their buffers into smaller polygons.

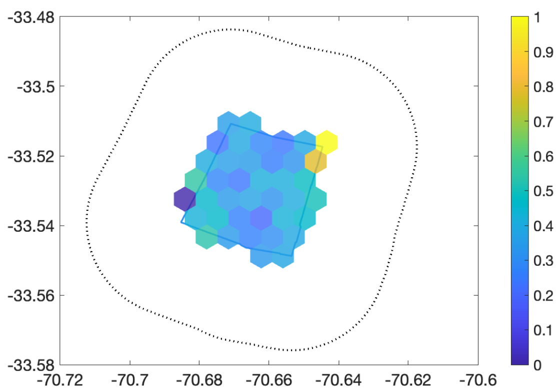
Panel A displays the socioeconomic composition of neighborhoods in Maipú within the buffer area, which coincides with the Municipality boundary. The polygons in the map show a maximum low-SeS share of 0.58 towards the west, while the minimum value is 0, observed in specific zones to the south, east, and north. As previously mentioned, Maipú is a Municipality that exhibits results similar to the city-wide average, partly due to its relative isolation from other Municipalities, reflected in the absence of a buffer zone in Maipú.

Panels B and C depict the results for La Cisterna, with Panel B focusing on the Municipality and Panel C encompassing the entire buffer area. La Cisterna tends to exhibit relative homogeneity in terms of socioeconomic composition, with the low-SeS share of the polygons ranging from 0.2 to 0.55, excluding polygons in the northeastern and southwestern corners of the map that contain fewer students and may not be representative. However, the composition of the neighboring areas, as shown in Panel C, reveals significant exposure to higher shares of low-SeS students in the west and southeast, and higher shares of no-SeS students in the north. This partly explains why results for La Cisterna, and other Municipalities under similar conditions, may differ from average results.

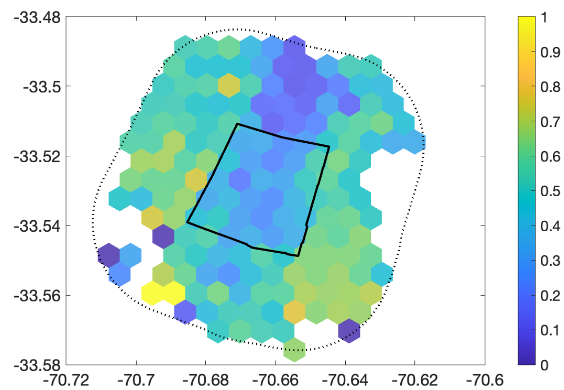
Figure A.3.2: Socioeconomic Composition of Neighborhoods in Maipú's and La Cisterna's Buffer



(a) Share of Low-SeS students in Maipú



(b) Share of Low-SeS students in La Cisterna (Municipality boundaries)



(c) Share of Low-SeS students in La Cisterna (Municipality+Buffer)

A.4 Costly Consideration Detailed Model to Data Mapping

Notation

Let's define $\mathcal{P}_i$ as the space of available programs to applicant i (within the Municipality + Buffer geography defined in Section 1.4), $\mathcal{C}_i$ as the set of alternatives considered by the applicant, and $\mathcal{R}_{\mathcal{C}_i}$ as the ordered set of programs considered, and acceptable, that is, descending on v_{ijt} with $v_{ijt} > 0$ (for simplicity we eliminate the time subscript in the expressions below).

In other words, $\mathcal{R}_{\mathcal{C}_i}$ is the partial ranking when only programs in $\mathcal{C}_i$ have been considered. Under this notation, we will only use ROL_i to express the final $\mathcal{R}_{\mathcal{C}_i}$ of i once it is not optimal for i to consider any more programs.

We use the notation $k_{r=l}$ to refer to the program ranked in the l position in a given partial ranking. $k_{r=|\mathcal{R}_{\mathcal{C}_i}|+1}$ refers to the outside option and has a utility $v_{ik_{r=|\mathcal{R}_{\mathcal{C}_i}|+1}} = 0$. For convenience we also define $v_{ik_{r=0}} \equiv \infty$.

Next, we need to introduce some notation for the normal distribution that models the uncertainty over ϵ_{ij} and $\xi_j^{g_i}$. To that end we use the conjugate priors defined for the implementation of the Gibbs Sampling (see Appendix A.6), assuming that applicants' expectations are determined by the same parameters. This implies that, as detailed in Appendix A.6, $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$, and $\xi_j^g \sim N(0, \sigma_{\xi^g}^2)$, and using these distributions we can further define $\sigma_{CV_i}^2 \equiv \sigma_\epsilon^2 + \sigma_{\xi^{g_i}}^2$.

$\phi(x)$ and $\Phi(x)$ represent the standard normal probability density and standard normal cumulative distribution functions evaluated at x. And the standardization of the relevant distributions is summarized with the notation $sb_{ij}^n \equiv \frac{v_{ik_{r=n}} - (v_{ij} - \theta_{ij})}{\sigma_{CV_i}}$ to simplify the expression below.

We furthermore use the convention $\prod_{l=1}^{0} \{1 - q_{ik_{r=l}}\} \equiv 1$ and note that given these definitions $sb_{ij}^{|\mathcal{R}_{i,\mathcal{C}_i}|+1} = \frac{-(v_{ij} - \theta_{ij})}{\sigma_{CV_i}}$, and $sb_{ij}^0 = \infty$ (and thus $\phi(sb_{ij}^0) = 0$ and $\Phi(sb_{ij}^0) = 1$).

As a result, the consideration value of program j, conditional on the partial ranking $\mathcal{R}_{i,\mathcal{C}_i}$ can be derived as follows:

First, consider the expected utility of a given partial ranking $\mathcal{R}_{i,\mathcal{C}_i}$:

$$\mathbb{E}[v_{i,\cdot}|\mathcal{R}_{i,\mathcal{C}_i}] = \sum_{n=1}^{|\mathcal{R}_{i,\mathcal{C}_i}|} \prod_{l=1}^{n-1} \{1 - q_{ik_{r=l}}\} \, q_{i,k_{r=n}} v_{i,k_{r=n}}$$

When deciding whether to consider a program j, the applicant additionally knows the observable part of the utility, which we can express as $v_{ij} - \theta_{ij}$. Depending on θ_{ij}'s value, v_{ij} can take any value above or below the utilities of programs already included in $\mathcal{R}_{i,\mathcal{C}_i}$. Therefore, all possibilities need to be considered.

Consideration Value

Let's start considering the possibility of v_{ij} being above all other programs in the partial ranking. That happens with probability $\Phi(sb_{ij}^0) - \Phi(sb_{ij}^1)$, and, using the properties of the normal distribution,[2] the expected value of v_{ij} if that is the case is given by

$$\mathbb{E}[v_{ij}|v_{ij} > v_{ik_{r=1}}] = v_{ij} - \theta_{ij} - \sigma_{CV}\frac{\phi(sb^0_{ij}) - \phi(sb^1_{ij})}{\Phi(sb^0_{ij}) - \Phi(sb^1_{ij})}$$

And combining both:

$$\begin{aligned}\mathbb{P}(v_{ij} > v_{ik_{r=1}})\mathbb{E}[v_{ij}|v_{ij} > v_{ik_{r=1}}] =\\ \left(\Phi(sb^0_{ij}) - \Phi(sb^1_{ij})\right)(v_{ij} - \theta_{ij}) - \sigma_{CV}\left(\phi(sb^0_{ij}) - \phi(sb^1_{ij})\right)\end{aligned}$$

Now, if j is ranked first, assignment to the program occurs only with probability q_{ij}. Moreover, what we are interested in is the expected change in the value of the partial ranking if j were to be considered, which implies also considering that including j in the portfolio would make assignment to the rest of the programs less likely. Adding these elements, we have that the change in the value of partial ranking is given by:

$$\begin{aligned}\mathbb{P}(v_{ij} > v_{ik_{r=1}})q_{ij}\left(\mathbb{E}[v_{ij}|v_{ij} > v_{ik_{r=1}}] - \mathbb{E}[v_{i,\cdot}|\mathcal{R}_{i,\mathcal{C}_i}]\right) =\\ q_{ij}\left(\Phi(sb^0_{ij}) - \Phi(sb^1_{ij})\right)(v_{ij} - \theta_{ij}) - \sigma_{CV}\left(\phi(sb^0_{ij}) - \phi(sb^1_{ij})\right) -\\ q_{ij}\left(\Phi(sb^0_{ij}) - \Phi(sb^1_{ij})\right)\sum_{n=1}^{|\mathcal{R}_{i,\mathcal{C}_i}|}\prod_{l=1}^{n-1}\left\{1 - q_{ik_{r=l}}\right\}q_{i,k_{r=n}}v_{i,k_{r=n}}\end{aligned}$$

Following this logic for the different regions in which the value of v_{ij} can affect the partial rankings' utility (when it has a positive value),[3] we can express the consideration value as follows:

Definition 3. ***Consideration Value***: $CV_{j,\mathcal{R}_{i,\mathcal{C}_i}}$

$$CV_{j,\mathcal{R}_{i,\mathcal{C}_i}} = \begin{aligned}&\underbrace{q_{ijt}\textstyle\sum_{n=0}^{|\mathcal{R}_{i,\mathcal{C}_i}|}\prod_{l=1}^{n}\left\{1 - q_{ik_{r=l}t}\right\}\begin{Bmatrix}[\Phi(sb^n_{ij}) - \Phi(sb^{n+1}_{ij})](v_{ij} - \theta_{ij}) - \\ \sigma_{CV_i}[\phi(sb^n_{ij}) - \phi(sb^{n+1}_{ij})]\end{Bmatrix}}_{\text{Expected utility of assignment to } j} - \\ &\underbrace{q_{ijt}\textstyle\sum_{n=1}^{|\mathcal{R}_{i,\mathcal{C}_i}|}\left\{(1 - \Phi(sb^n_{ij}))q_{ik_{r=n}t}v_{ik_{r=n}}\prod_{l=1}^{n-1}\left\{1 - q_{ik_{r=l}t}\right\}\right\}}_{\text{Expected counterfactual utility conditional on assignment to } j}\end{aligned}$$

[2]Specifically, if $x \sim N(\mu, \sigma^2)$, then:

$$\begin{aligned}\mathbb{P}(b > x > a) &= \Phi(\beta) - \Phi(\alpha)\\ \mathbb{E}[x|b > x > a] &= \mu - \sigma\frac{\phi(\beta) - \phi(\alpha)}{\Phi(\beta) - \Phi(\alpha)}\end{aligned}$$

where $\beta \equiv \frac{b-\mu}{\sigma}$ and $\alpha \equiv \frac{a-\mu}{\sigma}$.

[3]Those regions are $\left\{(\infty, v_{ik_{r=1}}), (v_{ik_{r=1}}, v_{ik_{r=2}}), ..., \left(v_{ik_{r=|\mathcal{R}_{i,\mathcal{C}_i}|}}, 0\right)\right\}$.

There are different ways to express $CV_{j,\mathcal{R}_{i,\mathcal{C}_i}}$, but an additional advantage of this formulation is that it is computationally convenient when computing the consideration value of multiple programs simultaneously.

To determine consideration, we can use the estimated utility parameters and consideration costs. The relation between the restrictions of the model and the data is expressed in result 2 and the empirical implementation to obtain partial rankings is presented in Appendix A.6.

A.5 Identification Discussion

To facilitate the discussion, we will start by simplifying equation 1.1 excluding applicants with family priorities and the coefficient on distance designed to account for the potential impact of the COVID-19 pandemic. We will also omit time subscripts for clarity:

$$\begin{aligned} v_{ij} &= \underbrace{x_j\beta^{g_i} + \xi_j^{g_i} + \epsilon_{ij}}_{u_{ij}} - d_{ij} \\ c_i &= c^{g_i} + \zeta_i \end{aligned}$$

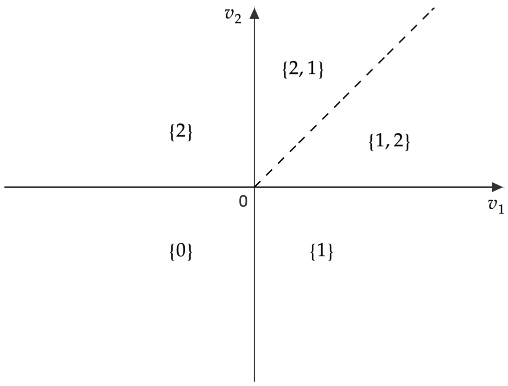
Identification of utility densities

The first part of the discussion in the costly consideration case, is that we could identify the density of u_{ij} conditional on consideration and conditional on $v_{ij} > 0$, by only considering applicants that included program j in their ROL, using the relative rank of the program to others in the portfolio of different applicants. Restricting ourselves to programs in ROLs enables us to use the non-parametric identification argument proposed by Agarwal and Somaini (2018). It's important to note that this restriction doesn't hinder identification in theory. In fact, with enough data variation, we can identify the same region of utilities ($\mathbf{u} > 0$) as in the full consideration model. However, due to data limitations, we identify utilities with less accuracy in practice.

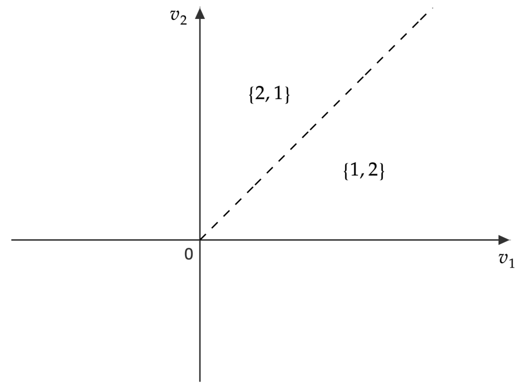
Figure A.5.1 presents a graphical explanation. Panels A.5.1a and A.5.1b display the indirect utility regions that correspond to different observed portfolios in the full consideration and in the costly consideration cases respectively. Additionally, Panels A.5.1c and A.5.1d explain how we can identify utilities by calculating the difference between the number of applicants selecting a portfolio at distances d_I and d_{IV} and the number of applicants selecting it at distances d_{II} and d_{III}.

Panel A.5.1c shows that, in the full consideration case, we can use portfolios with only one program (or even no programs) to aid in identifying the density of u. We are however restricted to utilities in the first quadrant due to the non-negative distance. On the other hand, Panel A.5.1d illustrates how we can identify utilities in the first quadrant in the costly

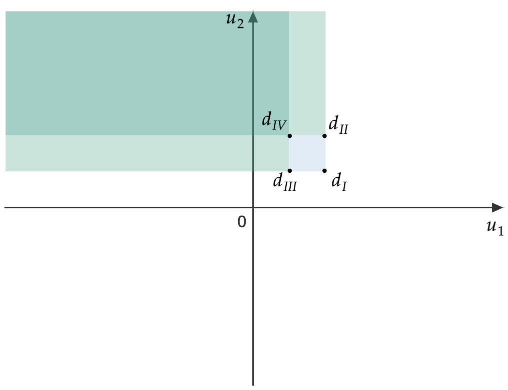
Figure A.5.1: Identification of the density of u using only information in ROLs



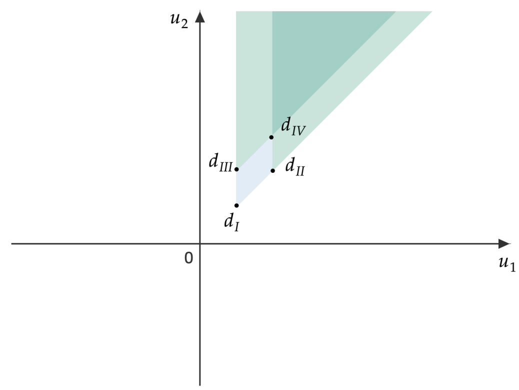
(a) Full consideration: Utility regions as a function of ROL chosen



(b) Costly consideration: Utility regions as a function only of programs in ROL chosen



(c) Shifts in distance between programs ranked and not ranked ($\{2, 0\}$ here) also identify the density of u



(d) Only shifts in distance between programs ranked (in this example $\{2, 1\}$) identify the density of u

consideration case with enough data variation using differences in distance vectors. The difference is that in the latter case identification in practice relies on fewer observations.

Separating θ_{ij} and $(v_{ij} - \theta_{ij})$

After identifying the density of $\mathbf{u}$, the next step would be to identify the density of consideration costs. Understanding the argument in Idoux (2022)'s costly ranking model is insightful to see the difference. In that model, programs with a ranking value above the ranking cost are eligible to enter the ROL during the portfolio formation phase, and the decision-making heuristic used selects the program with the highest indirect utility. Ranking values are a function of indirect utilities and assignment probabilities, as well as programs already included in the partial ranking. As a result, assuming that utility shocks are independent of ranking cost shocks ($\epsilon_{ij} \perp \zeta_i$) coupled with having identified the density of indirect utilities with other data variation, allows for identification of the ranking costs density using assignment probabilities.

However, in the costly consideration model, this is not sufficient. The consideration value of j depends on the indirect utilities of programs already included in the partial ranking, as well as on the probabilities of assignment to programs in the partial ranking and to program j, and on the correct identification of the observable $v_{ij} - \theta_{ij}$ part of indirect utilities. Therefore, to identify the density of consideration costs, we must separately identify θ_{ij} and v_{ij}. This is however not possible without an instrument that correlates with consideration and not with utilities, which we do not have (and also do not allow in our model). Estimated parameters thus depend on the specification of δ_j^g and observable program attributes in x_j. To clarify why, take the average of δ_j^g over individuals in a given socioeconomic group g to obtain a standard linear equation model:

$$\mathbb{E}[u_{ij} | i \in g, j \in ROL_i] = \delta_j^g = x_j \beta^g + \xi_j^g$$

Variation in utilities and program characteristics allows the estimation of β under the standard minimum least squares assumptions, with the normalization that ξ_j^g's distribution centers around zero and one of the programs having $\xi_j^g = 0$.[4]

In the implementation of the Gibbs sampling, to obtain β^g, we include a constant for each socioeconomic group in $\mathbf{x}$, taking advantage of it being estimated in a separate step of the Gibbs sampler from ξ_j^g, thus centering the distribution of ξ_j^g around zero.[5]

Identification of consideration costs

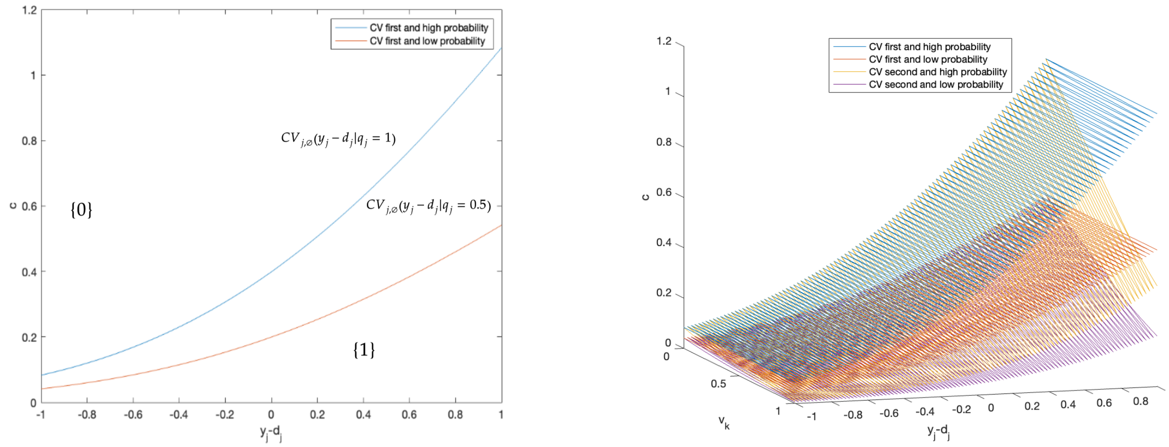
If θ_{ij} and v_{ij} were separately identified, the next part of the identification argument is the same as in Idoux (2022): consideration cost densities are identified by assumption $\epsilon_{ij} \perp \zeta_i$ and variation in assignment probabilities, which arises from variability in available seats, differences in applicant priorities, and changes in the composition of the expected applicant cohorts over time. As a result, to facilitate this identification, it is advisable to use application data spanning at least two years.

Figure A.5.2 provides a graphical representation of the identification argument for two cases: (1) program j is considered with an empty partial ranking, and (2) program k is ranked at the moment of j's consideration, implying that the applicant knows v_{ik}. In both cases, the figure displays the consideration value as a function of the observable part of the

[4]Formally, the distribution of ξ_j^g is centered around zero conditional on x_j^g. For simplicity, we do not include that in the notation.

[5]An alternative way to specify utilities is by grouping programs with similar observable characteristics, relaxing the linearity assumption implied by $x_j \beta^g$. Under this alternative, the observed utility corresponds to the average expected utility of programs with the same observables, while the unobservable part corresponds to the difference between that average and a given program's utility (plus the individual match effect ϵ_{ij}). However, this approach does not address bias generated by omitted observable program attributes and increases the number of parameters, making the interpretation of differences in preferences between socioeconomic groups less intuitive. Therefore, we use the more restrictive but simpler linear projection specification.

Figure A.5.2: Identification of the density of c using variation in assignment probability



(a) Consideration cost bounds for different levels of assignment probabilities and no other programs ranked so far

(b) Consideration cost bounds for different levels of assignment probabilities and another program already ranked

utility and, in panel A.5.2b, as a function of the other program's indirect utility. The figure varies the assignment probability q_{ij} from 1 to $1/2$.

We can observe that when v_{ik} is close to zero, considering program j for first or second in the partial ranking has little impact on its CV. However, as v_{ik} increases, the consideration value for program j drops and eventually intersects with $CV_{j,\varnothing|q_{ij}=1/2}$. The crucial point to note here is that the assumption that programs with higher consideration values are included first in the partial ranking enables us to select the consideration value that bounds the consideration cost.

The role of un-ranked programs

If a program is not ranked, its indirect utilities are either unbounded or bounded below zero, depending on whether or not it was considered. Therefore, the mass of indirect utilities above zero for all applicants that considered and ranked program j must be consistent with that below zero for all those that considered and did not. If too many individuals in group g consider but do not rank program j, the estimate for ξ_j^g will be pushed down, which in turn lowers $x_j\beta^g$ (albeit to a lesser extent). This leads to a lower estimate of applicants considering program j due to the smaller average $v_{ij} - \theta_{ij} - d_{ij}$ and smaller c_i values resulting from the lower average u_{ij}^g for those applicants ranking program j. Additionally, a smaller likelihood of ranking program j conditional on consideration due to the smaller average θ_{ij} resulting from a smaller ξ_j^g will lead to higher ϵ_{ij} values for those who did rank the program. The equilibrium result is unclear, but we want to emphasize that in practice estimated parameters are influenced by both including or not including a program in the ROL.

λ and $\Delta_{d,C19}$

To end this section, we discuss the identification of the parameters λ and $\Delta_{d,C19}$, which were omitted until now. The identification of λ depends on the correct specification of the model and assumes that the average utility differences between applicants with and without family priorities are constant. It is important to include this parameter to avoid any omitted variable bias problem in identifying ξ_j^g, as applicants with family priorities are expected to have higher indirect utilities in those programs. The identification of $\Delta_{d,C19}$ is given by the variation in preferences between 2019 and the later years.

A.6 Gibbs Sampling Detailed Explanation

Priors

Random coefficients

$$\begin{aligned}
\xi_j^g &\sim N(0, \sigma^2_{\xi^g}) \\
\epsilon_{ij} &\sim N(0, \sigma^2_{\epsilon}) \\
\zeta_i &\sim TN(0, \sigma^2_{\zeta}, -c^{g_i}, \infty)
\end{aligned}$$

The covariances of the random coefficients are assumed to have an Inverse-Wishart conjugate prior given by:

$$\begin{aligned}
\sigma^2_{\xi^g} &\sim IW(\tau_{\xi^g}, df_{\xi^g}) \\
\sigma^2_{\epsilon} &\sim IW(\tau_{\epsilon}, df_{\epsilon}) \\
\sigma^2_{\zeta} &\sim IW(\tau_{\zeta}, df_{\zeta})
\end{aligned}$$

Fixed coefficients

The priors on the fixed parameters are assumed to be:

$$\begin{aligned}
\lambda &\sim N(\mu_\lambda, \Sigma_\lambda) \\
\beta^g &\sim N(\mu_{\beta^g}, \Sigma_{\beta^g}) \\
\Delta_{d,C19} &\sim N(\mu_{\Delta_{d,C19}}, \sigma^2_{\Delta_{d,C19}}) \\
c^g &\sim TN(\mu_{c^g}, \sigma^2_{c^g}, 0, \infty)
\end{aligned}$$

The specific (proper) diffused priors used, following Abdulkadiroğlu et al. (2017), Rossi et al. (1996) and Idoux (2022) are:

$$
\begin{aligned}
\mu_{\lambda_{prior}} &= 0 \\
\Sigma_{\lambda_{prior}} &= 100\sigma_\epsilon^2 I_2 \\
\mu_{\beta^g_{prior}} &= 0 \\
\Sigma_{\beta^g_{prior}} &= 100\sigma_\epsilon^2 I_K \\
\mu_{\Delta_{d,C19}{}_{prior}} &= 0 \\
\sigma^2_{\Delta_{d,C19}{}_{prior}} &= 100\sigma_\epsilon^2 \\
\mu_{c^g_{prior}} &= 0 \\
\sigma^2_{c^g_{prior}} &= 100\sigma_\zeta^2
\end{aligned}
$$

Here, we include σ_ϵ^2 and σ_ζ^2 respectively in the priors to simplify the expression for the posteriors. This implies that the value of the prior changes with each Gibbs sampling iteration of σ_ϵ^2 and σ_ζ^2. Nevertheless, given that these values scale also the variance of the sample estimates, the priors remain diffuse and proper as long as these parameters are not "too large".

With respect to the covariances, the priors on their distributions are intuitively set to have an expected variance of 1 (and covariances of 0), with few observations, assumed to have informed this prior. Specifically and following the literature, we use:[6]

$$
\begin{aligned}
\tau_{\xi^g_{prior}} &= 3 \\
df_{\xi^g_{prior}} &= 3 \\
\tau_{\epsilon_{prior}} &= 100 \\
df_{\epsilon_{prior}} &= 100 \\
\tau_{\zeta_{prior}} &= 3 \\
df_{\zeta_{prior}} &= 3
\end{aligned}
$$

[6] Only for the case and the Municipality of Padre Hurtado do we alter these priors, placing more weight on that of ϵ. We set $df_{\epsilon_{prior}} = 500$. Given that Municipality has 25 programs, that implies weighting the prior at the equivalent of 20 applicants relative to the 502 in the estimation sample (i.e. 4%). Considering the larger than average results for σ_ϵ^2 obtained with the other models, we further increase the prior's variance, doubling the value of $\tau_{\epsilon_{prior}}$ relative to ${}_{\epsilon_{prior}}$, setting it at 1,000.

Posteriors

For a linear model $y = x\beta + \epsilon$ where $\epsilon_i \sim N(0, \sigma^2_\epsilon)$, the estimated parameters $\hat{\beta}$ have a multivariate normal distribution with mean and covariance given by:

$$\hat{\mu}_\beta = (x'x)^{-1}x'y$$
$$\hat{\Sigma}_\beta = \sigma^2_\epsilon (x'x)^{-1}$$

Given that, if we have a prior normal distribution on the β values, say

$$\beta_{prior} \sim N(\mu_{\beta_{prior}}, \Sigma_{\beta_{prior}}),$$

the posterior distribution of β given a sample $\{x, y\}$ is given by a multivariate normal distribution with mean and covariances given by:

$$\begin{aligned}
\Sigma_{\beta_{posterior}} &= \left(\Sigma^{-1}_{\beta_{prior}} + \frac{(x'x)}{\sigma^2_\epsilon} \right)^{-1} \\
\mu_{\beta_{posterior}} &= \Sigma_{\beta_{posterior}} \left(\Sigma^{-1}_{\beta_{prior}} \mu_{\beta_{prior}} + \frac{(x'x)}{\sigma^2_\epsilon} (x'x)^{-1} x'y \right) \\
&= \Sigma_{\beta_{posterior}} \left(\Sigma^{-1}_{\beta_{prior}} \mu_{\beta_{prior}} + \frac{(x'y)}{\sigma^2_\epsilon} \right)
\end{aligned}$$

For the Inverse-Wishart distribution on the other hand, if we have a prior

$$\sigma^2_{x_{prior}} \sim IW(\tau_{prior}, df_{prior}),$$

and N observations of x with mean $\overline{\mu}_x$, we have that:

$$\sigma^2_{x_{posterior}} \sim IW\left(\tau_{prior} + \sum_{i=1}^{N} (x_i - \overline{\mu}_x)^2, df_{prior} + N \right)$$

Therefore, the posterior distributions in our case are given by[7]:

$$\begin{aligned}
R_\lambda &\equiv v_{ijt} - x_{jt}\beta^{g_i} - \xi^{g_i}_j + (1 + \Delta_{d,C19} \mathbf{1}_{t>2019}) d_{ijt} \\
\Sigma_{\lambda_{posterior}} &= \left(\frac{1}{100\sigma^2_\epsilon} I_2 + \frac{(Family'Family)}{\sigma^2_\epsilon} \right)^{-1}
\end{aligned}$$

[7]For all parameters that depend on the socioeconomic group, the posterior is computed only using the observations that correspond to the group. That is, for example, $\beta^{g=SeS}$ includes only those applicants with low-SeS status.

$$
\begin{aligned}
&= \left(\frac{1}{100} I_2 + (Family'Family)\right)^{-1} \sigma_\epsilon^2 \\
\mu_{\lambda_{posterior}} &= \Sigma_{\lambda_{posterior}} \frac{(Family' R_\lambda)}{\sigma_\epsilon^2} \\
&= \left(\frac{1}{100} I_2 + (Family'Family)\right)^{-1} (Family' R_\lambda),
\end{aligned}
$$

$$
\begin{aligned}
R_{\beta^g} &\equiv v_{ijt} - Family_{ijt}\lambda - \xi_j^{g_i} + (1 + \Delta_{d,C19}\mathbf{1}_{t>2019}) d_{ijt} \\
\Sigma_{\beta^g{}_{posterior}} &= \left(\frac{1}{100} I_K + (x'x)\right)^{-1} \sigma_\epsilon^2 \\
\mu_{\beta^g{}_{posterior}} &= \left(\frac{1}{100} I_K + (x'x)\right)^{-1} (x' R_{\beta^g}),
\end{aligned}
$$

For i's in group g that have program j in their choice set $\mathcal{P}_{it}$ in year t:

$$
\begin{aligned}
R_{\xi_j^g} &\equiv v_{ijt} - Family_{ijt}\lambda - x_{jt}\beta^{g_i} + (1 + \Delta_{d,C19}\mathbf{1}_{t>2019}) d_{ijt} \\
\sigma^2_{\xi^g_{j,posterior}} &= \left(\frac{1}{\sigma^2_{\xi^g}} + \frac{\sum_{i=1}^N \mathbf{1}_{j\in\mathcal{P}_{it}, i\in g}}{\sigma_\epsilon^2}\right)^{-1} \\
\mu_{\xi^g_{j,posterior}} &= \sigma^2_{\xi^g_{j,posterior}} \frac{\sum_{i=1}^N R_{\xi_j^g} \mathbf{1}_{j\in\mathcal{P}_{it}, i\in g}}{\sigma_\epsilon^2}
\end{aligned}
$$

$$
\begin{aligned}
R_{\Delta_{d,C19}} &\equiv v_{ijt} - Family_{ijt}\lambda - x_{jt}\beta^{g_i} - \xi_j^{g_i} + d_{ijt} \\
\sigma_{\Delta_{d,C19}{}_{posterior}} &= \left(\frac{1}{100} + ((\mathbf{1}_{t>2019} d_{ijt})' \mathbf{1}_{t>2019} d_{ijt})\right)^{-1} \sigma_\epsilon^2 \\
\mu_{\beta^g{}_{posterior}} &= \left(\frac{1}{100} + ((\mathbf{1}_{t>2019} d_{ijt})' \mathbf{1}_{t>2019} d_{ijt})\right)^{-1} ((\mathbf{1}_{t>2019} d_{ijt})' R_{\Delta_{d,C19}}),
\end{aligned}
$$

$$
\begin{aligned}
R_{c^g} &\equiv c_i \mathbf{1}_{i\in g} \\
\sigma^2_{c^g{}_{posterior}} &= \left(\frac{1}{100} + N_g\right)^{-1} \sigma_\zeta^2
\end{aligned}
$$

$$\mu_{c^g{}_{posterior}} = \left(\frac{1}{100} + N_g\right)^{-1} (\mathbf{1}' R_{c^g}),$$

Posteriors for ϵ_{ijt} and ζ_i are then drawn from the prior distribution of those random coefficients, but using the bounds implied by result 2, which vary for each applicant and program in the case of ϵ_{ijt} and for each applicant in the case of ζ_i. Specifically:

$$\begin{aligned}\epsilon_{ijt} &\sim TN(0, \sigma^2_\epsilon, LB_{\epsilon_{ijt}}, UB_{\epsilon_{ijt}})\\ \zeta_i &\sim TN(0, \sigma^2_\zeta, -c^{g_i}, UB_{\zeta_i}).\end{aligned}$$

The posterior distribution of the covariances in turn is given by:

$$\begin{aligned}\sigma^2_{\xi^g_{posterior}} &\sim IW\left(3 + \sum_{j=1}^{J} \xi_j^{g\,2}, 3 + J\right)\\ \sigma^2_{\epsilon_{posterior}} &\sim IW\left(100 + \sum_{i=1}^{N}\sum_{j=1}^{J} \epsilon^2_{ijt}, 100 + N \times J\right)\\ \sigma^2_{\zeta_{posterior}} &\sim IW\left(3 + \sum_{i=1}^{N} (c_i - c^{g_i})^2, 3 + N\right)\end{aligned}$$

Gibbs Sampler Iterations

Gibbs Sampler Procedure in Costly Consideration Model

Initiate with

$$\begin{aligned}v_{ijt,0} &= |ROL_i| - r_{ij} && \forall j \in ROL_i\\ v_{ijt,0} &= -1 && \forall j \notin ROL_i,\end{aligned}$$

and $\lambda_0 = 0$, $\beta_0^{g_i} = 0$, $\Delta_{d,C19,0} = 0$, $\xi_{j,0} = 0$.

Sample covariances from the prior distributions:

$$\begin{aligned}\sigma^2_{\xi^g,0} &\sim IW(\tau_{\xi^g,prior}, df_{\xi^g,prior})\\ \sigma^2_{\epsilon,0} &\sim IW(\tau_{\epsilon,prior}, df_{\epsilon,prior})\\ \sigma^2_{\zeta,0} &\sim IW(\tau_{\zeta,prior}, df_{\zeta,prior})\end{aligned}$$

And sample $\zeta_{i,0} \sim TN(0, \sigma^2_{\zeta,0}, -c^g_0, UB_{\zeta_i,0})$ to get $c_{i,0} = c^g_0 + \zeta_{i,0}$. We set $c^g_0 = 0$ and $UB_{\zeta_i,0} = \infty$.

Then, iterate through the following steps:

1. Sample λ_1 from $N(\mu_{\lambda_{posterior}}, \Sigma_{\lambda_{posterior}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, $\beta_0^{g_i}$, $\Delta_{d,C19,0}$ and $\xi_{j,0}$.
2. Sample β_1^g from $N(\mu_{\beta^g_{posterior}}, \Sigma_{\beta^g_{posterior}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, $\Delta_{d,C19,0}$ and $\xi_{j,0}$.
3. Sample $\xi^g_{j,1}$ from $N(\mu_{\xi^g_{j,posterior}}, \Sigma_{\xi^g_{j,posterior}})$, given $\sigma^2_{\xi^g,0}$, $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β_1^g, $\Delta_{d,C19,0}$.
4. Sample $\Delta_{d,C19,1}$ from $N(\mu_{\Delta_{d,C19_{posterior}}}, \sigma^2_{\Delta_{d,C19_{posterior}}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β_1^g and $\xi_{j,1}$.
5. Sample covariances for:

$$\begin{aligned}\sigma^2_{\xi^g,1} &\sim IW(\tau_{\xi^g,posterior}, df_{\xi^g,posterior})\\ \sigma^2_{\epsilon,1} &\sim IW(\tau_{\epsilon,posterior}, df_{\epsilon,posterior})\end{aligned}$$

6. Now we proceed with the consideration process, identifying the set of considered programs of each applicant and the indirect utilities of the different programs. The process is as follows:

   a) Sample the utilities of programs j where the applicant has family priorities. If $j \notin ROL_i \Rightarrow v_{ijt,1} < 0$. For $j \in ROL_i$, we sample first the utility of the program with the highest ranking using the utility bounds implied by the previous iteration and so on, as we go on updating sampled applicant utilities.

      After this step, some applicants will have a non-empty partial ranking and some will have an empty one.

   b) For $k \in \{1, ..., max_i(ROL_i) + 1\}$:

      i. Include all applicants with $|\mathcal{R}_i| = k - 1$
      ii. Compute $CV_{j,\mathcal{R}_{i,\{j_{cr<k}\}}}$ for all j not yet considered (including those in the ROL_i).
      iii. If $k < |ROL_i| + 1$, identify the program $l \in ROL_i \setminus \mathcal{R}_i$ with the largest $CV_{j,\mathcal{R}_{i,\{j_{cr<k}\}}}$ and:
         A. Include l in $\mathcal{R}_i$, and sample its indirect utility according to the ranking order determined by ROL_i.
         B. Include all programs j not yet considered with $CV_{j,\mathcal{R}_{i,\{j_{cr<k}\}}} > CV_{l,\mathcal{R}_{i,\{j_{cr<k}\}}}$ in the set of considered programs and sample their indirect utility using the bound of unacceptable programs, that is $v_{ijt,1} < 0$
      iv. If $k = |ROL_i| + 1$:
         A. If $CV_{j,ROL_i} > c_{i,0} \Rightarrow v_{ijt,1} < 0$
         B. If $CV_{j,ROL_i} < c_{i,0} \Rightarrow v_{ijt,1} \in \mathbb{R}$

   Store the consideration ranking of programs considered in this as $cr_{ij,1} = k$.

7. Sample c_1^g from $TN(\mu_{c^g_{posterior}}, \sigma^2_{c^g_{posterior}}, 0, \infty)$ given $\sigma^2_{\zeta,0}$ and $c_{i,0}$

8. Sample $\sigma^2_{\zeta,1}$ from $IW(\tau_{\zeta,posterior}, df_{\zeta,posterior})$ given $c_{i,0}$ and c_1^g.

9. For each applicant i, identify the smallest empirical consideration value that bounds the largest possible value of c_i. Here, we use the term empirical to reflect that we only consider the restrictions implied by the observed choice data and not those implied by other programs sampled for consideration in step 6 above. Given that, the bound is given by:
$$CVB_{\zeta_i,1} \equiv min_{j \in ROL_i \cap \{j:\ Family_{ijt}=0\}} \left\{ CV_{j,\mathcal{R}_{i,\{jcr<cr_{ij,1}\}}} - c_1^{g_i} \right\}$$

10. Sample $\zeta_{i,1}$ (and thus $c_{i,1}$) from $TN(0, \sigma^2_{\zeta,1}, -c^{g_i}, CVB_{\zeta_i,1})$ given $v_{ijt,1}$, q_{ijt}, σ^2_ϵ, $\sigma^2_{\xi^g}$, λ_1, β_1^g, $\Delta_{d,C19,1}$, $\xi^g_{j,1}$, c^g and $CVB_{\zeta_i,1}$

Gibbs Sampler Procedure in Alternative Specifications

Full consideration

Note: We include random coefficients here to give the reader an example of how these are incorporated into the procedure, using the same notation of Abdulkadiroğlu et al. (2017). However, these coefficients are **not included** in the estimation.

Initiate with

$$\begin{aligned} v_{ijt,0} &= |ROL_i| - r_{ij} && \forall j \in ROL_i \\ v_{ijt,0} &= -1 && \forall j \notin ROL_i, \end{aligned}$$

and $\lambda_0 = 0$, $\beta_0^{g_i} = 0$, $\Delta_{d,C19,0} = 0$, $\gamma_{i,0} = 0$, $\xi_{j,0} = 0$.

Sample covariances from the prior distributions:

$$\begin{aligned} \Sigma_{\gamma,0} &\sim IW(\tau_{\gamma,prior}, df_{\gamma,prior}) \\ \sigma^2_{\xi^g,0} &\sim IW(\tau_{\xi^g,prior}, df_{\xi^g,prior}) \\ \sigma^2_{\epsilon,0} &\sim IW(\tau_{\epsilon,prior}, df_{\epsilon,prior}) \end{aligned}$$

Then, iterate through the following steps:

1. Sample λ_1 from $N(\mu_{\lambda_{posterior}}, \Sigma_{\lambda_{posterior}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, $\beta_0^{g_i}$, $\Delta_{d,C19,0}$, $\gamma_{i,0}$ and $\xi_{j,0}$.

2. Sample β_1^g from $N(\mu_{\beta^g_{posterior}}, \Sigma_{\beta^g_{posterior}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, $\Delta_{d,C19,0}$, $\gamma_{i,0}$ and $\xi_{j,0}$.

3. Sample $\xi^g_{j,1}$ from $N(\mu_{\xi^g_{j,posterior}}, \Sigma_{\xi^g_{j,posterior}})$, given $\sigma^2_{\xi^g,0}$, $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β_1^g, $\Delta_{d,C19,0}$ and $\gamma_{i,0}$.

4. Sample $\Delta_{d,C19,1}$ from $N(\mu_{\Delta_{d,C19_{posterior}}}, \sigma^2_{\Delta_{d,C19_{posterior}}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β^g_1, $\gamma_{i,0}$ and $\xi_{j,1}$.

5. Sample $\gamma_{i,1}$ from $N(\mu_{\gamma_{i,posterior}}, \Sigma_{\gamma_{i,posterior}})$, given $\Sigma_{\gamma,0}$, $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β^g_1, $\Delta_{d,C19,1}$ and ξ^g_1.

6. Sample covariances for:

$$\begin{aligned}\Sigma_{\gamma,1} &\sim IW(\tau_{\gamma,posterior}, df_{\gamma,posterior})\\ \sigma^2_{\xi^g,1} &\sim IW(\tau_{\xi^g,posterior}, df_{\xi^g,posterior})\\ \sigma^2_{\epsilon,1} &\sim IW(\tau_{\epsilon,posterior}, df_{\epsilon,posterior})\end{aligned}$$

7. Sample $v_{ijt,1}$ for ranked alternatives using the bounds implied by their ranking position, that is: $\infty > v_{ij_{r=1}} > ... > v_{ij_{r=|ROL_i|}} > 0$.

   For programs not included in the ranking, sample their utility in $\mathbb{R}^-$.

Costly ranking (as in Idoux (2022))

Initiate with

$$\begin{aligned}v_{ijt,0} &= |ROL_i| - r_{ij} && \forall j \in ROL_i\\ v_{ijt,0} &= -1 && \forall j \notin ROL_i,\end{aligned}$$

and $\lambda_0 = 0$, $\beta^{g_i}_0 = 0$, $\Delta_{d,C19,0} = 0$, $\xi_{j,0} = 0$.

Sample covariances from the prior distributions:

$$\begin{aligned}\sigma^2_{\xi^g,0} &\sim IW(\tau_{\xi^g,prior}, df_{\xi^g,prior})\\ \sigma^2_{\epsilon,0} &\sim IW(\tau_{\epsilon,prior}, df_{\epsilon,prior})\\ \sigma^2_{\zeta,0} &\sim IW(\tau_{\zeta,prior}, df_{\zeta,prior})\end{aligned}$$

And sample $\zeta_{i,0} \sim TN(0, \sigma^2_{\zeta,0}, LB_{i,0}, UB_{\zeta_i,0})$ to get $c_{i,0} = c^g_0 + \zeta_{i,0}$. We set $c^g_0 = 0$, $LB_{i,0} = 0$, and $UB_{\zeta_i,0} = \infty$.

Then, iterate through the following steps:

1. Sample λ_1 from $N(\mu_{\lambda_{posterior}}, \Sigma_{\lambda_{posterior}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, $\beta^{g_i}_0$, $\Delta_{d,C19,0}$ and $\xi_{j,0}$.

2. Sample β^g_1 from $N(\mu_{\beta^g_{posterior}}, \Sigma_{\beta^g_{posterior}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, $\Delta_{d,C19,0}$ and $\xi_{j,0}$.

3. Sample $\xi^g_{j,1}$ from $N(\mu_{\xi^g_{j,posterior}}, \Sigma_{\xi^g_{j,posterior}})$, given $\sigma^2_{\xi^g,0}$, $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β^g_1 and $\Delta_{d,C19,0}$.

4. Sample $\Delta_{d,C19,1}$ from $N(\mu_{\Delta_{d,C19_{posterior}}}, \sigma^2_{\Delta_{d,C19_{posterior}}})$, given $\sigma^2_{\epsilon,0}$, $v_{ijt,0}$, λ_1, β^g_1 and $\xi_{j,1}$.

5. Sample covariances for:

$$\sigma^2_{\xi^g,1} \sim IW(\tau_{\xi^g,posterior}, df_{\xi^g,posterior})$$
$$\sigma^2_{\epsilon,1} \sim IW(\tau_{\epsilon,posterior}, df_{\epsilon,posterior})$$

6. For programs in ROL_i, and starting with the programs ranked in first position[8]:

   a) Compute the utility bounds, identifying the more binding constraint between:

      i. Illustrating with a program ranked in third position:

$$\infty > v_{ik_{r=1}t,1} > v_{ik_{r=2}t,1} > v_{ijt,1} > v_{ik_{r=4}t,0} > \ldots > v_{ik_{r=|ROL_i|}t,0} > 0$$

      ii. $v_{ij_{r=\hat{r}}t,1} \geq \frac{c^{g_i}_{i,0}}{q_{ijt}\prod_{l=1}^{\hat{r}-1}\{1-q_{ik_{r=l}t}\}}$

      iii. For program j ranked in position $\hat{r}$ within ROL_i, and $k \in \{\mathcal{P}_i \setminus ROL_i\}$:

$$v_{ikt,0} \geq \frac{c^{g_i}_{i,0}}{q_{ikt}\prod_{l=1}^{\hat{r}-1}\{1-q_{im_{r=l}t}\}} \Rightarrow v_{ijt,1} > v_{ikt,0}$$

   b) Sample utilities from a truncated normal distribution using the bounds identified in the previous step.

7. For programs $k \in \{\mathcal{P}_i \setminus ROL_i\}$:

   a) Compute the utility bounds, identifying the more binding constraint between:

      i. $v_{ikt,1} < \frac{c^{g_i}_{i,0}}{q_{ikt}\prod_{l=1}^{|ROL_i|}\{1-q_{im_{r=l}t}\}}$

      ii. $\forall j \in ROL_i$ ranked in position $\hat{r}$:

$$v_{ikt,1} < max\left\{v_{ijt,1}, \frac{c^{g_i}_{i,0}}{q_{ikt}\prod_{l=1}^{\hat{r}-1}\{1-q_{im_{r=l}t}\}}\right\}$$

   b) Sample utilities from a truncated normal distribution using the bounds identified in the previous step

8. Sample c^g_1 from $TN(\mu_{c^g_{posterior}}, \sigma^2_{c^g_{posterior}}, 0, \infty)$ given $\sigma^2_{\zeta,0}$ and $c_{i,0}$

9. Sample $\sigma^2_{\zeta,1}$ from $IW(\tau_{\zeta,posterior}, df_{\zeta,posterior})$ given $c_{i,0}$ and c^g_1.

[8] Note that starting the Gibbs sampler with $c^{g_i}_{i,0} = 0$ guarantees that the bounds described here are always mutually consistent, given that the process always fits the ranking cost after utilities are updated. For example, if for a ranked program j we have that $q_{ijt} = 0$, then the ranking cost will remain equal to zero throughout the whole Gibbs sampling process.

10. Sample $\zeta_{i,1}$ (and thus $c_{i,1}$) from $TN(0, \sigma^2_{\zeta,1}, LB_{\zeta_i,1}, UB_{\zeta_i,1})$. Here, $LB_{\zeta_i,1}$ and $UB_{\zeta_i,1}$ are the more restrictive bounds obtained from the following inequalities:

    a) $\forall j \in ROL_i : c^{g_i}_{i,1} < v_{ij_{r=\hat{r}}t,1} q_{ijt} \prod_{l=1}^{\hat{r}-1} \{1 - q_{ik_{r=l}t}\}$

    b) $\forall k \in \{\mathcal{P}_i \setminus ROL_i\} : c^{g_i}_{i,1} > v_{ikt,1} q_{ikt} \prod_{l=1}^{|ROL_i|} \{1 - q_{im_{r=l}t}\}$

    c) For program j ranked in position $\hat{r}$ within ROL_i, and $k \in \{\mathcal{P}_i \setminus ROL_i\}$:

$$v_{ikt,1} \geq v_{ijt,1} \Rightarrow c^{g_i}_{i,1} > v_{ikt,1} q_{ikt} \prod_{l=1}^{\hat{r}-1} \{1 - q_{im_{r=l}t}\}$$

given $v_{ijt,1}$, q_{ijt}, σ^2_ϵ, $\sigma^2_{\xi^g}$, λ_1, β^g_1, $\Delta_{d,C19,1}$, $\xi^g_{j,1}$ and c^g.

A.7 Estimation Results Complement

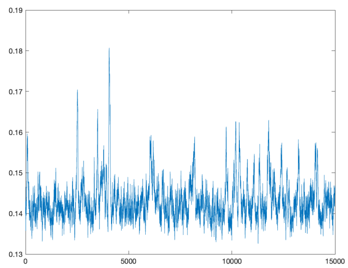
PSRF and traceplots

Table A.7.1: PSRF Range for Main Models Parameters

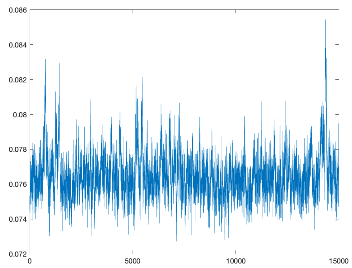
	Costly consideration		Costly ranking		Full consideration	
	min	max	min	max	min	max
λ_{SIB}	1	1.009	1	1.005	1	1.001
λ_{PW}	1	1.052	1	1.261	1	1.009
$\Delta_{d,C19}$	1	1	1	1	1	1
$\sigma^2_{\xi^{NSeS}}$	1.001	1.094	1	1.017	1	1.021
$\sigma^2_{\xi^{SeS}}$	1	1.066	1	1.003	1	1.027
σ^2_ϵ	1	1.014	1	1.013	1	1.013
c^{NSeS}	1	1.095	1	1.029		
c^{SeS}	1	1.049	1	1		
σ^2_{zeta}	1	1.08	1	1.01		

Note: In the costly ranking and full consideration models, the PSRF range is calculated only for the larger Municipalities of Maipú, Santiago, San Bernardo, La Pintana, La Florida, and Puente Alto. This is because, for computational efficiency, the other Municipality parameters were estimated with only one Gibbs Sampler chain, which reduced the burn-in period's computer time. Note that this approach was not used for the costly consideration model. In the next iteration of this analysis, we plan to include all Municipalities in the analysis.

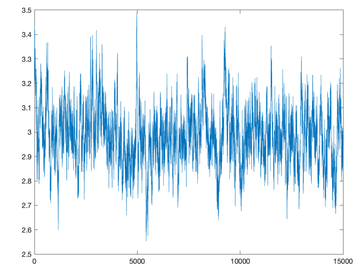
Figure A.7.1: Trace plots for σ_{ϵ}^{2}



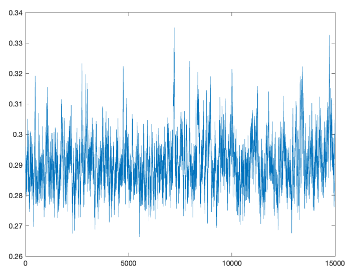
Cerrillos



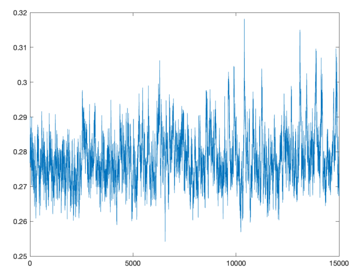
Cerro Navia



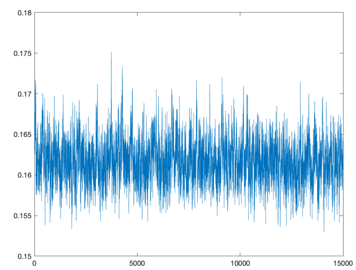
Colina



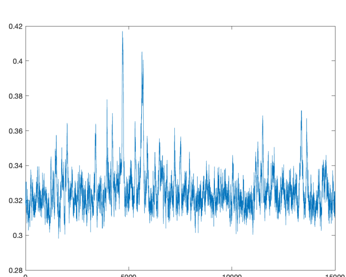
Conchali



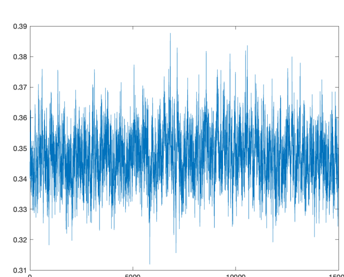
El Bosque



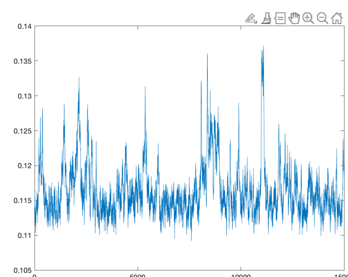
Estación Central



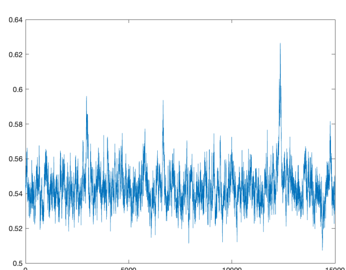
Huechuraba



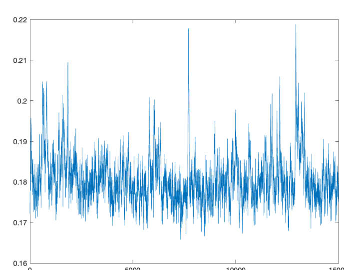
Independencia



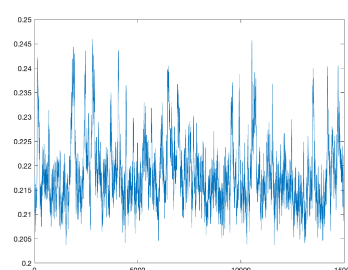
La Cisterna



La Florida

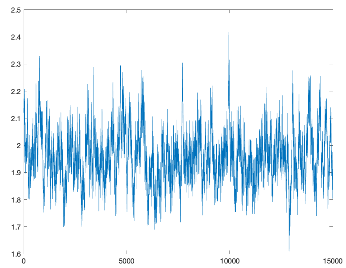


La Granja

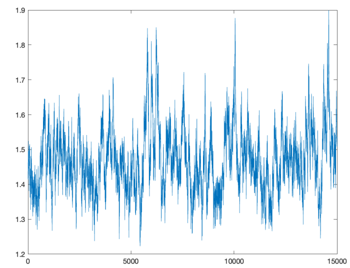


La Pintana

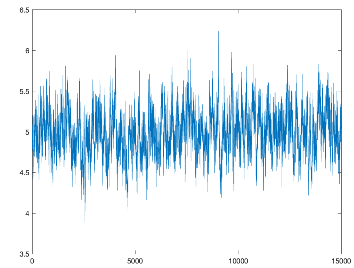
Trace plots for σ_ϵ^2



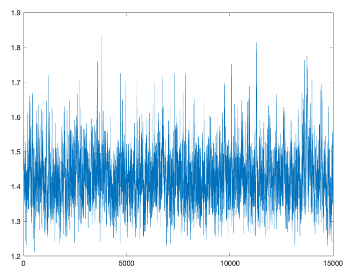
La Reina



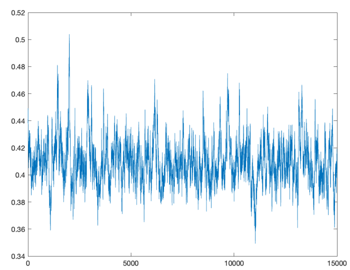
Lampa



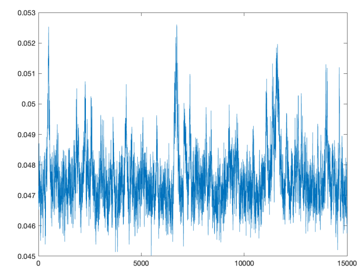
Las Condes



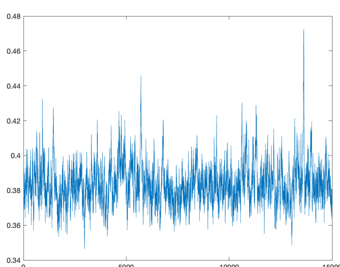
Lo Barnechea



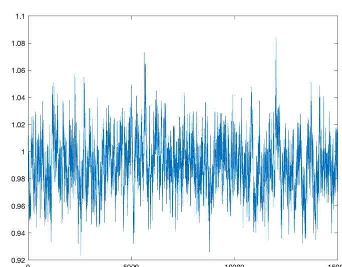
Lo Espejo



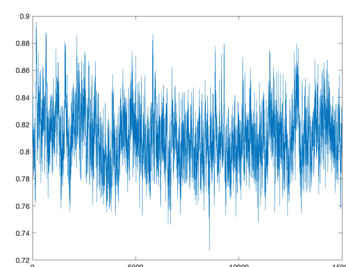
Lo Prado



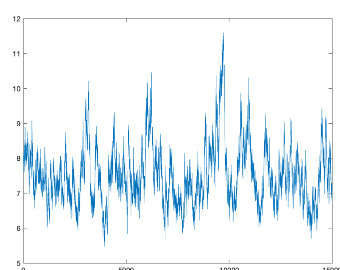
Macul



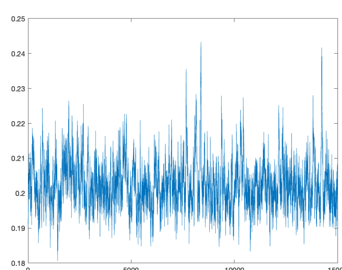
Maipú



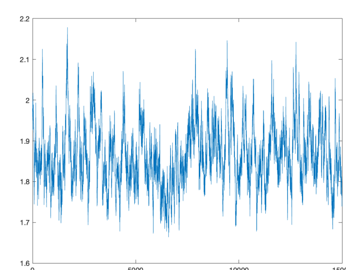
Ñuñoa



Padre Hurtado

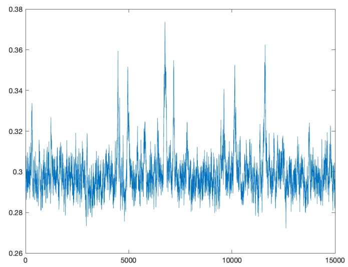


Pedro Aguirre Cerda

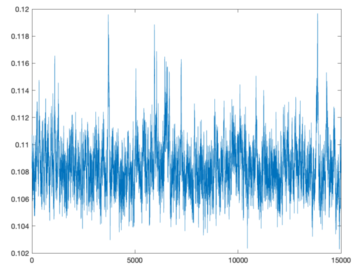


Peñalolén

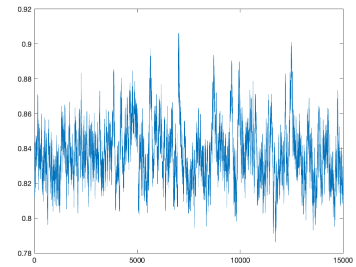
Trace plots for σ_ϵ^2



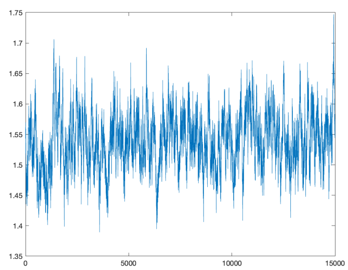
Providencia



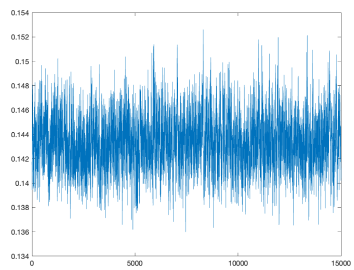
Pudahuel



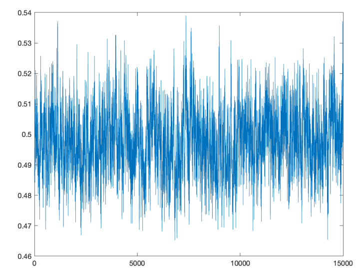
Puente Alto



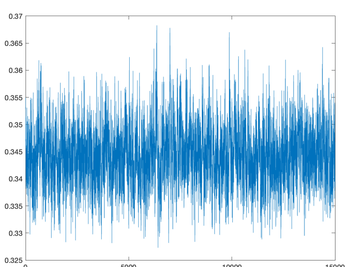
Quilicura



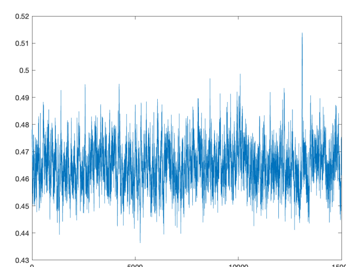
Quinta Normal



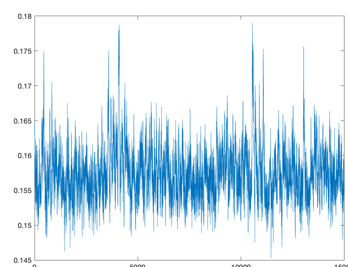
Recoleta



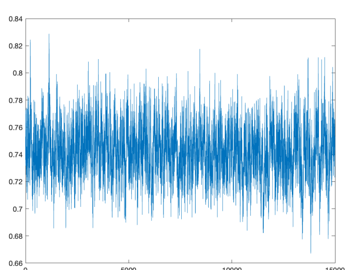
Renca



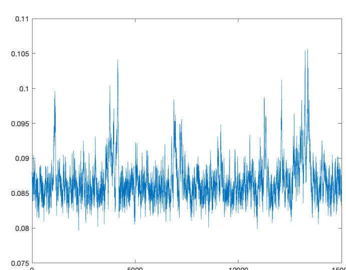
San Bernardo



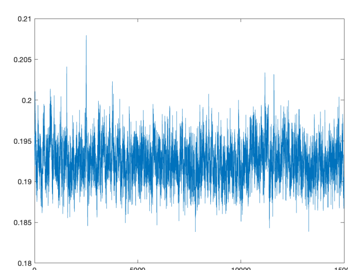
San Joaquín



San Miguel

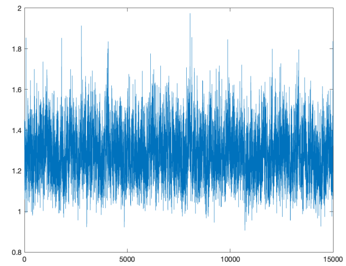


San Ramón



Santiago

Trace plots for σ_ϵ^2



Vitacura

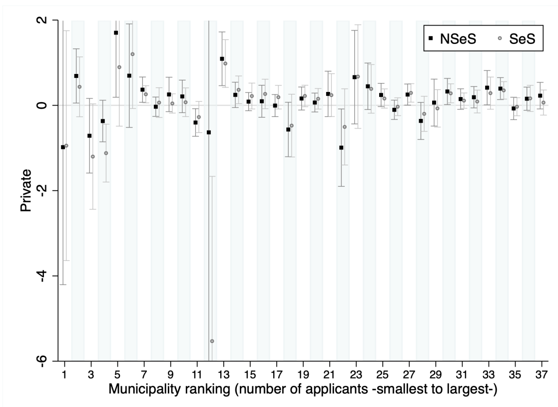
Additional Model Estimates

Table A.7.2: Summary of Estimated Parameters over Program Attributes

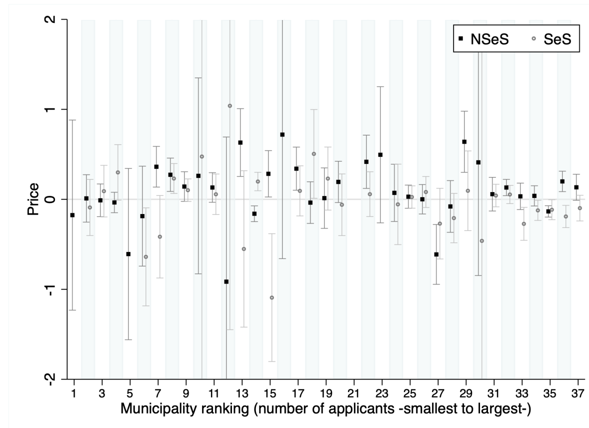
	Costly consideration		Costly ranking		Full consideration	
	wav	range	wav	range	wav	range
$\beta^{NSeS}_{Private}$	.143	-.992 - 1.703	.116	-2.316 - 2.486	.132	-1.534 - 1.822
	(1.2)	(-2.44 - 3.4)	(1.2)	(-2.82 - 3.89)	(1.2)	(-2.78 - 5.19)
$\beta^{SeS}_{Private}$	.051	-5.529 - 1.201	.042	-6.335 - 1.36	.045	-5.385 - 1.298
	(1.1)	(-3.24 - 3.43)	(1.1)	(-3.72 - 3.59)	(1.1)	(-3.15 - 3.35)
$\beta^{NSeS}_{HA\ Cat.}$	.766	-.48 - 2.509	.844	-.551 - 2.958	.751	-.446 - 2.554
	(3)	(-1.57 - 6.61)	(3.2)	(-1.49 - 7.23)	(3.1)	(-1.41 - 6.92)
$\beta^{SeS}_{HA\ Cat.}$	.664	-1.184 - 2.178	.708	-.767 - 2.373	.64	-1.012 - 2.294
	(2.9)	(-1.47 - 6.89)	(2.9)	(-2.08 - 6.99)	(2.8)	(-1.35 - 6.4)
$\beta^{NSeS}_{MA\ Cat.}$	.369	-.656 - 1.189	.418	-.398 - 1.649	.376	-.369 - 1.19
	(2.5)	(-1.57 - 4.98)	(2.5)	(-.95 - 5.35)	(2.5)	(-1.29 - 5.34)
$\beta^{SeS}_{MA\ Cat.}$	.391	-.54 - 2.703	.419	-.736 - 4.045	.368	-.569 - 2.537
	(2.4)	(-1.98 - 5.14)	(2.5)	(-2.35 - 6.69)	(2.5)	(-1.87 - 7.57)
β^{NSeS}_{Price}	.105	-.915 - .719	.214	-.582 - 1.453	.129	-.529 - .641
	(1.3)	(-4.17 - 3.69)	(2)	(-3.61 - 4.69)	(1.5)	(-4.35 - 3.63)
β^{SeS}_{Price}	-.073	-1.092 - 1.039	-.042	-1.257 - 1.839	-.085	-1.131 - 1.035
	(1.1)	(-3.03 - 3.78)	(1)	(-2.9 - 3.83)	(1.1)	(-3.92 - 3.7)

Note: weighted av. **t-stats in parenthesis** (if opposite sign from parameter set to zero)

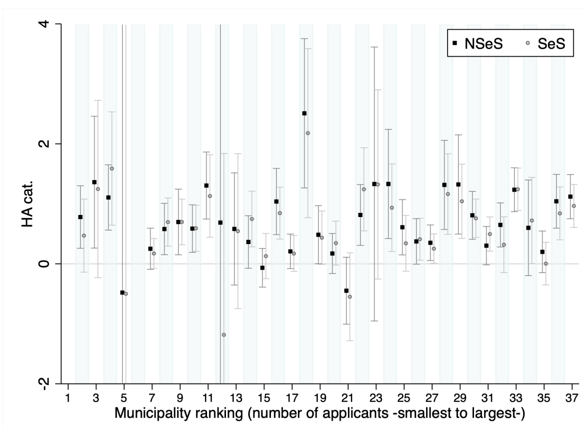
Figure A.7.5: Municipality level estimated parameter selection in costly consideration model



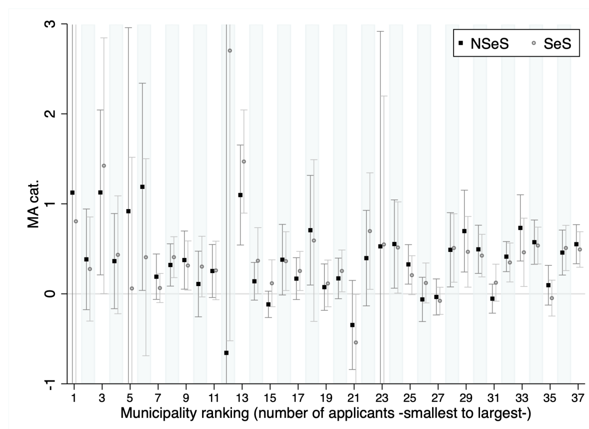
(a) Subsidized private school



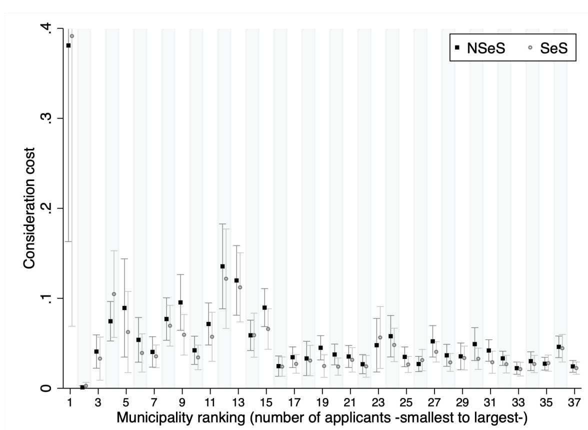
(b) Price (U.F.)



(c) HA cat.

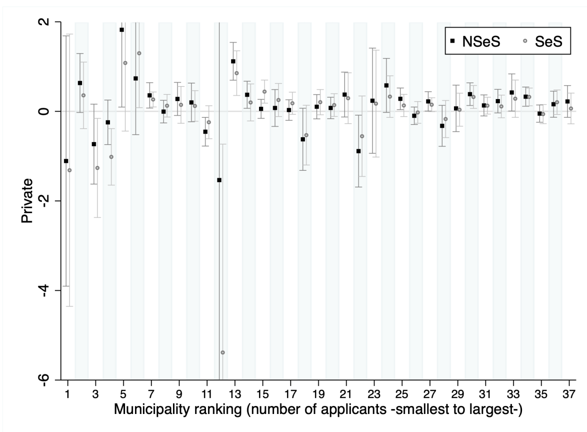


(d) MA cat.

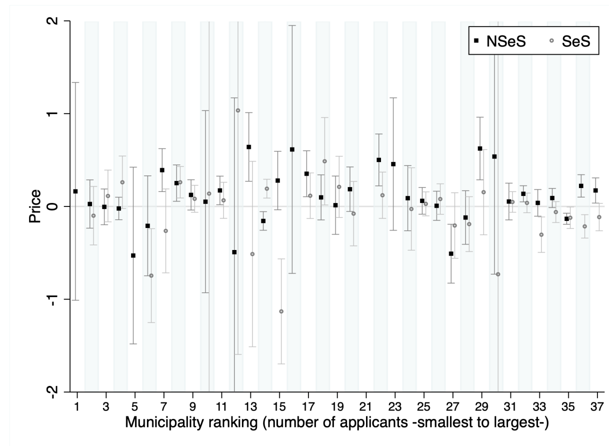


(e) Consideration cost

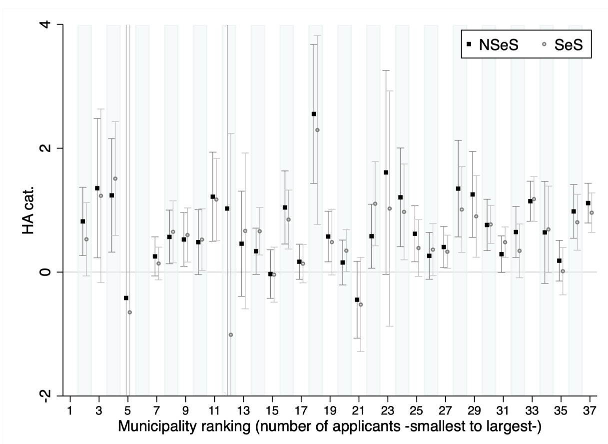
Figure A.7.6: Municipality level estimated parameter selection in full consideration model



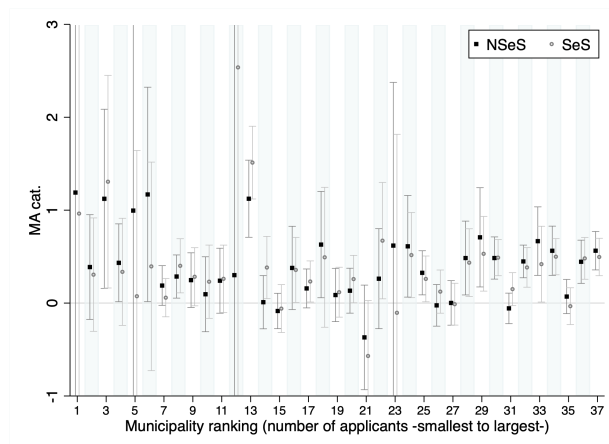
(a) Subsidized private school



(b) Price (U.F.)

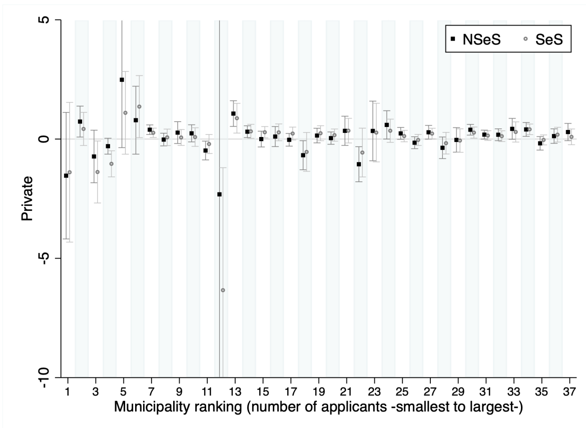


(c) HA cat.

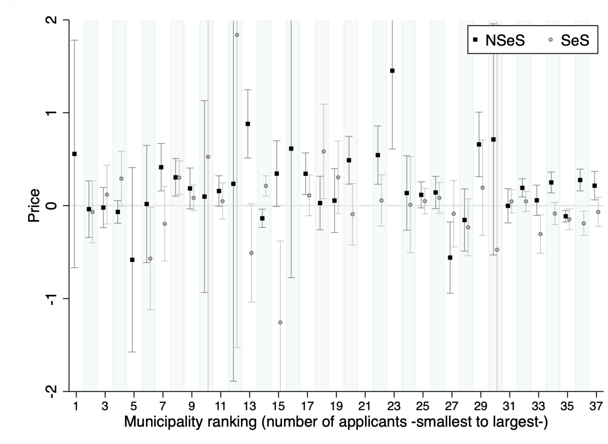


(d) MA cat.

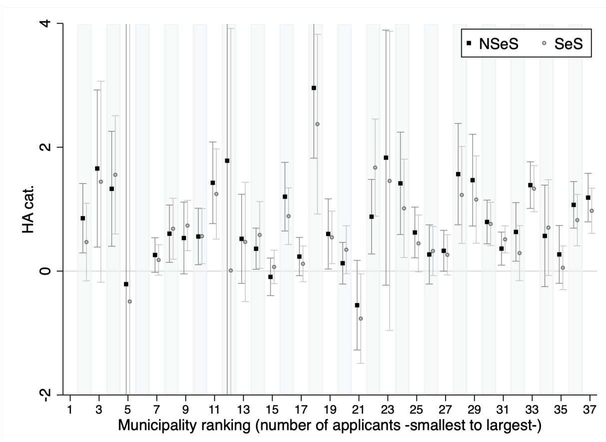
Figure A.7.7: Municipality level estimated parameter selection in costly ranking model



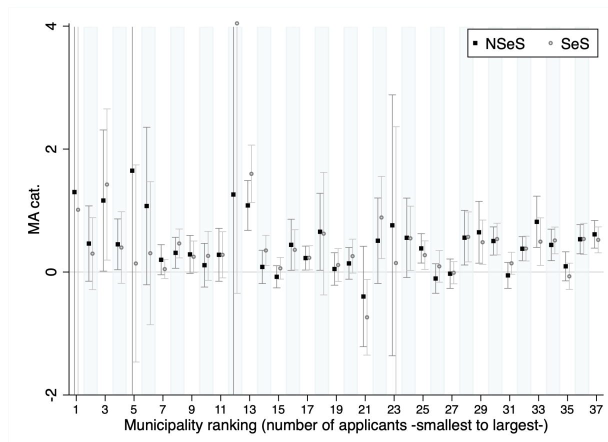
(a) Subsidized private school



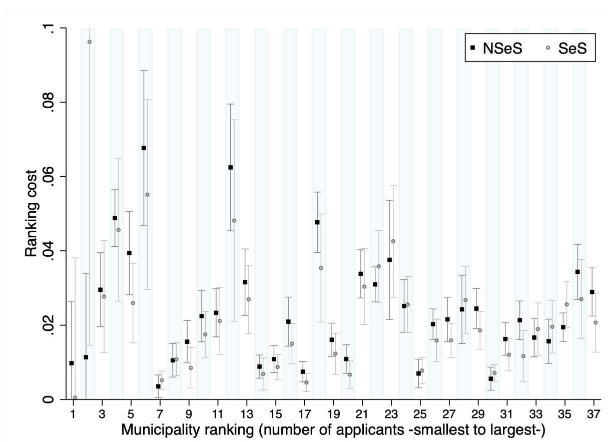
(b) Price (U.F.)



(c) HA cat.



(d) MA cat.



(e) Ranking cost

A.8 Counterfactual simulations

The counterfactual simulation process is relatively straightforward, but there are some details, particularly in the simulation of random coefficients that are important to understand. First, as in Idoux (2022), in our simulations, we generate random coefficients consistent with ROLs. This is done in the three models considered, but naturally only for the baseline case where the socioeconomic quota is at the 15% level in place when applications were submitted and without any changes in applicant locations or preferences. This implies that in the baseline case, our applications are by construction exactly identical to those actually reported. These simulated random coefficients consistent with the baseline scenario are then used to generate counterfactual applications when changes in assignment probabilities, preferences, or distance to school are introduced. Only in the case of the minimum benchmark segregation, where we simulate with equal preferences do we introduce a new and fully random preference shock (ϵ_{ijt} and ζ_{it}) for each applicant-program combination. The idea is to remove all influence from preference and model specification for this benchmark.

There are specifically two details that are important to keep in mind when reviewing our procedure: First, for ranked programs, the order in which random coefficients are drawn matters. The reason is that indirect utilities of other ranked programs affect the bounds of the draws. Using the simpler full consideration case to exemplify, the program (in the ROL) drawn first will have an ϵ_{ij} consistent with an indirect utility bounded between infinity and zero. If that program was ranked first, the value obtained will imply an upper bound for all other ranked programs. If it was ranked last, it will imply a lower bound. To tackle this issue, we randomize the order in which utilities for ranked programs are drawn. Second, for the consideration cost and the ranking cost models, sampled utilities will bound the cost values from above, and those costs will then bound the indirect utilities of the programs not included in ROLs.

Counterfactual procedure (costly consideration)

To simplify, in the description of the process, we omit programs surely considered. They are included in the same way as in step 6 of the Gibbs sampler described in Appendix A.6.

1. Obtain indirect utility and consideration cost shocks consistent with ROLs under the baseline.

   a) Use estimated parameters for λ, β^g, ξ^g, $\Delta_{d,C19}$, c^g and variances σ^2_ϵ, σ^2_ζ

   b) Restricting now only to $j \in ROL_i$:

      i. Compute $CV_{j,\varnothing}$ using the estimated observable part of the indirect utility $(y_{ij} - (1 + \Delta_{d,C19}\mathbf{1}_{t>2019})d_{ij})$

      ii. Select the program with the highest $CV_{j,\varnothing}$ and sample ϵ_{ij} consistent with $v_{ij} > 0$

iii. Continue with the remaining programs $k \in ROL_i \setminus \{j\}$, selecting the largest $CV_{k,\{j\}}$ and so on, sampling ϵ_{ik} consistent with the relative ranking of the programs included in the corresponding partial ranking.
Note: It is important to store the consideration values obtained in this step to use them below in step 1d.

c) The smallest consideration value will be that of the program drawn last ($l_{cr=|ROL_i|}$). Using that program, compute the consideration cost shock ζ_i consistent with $c_i \in \left[0, CV_{l_{cr=|ROL_i|}, ROL_i \setminus \{l_{cr=|ROL_i|}\}}\right)$.

d) Using c_i:

i. Compute, $\forall j \notin ROL_i$ the value $CV_{j,\varnothing}$ and label as considered all programs where $CV_{j,\varnothing} > CV_{k_{cr=1},\varnothing}$.

ii. Follow the same logic for all programs not yet considered and not included in ROL_i, comparing their consideration value given the corresponding partial ranking with that of the program that was identified as included in the ROL in that step of consideration.

iii. After the last step of comparison with the consideration value of ranked programs (the step comparing all remaining programs with $CV_{k_{cr=|ROL_i|}, ROL_i \setminus \{k_{cr=|ROL_i|}\}}$), label as considered all remaining programs for which $CV_{j,ROL_i} > c_i$

e) Sample the ϵ_{ij} of programs $j \notin ROL_i$ depending on their consideration. That means that the sampled ϵ_{ij} has to be consistent with:

$$\begin{aligned} j \notin ROL_i \text{ and } j \text{ considered} &\Rightarrow v_{ij} < 0 \\ j \notin ROL_i \text{ and } j \text{ not considered} &\Rightarrow v_{ij} \in \mathbb{R} \end{aligned}$$

2. For any given counterfactual affecting preferences, distance to school, or initial rational expectation assignment probabilities, use estimated preference parameters, simulated utility and consideration cost shocks, and Result 2 to identify which programs are considered and included in the applicant's reported portfolio.

3. Given these applications, obtain new rational expectation assignment probabilities, as detailed in Appendix A.10.

4. Iterate over steps 2 and 3 until educational segregation is moving within a range of values.

Counterfactual procedure in alternative specifications

Full consideration

1. Obtain random utility shocks consistent with reported ROLs under the baseline.

a) Use estimated parameters for λ, β^g, ξ^g, $\Delta_{d,C19}$ and σ_ϵ^2.

b) Restricting now only to $j \in ROL_i$:

   i. We draw one of the ranked programs at random and simulate its ϵ_{ij} consistent with $v_{ij} \in \mathbb{R}^+$

   ii. We then draw a second program and now sample its ϵ_{ik} consistent with its relative ranking to program j drawn in the first step. This implies:

$$r_k < r_j \Rightarrow v_{ik} \in (v_{ij}, \infty)$$
$$r_k > r_j \Rightarrow v_{ik} \in (0, v_{ij})$$

   iii. Continue with the rest of the ranked programs finding the correct upper and lower bounds

c) For $j \notin ROL_i$, simply sample ϵ_{ij} consistent with $v_{ij} < 0$

2. For any given counterfactual affecting preferences or distance to school, use estimated preference parameters and simulated utility shocks to identify which programs are included in the counterfactual ROL_i ($v_{ij} > 0$), and in what order.

Costly ranking (as in Idoux (2022))

1. Obtain utility and ranking cost shocks using the reported ROLs of the baseline scenario.

   a) Use estimated parameters for λ, β^g, ξ^g, $\Delta_{d,C19}$, c^g and variances σ_ϵ^2, σ_ζ^2

   b) Restricting now only to $j \in ROL_i$, as in the full consideration case:

      i. We draw one of the ranked programs at random and simulate its ϵ_{ij} consistent with $v_{ij} \in \mathbb{R}^+$

      ii. We then draw a second program and now sample its ϵ_{ik} consistent with its relative ranking to program j drawn in the first step. This implies:

$$r_k < r_j \Rightarrow v_{ik} \in (v_{ij}, \infty)$$
$$r_k > r_j \Rightarrow v_{ik} \in (0, v_{ij})$$

      iii. Continue with the rest of the ranked programs finding the correct upper and lower bounds

   c) We now compute the expected ranking value that each of these ranked programs had when included in the ranking (i.e. conditional on the partial portfolio which in this case is known due to the nature of the portfolio formation heuristic). If the program j is ranked in position $\hat{r}$, this value is given by[9]:

$$RV_{ij} = v_{ij} q_{ij} \prod_{l=1}^{\hat{r}-1} \{1 - q_{ik_{r=l}}\}$$

d) We now sample the ranking cost shock of the applicant, such that it is consistent with being equal or larger than zero, and:

$$c_i < min_{j \in ROL_i} RV_{ij}$$

e) Using the ranking cost, we can now bound the indirect utilities of all programs not ranked by applicant i. This can be written in short as, $\forall j \in \mathcal{P}_i \setminus ROL_i$, ϵ_{ij} has to be consistent with:

$$v_{ij} < min \left\{ min_{k \in ROL_i} \{v_{ik}\}, \frac{c_i}{q_{ij} \prod_{r=1}^{|ROL_i|} \{1 - q_{ik_r}\}} \right\}$$

2. For any given counterfactual affecting preferences, distance to school, or initial rational expectation assignment probabilities, use estimated preference parameters, simulated utility and ranking cost shocks, and Result 1, or more generally the portfolio formation process detailed in Idoux (2022) to obtain new applications.

3. Given these applications, obtain new rational expectation assignment probabilities, as detailed in Appendix A.10.

4. Iterate over steps 2 and 3 until educational segregation is moving within a range of values.

A.9 Deferred Acceptance Algorithm and Application Process Description

The algorithm implemented in Chile is based on the deferred acceptance mechanism proposed by Gale and Shapley (1962), with some context-specific features explained in detail in Correa et al. (2019). Here we summarize the process:

Step 0: i) Collect vacancies at each academic program and divide them into the quotas described in Table 1.1.

Step 0: ii) Collect all applications, priorities and family relations between applicants (blood-related siblings as well as family-linked applications).

[9]Remember the convention used throughout the chapter that $\prod_{l=1}^{0} x_l = 1$

Step 0: iii) Get lottery draws for each applicant at each school.

Note: In Chile, lotteries are drawn at the school and family levels. That is: all applicants belonging to a family get the same lottery draw (and are then ordered within the family), and relative orders at all programs of a given school are the same for any given pair of applicants.

Step 1: Starting with the highest grade:

a) Identify the quota-ordering in which each applicant is processed in each school program.

b) Temporarily assign, up to the available vacancies in a given program-quota, applicants that have that program-quota first on their applications list. If there are more applicants than vacancies available at that program-quota combination, select the applicants with the highest priority and lottery number combination.

   Note: If an applicant has secured enrollment priority, she is assigned to the program irrespective of its capacity. That means the assignment is forced beyond program capacity in those cases.

c) All applicants temporarily unassigned that have been processed in all their program-quota combinations exit the algorithm and are labeled as unassigned.

   Those that have not been processed in at least one of the program-quota combinations in their ROL are now considered at the next program-quota in their list.

d) Among all applicants temporarily assigned to a given program-quota, and all those included for consideration in the previous step, those with the highest priority-lottery combination up to capacity are (or remain) temporarily assigned to that program quota.

e) The process continues until no temporarily unassigned applicant has remaining program-quota combinations at which she has not been processed in their ROL.

   Temporarily assignments are then considered permanent.

Step 2: For the highest grade not yet processed:

a) Identify all applicants in the current grade with siblings assigned in higher grades to schools shared in their ROLs. Update the priority status of those applicants in all programs in their ROLs offered by those schools to the maximum between the current status and dynamic sibling priority.

b) Identify all applicants in the current grade with family-linked applicants assigned in higher grades to schools shared in their ROLs. Update the ROL order of those applicants using the lexicographic criterion of placing first programs affected by the family link and then programs originally ranked higher in the applicant's ROL.

c) Repeat the process described in the previous step for the highest grade with the grade currently processed.

Step 3: After the smallest grade, the algorithm finishes.

A.10 Process to obtain rational expectation assignment probabilities (ratex)

Ratex estimation challenges

As explained in Appendix A.9, on top of vacancies, an applicant's assignment to a given program depends on: i) her application and that of other applicants, ii) her priorities relative to those of other applicants, and iii) the lottery assigned to that applicant in that program. It is important to note that the lottery, or more importantly, the relative ranking of applicants is the same in all programs offered by the same school.

Applicants know their priority ordering but are uncertain about their lottery number and that of other applicants. But additionally, applicants are uncertain about the congestion levels that will exist in equilibrium in a given program, as they also depend on the characteristics of the pool of applicants participating in the process in a given year and on the applications submitted by them. As a consequence, the ex-post observed pool of applicants is ex-ante uncertain, which increases the dispersion in estimated or expected assignment probabilities beyond that generated by lotteries. In simple, if more children residing in a given neighborhood participate in the system in a given year, schools in that neighborhood will likely have smaller ratex. To accommodate both sources of uncertainty, the ratex estimation process described below samples both lotteries and the pool of applicants in each simulation.

Before describing our procedure in detail, we first give some context on the challenge of estimating the assignment probabilities to a program when its vacancies are divided into different quotas.

Under a given assignment, determined by a given set of applications and lotteries, computing the probability of assignment that an additional applicant i would have had in a given program j -if ranked first- is straightforward: we can simply look at the priorities and lotteries of assigned applicants and ask: What lottery number would the additional applicant require to displace one of the currently assigned ones? The answer to this question is:

- If the program has remaining vacancies, any lottery number suffices ⇒ one.
- If in any quota of the program, the applicant with the worst priority has a worse priority than the new applicant ⇒ one.

- If, on the other hand, in all the quotas of the program, the applicant with the worst priority has a better priority than the new applicant ⇒ zero.
- If, in one or more quotas, the applicant with the worst priority has the same priority as the new applicant ⇒ the assignment likelihood is determined by the probability of obtaining a lottery number better than that of at least one of them.

This response has two problems. First, the probability computed depends on the lottery draws, particularly that of the assigned applicant with the worst priority. To get a more precise estimate, we would want to generate several lottery draws (and assignments). Conceptually, we could tackle this problem computing the expected lottery value of the marginal applicant which corresponds to the expected value of the ordered statistic of all applicants in the same priority group that were processed to that quota. Second, and more importantly, the cutoff value doesn't exactly answer the question of interest for us, which is what were the assignment probabilities of applicants with different priorities in a given year. Instead, it gives us the probability that an additional applicant would face.

These problems are well illustrated with a simple example on which we will continue to expand throughout this section: imagine for now that we only have one quota with one vacancy and two applicants with the same priority participating in the process. The probability of assignment that we would want to obtain is 50%, but using the cutoff value we first get a random (problem one), but additionally (and assuming a uniform distribution of lotteries), we get on average an assignment probability of 33.3% (problem two).

To tackle these issues, Arteaga et al. (2021) implement a process that does not rely on cutoff values. They instead identify the priority of the assigned applicant with the worst priority and using the pool of applicants processed at the program with that priority compute the ratio of assigned applicants to processed applicants. In our example above, we have one assigned applicant and two processed applicants, which implies that the ratio is 50%. The problem with this approach however is that it is not robust to the variations in priorities between applicants processed at the different quotas of a given program. To see this, let's expand on the previous example. Imagine now that we have two quotas, the low-SeS and the regular one, with one vacancy each, and four applicants in total, two with low-SeS priority (and no other priorities), and two with parent school-worker priority (and no other priorities). This implies, using the ordering described in Table 1.1, that one low-SeS applicant will be assigned to the first quota, and one of the applicants with parent school-worker priority will be assigned to the other. Using the procedure of Arteaga et al. (2021), we get that applicants with only low-SeS priority of 50% and applicants with only parent school-worker priority have also 50%. This is correct. The problem, however, is computing the expected assignment probability of an applicant that has both priorities, as she could receive an assignment in either quota. If using her priorities, we compute the maximum between quotas, we get 50%, which is biased downward. If we instead process her sequentially, assuming lottery independence between quotas, we get an assignment probability of $50\% + (100\% - 50\%) * 50\% = 75\%$. This value overestimates her assignment likelihood because the applicant is not assigned to the low-SeS quota when her lottery number (which is the same as the other quota!) is low.

To tackle the problem explained in the previous paragraph, which arises due to the heterogeneity in priorities, relying on cutoffs is preferable as they provide us with an empirical estimate of the correct distribution of assignment probabilities for all groups simultaneously. In our example above, the relevant assignment probability for an applicant with both priorities is the minimum cutoff between both quotas, which implies an expected probability of assignment above the $33,3\%$ computed for the one quota case (now we are computing the expected minimum of two separate maximums in two draws). However, we are still faced with the problem that the cutoff characterizes the assignment probabilities of a new additional applicant - a fifth applicant in our expanded example -. Note that this problem is diminished as the number of vacancies and applicants increases, but we can additionally do better.

Our method to approximate relevant ratex using cutoffs is to some extent similar to that of Agarwal and Somaini (2018) in that they identify cutoff values that maintain the overall assignment unchanged. We instead use a more direct approximation based on the formula of the expected value of the n-statistic in scenarios with independent draws from a uniform distribution. Omitting the quotas and priorities for a moment, if we have V vacancies in a program and A applicants being processed for assignment, the expected assignment probability implied by the expected cutoff value is given by $^{V}/_{(A+1)}$. Therefore, if we use the correction factor $^{(A+1)}/_{A}$, we get the desired probability $^{V}/_{A}$. Following this logic, our approach consists in using a correction factor at the quota level that is given by the ratio between the number of applicants assigned to that quota with the lowest priority and all applicants processed to that quota with the same priority (i.e. those with the same priority in the quota that are assigned to it plus those unassigned to the program, plus those assigned to the same program, but in a quota that was for them processed later following the process described in Appendix A.9).

A final point to consider that is unrelated to the computation of ratex, but that is related to their use in our preference estimation and counterfactuals is that, as explained in Appendix A.9, lotteries are drawn at the school rather than program level. This implies that the possibility of assignment to different programs offered by the same school is correlated. We do not however take this correlation into account and thus indirectly assume that lotteries are drawn at the program level, because i) applicants are unlikely to know this fact, ii) integrating this correlation complicates the estimation and counterfactual generating processes significantly, and iii) if this correlation is taken into account, so should the fact that consideration (or ranking) costs are likely to be smaller when adding a second program from the same school. Given that this force operates in the opposite direction, they are compensated to some extent.

Process Description

The ratex simulation process is as follows:

Step 0: For computational efficiency, we fix the assignment of all applicants in grades above PreK. It would be more accurate to sample families, including applicants in PreK and higher grades, but it implies increasing the computational burden significantly for a small gain in accuracy.

For each simulation:

Step 1: In each Municipality separately, using the number of applicants in a given year as target, we sample **with repetition** the simulated cohort of applicants. Note here that the share of low-SeS applicants is allowed to vary between simulations.

Sampled applicants have the ROL either corresponding to their original application or to that generated in a given counterfactual. The same goes for priorities. The only difference in the simulation is that some applicants are absent, and some are repeated.

Step 2: For the simulated cohort of applicants, we sample new lotteries at the school level (independently of family-related applicants in higher grades).

Step 3: Using the simulated applicants and applications, we now simulate assignment using the available vacancies in each program-quota, depending on which counterfactual is being implemented (0% low-SeS quota, 15% or the share of low-SeS applicants in the Municipality).

Step 4: For each simulated assignment, we compute the simulated probability of assignment of any given combination of priorities - which we label as priority profile - in a program as follows:

a) If all applicants assigned in the simulation to that program have a higher priority in each of the quotas than the priority profile being evaluated, the simulated probability of assignment is zero.

b) If in at least one of the quotas, one of the assigned applicants has a priority below that of the priority profile, then the simulated probability of assignment is one.

c) If in all quotas, all assigned applicants have a priority number equal or better to that of the priority profile being evaluated, and if in at least one of the quotas the assigned applicant with the lowest priority number has the same one as the priority profile evaluated in that quota - the marginal priority -, the probability of assignment is given by the maximum probability of assignment to any such quotas, which is computed using the corrected cutoff value described above.

Using $N_{MP,q}$ to express the total number of applicants with the "marginal priority" processed in a given quota during the simulation, and using $\mathcal{Q}_{MP}$ to express the quotas in which the marginal priority is the same as in the priority profile, we can express the value as:

$$max_{q \in \mathcal{Q}_{MP}} \mathbb{P}(lottery\ number \succ cutoff\ lottery_q) \times \frac{N_{MP,q} + 1}{N_{MP,q}}$$

Remember that processed applicants are those assigned to the quota plus those that were processed but did not receive an assignment in that quota (they were however maybe assigned in another quota in that program but processed after and not before the one being evaluated).

Step 5: We run 100 simulated assignments and average out the probabilities obtained in each of them.

A.11 Educational Segregation in Chile

As explained in (Bellei, Valenzuela, and De los Rios, 2010), educational segregation in Chile received increasing attention during the late 2000s. While some studies reached different conclusions depending on the methodology used,[10] international comparisons such as the ones highlighted in (Bellei et al., 2010), driven in part by Chile's increasing development levels, raised public awareness of the issue. In particular, Chile's level of educational segregation was found to be among the highest when compared with the 57 countries that participated in the 2006 PISA exams. Subsequently, when Chile was included in the OECD group of countries in 2010, it was revealed that it had the highest level of educational segregation among the group.

Regarding the roots of observed educational segregation, residential sorting is a consensus-relevant contributing factor, which is why residency-based priorities are not included in the SAE system. Nevertheless, the level of residential segregation has been below that of educational segregation, suggesting the presence of other contributing factors (Carrasco, Mizala, Contreras, Santos, Elacqua, Torche, Flores, and Valenzuela, 2014). School selection practices had been identified as one likely relevant factor prior to the Inclusion Law Reform (Bellei et al., 2010; Carrasco et al., 2014). In a survey of school principals, Carrasco, Bogolasky, Flores, San Martin, and Gutierrez (2015) argue that selection practices were openly used in admissions, with high levels of sophistication, mainly concentrated in voucher schools, including those that had an agreement with the Ministry imposing tighter restrictions under the SEP Law. They conclude that the law restricting such practices (The Ministry of Education's Decree Law 196 from 2006) was ineffective as it was not tightly enforced, and the system's incentives on schools to obtain good test results pushed them to be more selective, countering the spirit of the Decree.

Although selection practices in public and voucher schools have been eliminated under the new CCAS, as discussed earlier, it appears that educational segregation in Chile has not

[10] More specifically, Valenzuela, Bellei, and de los Ríos (2009) argued that educational segregation was high and had experimented a slight increase, while Elacqua (2009) argued that the level of educational segregation was moderate, on par with developed countries, and that it had been decreasing.

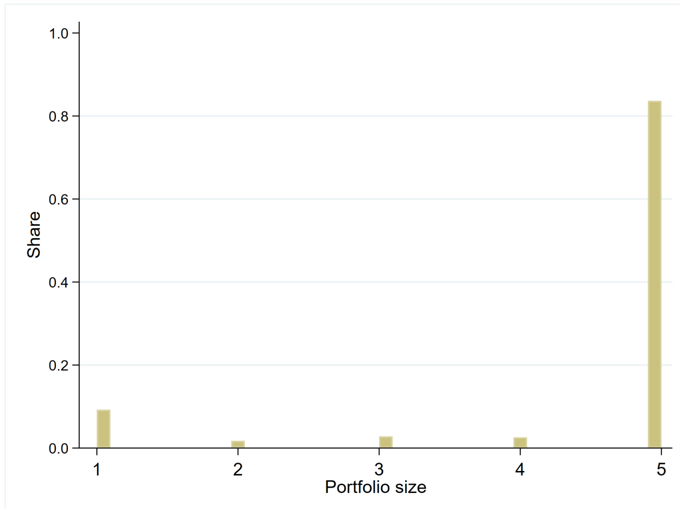
decreased significantly. This underscores the need to prioritize addressing the underlying drivers of segregation and devising effective policies to reduce it, which should be a key focus of policy discussions.

Appendix B

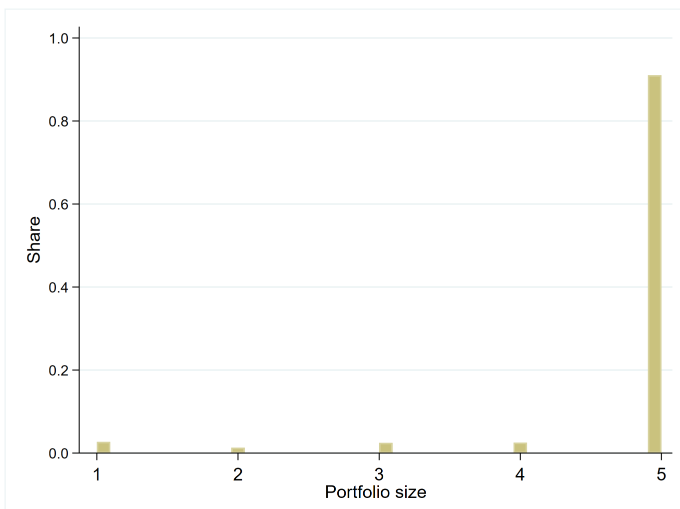
Appendix: The Potential of Smart Matching Platforms in Teacher Assignment: The Case of Ecuador

B.1 Additional Figures

Figure B.1.1: Portfolio size



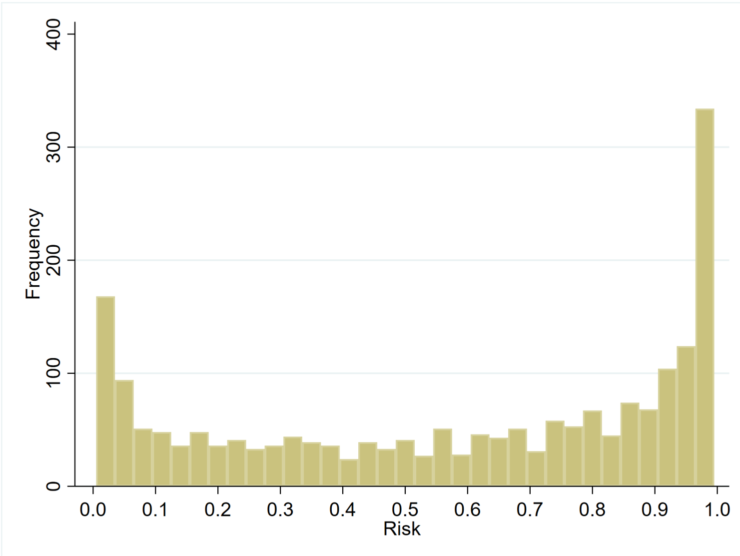
(a) Pre-validation period



(b) Post-validation period

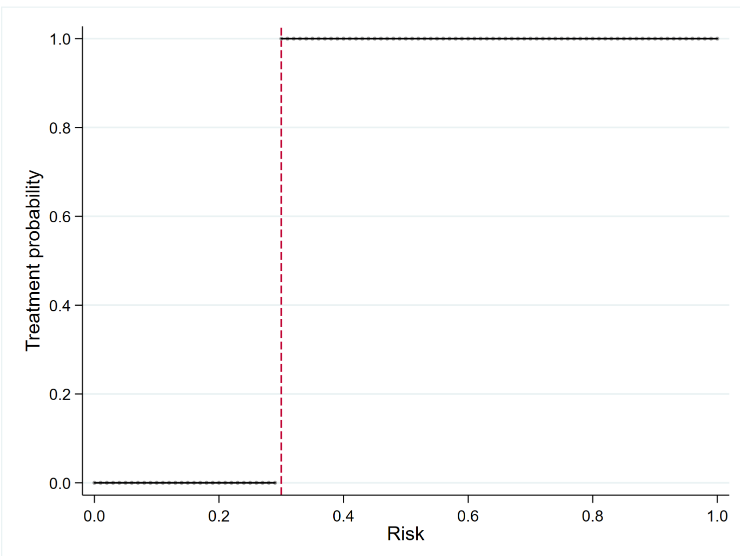
Note: Distribution of portfolio size pre and post-validation period. The sample is limited to the applicants who received the personalized report.

Figure B.1.2: Risk distribution for compliers



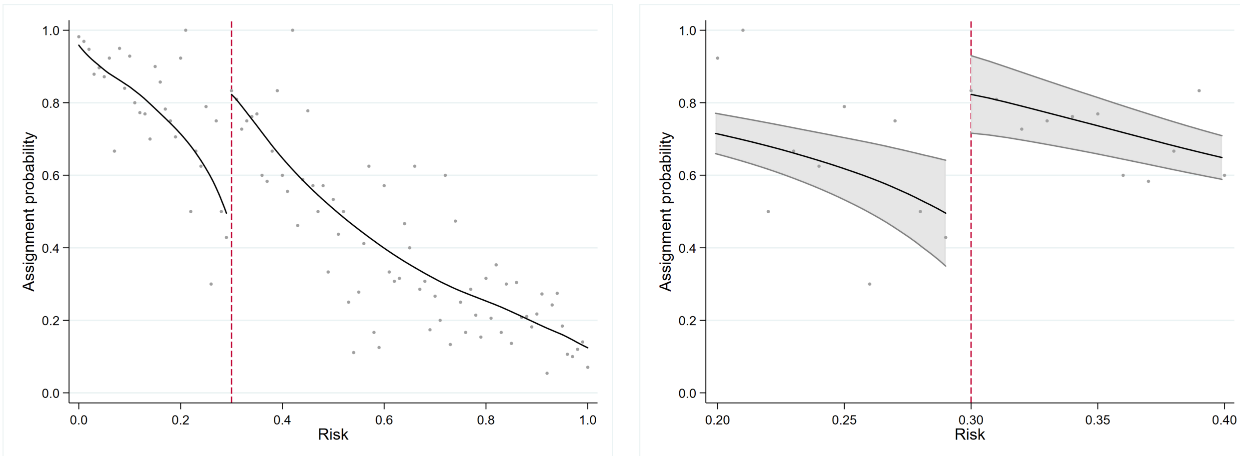
Note: 2,051 observations. The extremes are omitted.

Figure B.1.3: Treatment Probability



Note: Total observations: 3,653. The bin that contains 0 consists of 727 observations. The bin that contains 1 consists of 1078 observations. The rest of the bins consist, on average, of 18.7 observations. The size of the bins is 0.01.

Figure B.1.4: RDD results on assignment probability



(a) Assignment probability (b) Assignment probability with CI

Note: Figure (a) plots the probability of obtaining an assignment using linear polynomials. Figure (b) plots the same but within the optimal bandwith and with confidence intervals. Total observations are 3,653. The size of the bins is 0.01. The bin that contains 0 consists of 727 observations. The bin that contains 1 consists of 1,078 observations. The rest of the bins consist, on average, of 18.7 observations.

Figure B.1.5: RDD on assignment probability comparing partial and final applicants

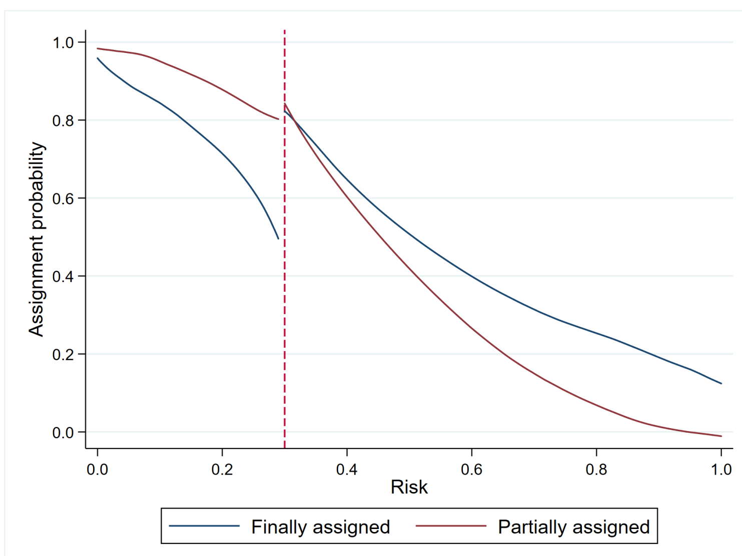
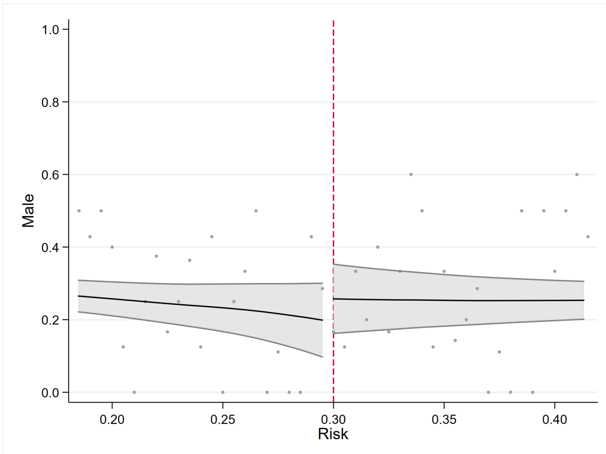
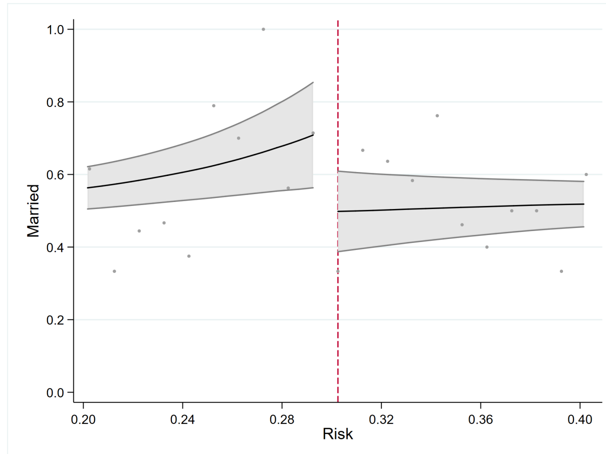


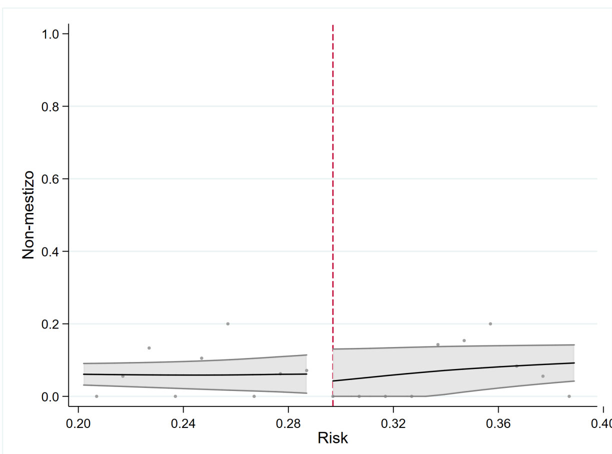
Figure B.1.6: Balance near the threshold



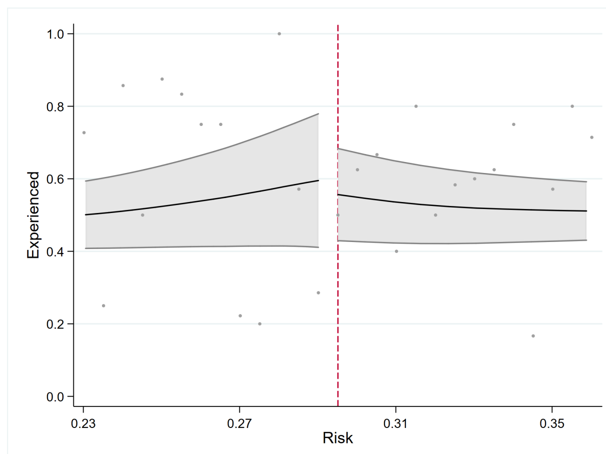
(a) Male



(b) Married



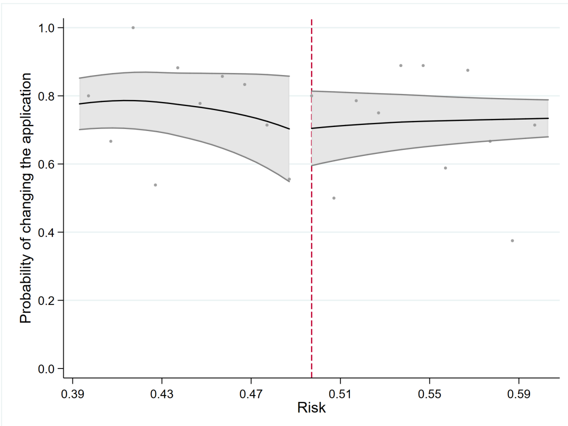
(c) Non-mestizo



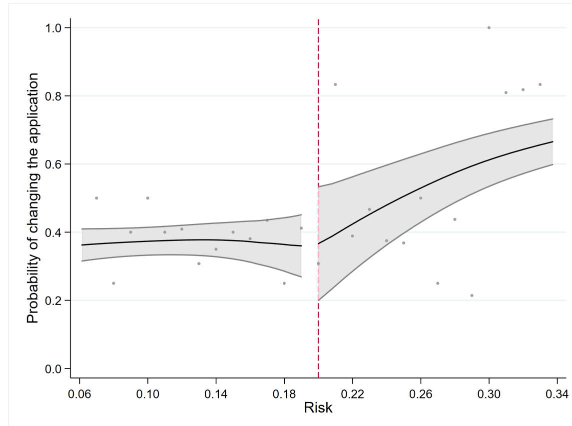
(d) Experienced

Note: Figure (a) plots the probability of being male. Figure (b) plots the probability of being married. Figure (c) plots the probability of being non-mestizo. Figure (d) plots the probability of being experienced.

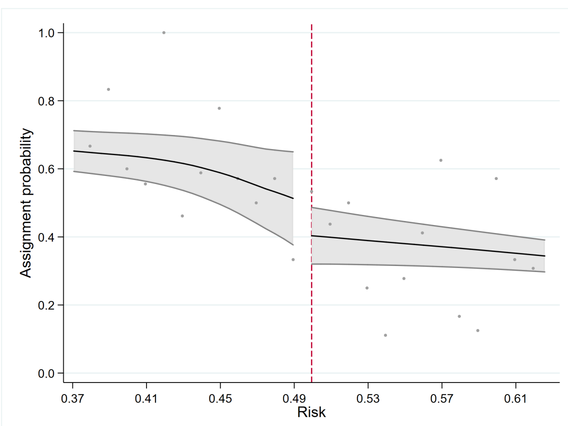
Figure B.1.7: Placebo test



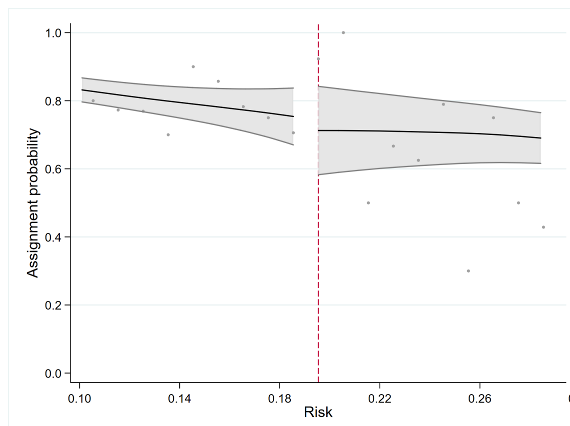
(a) Probability of changing the application with fake cutoff at 0.5 risk level



(b) Probability of changing the application with fake cutoff at 0.2 risk level



(c) Assignment probability with fake cutoff at 0.5 risk level



(d) Assignment probability with fake cutoff at 0.2 risk level

Note: Figure (a) plots the probability of changing the application with a fake cutoff at risk level 0.5. Figure (b) plots the probability of changing the application with a fake cutoff at risk level 0.2. Figure (c) plots the assignment probability with fake cutoff at risk level 0.5. Figure (d) plots the assignment probability with fake cutoff at risk level 0.2.

Figure B.1.8: Quality of reassigned teachers

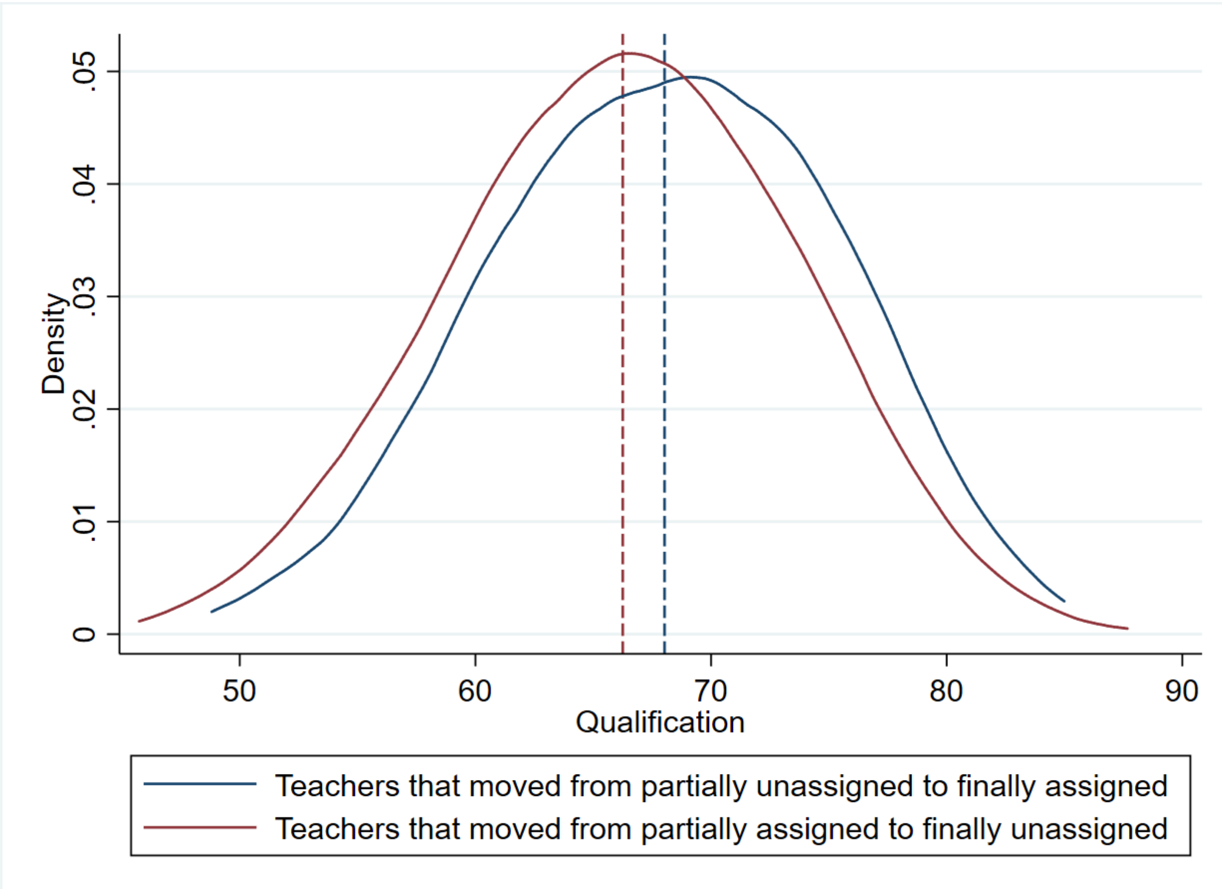
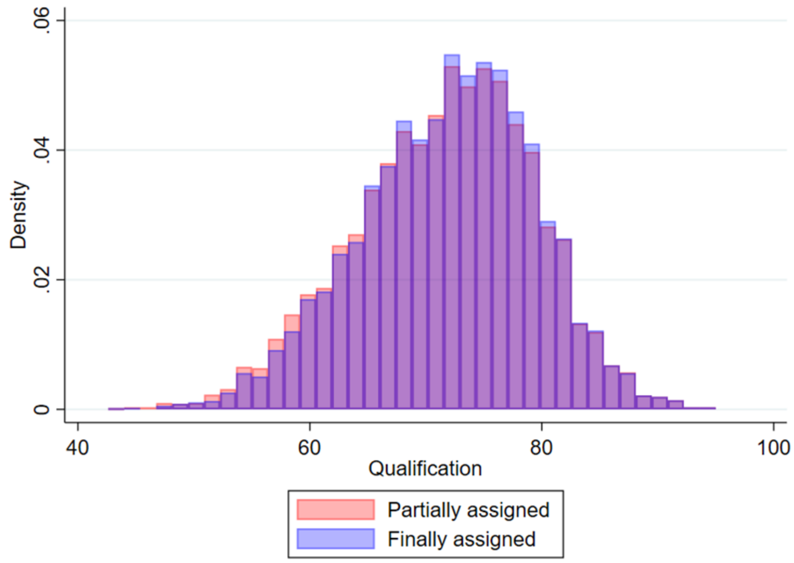
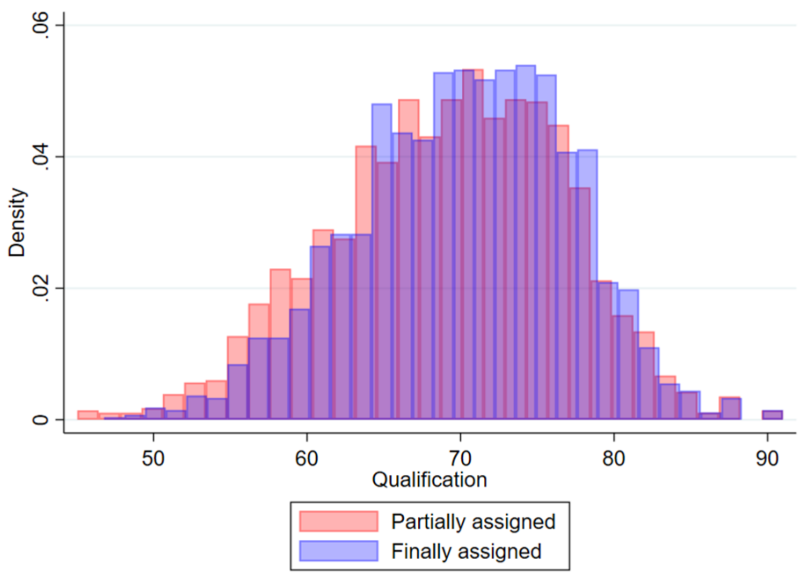


Figure B.1.9: Scores of assigned applicants pre- and post-validation period



(a) All vacancies assigned pre- and post-validation



(b) All vacancies assigned to different applicants pre- and post-validation

Note: Figure (a) presents the distribution of scores for vacancies that had someone assigned both pre- and post-validation. Figure (b) presents only the vacancies where the assigned teacher is different in the post-validation assignment.

B.2 Additional Tables

Table B.2.1: Pre-Bonus Scoring System

Merits		**Oposition**	
Criteria	Maximum score	Criteria	Maximum score
Academic Background	20	Specific knowledge test	40
Work experience	10	Mock lecture	25
Publications	3		
Continuous training	2		
Total weight	**35**	**Total weight**	**65**

Table B.2.2: Shares by specialties

Specialty	**Share (%)**
Basic General Education (Egb) From 2nd to 7th grade	22.33
Initial education	17.54
Mathematics Basic General Education (Egb From 8th To 10th grade)	9.15
Social Studies Basic General Education (Egb From 8th To 10th grade)	6.96
Entrepreneurship and Management General Unified High School (Bgu)	6.72
Natural Sciences Basic General Education (Egb From 8th To 10th grade)	6.58
Fip: Accounting	4.62
English	3.59
Education for Citizenship General Unified High School (Bgu)	3.54
Language and literature General Unified High School (Bgu)	3.23
Physical Education 2nd grade Egb to Bgu	2.99
Biology General Unified High School (Bgu)	2.10
Artistic and Aesthetic Education 2º grade Egb to Bgu	1.93
Fip: Computing	1.83
Special education	1.36
Chemistry General Unified High School (Bgu)	1.01
Fip: Agricultural production	0.88
History General Unified High School (Bgu)	0.71
Physics General Unified High School (Bgu)	0.67
Fip: Sales and Tourist Information	0.54
Fip: Electromechanics	0.47
Philosophy General Unified High School (Bgu)	0.43
Fip: Consumer electronics	0.43
Fip: Music	0.40

Table B.2.3: Sensitivity test

	(1) BW 0.3	(2) BW 0.25	(3) BW 0.2	(4) BW 0.15	(5) BW 0.1
Panel A. Any modification					
RDD estimate	0.387	0.388	0.407	0.429	0.425
	(0.065)	(0.074)	(0.086)	(0.096)	(0.115)
Total observations in BW	941	650	490	384	244
Panel B. Add any					
RDD estimate	0.437	0.469	0.515	0.518	0.538
	(0.064)	(0.072)	(0.084)	(0.092)	(0.109)
Total observations in BW	941	650	490	384	244
Panel C. Add any from recommendations					
RDD estimate	0.661	0.634	0.604	0.549	0.449
	(0.087)	(0.104)	(0.126)	(0.135)	(0.210)
Total observations in BW	441	334	256	201	127
Panel D. Assigned					
RDD estimate	0.215	0.286	0.271	0.317	0.359
	(0.063)	(0.072)	(0.083)	(0.095)	(0.123)
Total observations in BW	941	650	490	384	244
Panel E. Assigned in recommendation					
RDD estimate	0.491	0.461	0.422	0.354	0.361
	(0.063)	(0.071)	(0.079)	(0.085)	(0.119)
Total observations in BW	680	442	330	256	161

Note: Robust standard errors in parentheses. All columns only consider model (1) from Table 2.3 using different arbitrary BW. All estimates control for specialty, sex, marital status, and region. Panel C is conditional on having added something to the application. Panel E is conditional on having been assigned.

Table B.2.4: Balance tests

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Panel A. Male					
RDD estimate	-0.099	-0.062	-0.069	-0.035	-0.085
	(0.098)	(0.091)	(0.129)	(0.119)	(0.150)
Left BW	0.113	0.096	0.075	0.063	0.113
Right BW	0.113	0.221	0.075	0.147	0.113
Total observations in BW	267	384	180	245	267
Panel B. Non-mestizo					
RDD estimate	-0.062	-0.058	-0.100	-0.046	-0.039
	(0.066)	(0.070)	(0.073)	(0.076)	(0.098)
Left BW	0.097	0.095	0.064	0.063	0.097
Right BW	0.097	0.219	0.064	0.145	0.097
Total observations in BW	238	366	152	245	238
Panel C. Married					
RDD estimate	-0.151	-0.153	-0.245	-0.156	-0.430
	(0.137)	(0.113)	(0.170)	(0.150)	(0.194)
Left BW	0.103	0.099	0.068	0.066	0.103
Right BW	0.103	0.223	0.068	0.148	0.103
Total observations in BW	248	384	170	255	248
Panel D. Experienced					
RDD estimate	0.170	0.300	0.407	-0.082	0.396
	(0.180)	(0.155)	(0.260)	(0.215)	(0.281)
Left BW	0.063	0.048	0.042	0.032	0.063
Right BW	0.063	0.190	0.042	0.126	0.063
Total observations in BW	152	276	94	180	152

Note: Robust standard errors in parentheses. This table reports parametric estimates using different strategies to calculate the optimal bandwidth and different types of polynomials. (1) is estimated using a linear polynomial and the BW is calculated using the "one common MSE-optimal". (2) is estimated using a linear polynomial and the BW is calculated using the "two different MSE-optimal" that calculates two different BW below and above the cutoff. (3) is estimated using a linear polynomial and the BW is calculated using the "one common CER-optimal" bandwidth selector. (4) is estimated using a linear polynomial and the BW is calculated using the "two different CER-optimal" that calculates two different BW below and above the cutoff. (5) is estimated using a quadratic polynomial and the BW is calculated using the "one common MSE-optimal" method. All estimates control for specialty, sex, marital status, and region. Panel C is conditional on having added something to the application. Panel E is conditional on having been assigned.

Table B.2.5: Determinants of the probability of adding a recommendation

	(1)	(2)	(3)
Coefficients of interest			
Recommendation in teacher's province	0.230	1.320	2.516
	(0.008)	(0.039)	(0.076)
Another province in pre-validation applications	0.184	0.962	1.819
	(0.012)	(0.046)	(0.085)
Rural institution	-0.011	-0.065	-0.120
	(0.004)	(0.028)	(0.056)
Original application size	0.003	0.018	0.036
	(0.001)	(0.011)	(0.021)
$\frac{Score - Mean\ score\ specialty}{Sd.specialty}$	-0.025	-0.210	-0.423
	(0.003)	(0.025)	(0.050)
Male	0.014	0.101	0.167
	(0.005)	(0.036)	(0.072)
Number of observations	19,783	19,694	19,694
Controls			
Number of recommendations	Yes	Yes	Yes
Specialty	Yes	Yes	Yes
Teacher's Province	Yes	Yes	Yes

Note: Robust standard errors in parentheses. This table reports estimates of the probability of adding one of the recommendations included in the personalized report, during the validation period. We include only teachers in the treatment group that opened the personalized report and include different controls to try to isolate the coefficient on recommendations in the applicant's province or other provinces included in the pre-validation applications, as well as on rural alternatives, in line with the process used to create the recommendations in the first place.

(1) is estimated using a linear probability model, while (2) is estimated using a probit model, and (3) is estimated using a logit model.

Table B.2.6: Percentage of assigned vacancies

	(1) Partial assignment + new applicants	(2) Final assignment
Total assignment	6,839	6,904
Unfilled vacancies	1,170	1,105
Treatment compliers (% assignment)	9.07	20.32
Control compliers (% assignment)	96.83	90.95
New applicants (% assignment)	30.78	28.02

Note: Column (1) considers partial applicants and the new applicants that appeared after the validation period. Column (2) considers new applicants and partial applicants with modifications after the validation period.

B.3 Description of the bonus score

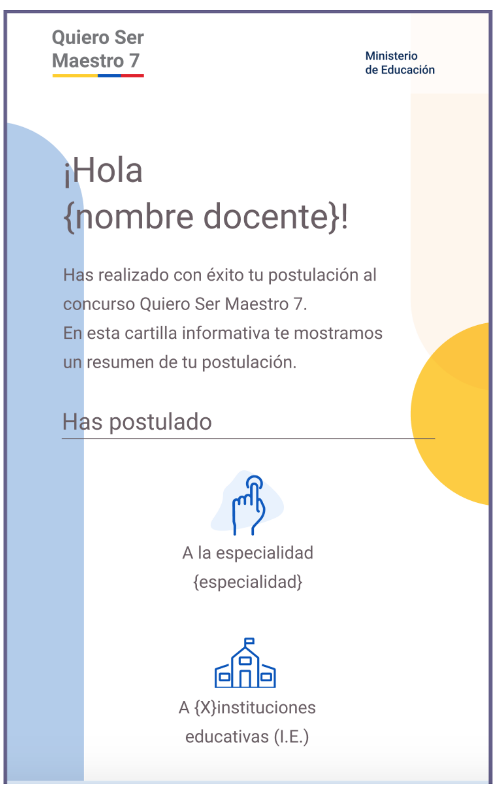
In the QSM7 contest, the bonus score was calculated using the following criteria:

1. 2 points for each of the following:
   - Applicants residing in the "educational circuit" where the institution offering the vacancy is located.
   - Applicants that present proof of a non-limiting disability.
   - Applicants currently residing abroad in "migration" status for at least one year.
   - Applicants choosing "fiscomisional" institutions (which are private institutions receiving government funds to complement public alternatives).
   - Applicants who already served their one year mandatory rural service.
   - Applicants from indigenous, Afro-Ecuadorian or Montubio ethnic groups.
   - Applicants demonstrating status as a "person returned to Ecuador."
   - Applicants residing in rural localities within a 40km radius of the Ecuadorian border.
2. 1 point for each of the following criteria:
   - Applicants currently serving under an occasional, definitive or provisional contract in public schools.
   - Applicants who are a "former community teacher."
3. Additional criteria:
   - 10% score bonus over the pre-bonus score for applicants demonstrating the status of "heroe" (hero) according to the corresponding law.

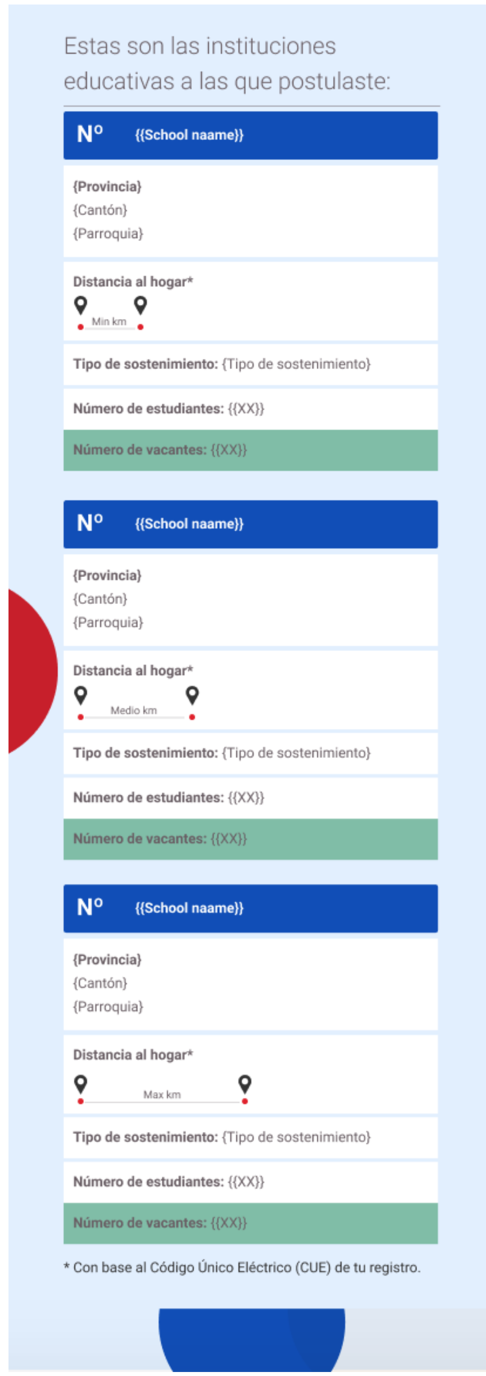
- 5% score bonus over the pre-bonus score for applicants demonstrating the status of "former *combatiente*" (former combatant) according to the corresponding law.
- 6 points to applicants residing in the Galápagos province and applying to a school within that province.

B.4 Personalized Report Outline

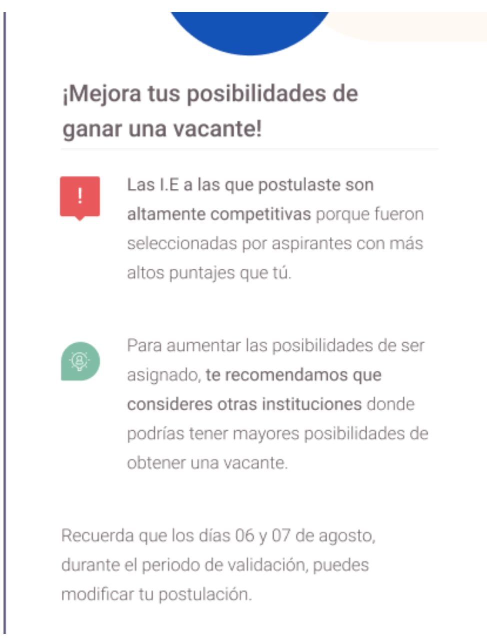
Figure B.4.1: Personalized Report Outline



(a) Section 1: Welcome

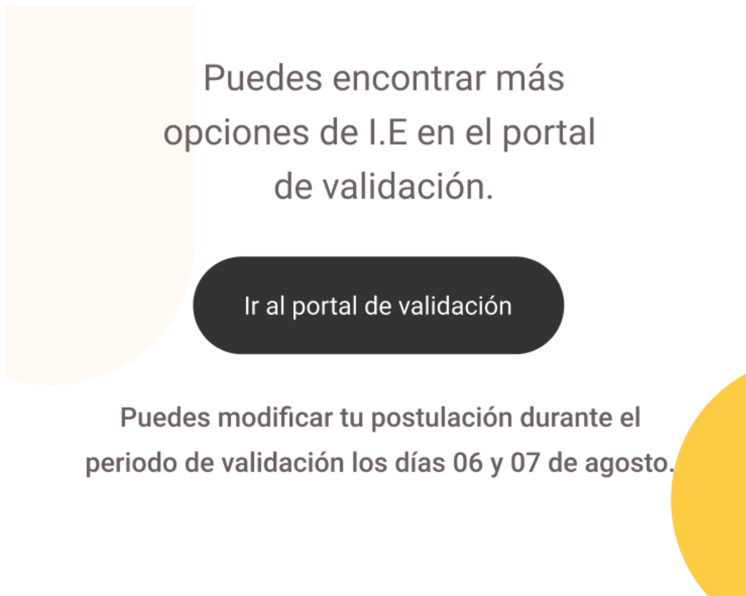


(b) Section 2: Your Portfolio



(c) Section 3: Non-assignment Warning

(d) Section 4: Recommendations



(e) Section 5: Link to Application Webpage

B.5 Survey Results

The survey was implemented after the application period but before the results of the contest were published. It was distributed via email to all teacher candidates and aimed to measure different dimensions of the process, as well as beliefs regarding assignment and awareness of available alternatives within an applicant's specialization.

- 11,948 teachers participated in the survey. On average, they rated the application process at 6.96 on a scale of 1 to 10.

Table B.5.1: Survey: Evaluation of Application Process

	Mean	**Standard deviation**	**Total**
Vacancy search	6.85	2.49	11,609
Information about educational institutions	6.93	2.47	11,609
On average, which grade would you give to the application process?	6.96	2.38	11,609

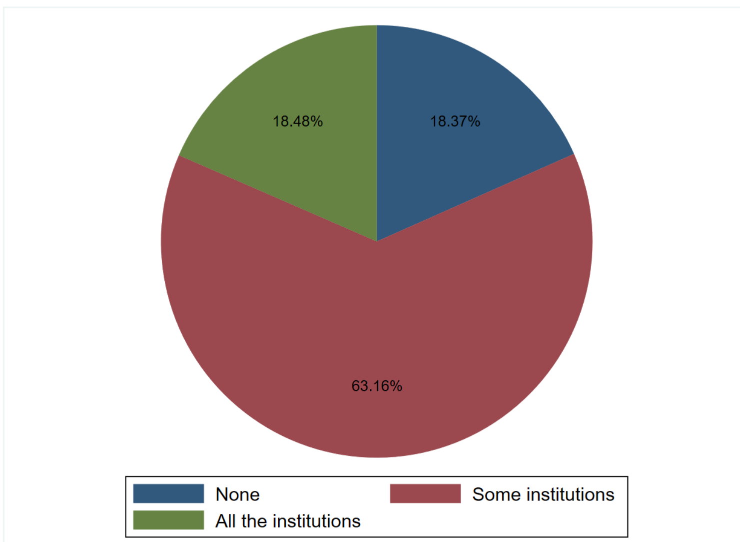
- If only the best teachers are considered (those who are above the 75th percentile of the distribution), the average of the evaluation rises to 7.17.

Table B.5.2: Survey: Evaluation of Application Process for Only Top 75th Score Percentile Teachers

	Mean	Standard deviation	Total
Vacancy search	7.05	2.45	2902
Information about educational institutions	7.14	2.44	2902
On average, which grade would you give to the application process?	7.17	2.28	2902

- Most of the teachers did not have a clear idea about the institutions to which they were going to apply: 18.3% did not have any institution in mind, 63% only had some in mind, and just 18.4% knew all or almost all of them.

Figure B.5.1: Survey: Answer to the question: "Did you have in mind which educational institutions you wanted to work at?"



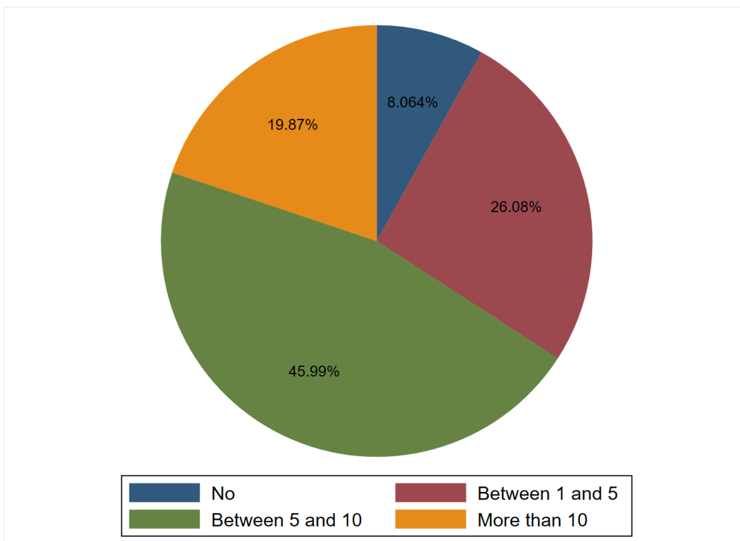
- 69% claim to have received the personalized report and, on average, they rated the report at 8.22 on a scale of 1 to 10.

Table B.5.3: Survey: Evaluation of Personalized Report

	Mean	Standard deviation	Total
Ease of access to the link	8.32	1.82	8,067
Design and clarity of the personalized report	8.28	1.82	8,067
Usefulness of the information presented	8.17	1.93	8,067
Usefulness of the recommendations received	7.55	2.36	3,286
Clarity of the message	7.70	2.30	3,400
On average, which grade would you give to the personalized report?	8.22	1.86	8,067

- For those who answered the question about what information they would like to receive in the personalized report, 34.8% stated that they would like to receive more information about the educational institutions to which they applied.
- 82% say they want more information about their chances of getting assigned.
- 15% want more information about the institutions they did not apply to.
- For those not assigned, 55% state they would have wanted more information about their chances of assignment.
- 91.94% of the teachers would have liked to apply to more educational institutions.

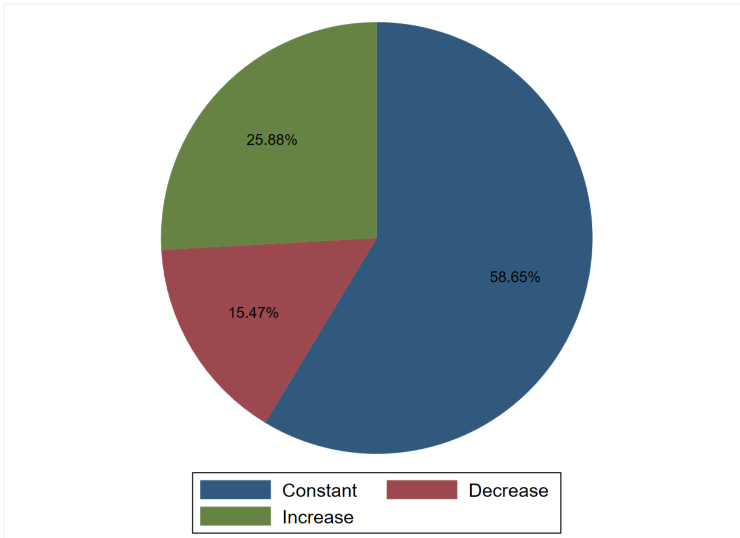
Figure B.5.2: Survey: Interest in Applying to More Alternatives



- 33% did not apply to their preferred educational institutions because they thought they would not get assigned. Of those, 58% said that they wanted more information about probabilities of assignment.
- The main reason why applicants applied for fewer than 5 options was because the system did not display more vacancies in their specialty.
- 16% stated that it was difficult to find other institutions to apply to.
- 2% preferred to not be assigned to a position rather than apply to the available alternatives.
- 13.1% were sure that they were going to get assigned; however, only 25% of these applicants were finally assigned.

- Most of the teachers did not change their beliefs after receiving the personalized report.

Figure B.5.3: Survey: Change in Assignment Belief After Receiving Personalized Report



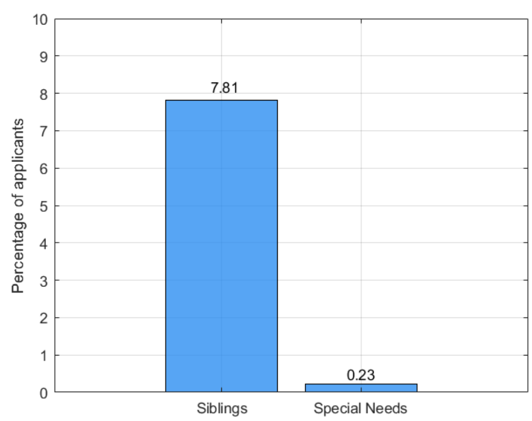
- Most of the teachers that did not change their application despite receiving the personalized report were confident about their assignment probabilities.
- Teachers were asked how satisfied they would feel if they were placed at the first-ranked school on their application, if they were placed at the last-ranked school, or if they were not placed. Most of the teachers stated that they would feel satisfied if placed at any option, while 81% would be unsatisfied with non-placement.

Appendix C

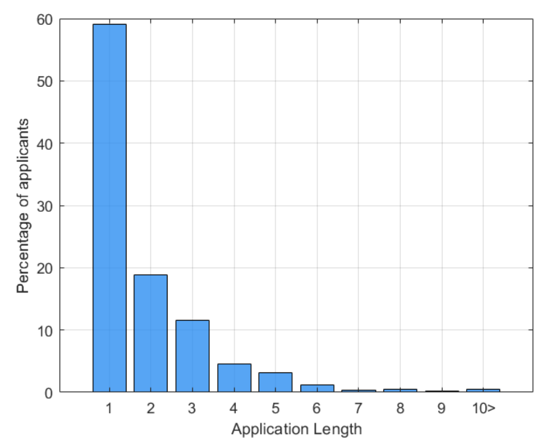
Appendix: The Welfare Effects of including Household Preferences in School Assignment Systems: Evidence from Ecuador

C.1 Additional Figures

Figure C.1.1: Distribution of Declared Applicant Priorities and Ranked Ordered List Size

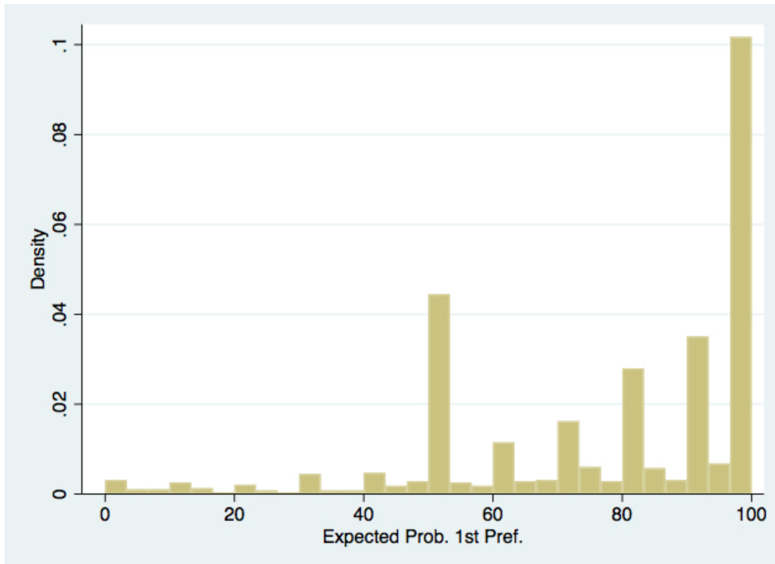


Priorities declared by applicants



Application length

Figure C.1.2: Perceived Probability of Admission to 1st Preference



Note: These responses were obtained in an online survey carried out after the end of the application period but before assignment results were communicated (to avoid biasing responses).

Figure C.1.3: Ranking Assigned: DA and Distance Mechanism

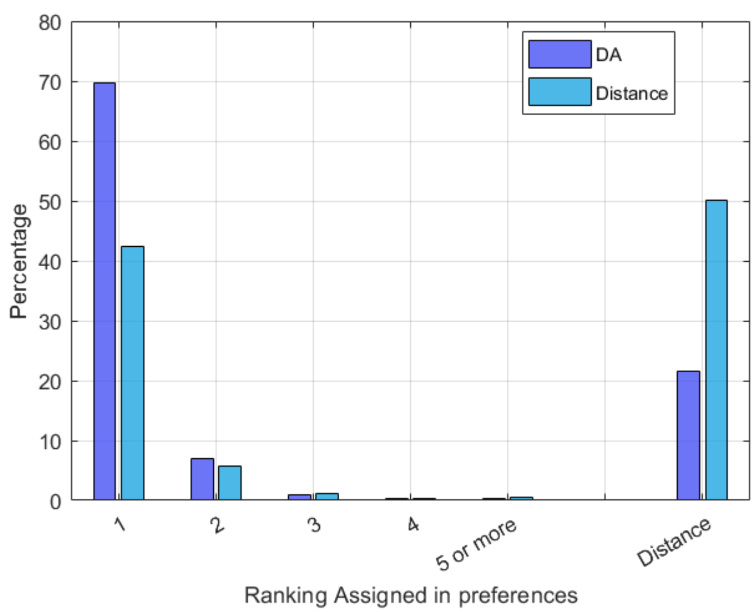
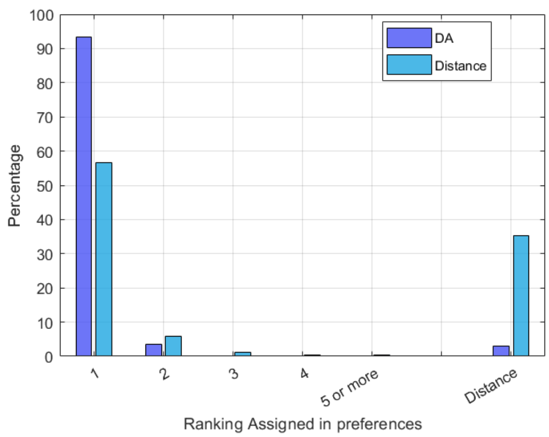
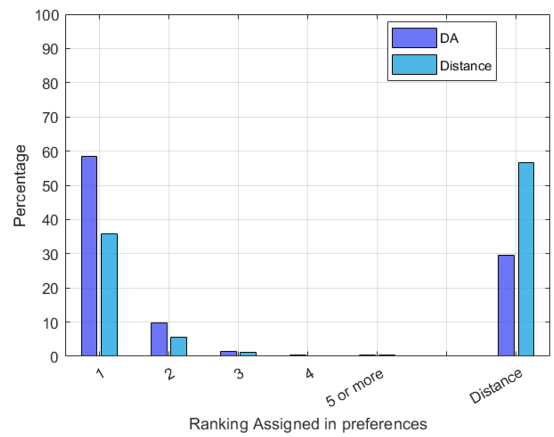


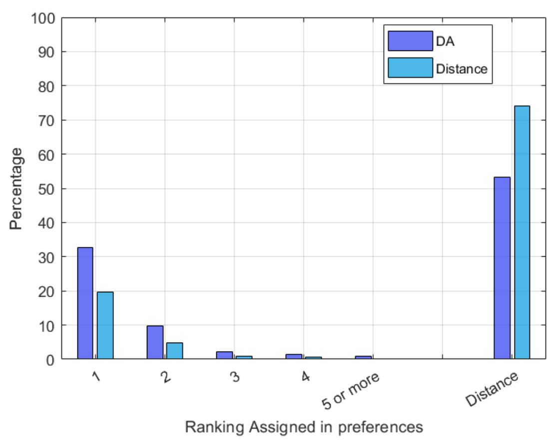
Figure C.1.4: Ranking Assigned by Grade: DA and Distance Mechanism



Pre-School 1

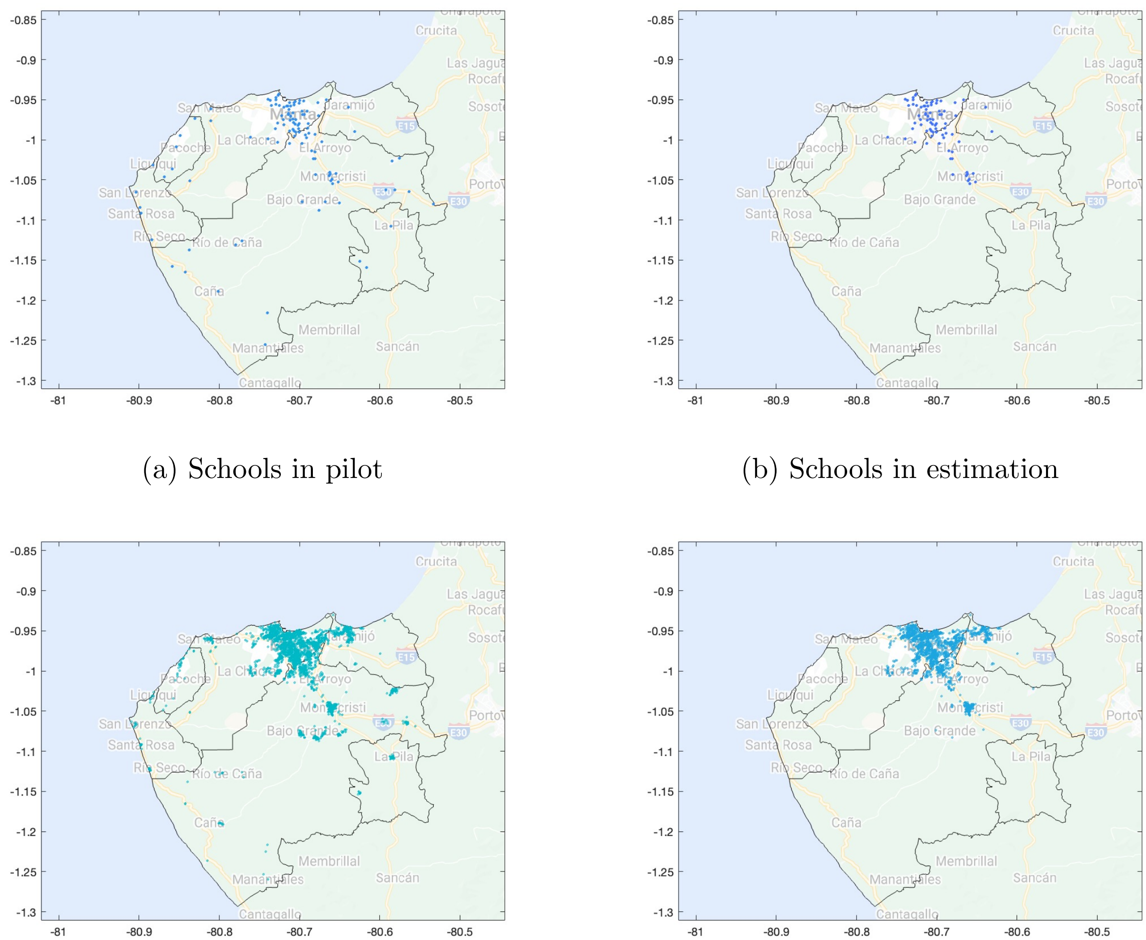


Pre-School 2



Primary 1

Figure C.1.5: School and Applicants in Pilot and in Welfare Estimation



(a) Schools in pilot

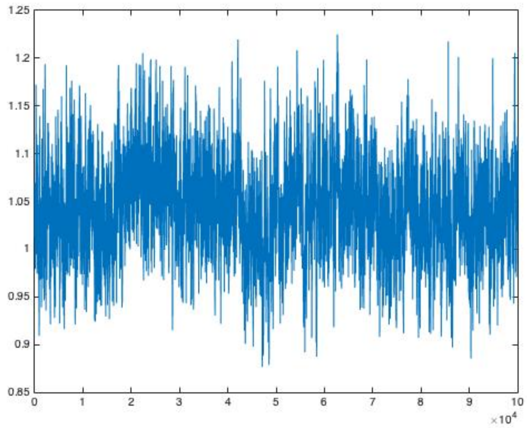
(b) Schools in estimation

(c) Applicants in pilot

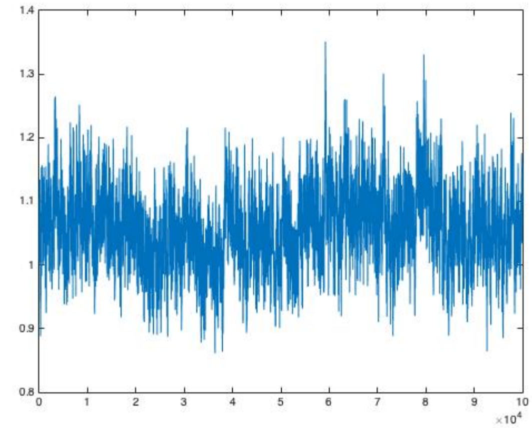
(d) Applicants in estimation

Note: A very small proportion of applicants from the urban area of Manta, where the estimation sample was located, were not included in the estimation process as they had included schools located outside of this area in their applications.

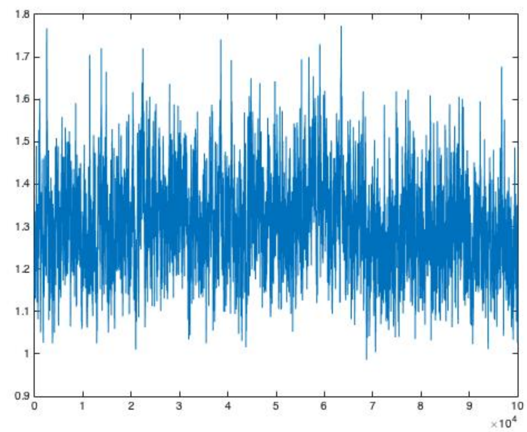
Figure C.1.6: Trace Plots σ_ϵ in Main Specification



Pre-School 1

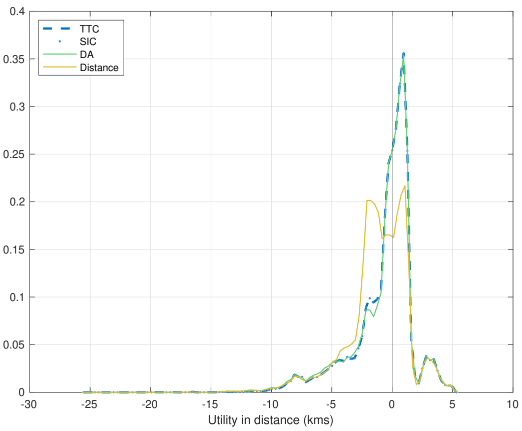


Pre-School 2

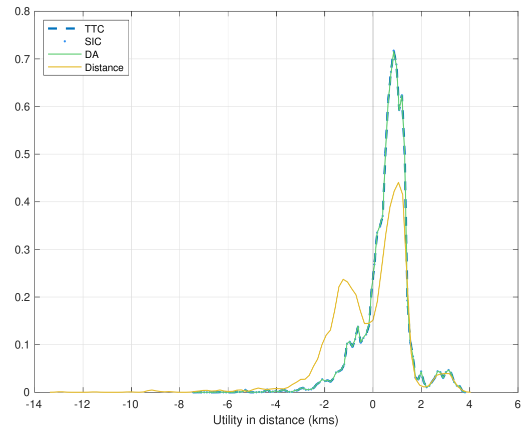


Primary 1

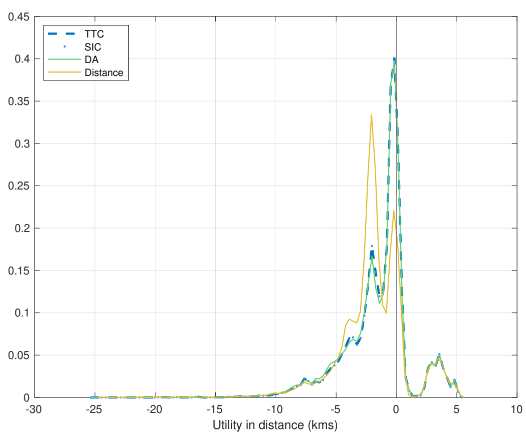
Figure C.1.7: Welfare Distribution



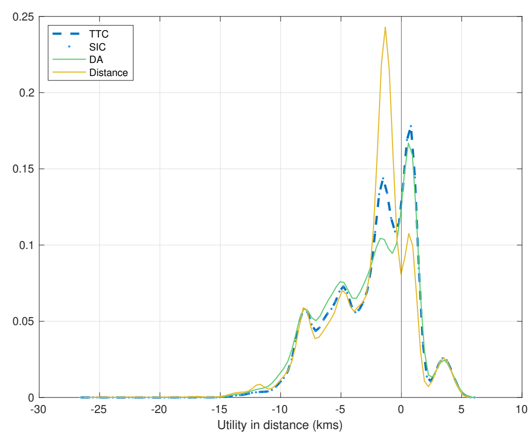
All applicants



Pre-School 1



Pre-School 2



Primary 1

Note: In this figure, we plot the utilities obtained with our model when using the scale normalizations $\overline{\delta} \equiv 0$ and -1 as the average disutility from each linear km of distance between the school and the reported location of the family. The level of utility is not relevant, as it depends on the normalization. However, the mass from the utility distribution when using the distance-centric algorithm being shifted to the left is relevant, as it indicates how the relative distributions of utilities compare, and lead to the average differences presented in Table 3.4.

Figure C.1.8: Figure 5 of Abdulkadiroğlu et al. (2017)

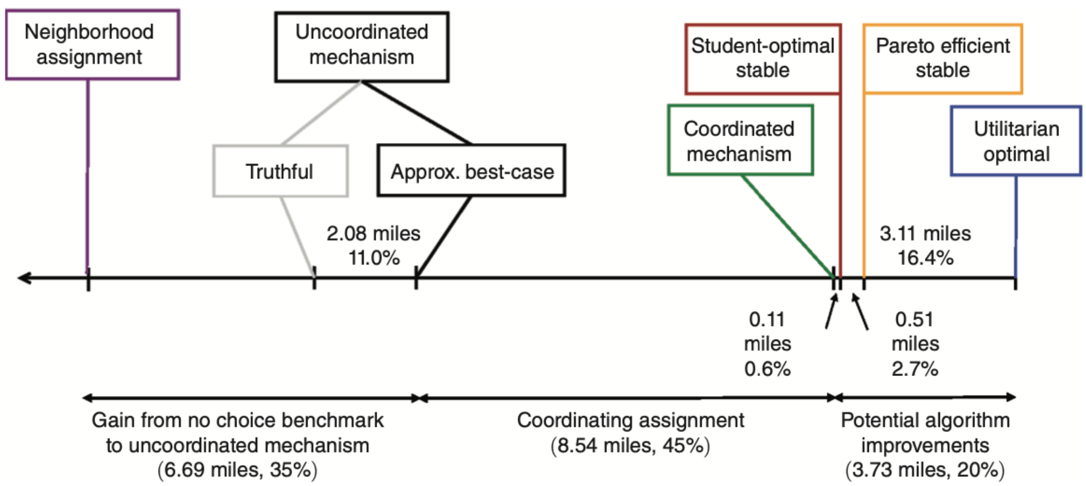


FIGURE 5. COORDINATING ASSIGNMENTS VERSUS ALGORITHM IMPROVEMENTS

C.2 Additional Tables

Table C.2.1: Comparative Statistics for Guayaquil and Manta

	Guayaquil	**Manta**
Total population	2,291,158	221,122
Population 3-5 years old (% of population 3-17 years old)	19.7	19.1
Minors in the school system (% of population 3-17 years old)	78.9	80.2
Average Mother's Education (of minors 3-17 years old)	11.3 years	10.8 years
Total schools	885	153
Share of public schools	54%	43%
Share of private schools	44%	54%
Share of "fiscomisional" schools	2%	3%
Total enrollment	687,046	86,455
Share of enrollment in public schools	57%	67%
Share of enrollment in private schools	40%	25%
Share of enrollment in "fiscomisional" schools	4%	8%

Table C.2.2: Estimates and Potential Scale Reduction Factors: Main Specification

	Pre-School 1		Pre-School 2		Primary 1	
Estimate	Mean (SD)	PSRF	Mean (SD)	PSRF	Mean (SD)	PSRF
Mean(ξ_j)	0.084		-0.915		0.055	
	(0.157)		(0.250)		(0.296)	
λ	2.962	1	4.402	1.001	3.698	1
	(0.354)		(0.527)		(0.804)	
σ_ξ	0.268	1.007	1.227	1.027	0.486	1
	(0.060)		(0.410)		(0.123)	
σ_ϵ	0.967	1.015	0.935	1.001	1.238	1
	(0.042)		(0.046)		(0.093)	
σ_γ	0.294		0.307		0.145	
	(0.055)		(0.073)		(0.029)	
Tot. schools	55		57		54	
Tot. students	1,098		885		389	

Table C.2.3: Mechanism Comparison - Results Pre-School 1

Panel A: All Applicants		
Assigned in:	DA	DC
Any preference	1,654	1,102
	(96.95%)	(64.60%)
First preference	1,592	966
	(93.32%)	(56.62%)
Average assignment distance	0.87km	0.52km
Panel B: Applicants maintaining 1st preference (985)		
Assigned in:	DA	DC
Any preference	965	955
	(97.97%)	(96.95%)
First preference	942	946
	(95.63%)	(96.04%)

Table C.2.4: Mechanism Comparison - Results Pre-School 2

Panel A: All Applicants		
Assigned in:	DA	DC
Any preference	1,199	737
	(70.45%)	(43.30%)
First preference	93.32	609
	(58.52%)	(35.78%)
Average assignment distance	1.32km	1.08km
Panel B: Applicants maintaining 1st preference (959)		
Assigned in:	DA	DC
Any preference	708	607
	(73.83%)	(63.30%)
First preference	618	569
	(64.44%)	(59.33%)

Table C.2.5: Mechanism Comparison - Results Primary 1

Panel A: All Applicants		
Assigned in:	DA	DC
Any preference	265	147
	(46.74%)	(25.93%)
First preference	185	111
	(32.63%)	(19.58%)
Average assignment distance	2.56km	2.29km
Panel B: Applicants maintaining 1st preference (281)		
Assigned in:	DA	DC
Any preference	144	112
	(51.25%)	(39.86%)
First preference	116	103
	(41.28%)	(36.65%)

Table C.2.6: Assignment In and Out Preferences under DA and DC Algorithms

Panel A: Share of Applicants With the Same of Different Assignment			
	Pre-School 1	Pre-School 2	Primary 1
Applicants assigned to same schools under DA and DC	967 (56.68%)	801 (47.06%)	261 (46.03%)
Applicants assigned to different schools under DA and DC	739 (43.32%)	901 (52.94%)	306 (53.97%)
Panel B: Outcomes of Each Group			
B.1.: Applicants assigned to same schools under DA and DC			
	Pre-School 1	Pre-School 2	Primary 1
Both DA and DC in preferences	941 (97.31%)	630 (78.65%)	113 (43.30%)
Both DA and DC out of preferences	26 (2.69%)	171 (21.35%)	148 (56.70%)
B.2.: Applicants assigned to different schools under DA and DC			
	Pre-School 1	Pre-School 2	Primary 1
DA in preferences and DC out of preferences	569 (77.00%)	478 (53.05%)	125 (40.85%)
DA out of preferences and DC in preferences	17 (2.30%)	16 (1.78%)	7 (2.29%)
Both DA and DC in preferences	144 (19.49%)	91 (10.10%)	27 (8.82%)
Both DA and DC out of preferences	9 (1.22%)	316 (35.07%)	147 (48.04%)

Table C.2.7: Transfer Requests and Changes Between Assignment in Enrollment

Aftermarket outcome	All Grades	Pre-School 1	Pre-School 2	Primary 1
Transfer requests	380	40	196	144
Enrollment ≠ Assignment	585	89	282	214
Transfer request & Enrollment ≠ Assignment	360	38	192	130
Applicants with enrollment data	**3,907**	**1,680**	**1,667**	**560**
Total Applicants	**3,985**	**1,710**	**1,705**	**570**

Table C.2.8: The Effect of Preference Assignment on Aftermarket Movements

Panel A: Congested 1st Pref.			
Probability (%)	(1)	(2)	(3)
β_0	4.31	8.65	9.19
	(0.85)	(1.01)	(1.01)
β_1	18.90	24.27	22.87
	(1.62)	(1.92)	(1.90)
Observations	**2,314**		
FE categories	**127**		
Panel B: Congested Last Pref.			
Probability (%)	(1)	(2)	(3)
β_0	4.14	9.17	9.61
	(0.91)	(1.10)	(1.10)
β_1	11.28	13.36	12.77
	(1.24)	(1.49)	(1.49)
Observations	**2,681**		
FE categories	**128**		

Notew: Column (1) shows the probability of a transfer request, Column (2) shows the probability of having a different assigned school and school of enrollment, and Column (3) shows the probability of either of the outcomes.

Standard errors in parenthesis

C.3 Robustness Checks

Table C.3.1: Estimates and Potential Scale Reduction Factors. Main Specification without Random Coefficients

	Pre-School 1		Pre-School 2		Primary 1	
Estimate	Mean (SD)	PSRF	Mean (SD)	PSRF	Mean (SD)	PSRF
Mean(ξ_j)	0.131		-1.480		0.075	
	(0.532)		(0.693)		(0.606)	
λ	3.486	1	4.719	1	4.301	1
	(0.396)		(0.561)		(0.919)	
σ_ξ	0.350	1.001	2.766	1	0.497	1
	(0.082)		(0.747)		(0.131)	
σ_ϵ	1.403	1	1.314	1	1.702	1
	(0.054)		(0.061)		(0.121)	
Tot. schools	55		57		54	
Tot. students	1,098		885		389	

Table C.3.2: Estimates and Potential Scale Reduction Factors. Main Specification without Siblings

	Pre-School 1		Pre-School 2		Primary 1	
Estimate	Mean (SD)	PSRF	Mean (SD)	PSRF	Mean (SD)	PSRF
Mean(ξ_j)	0.075		-1.116		0.047	
	(0.160)		(0.25)		(0.309)	
σ_ξ	0.269	1	1.642	1.018	0.495	1.001
	(0.061)		(0.525)		(0.130)	
σ_ϵ	0.991	1.003	0.886	1	1.270	1
	(0.042)		(0.045)		(0.097)	
σ_γ	0.292		0.266		0.144	
	(0.056)		(0.058)		(0.029)	
Tot. schools	55		57		54	
Tot. students	1,021		839		345	

Table C.3.3: Differences in Welfare: Student-Optimal vs. DC and DA algorithms. Specification without Random Coefficients

Panel A: All simulated applicants

	Pre-School 1		Pre-School 2		Primary 1	
Measure	Dist	DA	Dist	DA	Dist	DA
Δ Mean utility (km)	-0.773	-0.003	-0.456	-0.071	-0.317	-0.296
$({}^{\Delta Mean\ utility}/_{\sigma_{Ut.\ FB}})$	-0.750	-0.003	-0.177	-0.027	-0.091	-0.085

Panel B: Applicants with different assignments across algorithms

	Pre-School 1		Pre-School 2		Primary 1	
Measure	Dist	DA	Dist	DA	Dist	DA
Δ Mean utility (km)	-1.667	-0.006	-0.707	-0.109	-0.415	-0.388
$({}^{\Delta Mean\ utility}/_{\sigma_{Ut.\ FB}})$	-1.631	-0.006	-0.302	-0.047	-0.124	-0.116

Note: Δ Mean utility (km) is measured computing $u_{i,j(\mu)} - u_{i,j(TTC)}$, where $j(\mu)$ represents the school to which individual i is assigned under mechanism μ. We then compute average utilities for each algorithm and simulation and finally compute the average for each algorithm across simulations. $\frac{\Delta Mean\ utility}{\sigma_{Ut.\ FB}}$ simply uses the utility variance under the TTC mechanism to scale this difference in each simulation. This is done to facilitate extrapolations to other contexts.

Table C.3.4: Differences in welfare: Student-optimal vs DC and DA algorithms. Specification without Siblings

Panel A: All simulated applicants						
	Pre-School 1		Pre-School 2		Primary 1	
Measure	Dist	DA	Dist	DA	Dist	DA
Δ Mean utility (km)	-0.661	-0.004	-0.331	-0.073	-0.170	-0.339
$({}^{\Delta Mean\ utility}/_{\sigma_{Ut.\ FB}})$	-0.835	-0.005	-0.169	-0.037	-0.053	-0.105
Panel B: Applicants with different assignments across algorithms						
	Pre-School 1		Pre-School 2		Primary 1	
Measure	Dist	DA	Dist	DA	Dist	DA
Δ Mean utility (km)	-1.439	-0.008	-0.517	-0.113	-0.222	-0.441
$({}^{\Delta Mean\ utility}/_{\sigma_{Ut.\ FB}})$	-1.603	-0.009	-0.243	-0.053	-0.069	-0.136

Note: Δ Mean utility (km) is measured computing $u_{i,j(\mu)} - u_{i,j(TTC)}$, where $j(\mu)$ represents the school to which individual i is assigned under mechanism μ. We then compute average utilities for each algorithm and simulation and finally compute the average for each algorithm across simulations. $\frac{\Delta Mean\ utility}{\sigma_{Ut.\ FB}}$ simply uses the utility variance under the TTC mechanism to scale this difference in each simulation. This is done to facilitate extrapolations to other contexts.

Table C.3.5: Priorities and Assignments in DA: Potential Improvements for SIC and TTC

Panel A: Ranking of schools where an applicant has sibling priority(*)			
	Pre-School 1	Pre-School 2	Primary 1
1st preference	100	157	42
2nd preference	7	3	1
3rd preference	1	2	0
Panel B: Ranking of DA assignments for applicants with sibling priority below 1st preference			
	Pre-School 1	Pre-School 2	Primary 1
1st preference	8	3	0
2nd preference	0	2(**)	1(***)
3rd preference	0	0	0

Note: None of the potential applicants that could participate in an improvement cycle (Panel B) coincide in the programs to which they were applying, such that no cycles were attainable.
(*) Panel A shows the highest ranked program where applicants have a sibling priority. If an applicant has priority in both the 1st and 2nd preference, they will only appear in the 1st preference in this table.
(**) One of these two applicants had sibling priority in their second preference, and the other had sibling priority in their third preference.
(***) This applicant had sibling priority in their second preference.

C.4 Phases of the Distance-Centric Algorithm Implementation Process

The overall process started with the Preparation Phase, in which the Ministry of Education updated all school supply information (i.e., location, available spaces, closure or opening of educational programs, etc.).

In the second, or Registration Phase, families registered their children on a website in order to be granted a spot in a public school. Legal guardians needed to indicate the type of registration (individual or sibling group), the grade level to be attended, any older siblings already enrolled in the public school system, special educational needs, and nationality. They also provided their electricity bill number so as to be geo-located.

This was followed by the Assignment Phase and then the Consultation Phase, during which time families could enter the website to see their school assignments. Finally, the fifth and sixth phases consisted of the School Change Petitions Phase and Continuous Enrollment. Applicants could ask to change schools if there were spaces available, and they could also enroll in a given school once the academic year had already started.

C.5 Full Utility Specification

$$
\begin{aligned}
u_{ij} &= S_{ij}\lambda + \delta_j - d_{ij} + \gamma_i d_{ij} + \epsilon_{ij} \\
\delta_j &= \overline{\delta} + \xi_j \\
\overline{\delta} &\equiv 0
\end{aligned}
$$

$$
\begin{aligned}
\lambda &\sim \mathcal{N}(0, \sigma_\lambda) \\
\gamma_i &\sim \mathcal{N}(0, \sigma_\gamma) \\
\xi_j &\sim \mathcal{N}(0, \sigma_\xi) \\
\epsilon_{i,j} &\sim \mathcal{N}(0, \sigma_\epsilon) \\
\sigma_\gamma &\sim IW(\tau_\gamma, df_\gamma) \\
\sigma_\xi &\sim IW(\tau_\xi, df_\xi) \\
\sigma_\epsilon &\sim IW(\tau_\epsilon, df_\epsilon)
\end{aligned}
$$

We follow Rossi et al. (1996) and Abdulkadiroğlu et al. (2017) in using disperse priors. The only exception is the use of a smaller τ_γ, given that in this context it is reasonable to impose a smaller prior on the mean variance of the parameter, considering that $\gamma_i > 1$ would imply that a family actually prefers schools farther away from home. Specifically, we use $\sigma_\lambda = 100$, $\tau_\gamma = \frac{1}{5}(3 + n_{schools})$, $df_\gamma = 2(3 + n_{schools})$,[1] $\tau_\xi = 1$, ${}_\xi = 2$, $\tau_\epsilon = 3 + n_{schools}$, and $df_\epsilon = 3 + n_{schools}$.

C.6 DA-SIC and TTC Equivalence in our Context

As shown in Table C.3.5, there is no potential for priority trading cycles.

The Top Trading Cycles (TTC) algorithm includes the possibility of trading priorities between applicants, which happens when they prefer the alternatives in which they do not have the priority more than ones in which they do, and are thus "willing to trade" the priority. In other words, TTC has the potential to provide improvements over SIC, when there is not a complete correlation between priorities and preferences. In our case, for the priority at declared preferences (over non-preferences imputed by distance), the correlation is one since these are always ranked higher. Thus, the only possibility for the TTC algorithm to improve over the SIC algorithm is to find trades involving the static sibling priority. However, as shown in Table C.3.5 (and explained in the footnote), that is not feasible.

To illustrate this, imagine a system with two schools (A and B), both with only one vacancy, and three applicants(i, j and k). i has priority in A but prefers B over A.j has priority in school B, but prefers A over B. k has priority in both schools, prefers A over B, and has the worst lottery number of the system. The result of the DA and SIC assignment would be i assigned to A and j assigned to B. The TTC algorithm would allow them to trade their priorities and switch their assignments. With that assignment switch, applicant k is now unassigned but has a higher priority in both schools that rejected him (higher priority pre-trade, of course). Such a situation can only arise when the correlation between preference and priority is not one, thus leaving room to trade the priority and get a better assignment.

[1] This implies that the mean of the σ_γ prior is 0.1.