# Segregation and affirmative action in school choice *

Juan F. Escobar† Leonel Huerta‡

July 12, 2024

**Abstract**

Segregation in schools is prevalent in cities around the world. We analyze the impact of affirmative action policies commonly used in centralized school choice on segregation and efficiency. In a large market model, we show that minority reserves–which guarantee a number of seats to minority students–are an effective tool for reducing segregation in schools. Minority reserves also increase the number of students assigned to their first preferences and improve efficiency. The cost of increasing minority reserves is leaving more students assigned to unattractive schools. The theoretical predictions are confirmed by simulations using data from school choice programs in Chile.

## 1 Introduction

Many school choice programs around the world use centralized procedures to assign students to schools. Based on the Gale-Shapley deferred acceptance algorithm (Gale and Shapley 1962), these procedures result in assignment processes that are considered successful by scholars and policymakers. Yet, our understanding of the impact that affirmative action policies have on segregation and different market outcomes is rather limited, despite the fact that segregation in schools is a key societal and economic problem. Indeed, segregation in schools impacts both learning outcomes and social attitudes (Karsten 2010, Rao 2019) and, as policy debates recognize, is in part determined by the algorithm

*All data used in this work is publicly available from the Chilean Ministry of Education.

†Department of Industrial Engineering, University of Chile, Beauchef 851, Santiago, Chile. E-mail: juanescobar@uchile.cl.
‡Department of Industrial Engineering, University of Chile, Beauchef 851, Santiago, Chile. E-mail: leonel.huerta@uchile.cl.

used to admit students.[1] This paper explores some of the tradeoffs policy makers face when reducing segregation in school choice programs using affirmative action policies.

We focus on the impact of minority reserves on segregation and efficiency in school choice programs. Minority reserves guarantee a given number of seats to minority students and respect each school ranking otherwise (Hafalir, Yenmez, and Yildirim 2013, Ehlers, Hafalir, Yenmez, and Yildirim 2014, Echenique and Yenmez 2015). These reserves are a simple and transparent way to include diversity considerations into centralized school choice programs. We use simulations from the Chilean centralized school choice system and derive theoretical results to understand the impact that minority reserves have on several market outcomes, including segregation, the number of students assigned to their top schools, and efficiency.

Our analysis is motivated by data and simulations from the centralized school choice system used in Chile. The Chilean system reserves 15% of the seats in each school to socially disadvantaged minority students and assigns students to schools using the Gale-Shapley deferred acceptance algorithm. By simulating the system for different minority reserves, we show that reserves can be an important tool to reduce segregation in schools. More subtly, minority reserves increase the number of students assigned to their top schools and also improve the overall efficiency of the final assignment. However, the simulations reveal that the main cost of minority reserves is to leave more students assigned to less attractive schools or unassigned.

To understand the impact of minority reserves on important market outcomes, we explore a large market model in which a continuum of students apply to a finite number of schools (Abdulkadiroglu, Che, and Yasuda 2015, Azevedo and Leshno 2016). A student is either regular or minority. Schools fall under two tiers, 1 and 2. Each school ranks students randomly. Tier 1 schools are popular and over-demanded, while tier 2 schools are unpopular and under-demanded. Our model assumes that regular students apply more intensely to tier 1 schools than minority students. As we show, this assumption is spported by data from the Chilean school choice system. It is also consistent with evidence from school systems in Europe and the US (Hastings, Kane, and Staiger 2009, Laverde 2020, Oosterbeek, Sóvágó, and van der Klaauw 2021).

We demonstrate that minority reserves benefit minority students and reduce segregation, unless an excessive number of seats are reserved for minority students. An important insight from our main comparative statics results is that the effect of minority reserves on efficiency measures depends critically on market specifics. Our main result shows that in markets where popular schools have

[1]For example, segregation in the current New York City centralized school match is an issue of intense debate and some proponents argue that the city should modify the algorithm used to assign students. See the New York Times story at https://www.nytimes.com/2021/03/09/nyregion/nyc-schools-segregation-lawsuit.html. See also the discussion in Alvin Roth's blog at https://marketdesigner.blogspot.com/2019/04/should-nyc-school-choice-diversify.html.

relatively few seats and minority students apply less frequently to these popular schools, increasing minority reserves results in more students assigned to their top-choice schools. In contrast, in markets with abundant capacity or where minority and regular students concentrate their applications on different popular schools, increasing minority reserves decreases the number of students assigned to their top-choice schools.

We also measure the inefficiency of the assignment by the number of students in Pareto improving pairs. Two students form a Pareto improving pair if, by swapping schools, both are better off. We show conditions under which increasing minority reserves improves the efficiency of the system by reducing the number of students in Pareto improving pairs.

The theoretical analysis additionally exposes the costs of reducing segregation in schools by using minority reserves. We show that increasing minority reserves leaves more students assigned to less attractive schools. Formally, we prove that as minority reserves increase, the cumulative rank distributions cross and, as a result, cannot be compared in the first order stochastic dominance sense. In particular, increasing the total number of students assigned to unattractive schools is an important consequence of a rise in minority reserves.

Our comparative statics results are the main theoretical contribution of the paper. Although our model is stylized, each of our theoretical results links up with real-world data from the school assignment processes in Chile. We use our theory and simulations to show that the design of minority reserves is an important and distinctive policy decision.

Segregation in schools is pervasive in cities around the world. Centralized school choice programs are often seen as providing equal access to schools. However, recent research shows that systematic differences in the application patterns of different groups may limit the efficacy of centralized school choice programs at reducing social, ethnic, or racial segregation in schools (Laverde 2020, Kutscher, Nath, and Urzua 2020, Son 2020, Oosterbeek, Sóvágó, and van der Klaauw 2021).[2] We build from the premise that centralized school choice alone may not be enough to integrate schools and explore the trade-offs faced when designing minority reserves.

Our analysis has important practical implications. The number of students assigned to their top schools is typically reported by districts implementing centralized school choice platforms (Featherstone 2020). Moreover, authorities in New York and Boston have made explicit algorithmic decisions to maximize the number of students in their top schools (Abdulkadiroğlu, Pathak, and Roth 2009, Abdulka-

[2]Laverde (2020) shows that in some dimensions the outcome of the centralized school choice program in Boston is similar to the outcome generated by an assignment based on proximity between residences and schools. Kutscher, Nath, and Urzua (2020) show that the introduction of the centralized school choice program in Chile has had a limited impact on segregation in schools. Oosterbeek, Sóvágó, and van der Klaauw (2021) show that differences in application patterns explain a substantial fraction of segregation in secondary schools in Amsterdam.

diroglu, Pathak, Roth, and Sönmez 2006).[3] The number of Pareto improving pairs in the assignment is an inefficiency measure authorities in Amsterdam have also looked at (Ashlagi and Nikzad 2020). Thus our results are of interest to policymakers who may combat segregation in schools, increase the number of families obtaining their top choices, reduce the number of applicants in Pareto improving pairs, but incur the costs of leaving more students assigned to unattractive schools.

Abdulkadiroğlu and Sönmez (2003) apply matching theory to school choice problems. The school choice literature has shown that the design of matching mechanisms involves complex tradeoffs (Gale and Shapley 1962, Roth and Sotomayor 1990, Che and Tercieux 2019, Leshno and Lo 2021). Reducing segregation in schools is another important desideratum and our results expose new forces that seem important for practical implementations of matching algorithms.

Our paper contributes to the extensive literature on matching problems with diversity considerations, including work by Abdulkadiroğlu (2005), Kojima (2012), Westkamp (2013), Hafalir, Yenmez, and Yildirim (2013), Ehlers, Hafalir, Yenmez, and Yildirim (2014), Echenique and Yenmez (2015), Kominers and Sönmez (2016), Fragiadakis and Troyan (2017), Dur, Kominers, Pathak, and Sönmez (2018), Nguyen and Vohra (2019), Dur, Pathak, and Sönmez (2020), Aygün and Turhan (2020), Pathak, Sönmez, Ünver, and Yenmez (2020), Rios, Larroucau, Parra, and Cominetti (2021), Sönmez and Yenmez (2022), and Aygun and Bó (2021). Throughout the paper, we borrow some definitions and concepts from these works. In particular, our formulation of minority reserves and soft bounds follow Hafalir, Yenmez, and Yildirim (2013). Hafalir, Yenmez, and Yildirim (2013) also show that the introduction of minority reserves favors at least one minority student and, under strong restrictions on priorities and preferences, that all minority students are better off when reserves are introduced. An important difference between the work by Hafalir, Yenmez, and Yildirim (2013) and ours is that we explore the impact of minority reserves on outcomes that are key for policy design but have been neglected by previous research, such as segregation, the rank distributions of assignments, and the number of applicants in Pareto improving pairs.

We obtain comparative statics results in a large market model to provide guidance on market design questions. This paper is thus related to the growing literature using tractable large market models to shed light on new market design issues; see Abdulkadiroglu, Che, and Yasuda (2015), Azevedo and Leshno (2016), Ashlagi and Nikzad (2020), Che and Tercieux (2019), Leshno and Lo (2021).

The rest of the paper is organized as follows. Section 2 describes the Chilean setting and shows

[3] As Abdulkadiroğlu, Pathak, and Roth (2009) observe: "The greater number of students obtaining one of their top choices in a similar simulation and in the first year of submitted preference data convinced New York City to employ a single tiebreaker in their assignment system." When discussing the Boston school choice experience, Abdulkadiroglu, Pathak, Roth, and Sönmez (2006) argue that "the ability to tell the public that a high proportion of students receive their top choices may be a reason for the widespread popularity of the Boston mechanism."

some motivating simulations. Section 3 presents our model. Section 4 introduces minority reserves and provides our main comparative statics results. Section 4 also presents variations of our model. Section 5 discusses our findings. Section 6 concludes. The Appendix contains supporting material.

# 2 Motivation

Chile initiated a gradual transition from its decentralized voucher-based school choice system to a centralized model in 2016. Before 2016, the system worked as a decentralized voucher market in which schools selected applicants based on non-transparent criteria. According to policymakers and politicians that pushed for reform, the decentralized nature of the system was the culprit for the high levels of segregation in Chilean schools. As noted by Kutscher, Nath, and Urzua (2020) and Honey and Carrasco (2022), the new centralized system has had a modest impact on segregation in Chilean schools. Naturally, this situation raises significant questions about the design of algorithms and policy tools, and the trade-offs policymakers must solve, to address school segregation.

## 2.1 Centralized school choice in Chilean cities

To get admission to schools, students access a platform and fill a rank order list. The law regulating admission reserves 15% of seats in each school to socially disadvantaged minority students.[4] The minority reserve policy is a soft target that the design tries to achieve but may fail if the target is too high (Hafalir, Yenmez, and Yildirim 2013). The reserve policy is an explicit attempt to promote social inclusion in schools.

Schools rank students using a variety of criteria, but many of them are relevant for a small fraction of the applicants. Many students cannot be ranked by schools simply using any of the priority criteria.[5] In those cases, each school runs a lottery over its whole set of applicants. Students are then assigned to schools by running a Gale-Shapley deferred acceptance algorithm with minority reserves (Gale and Shapley 1962, Hafalir, Yenmez, and Yildirim 2013). The assignment algorithm runs as follows:

Step 1: Each student proposes to her first choice. Each school tentatively assigns seats to its proposers, following the priority and lottery orders, and reserving 15% of its seats to minority students. Remaining proposers are rejected.

Step $k$: Each student rejected in the previous step proposes to her next best choice. Each school considers

[4] A student is considered minority if her social background impairs her learning process and educational outcomes. See https://sep.mineduc.cl/alumnos-prioritarios-preferente/ for details.

[5] For details on the Chilean system, see Correa, Epstein, Escobar, Rios, Bahamondes, Bonet, Epstein, Aramayo, Castillo, and Cristi (2022).

the students it has been holding together with its new proposers and tentatively assigns its seats following the priority and lottery orders and the minority reserves. Any remaining proposers are rejected. Go to Step $k + 1$.

The algorithm terminates either when there are no new proposals or when all rejected students have exhausted their preference lists.

We focus on the admission process for Pre-Kindergarten during 2019 in the main urban centers in Chile: Valparaíso, Concepción, Santiago.[6] Appendix B provides details about the data. Table 7 presents summary statistics for each market. Note that in each of the cities, the percentage of minority students far exceeds the current minority reserve of 15%.[7]

Table 1: Valparaiso, Concepcion and Santiago markets

| | Valparaíso | Concepción | Santiago |
|---|---|---|---|
| Number of schools | 275 | 250 | 1,214 |
| Total capacity (seats) | 8,754 | 9,199 | 56,331 |
| Number of students | 6,819 | 7,523 | 49,108 |
| Minority students | 2,994 (43.91%) | 3,233 (42.97%) | 18,399 (37.47%) |

## 2.2 Minority reserves and market outcomes

We now present simulation results for Valparaíso, Concepción and Santiago. For each market, we run the algorithm used by the Ministry of Education for different minority reserves.[8] Concretely, for each market and for each $f \in \{0, \ldots, 100\}$, we run independent simulations of the algorithm where the minority reserve in each school equals $f$% of its seats. Since reserves are soft, when minorities demand for a popular school is below the target, those seats are open to other students. For example, suppose that $f = 70\%$, a school has 100 seats, but only 10 minority students list that school. In this case, at least 90 seats will be available to regular students.

We assume that variations in the algorithm do not change applications. This assumption is justified since in Chile applicants are allowed to submit rank order lists of arbitrary length and the deferred acceptance algorithm with minority reserves is strategy-proof (Hafalir, Yenmez, and Yildirim 2013).

[6] In Appendix D.2, we also run simulations for 9th grade and obtain similar results. We have also obtained similar results in simulations using data from smaller Chilean cities.

[7] Each of our markets include some rural areas in which the supply of schools is limited and therefore naturally families apply to one or two schools. However, virtually all students in our sample live in urban centers.

[8] In particular, our simulation considers all the criteria used by the Ministry of Education to rank students in each school, including special needs and sibling priority.

To measure segregation, we compute the Duncan segregation index (Duncan and Duncan 1955):

$$D = \frac{1}{2}\sum_{c\in C}\left|\frac{\text{Regular students in } c}{\text{Total number regular students}} - \frac{\text{Minority students in } c}{\text{Total number minority students}}\right|$$

The Duncan index measures the fraction of minority students that need to be reassigned so that every school has the same ratio of students of each group. The Duncan index equals 0 under perfect integration, and 1 under perfect segregation.

For each allocation, we compute the Duncan segregation index, the fraction of students assigned to their top schools, and the fraction of students assigned to their top three choices. Figures 1 and 2 illustrate our simulations.

The simulations show three important results. First, schools are segregated. School choice alone is not enough to eliminate segregation in Chilean schools, even when schools rank the majority of their students by running independent lotteries. We have also simulated the system by generating random priorities (rather than the actual, almost-random, priorities used by schools in Chile), and have obtained approximately the same results.[9] Systematic differences in application patterns of regular and minority students are the key drivers of segregation in schools.

Second, segregation is U-shaped. As Figure 1 shows, the Duncan index is minimized close to the fraction of minority students in each city.[10]

Third, as the minority reserve increases, more students are assigned to their top schools. Yet, increasing reserves also leaves fewer students assigned to their top three schools (and thus more students assigned to schools ranked 4 or worse).

[9] In our simulations with random priorities and a reserve of 15%, the Duncan index in Santiago equals 0.304 (0.002). As shown in Table 3, when using actual priorities, the Duncan index in Santiago equals 0.303 (0.002).

[10] Our theoretical results show that segregation is minimized at a reserve that equals the fraction of minority students in the market. In our simulations, in contrast, segregation is minimized slightly to the right of our theoretical prediction. Two reasons explain this. First, around 2% of seats in each market are reserved for students with special needs. The fraction of seats reserved for minorities is computed excluding the special needs seats and minorities are underrepresented among students with special needs. Second, to compute the Duncan index we are ignoring unassigned students. Minorities are underrepresented among unassigned students.

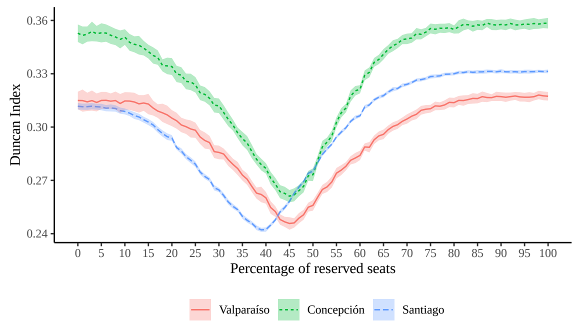
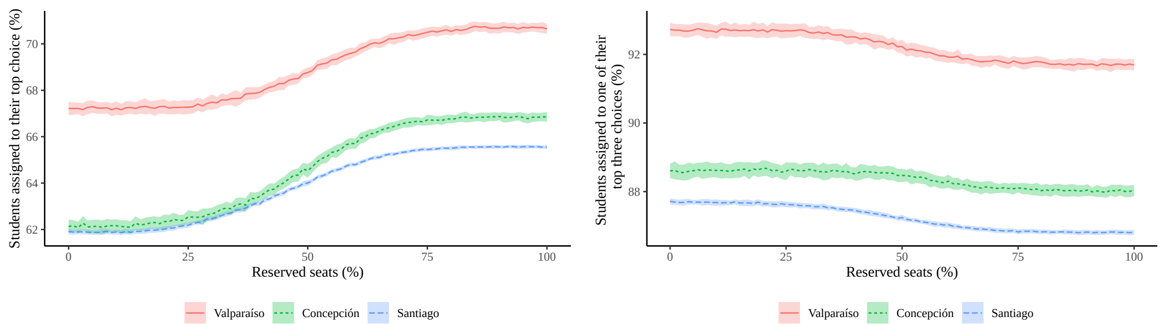



Figure 1: Duncan segregation index.




(a) Percentage of students assigned to top schools.

(b) Percentage of students assigned to one of their top three schools.

Figure 2: Students assigned to their top choices and top three choices.

## 2.3 Schools and application patterns

In this Subsection, we describe the differences in application patterns between minority and non-minority students.

### 2.3.1 Popular schools

Schools face different demand levels. Given students' preferences, we measure the popularity of each school $c$ as the ratio between the number of students for whom school $c$ is their top choice and the capacity of school $c$. More formally, let $p(c)$ be the number of students that list school $c$ as their top choice and let $q_c$ be the number of seats that school $c$ has. The *popularity* of school $c$ is given by $pop(c) = \frac{p(c)}{q_c}$. Table 2 shows the popularity of schools in our markets. We say that school $c$ is *popular* if $pop(c) \geq 1$. A popular school will fill its seats under different variations of the deferred acceptance algorithm. As shown in Table 2, close to 1/4 of schools in Santiago and Valparaíso are popular schools.

| | Valparaíso | Concepción | Santiago |
|---|---|---|---|
| Median | 0.53 | 0.44 | 0.60 |
| Mean | 0.74 | 0.67 | 0.80 |
| Third quartile | 0.96 | 0.86 | 1.03 |

Table 2: Popularity of schools

In Appendix B.2, we show that all results in this Section (including Figures 3 and 4 below) extend to alternative definitions of popular schools.

### 2.3.2 Differences in application patterns

For each student $s$ in our database, we define her *application intensity* as the number of popular schools in the top of her application. Formally, letting $(c_1, c_2, \ldots, c_L)$ be the rank order list of student $s$, we define her application intensity as

$$l(s) = \begin{cases} \max\{i \mid c_1, \ldots, c_{i-1}, c_i \text{ are all popular schools.}\} & \text{if } c_1 \text{ is popular} \\ 0 & \text{if not.} \end{cases}$$

The application intensity $l(s)$ of student $s$ is the number of popular schools student $s$ competes for. Since school $c_{l(s)+1}$ is not popular, if student $s$ is rejected from all her top $l(s)$ schools then she will likely get accepted to school $l(s)+1$ in her list. In particular, whether student $s$ ranks other popular schools below $c_{l(s)+1}$ is likely to be irrelevant for the assignment.

For each group of students $t \in \{r, m\}$ ($r$ for regular and $m$ for minority) and $l \in \mathbb{N}$, we compute the *empirical distribution function of application intensities* as

$$\hat{\Pi}_t(l) = \frac{|\{s \text{ is in group } t \mid l(s) \leq l\}|}{|\{s \text{ is in group } t\}|}. \tag{2.1}$$

In each city, $\hat{\Pi}_t(l)$ is the fraction of group $t$ students with application intensities weakly less than $l$. Figure 3 shows the distributions of application intensities. Notably, in each city, $\hat{\Pi}_r$ is larger than $\hat{\Pi}_m$ in the first order stochastic dominance sense. Thus minority students apply with lower intensity to popular schools than regular students.[11] We also compute the distributions of application intensities

[11] We are agnostic about why minority students apply less to popular schools. Online Appendix F shows that minority students tend to live farther away from popular schools. We also show that students attending schools with higher popularity tend to perform better in standarized tests.

conditional on $l \geq 1$ as

$$\hat{\Pi}_t(l \mid l \geq 1) = \frac{|\{s \text{ is in group } t \mid 1 \leq l(s) \leq l\}|}{|\{s \text{ is in group } t \text{ and } l(s) \geq 1\}|}.$$

As shown in Figure 4, $\hat{\Pi}_r(\cdot \mid l \geq 1)$ dominates $\hat{\Pi}_m(\cdot \mid l \geq 1)$. Even restricting attention to students that apply first to popular schools, regular students apply more to popular schools.[12]

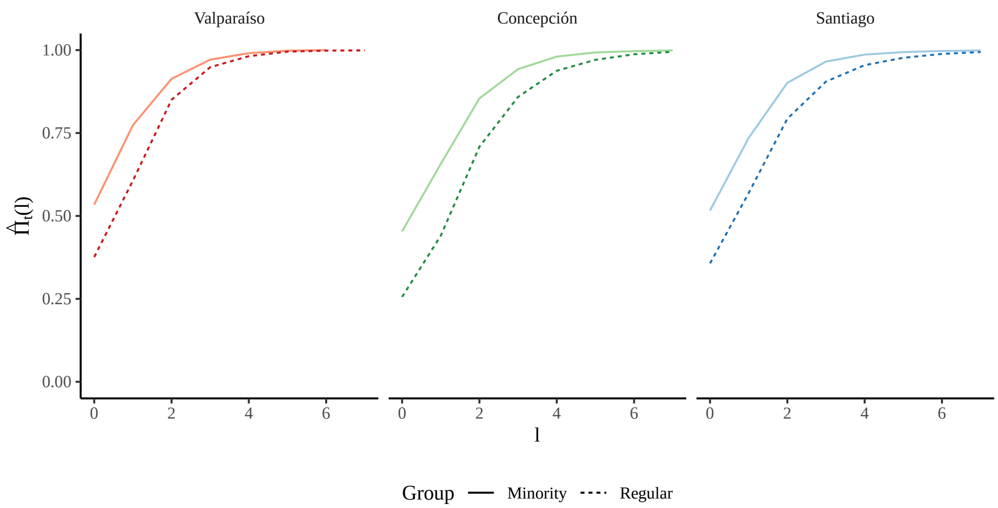



Figure 3: Empirical distributions of application intensities $\hat{\Pi}_t(\cdot)$ for each group $t$.

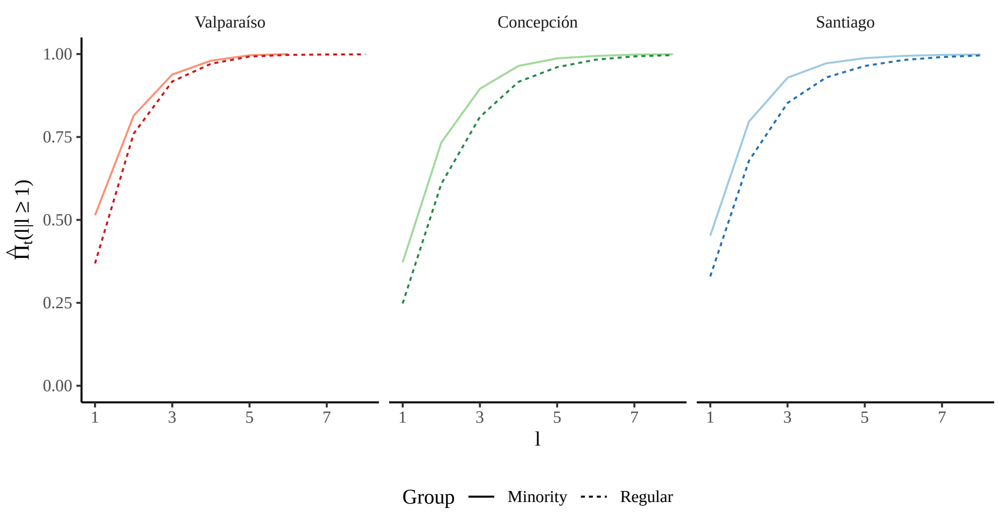



Figure 4: Empirical distributions of application intensities $\hat{\Pi}_t(\cdot \mid l \geq 1)$ conditional on $l \geq 1$.

[12] Appendix C shows that the higher the popularity of a school, the higher the fraction of regular students that demand that school.

# 3 Model

In this Section, we introduce a tractable large market model. Our goal is to better understand the impact of minority reserves on several market outcomes. In markets in which the final allocation is the result of the deferred acceptance algorithm, raising reserves causes rejections that, by triggering new applications and chains of rejections, make it difficult to derive comparative statics results. Our large market model is particularly suitable to deal with these difficulties.

## 3.1 Environment

We consider a school choice problem with a continuum of students and a finite number of schools (Abdulkadiroglu, Che, and Yasuda 2015, Azevedo and Leshno 2016). There is a measure 1 of regular students ($r$), and a measure $\beta > 0$ of minority students ($m$).[13] A student is characterized by $s = (t, x) \in (\{r\} \times [0, 1]) \cup (\{m\} \times [0, \beta])$. The set of all students is denoted $S$.

The set of schools is $C = \{1, \ldots, n, n + 1, \ldots, n + N\}$. Schools $c \in C_1 = \{1, \ldots, n\}$ are tier 1 , while schools $c \in C_2 = \{n + 1, \ldots, n + N\}$ are tier 2. Tier $i$ schools have capacity $k_i$.

For $l \in \{0, 1, \ldots, n\}$, we define the set $Z(l)$ of complete and transitive preferences over schools such that the $l$-most prefered schools are all tier 1 schools, but the school ranked $l + 1$ is tier 2. [14] In particular, $Z(n)$ is the set of all preferences $\succ$ such that for all $c_1 \in C_1$ and all $c_2 \in C_2$, $c_1 \succ c_2$.

We have specified preferences that rank all schools. In practice, families do not rank all schools. We could specify the set $Z(l)$ as the set of all preferences ranking only $l + 1$ schools, such that the $l$-most prefered schools are all tier 1, but the school ranked $l + 1$ is tier 2. All our results hold for this alternative model.

For $t \in \{r, m\}$ and $l \in \{0, \ldots, n\}$, a fraction $\pi_t(l) \in [0, 1]$ of group $t$ students have preferences uniformly distributed over $Z(l)$, with $\sum_{l=0}^{n} \pi(l) = 1$. In particular, a fraction $\pi_t(0)$ of type $t$ students list a tier 2 school as first choice. We denote by $\Pi_t(l)$ the fraction of type $t$ students that list $l$ or less tier 1 schools:

$$\Pi_t(l) = \sum_{l' \leq l} \pi_t(l').$$

We call $\Pi_t$ the *distribution of application intensities* for type $t$. The distribution $\Pi_t$ is the theoretical analog of the empirical distribution $\hat{\Pi}_t$ constructed in (2.1). For both types, the preference profile of a student $(t, x)$ is entirely determined by $(t, x)$.[15]

[13]Strictly speaking, the total number of minority students could exceed regular students. As will be clear later, all what matters for our results is that minority students are under-represented in over-demanded schools.

[14]In particular, for a preference that belongs to $Z(l)$, with $l \leq n - 1$, the school ranked $l + 2$ could be tier 1 or tier 2.

[15]For example, for each $t \in \{r, m\}$, we can divide the interval $[0, 1]$ in a finite number of disjoint intervals, each of them corresponding to a preference profile in $\cup_{l=0}^{L_t} Z(l)$.

Tier 1 schools are over demanded, but the total capacity of the market exceeds total demand. We thus assume that $nk_1 \leq (1 - \Pi_r(0)) + \beta(1 - \Pi_m(0))$ and $nk_1 + Nk_2 > 1 + \beta$.

We assume that $\Pi_r$ first order stochastically dominates $\Pi_m$. In other words, for all $l \in \{0, 1, \ldots, n\}$,

$$\Pi_r(l) \leq \Pi_m(l). \tag{3.1}$$

This assumption is motivated by Figure 3. Condition (3.1) captures the idea that minority students are less likely to apply to popular and high-demand schools (Hastings, Kane, and Staiger 2009). This assumption is relevant in centralized school choice programs in cities in Europe and the US. For example, Laverde (2020) shows that white families are more likely than Black and Hispanic families to rank high-achievement schools in Boston Public Schools choice system. Oosterbeek, Sóvágó, and van der Klaauw (2021) also provide evidence of heterogeneity in school preferences between students from different backgrounds in the Amsterdam centralized system. Letting $L_t = \min\{L \mid \Pi_t(L) = 1\}$ and noting that $\Pi_r$ dominates $\Pi_m$, we deduce that $L_m \leq L_r$.

An implicit assumption in our model is that all popular schools tend to be more demanded by regular students. In practice, some popular schools are particularly attractive for minority students. However, Figure 9 in Appendix C shows a significant correlation between the popularity of a school and the fraction of regular students that demand that school. So, our assumption is a good approximation at least for the Chilean system. Moreover, Section 4.4.1 shows that when some popular schools are attractive for minority students (so that the popularity of a school does not correlate with the fraction of regular students that find the school more attractive), our comparative statics results change radically and cannot explain the patterns exposited in Figure 2.

In many school choice systems, schools rank students independently and uniformly. We allow some more generality and assume that ranks are not necessarily uniform. A student $s = (t, x)$ at each school $c$ draws $\omega_c^s \in [0, 1]$ independently from the cumulative distribution $G_t$ on $[0, 1]$, with derivative $G_t' = g_t > 0$. A higher number $\omega_c^s$ implies that the student has higher priority in school $c$. We will refer to $\omega_c^s$ as the *score* that student $s$ has in school $c$. We assume that regular students tend to have higher scores than minority students so that $G_r$ dominates $G_m$ in the first order stochastic sense: $G_r(\omega) \leq G_m(\omega)$ for all $\omega \in [0, 1]$. The assumption that $G_r(\omega) \leq G_m(\omega)$ is appropriate in school choice programs in which schools rank students according to academic performance, or in school systems in which siblings or children whose parents work in the school have higher priority. Under all these criteria, a minority student is weakly less likely to be highly ranked in a school than a regular student. In many school choice programs in the world (including the Chilean system), schools rank all students uniformly. Obviously, the case of random priorities in school choice is a special case of our model which is obtained by setting $G_r = G_m$ equals the uniform distribution on $[0, 1]$.

Our two-tier model is natural in school choice applications, in which a tier 1 school tends to be more attractive than a tier 2 school for all students. However, a given minority student is less likely than a regular student to apply to a tier 1 school. When $\beta = 0$ and $\Pi_r$ is uniformly distributed in $\{1, \ldots, n\}$, our model has only one group and all students in the group prefer all tier 1 schools over any tier 2 school. In this case, our model is analogous to the limit models in Che and Tercieux (2019) and Ashlagi and Nikzad (2020). We extend these models to accomodate different groups of students in the market and distinct application patterns.

## 3.2 Matchings and cutoffs

A matching is a function $\mu \colon S \cup C \to C \cup 2^S$ such

i. For all $s \in S$, $\mu(s) \in C$;

ii. For all $c \in C_i$, $\mu(c) \subseteq S$ with $|\{s \mid \mu(s) = c\}| \leq k_i$;

iii. For all $c \in C$ and all $s \in S$, $\mu(s) = c$ iff $s \in \mu(c)$.

The first condition says that each student is assigned to a school, the second condition says that each school is assigned to a measure of students that does not exceed its capacity, the third condition says that a student is assigned to a school iff the school is assigned to that student. A matching $\mu$ is *stable* if for all $c \in C_i$ and all $s = (t, x) \in S$ with $c \succ_s \mu(s)$, the following two conditions hold: (i) $|\{s \mid \mu(s) = c\}| = k_i$; and (ii) $\omega_c^s < \omega_c^{s'}$ for all $s' = (t', x')$ with $\mu(s') = c$. Intuitively, a matching is stable if there is no pair $(s, c)$ that can block the matching.[16]

Following Abdulkadiroglu, Che, and Yasuda (2015) and Azevedo and Leshno (2016), we can characterize a stable matching by means of admission cutoffs $p_c \in [0, 1]$, for all $c \in C$. A cutoff $p_c$ determines the lowest lottery number $\omega_c$ that a student can have to be admitted to school $c$. The highest the cutoff $p_c$, the harder it is to get to school $c$. Two observations simplify the characterization of cutoffs. First, schools within the same tier are symmetric and therefore $p_c = p_{c'}$ for all $c, c' \in C_i$.[17] Second, in any stable matching a tier two school will have excess capacity and therefore its cutoff will equal 0. We can therefore characterize a stable matching by means of a single cutoff $p$ that clears the

[16] This notion defines a stable matching up to a measure 0 set of students. The market outcomes we analyze are not altered by this ambiguity. See Azevedo and Leshno (2016) for discussion.

[17] If two tier 1 schools had different cutoffs, one of them would have excess demand or excess capacity. If the market clearing condition is satisfied in the school with the highest (resp. lowest) cutoff, then the school with the smallest (resp. highest) cutoff has excess demand (resp. capacity). As a result, the matching would not be stable. This observation also applies to the model with minority reserves studied in Section 4.

market for tier 1 schools:

$$\sum_{l=1}^{L_r}\pi_r(l)\sum_{q=1}^{l}\frac{1}{n}G_r(p)^{q-1}(1-G_r(p))+\beta\sum_{l=1}^{L_m}\pi_m(l)\sum_{q=1}^{l}\frac{1}{n}G_m(p)^{q-1}(1-G_m(p))=k_1. \tag{3.2}$$

The left hand side in equation (3.2) is the demand for a school $c\in C_1$ when the admission cutoff in all schools is $p$. The first term on the left hand side of (3.2) is the demand for school $c$ of regular students. For each school $c\in C_1$ and $l\in\{1,\ldots,L_r\}$, a fraction $\pi_r(l)$ of regular students will rank $l$ schools. A regular student ranking $l$ tier 1 schools will rank school $c$ in the $q$-th position with probability $1/n$, for $q\in\{1,\ldots,l\}$. A student that ranks school $c$ in the $q$-th position will demand school $c$ if her scores in schools ranked above $c$ are below the cutoffs (which happens with probability $G_r(p)^{q-1}$) but her score in school $c$ is above the cutoff (which happens with probability $1-G_r(p)$). The second term on the left of (3.2) is the demand for school $c$ from minority students and is analogously computed.

For each $t\in\{r,m\}$ and function $f\colon\{0,\ldots,L_t\}$, we write $\mathbb{E}_t[f(l_t)]=\sum_{l_t=0}^{L_t}\pi_t(l_t)f(l_t)$. A solution $\bar{p}\in[0,1]$ to equation (3.2) also solves:

$$\frac{1}{n}\mathbb{E}_r[1-(G_r(\bar{p}))^{l_r}]+\frac{\beta}{n}\mathbb{E}_m[1-(G_m(\bar{p}))^{l_m}]=k_1 \tag{3.3}$$

It is relatively simple to show that equation (3.3) has a unique solution $\bar{p}\in]0,1[$ (for which in general there is no closed form solution). Naturally, $\bar{p}$ increases when the supply of tier 1 schools, $nk_1$, decreases. The market clearing cutoff also increases when the distribution $\Pi_t$ increases in the first order stochastic dominance sense.

In the unique stable matching, minority students are underrepresented in tier 1 schools. Indeed, the ratio of minority to regular students in the whole population equals $\beta$, while the ratio of minority to regular students assigned to a tier 1 school is

$$\frac{\beta\mathbb{E}_m[1-(G_m(\bar{p}))^{l_m}]}{\mathbb{E}_r[1-(G_r(\bar{p}))^{l_r}]}<\beta.$$

To see this, note that $\mathbb{E}_m[1-G_m(\bar{p})^{l_m}]\leq\mathbb{E}_m[1-G_r(\bar{p})^{l_m}]<\mathbb{E}_r[1-G_r(\bar{p})^{l_r}]$. The first inequality follows since a minority student tends to have lower scores so $G_m(\bar{p})\geq G_r(\bar{p})$. The second inequality follows since $\Pi_m$ is dominated by $\Pi_r$ and $1-(G_r(\bar{p}))^l$ is increasing in $l$. These forces combine to result in school segregation.

# 4 Minority reserves, segregation and efficiency

We now introduce minority reserves and present our main results. Our main results show that the application patterns exposited in Figures 3 and 4 are a key driver behind the simulations illustrated in Figures 1 and 2. Subsection 4.1 introduces and characterizes stable matchings under minority reserves. Subsections 4.2 and 4.3 state and discuss our comparative statics propositions. Subsection 4.4 presents some variations of our results.

## 4.1 Stable matching under minority reserves

A minority reserve ensures that whenever the number of minority students in a school $c$ is below the reserve, all other minority students must be assigned to schools that they strictly prefer to $c$. We adapt Hafalir, Yenmez, and Yildirim (2013) to model minority reserves as follows. Let $\rho = (\rho_1, \rho_2)$ be a vector of minority reserves in tier 1 and tier 2 schools. A matching $\mu$ is *stable under reserves* $\rho$ if for all $c \in C_i$ and all $s = (t, x) \in S$ with $c \succ_s \mu(s)$, the following three conditions hold:

i. $|\{s \mid \mu(s) = c\}| = k_i$;

ii. if $|\{s' = (t', x') \mid \mu(s') = c, t' = m\}| \geq \rho_i$, then $\omega_c^s < \omega_c^{s'}$ for all $s' = (t', x')$ with $\mu(s') = c$; and

iii. if $|\{s' = (t', x') \mid \mu(s') = c, t' = m\}| < \rho_i$, then $t = r$ and $\omega_c^s < \omega_c^{s''}$ for all $s'' = (r, x'') \in \mu(c)$.

A matching is stable under reserves $\rho$ if whenever a student $s$ would like to move to another school $c$, that school is filling its seats, it is admitting students having higher priority and exceeding the minority reserves, and if it is not exceeding the minority reserves then $s$ is a regular student having a score below the lowest score of regular students assigned to $c$. Note that when $\rho \equiv 0$, a matching is stable under reserves $\rho$ iff it is stable.

A matching $\mu$ that is stable under reserves $\rho$ always exists. It can be computed by the deferred acceptance algorithm by either properly defining a choice function or by making a copy of each school that targets minority students (Hafalir, Yenmez, and Yildirim 2013). Note that since our model has a continumm of students, the deferred acceptance algorithm need not converge in finite time (Abdulkadiroglu, Che, and Yasuda 2015).

We now characterize the unique stable matching under reserves $\rho$. First note that if $\rho_1 < \frac{1}{n}\mathbb{E}_m[1 - G_m(\bar{p})^{l_m}]$, then the stable matching characterized by cutoffs $\bar{p}$ is stable under reserves $\rho$. This simply follows from the observation that the minority reserve $\rho_1$ is already filled in tier 1 schools and therefore Conditions ii. and iii. in the definition of stability under reserves are equivalent to Condition ii in the definition of stability. Second, note that when $\rho_1 > \min\{\frac{\beta}{n}(1 - \pi_m(0)), k_1\}$, the reserve either is above the number of minority students that demand the school, or exceeds the capacity of the school. We thus define the set where reserves have a nontrivial impact on the final assignment: $R = [\frac{1}{n}\mathbb{E}_m[1 - G_m(\bar{p})^{l_m}], \min\{\frac{\beta}{n}(1 - \pi_m(0)), k_1\}]$.

Take a reserve $\rho_1 \in R$. We can characterize stability under reserves by means of cutoffs $p_c^t$ that depend on the school $c$ and the types $t \in \{r, m\}$ of the applying students. Similar to the analysis in Subsection 3.2, we can restrict attention to cutoffs such that $p_c^t = p_{c'}^t$ for all $c, c' \in C_1$ and $p_c^t \equiv 0$ for all $c \in C_2$ and all $t$. It is therefore enough to characterize the cutoffs $p_m$ and $p_r$, with $p_m \leq p_r$, that minority and regular students face in tier 1 schools. First, the market clearing condition can be written as:

$$\frac{1}{n}\mathbb{E}_r[1-(G_r(p_r))^{l_r}] + \frac{\beta}{n}\mathbb{E}_m[1-(G_m(p_m))^{l_m}] = k_1. \tag{4.1}$$

This is similar to equation (3.2), but in this market clearing condition different groups face different cutoffs. Second, the minority reserve condition must hold. Since $\rho_1 \geq \frac{1}{n}\mathbb{E}_m[1-G_m(\bar{p})^{l_m}]$, the reserve must bind and therefore the number of minority students in a tier 1 school equals the reserve:

$$\frac{\beta}{n}\mathbb{E}_m[1-G_m(p_m)^{l_m}] = \rho_1. \tag{4.2}$$

These two conditions have a unique solution $p_m$ and $p_r$. Figure 5 illustrates how these cutoffs are determined. Note that increasing $\rho_1$ moves the minority reserve condition (4.2) to the left in Figure 5. So, after an increase in minority reserves, $p_m$ decreases and $p_r$ increases. Increasing $\rho_1$ makes the access to tier 1 schools easier for minority students and harder for regular students.[18] We denote by $\mu_\rho$ the stable matching under reserves $\rho$.

The main focus of the paper is the impact of reserves $\rho$ on several market outcomes. Note that since tier 2 schools have excess capacity, $\rho_2$ is irrelevant for the allocation. We explore the role of reserves by stating several comparative statics results with respect to $\rho_1$.

## 4.2 Segregation

There are several ways to measure segregation in schools, but one of the the most common ones is the Duncan index (Duncan and Duncan 1955). Given a matching $\mu$, the *Duncan index* $D_\mu$ is defined by

$$D_\mu = \frac{1}{2}\sum_{c=1}^{n+N}\left|\eta_\mu^r(c) - \frac{\eta_\mu^m(c)}{\beta}\right| \in [0,1]$$

where $\eta_\mu^t(c)$ is the mass of students of type $t$ assigned to school $c$ in the matching $\mu$. The index equals 0 under perfect integration, where each school is filled by exactly the same number of students of each type. More generally, the Duncan index can be interpreted as the mass of regular students that would need to be moved to different schools so that every school had the same proportions of students of

[18] This means that for each tier 1 school, there are some regular students that would like to be assigned to that school, will be rejected but have higher scores than some minority students that have been accepted in the school.

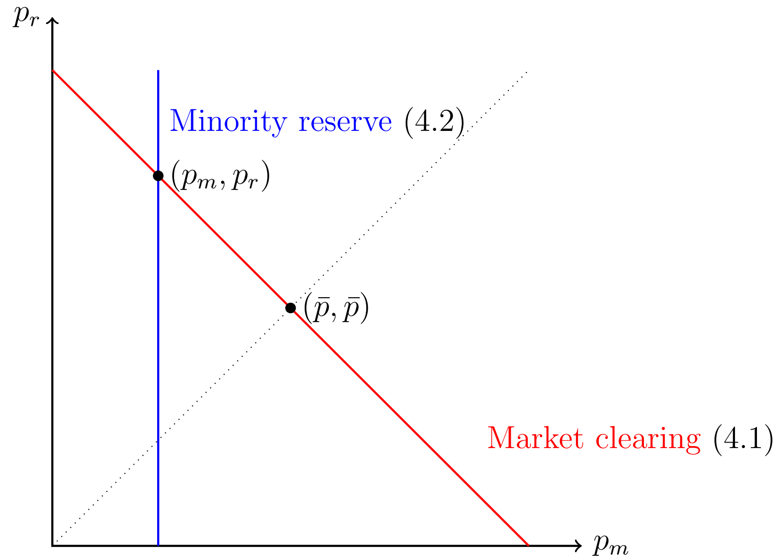



Figure 5: The market clearing condition and the minority reserve condition determine cutoffs $p_r$ and $p_m$. The cutoff $\bar{p}$ is in the intersection of the market clearing condition and the 45 degree line.

each group.

Given a reserve $\rho \in R$, we denote $D(\rho) = D_{\mu_\rho}$.

**Proposition 1.** *$D(\rho_1)$ is nonincreasing over $\rho_1 < \frac{\beta}{1+\beta}k_1$ and is non-decreasing over $\rho_1 > \frac{\beta}{1+\beta}k_1$.*

This result shows that reserves have an impact on segregation in schools. The Duncan segregation index is minimized when the fraction of seats reserved to minority students, $\rho_1/k_1$, equals the share of minority students in the population, $\beta/(1+\beta)$. Actually, in the proof we show a slightly stronger result: Segregation in each school $c$, $\left|\eta^r_{\mu_{\rho_1}}(c) - \frac{\eta^m_{\mu_{\rho_1}}(c)}{\beta}\right|$, is non-increasing over $\rho_1 < k_1\frac{\beta}{1+\beta}$ and non-decreasing over $\rho_1 > k_1\frac{\beta}{1+\beta}$. Intuitively, when $\rho_1 < k_1\frac{\beta}{1+\beta}$, minority students are underrepresented in tier 1 schools and overrepresented in tier 2 schools, and increasing $\rho_1$ moves minority students from tier 2 to tier 1 schools. This stronger property also implies that the index we actually use to measure segregation in our model is rather irrelevant for the Proposition.[19]

## 4.3 Rank distribution and efficiency

We now explore how $\rho_1$ impacts the efficiency of the assignment. Obviously, changing $\rho_1$ does not Pareto improve the assignment for students. Therefore, we evaluate changes to the assignment using two efficiency measures.

[19] Proposition 1 and our field evidence also apply to alternative segregation indexes, such as the ones discussed by Hutchens (2004) or Frankel and Volij (2011). See Online Appendix E.

The first measure is the rank distribution of students, which is a function that, for each $q \in \{1, \ldots, L_r + 1\}$, yields the fraction of students assigned to one of their $q$ most preferred schools. Our second measure is the number of students that belong to a Pareto improvement pair. As we discuss in the Introduction, both of these measures are important in practical implementations of centralized school choice algorithms.[20] The main results in this Subsection characterize how $\rho_1$ changes the rank distribution of students and the fraction of students in Pareto improving pairs.

**Rank distribution.** A type $t$ student ranks at most $L_t$ tier 1 shools and, since tier 2 schools always have free slots, type $t$ students are assigned to one of their $(L_t + 1)$-most preferred schools. The share of type $t$ students assigned to their $q$-th preference is

$$f_t(q) = \sum_{l=q}^{L_t} \pi_t(l) G_t(p_t)^{q-1}(1 - G_t(p_t)) + \pi_t(q-1) G_t(p_t)^{q-1}$$

for $q \in \{1, \ldots, L_t\}$. The first term represents all type $t$ students that applying to $q$ or more tier 1 schools get accepted in their $q$-th preference. The second term represents type $t$ students that applying to $q-1$ tier 1 schools are assigned to a tier 2 school. Note that for $q = L_t+1$, $f_t(L_t+1) = \pi_t(L_t) G_t(p_t)^{L_t}$. The cumulative rank distribution for type $t$ students is thus

$$F_t(q) = \sum_{q' \leq q} f_t(q') = \begin{cases} 1 - G_t(p_t)^q (1 - \Pi_t(q-1)) & \text{if } q \leq L_t, \\ 1 & \text{if } q = L_t + 1. \end{cases}.$$

To understand this formula intuitively, note that the mass of students assigned to schools ranked $q+1, q+2 \ldots$ is the fractions of students applying to $q$ or more schools (which happens with probability $1 - \Pi_t(q-1)$) and rejected in $q$ of them (which happens with probability $G_t(p_t)^q$). Thus, for $q \leq L_t$, $F_t(q) = 1 - G_t(p_t)^q(1 - \Pi_t(q-1))$. We will sometimes emphasize the dependence of these distributions on $\rho_1$ by writing $F_t(q, \rho_1)$.

**Lemma 1.** *Take $\rho_1 \in R$. Then, $\frac{\partial}{\partial \rho_1} F_m(q, \rho_1) > 0$ for all $q \leq L_m$ and $\frac{\partial}{\partial \rho_1} F_r(q, \rho_1) < 0$ for all $q \leq L_r$.*

This lemma says that increasing $\rho_1$ reduces (in the first order stochastic dominance sense) the cumulative rank distribution for minority students and increases the rank distribution of regular students. In other words (and not surprisingly), increasing reserves improves outcomes for minorities, at the expense of non-minorities.

[20] Obviously, both measures are imperfect. For example, it is entirely possible that under some matching very few students belong to Pareto improving pairs, but those improvements are very significant. To analyze this possibility, one would need to make additional assumptions to estimate utility functions. In contrast, our efficiency measures are based only on observable data. The two approaches are complementary.

Hafalir, Yenmez, and Yildirim (2013) derive comparative statics results with respect to minority reserves. Their Theorem 2 shows, in a general matching model, that the introduction of minority reserves favors at least one minority student. They also provide strong restrictions on priorities and preferences such that all minority students are better off when reserves are introduced. Lemma 1 thus complements these result.

Our main results explore the impact of minority reserves on the overall efficiency of the assignment. We thus define the total cumulative rank distribution as

$$F(q,\rho_1) = \frac{1}{1+\beta}\Big(\beta F_m(q,\rho_1) + F_r(q,\rho_1)\Big)$$

which measures the fraction of students assigned to one of their top $q$ schools. Determining the impact of $\rho_1$ on $F(q,\rho_1)$ is not obvious since Lemma 1 shows that reserves favor minorities but hurt regular students. The following is the first main result in the paper.

**Proposition 2.** *a. Assume that* $\mathbb{E}_r[l_r \mid l_r \geq 1] > \mathbb{E}_m[l_m \mid l_m \geq 1]$ *and define*

$$\bar{K} = \max\Big\{K \in [0,1] \mid \mathbb{E}_r[l_r(1 - \frac{K}{1-\pi_r(0)})^{l_r} \mid l_r \geq 1] \geq \mathbb{E}_m[l_m \mid l_m \geq 1]\Big\} \in ]0, 1-\pi_r(0)[.$$

*Then, for all* $nk_1 < \bar{K}$ *and all* $\rho_1 \in R$:

$$\frac{\partial F}{\partial \rho_1}(1,\rho_1) > 0.$$

*b. There exists* $\hat{K} = \hat{K}(\Pi_r, \Pi_m, \beta) < (1-\pi_r(0)) + \beta(1-\pi_m(0))$ *such that for all* $nk_1 > \hat{K}$ *and all* $\rho_1 \in R$:

$$\frac{\partial F}{\partial \rho_1}(1,\rho_1) < 0.$$

*c. Assume that* $L_r > L_m$. *Then, for all* $q \geq L_m + 1$:

$$\frac{\partial F}{\partial \rho_1}(q,\rho_1) < 0.$$

Proposition 2 describes the impact of minority reserves on the rank distribution. In particular, it shows that the impact of minority reserves on the fraction of students assigned to their top schools depends critically on the specifics of the market.

Proposition 2 part a shows that in markets where popular schools have relatively few seats ($nk_1$ is below a threshold $\bar{K}$) and minority students apply less frequently to these popular schools ($\mathbb{E}_r[l_r \mid l_r \geq 1] > \mathbb{E}_m[l_m \mid l_m \geq 1]$), increasing minority reserves results in more students assigned to their top-choice schools: $\frac{\partial F}{\partial \rho_1}(1,\rho_1) > 0$.

As Example 2 below shows, even when we assume condition (3.1), the assumption $\mathbb{E}_r[l_r \mid l_r \geq 1] > \mathbb{E}_m[l_m \mid l_m \geq 1]$ is critical to establish Proposition 2. Naturally, this condition is motivated by Figure 4.

In contrast, Proposition 2 part b shows that in markets with abundant capacity ($nk_1 > \hat{K}$), raising reserves decreases the mass of students assigned to their first preferences. Recall that our working assumption is that popular schools are congested so that $nk_1 \leq (1 - \pi_r(0)) + \beta(1 - \pi_m(0))$. Since $\hat{K} < (1 - \pi_r(0)) + \beta(1 - \pi_m(0))$, Proposition 2 part b applies to a non empty set of capacities where popular schools are congested.

Finally, Proposition 2 part c shows that increasing $\rho_1$ also increases the mass of students assigned to schools that are not highly ranked.

Proposition 2 parts a and b show that the total capacity of popular schools, by determining market congestion, has important consequences for the impact of minority reserves on the fraction of students assigned to their top schools. To grasp intuition, consider first the case in which capacity is abundant (as in part b) so that most regular students are assigned to their top schools. The marginal student assigned to each popular school has a relatively low score. Ann is a marginal regular student in some popular school $c$. After increasing the minority reserve in $c$, Ann will be displaced from $c$ by some minority student and will compete for seats in her second school $c' \neq c$. In school $c'$, Bob is the marginal regular student and probablu school $c'$ was his top school. Since scores are independent and Bob has a low score in $c'$, Ann has a significant probability of replacing Bob in $c'$.[21] After Ann replaces Bob in $c'$, Bob will apply to his second school $c''$. In that school, he has a significant probability of displacing the marginal regular student who with significant probability ranked $c''$ in the top. An so on. All in all, increasing the minority reserve in $c$ triggers competition in all other schools and causes a chain of rejections and new applications that reduces the number of regular students assigned to their top schools.

Consider now the case in which capacity is scarce (as in part a) so that the scores of regular students accepted to popular schools are relatively high. Suppose Ann is a marginal regular student in a popular school $c$. Ann is a regular student applying to many popular schools and, since cutoffs in popular schools are high, school $c$ is probably not her top choice. After increasing the minority reserve in $c$, Ann will be displaced from $c$ by some minority student. The minority student replacing Ann in $c$ applies to popular schools with low intensity and thus he is more likely to rank $c$ as his top school than Ann. Ann, in turn, will compete for a seat in $c'$. If $c'$ is a popular school, the marginal student in $c'$ has a relatively high score. As a result, Ann is unlikely to get accepted to $c'$ by displacing a

[21]The probability that Ann has a higher score than Bob approaches 1/2 as $nk_1 \to (1 - \pi_r(0)) + \beta(1 - \pi_m(0))$. Thus, a fraction approaching 1/2 of all regular students displaced from $c$ gets accepted to their second school.

student who ranked $c'$ first.[22] All in all, increasing the minority reserve in $c$ impacts the total number of students assigned to their top school by replacing Ann (who does not rank $c$ top) by a student for whom $c$ is the top school.

We now provide intuition for Proposition 2 part c. To see why increasing $\rho_1$ decreases $F(q)$ for $q \in \{L_m+1,\ldots,L_r\}$ , note that all minority students are assigned to one of their top $L_m+1$ schools. As Lemma 1 shows, the cumulative rank distribution for regular students is decreasing in $\rho_1$. As a result, for $q \in \{L_m+1,\ldots,L_r\}$, as $\rho_1$ raises, more students are assigned to schools they rank $q$ or worse. Example 1 shows that when $L_r = L_m = L$, the fraction of students assigned to schools ranked $L$ or worse may be decreasing in $\rho_1$. However, in the Appendix we extend Proposition 2 part c and show that even when $L_r = L_m = L$, $F(L)$ is decreasing in $\rho_1$ provided $\pi_m(L)$ is small enough. See Proposition 6 in Appendix A. Thus, the main driver of this comparative statics result is the fact that minorities are much less likely than regular students to apply to popular schools with very large intensities.

**Efficiency.** We now discuss the efficiency impact of minority reserves. Increasing reserves does not Pareto improve the assignment. A higher minority reserve does not improve the rank distribution of the assignment either.[23] We thus measure the efficiency of the assignment by the number of students in Pareto improving pairs.

Given a matching $\mu$, students $s = (t, x)$ and $s' = (t', x')$ form a *Pareto improving pair* if $c' = \mu(s') \succ_s c = \mu(s)$ and $c \succ_{s'} c'$. In this case, we say that $s$ is in a Pareto improving pair. Let $P(\rho_1)$ be the total measure of students $s$ that are in a Pareto improving pair. Arguably, $P(\rho_1)$ measures the inefficiency of the matching. The following proposition shows that minority reserves have an unambiguous effect on $P(\rho_1)$.

**Proposition 3** (Pareto improvements)**.** *Under the conditions of Proposition 2 part a (resp. part b), $P(\rho_1)$ is decreasing (resp. increasing) in $\rho_1$.*

When $\rho_1$ increases and the capacity of tier 1 schools is low enough, fewer students are in a Pareto improving pair. Thus, a higher reserve increases the efficiency of the matching in congested markets. In the proof, we show that a student $s$ can Pareto improve by switching school iff $s$ is assigned to a tier 1 school that is not her top choice. Thus, Proposition 3 follows immediately from Proposition 2. Our characterization also implies that the set of students in a Pareto improving pair coincides with the set of students in Pareto improving cycles.

[22] The probability with which Ann displaces the marginal student in $c'$ goes to 0 as $k_1 \to 0$.

[23] Featherstone (2020) explores rank efficiency in assignment problems.

**Examples.** To conclude this Section, we provide two examples showing that our main comparative statics results do not hold if we relax some of the assumptions in Propositions 2.

We first show that Proposition 2 part c need not hold when $L_r = L_m$.

**Example 1.** *Consider a model in which $\beta = 1$, $G_r(x) = G_m(x) = x$ and for $L \geq 3$*

$$\pi_r(0) = \ldots, = \pi_r(L-2) = 0 \quad \pi_r(L-1) = \pi_r(L) = 1/2$$

$$\pi_m(0) = 0 \quad \pi_m(1) = \pi_m(L) = 1/2 \quad \pi_m(2) = \cdots = \pi_m(L-1) = 0.$$

*In this model, regular students include more popular schools in their applications than minorities, but $L_r = L_m = L$. We show that when minority reserves increase, more students are assigned to schools they rank L or worse.*

*The market clearing cutoffs for regular and minority students solve:*

$$1 - \frac{1}{2}\left(p_r^{L-1} + p_r^L\right) = n(k_1 - \rho_1) \quad 1 - \frac{1}{2}\left(p_m + p_m^L\right) = n\rho_1$$

*In this model, all agents are asigned to one of their top $L+1$ schools so $F(L+1) = 1$. The fraction of agents assigned to one of their top L preferences is*

$$F(L) = 1 - \frac{1}{2}(p_r^L + p_m^L).$$

*It is simple to see that $\frac{\partial F(L)}{\partial \rho_1} > 0$ if and only if*[24]

$$(L-1)\pi_r(L-1)p_r^{L-1} + L\pi_r(L)p_r^L > \pi_m p_m + L\pi_m(L)p_r^L \tag{4.3}$$

*When $\rho_1 \in R$ is close to $\frac{1}{n}\mathbb{E}_m[1 - \bar{p}^{l_m}]$ (where $\bar{p}$ is the market clearing cutoff in the model without cutoffs), $p_r$ is close to $p_m$ (and close to $\bar{p}$) and (4.3) holds for all $L \geq 3$. It follows that more students are assigned to one of their top L schools when $\rho_1$ increases, for all $\rho_1 \in R$ close to $\frac{1}{n}\mathbb{E}_m[1 - \bar{p}^{l_m}]$.*

We now show that the assumption that the conditional expectation of the application intensities is higher for regular students is key for Proposition 2 part a.

**Example 2** (Expected number of popular schools)**.** *Consider $\beta = 1$ and the following distributions of application intensities:*

$$\pi_m(0) = \alpha, \ \pi_m(1) = \cdots = \pi_m(L-2) = 0, \ \pi_m(L-1) = 1 - \alpha$$

[24] This expression follows immediately from the proof of Proposition 2.

*and*

$$\pi_r(0) = 0,\ \pi_r(1) = \alpha,\ \pi_r(2) = \cdots = \pi_r(L-1) = 0, \pi_r(L) = 1-\alpha.$$

*A fraction $\alpha$ of minority students ranks first a tier 2 school, while a fraction $(1-\alpha)$ lists $L-1$ randomly chosen tier 1 schools. Analogously, a fraction $\alpha$ of regular students rank first one tier 1 school followed by a tier 2 school, and a fraction $(1-\alpha)$ lists $L$ tier 1 schools. We assume $L \leq n$. Clearly, $\Pi_r$ dominates $\Pi_m$. When $\alpha > \frac{1}{L-1}$, $L-1 = \mathbb{E}_m[l_m \mid l_m \geq 1] > \mathbb{E}_r[l_r \mid l_r \geq 1] = \alpha + (1-\alpha)L$. Morever, using equation (A.2) in the Appendix, it is relatively simple to show that $\frac{\partial F(1)}{\partial \rho_1} < 0$ when $\alpha + (1-\alpha)L < (L-1)(1-\rho_1 n)$. When $nk_1 < 1 - \frac{(1-\alpha)L+\alpha}{L-1} \in ]0,1[$, it follows that for all $\rho_1 \in R$, $\frac{\partial F(1)}{\partial \rho_1} < 0$.*

## 4.4 Extensions

We now discuss variations of our model and results. Subsection 4.4.1 explores a model in which preferences are polarized in the sense that minorities and regular students concentrate their applications over different sets of popular schools. We show that polarized preferences result in segregation, but the comparative statics with respect to reserves is different from our main model. Subsection 4.4.2 explores a model in which regular students rejected from popular schools get admission to an underdemanded school minority students cannot afford. For example, if regular students are rejected, then they could attend a private school outside the centralized system. We prove that even in this scenario, our main results hold.

We also explore some alternative algorithmic decisions. In Subsection 4.4.3, we introduce a double reserve policy (Echenique and Yenmez 2015) and argue that, while this policy promotes inclusion with polarized preferences, it has no impact in our main model. Subsection 4.4.4 shows how our results apply when the affirmative action policy is implemented by setting aside seats (Dur, Kominers, Pathak, and Sönmez 2018).

### 4.4.1 Polarized preferences

In our main model, regular students concentrate their applications on high demand schools, while minority students apply with lower intensity to overdemanded schools. In theory (but not in our field data), segregation could arise because minorities and regular students concentrate their applications on different sets of overdemanded schools. The following example shows that under this type of preferences, our comparative statics results need not hold. Indeed, we show that the number of students assigned to their top school need not increase with reserves. As a result, for minority reserves to improve efficiency, it is not enough that distinct groups have different preferences.

**Example 3** (Polarized prefences)**.** *We restrict our main model to $n = 2$, $\beta = 1$, $k_1 \leq 1$, $k_2 = 2$, but now we assume preferences are given by*

$$r : c_1 \succ c_2 \succ c_3 \quad m : c_2 \succ c_1 \succ c_3.$$

*Schools rank students uniformly and independently. In this setup, while all students prefer tier 1 schools over the tier 2 school (school $c_3$), minority students prefer $c_2$ to $c_1$ while regular students prefer $c_1$ to $c_2$.*

*When no reserve is imposed, it is relatively simple to find the cutoff $\bar{p} = \sqrt{1-k_1}$ for each tier 1 school. In the stable matching without reserves, minority students are underrepresented in $c_1$, and a fraction $F(1) = 1 - \sqrt{1-k_1}$ of all students are assigned to their top school.*

*Now, we impose a reserve $\rho_1 \in [\sqrt{1-k_1}\big(1-\sqrt{1-k_1}\big), k_1]$. We can characterize the stable matching by solving the market clearing conditions:*

$$k_1 - \rho_1 = 1 - p_1^r \quad \rho_1 = p_2(1 - p_1^m) \quad k_1 = 1 - p_2 + p_1^r(1 - p_2)$$

*where $p_1^r$ (resp. $p_1^m$) is the cutoff faced by regular (resp. minority) students in school $c_1$. Solving the system of equations, we deduce that the fraction of students assigned to their top school is*

$$F(1, \rho_1) = \frac{k_1 - \rho_1}{2} + \frac{1}{2}\frac{k_1}{2 - k_1 + \rho_1}.$$

*The function $F(1, \rho_1)$ is decreasing in $\rho_1$.*

The example shows that when both groups of students concentrate their applications in different schools, imposing a reserve reduces the number of students assigned to their top schools.[25] There are two forces behind this result. First, after the reserve is imposed in $c_1$, regular students are replaced by minority students for whom $c_1$ is not their most preferred school. Second, displaced regular students demand school $c_2$ and thus $1 - p_2$ decreases. As a result, fewer minority students are assigned to school $c_2$.

A model with polarized preferences is not a good description of application patterns in Chile. When preferences are polarized as in Example 3, in a stable matching (without reserves) minority students should be overrepresented in many popular schools. While it is true that in the data some popular schools are particularly attractive for minorities, those schools are the exception. Indeed, Appendix C shows that popular schools tend to have a low fraction of minority students.[26]

[25] Note that this holds for all $k_1 \leq 1$. In particular, it holds even if the market is slack.

[26] Additionally, the comparative statics of a model with polarized preferences does not match the simulations using Chilean data, as shown in Sections 4 and 4.4.3.

### 4.4.2 Polarized preferences II

Our main model assumes that regular and minority students rejected in tier 1 schools apply uniformly to tier 2 schools. However, it is possible that some regular students that do not get into an overdemanded school may choose a private school that minorities cannot afford. We use our setup to explore this possibility.

We now assume there are only two tier 2 schools, $c_m$ and $c_r$ (so $N = 2$). Both $c_r$ and $c_m$ have excess capacity. A type $t$ student with a preference profile in $Z(l)$ will rank school $c_t$ in the $l+1$ position. All the other details of the models are unchanged. School $c_t$ is an undersubscribed school that is demanded exclusively by type $t$ students. We thus interpret $c_r$ as a private undersubscribed school only regular students can afford.

The introduction of reserves moves minority students from $c_m$ to tier 1 schools, while regular students move from tier 1 schools to $c_r$. As a result of a reserve policy, minorities will increase their under-representation in tier 1 schools, while reduce their over-representation in school $c_m$. Regular students will decrease their over-representation in tier 1 schools, but increase even more their over-representation in school $c_r$. Thus, it is not immediately obvious what impact minority reserves have on segregation. However, for $\rho_1 \in R$, the Duncan index can be computed as

$$D(\rho_1) = \frac{1}{2}\Big\{n|k_1 - \rho_1 - \frac{\rho_1}{\beta}| + |1 - (k_1 - \rho_1)n| + |\frac{\beta - n\rho_1}{\beta}|\Big\}.$$

and is minimized at $\rho_1 = \frac{\beta}{1+\beta}k_1$.[27] Proposition 1 and all the main results in Section 4.3 also hold in this model.

### 4.4.3 Double reserves

Another measure that can be used to promote integration in schools is to reserve seats for *both* types of students (Echenique and Yenmez 2015). We consider a double reserve policy such that in each tier 1 school, $\frac{\beta}{1+\beta}k_1$ seats are reserved to minority students and $\frac{1}{1+\beta}$ seats are reserved to regular students.

The double reserve policy will promote integration in Example 3. Indeed, the Duncan index with minority reserves $\frac{\beta}{1+\beta}k_1$ will be higher than the Duncan index under the double reserve policy. Intuitively, a double reserve policy allows regular students to get accepted in the popular school $c_2$ where minorities are overrepresented.

In contrast, moving from a minority reserve of $\frac{\beta}{1+\beta}k_1$ to a double reserve policy does not change the assignment in our main model. It is relatively simple to show that given a minority reserve $\frac{\beta}{1+\beta}k_1$,

[27] This need not be the case for other segregation measures since reserves exacerbate the representation of regular students in $c_r$.

in each school the fraction of regular students equals $\frac{1}{1+\beta}$ and thus introducing an additional reserve for regular students does not give any advantage to them. Importantly, Section 5.1.1 confirms this theoretical finding using our field data.

### 4.4.4 Set asides

We have interpreted the affirmative action policy as a minimum guarantee for minority students. As noted by Dur, Kominers, Pathak, and Sönmez (2018), an alternative interpretation of an affirmative action policy is to set aside seats for minority students. Under a set aside policy, a school assigns first the $k_1 - \rho_1$ open seats and reserves the remaining $\rho_1$ seats for minority students. In this Subsection, we extend our results to this alternative affirmative action implementation.

To characterize a stable matching under set asides, we again consider cutoffs $p_r^{SA}$ and $p_m^{SA}$ that apply to regular and minority students in tier 1 schools under the set aside policy. The market clearing and reserve conditions for a set aside policy are

$$\frac{1}{n}\mathbb{E}_r[1 - G_r(p_r^{SA})^{l_r}] + \frac{\beta}{n}\mathbb{E}_m[1 - G_m(p_m^{SA})^{l_m}] = k_1$$

and

$$\frac{\beta}{n}\mathbb{E}_m[G_m(p_r^{SA})^{l_m} - G_m(p_m^{SA})^{l_m}] = \rho_1. \tag{4.4}$$

Equation (4.4) is the set aside condition. Motivated by Dur, Kominers, Pathak, and Sönmez (2018), the set aside condition says that the number of minority students with scores below the regular cutoff $p_r^{SA}$ and that get admitted to a school should equal the reserve $\rho_1$. In contrast to minority reserves, under this interpretation of the affirmative action policy, the number of minority students effectively admitted to a tier 1 school exceeds the reserve $\rho_1$. We define by $F^{SA}(q, \rho_1)$ as the fraction of students assigned to a school ranked $q$ or better under a set aside affirmative action policy $\rho_1$. Analogously, we define $P^{AS}(\rho_1)$ as the total measure of students than can Pareto improve in the matching with set aside reserves $\rho_1$.

As Dur, Kominers, Pathak, and Sönmez (2018) show, in terms of the final assignment of students, the precedence order with which reserves are processed has an impact similar to adjusting reserve sizes. We derive a similar result in our framework. In contrast to the analysis by Dur, Kominers, Pathak, and Sönmez (2018), our focus is on the proportion of students assigned to their top schools and the number of students in Pareto improving pairs.

**Proposition 4.** *Under the conditions of Proposition 2,* $F^{SA}(1, \rho_1) > F(1, \rho_1)$ *and* $P^{SA}(\rho_1) < P(\rho_1)$.

Fewer minority students are assigned to tier 1 schools under minority reserves than under set aside. Changing the interpretation of the affirmative action policy from minority reserves to set asides increases the number of students assigned to their top schools and reduces the number of students who can Pareto improve by switching schools. Obviously, compared to the minority reserve policy, the set aside policy may or may not reduce segregation by placing more minority students in tier 1 schools.

The following result shows that the main insights from Propositions 2 and 3 extend to the set aside policy.

**Proposition 5.** *There exists $\bar{k} \in ]0,1[$ such that for all $k_1 < \bar{k}$ and all $\rho_1 < \bar{k}$,*

$$\frac{\partial F^{SA}(1,\rho_1)}{\partial \rho_1} > 0 \text{ and } \frac{\partial P^{SA}(\rho_1)}{\partial \rho_1} < 0.$$

The main intuitions and arguments behind this result are similar to the ones in Subsection 4.3 and are therefore omitted. Section 5.1.2 simulates the set-aside policy and confirms all these theoretical predictions.

# 5 Discussion

Our results are relevant for policy discussion. For each of our cities (Valparaiso, Concepción, Santiago), Table 3 reports outcomes with no reserves, reserves equal to 15% (as currently determined by the Law), $75\%, 100\%$ and equal to the fraction of minority students in the market.[28] As can be seen, the simulations confirm each of our theoretical results in Section 4. The simulations also confirm the most subtle theoretical prediction from the model, Proposition 2 and Proposition 3.[29] Table 3 shows that the total number of student assigned to their top school and the total number of students assigned to schools that are not very attractive (ranked fourth or below) move precisely as predicted by our theory.[30]

The simulations thus show that policymakers may integrate schools, increase the proportion of applicants obtaining their top choices, reduce the number of students in Pareto improving pairs, but incur the costs of leaving more students to relatively unattractive schools. In our quantitative exercise, minority reserves can reduce segregation by more than 20% in each of the three cities. At

[28]The reserve policy of 15% has been in place and never modified since the inception of the centalized system in 2016.

[29]Even though in our theoretical framework all the students are assigned, in the simulations we have computed the fraction of unassigned students as a measure of students whose assignment is unattractive. The fact that more students are unassigned as we increase the reserve is related to the result in Proposition 2 that the fraction of students assigned to schools that are not highly ranked increases with the reserve.

[30]In Appendix D.1, we expand capacities and drop students to obtain a slack market. As predicted by Proposition 2 part b, in a slack market reserves decrease the number of students assigned to top schools.

the same time, increasing minority reserves leaves slightly more students assigned to their top schools and reduces the proportion of students in Pareto improving pairs by close to 15%. The main cost of a reduction in segregation is an increase in the number of unassigned students. For example, Table 3 shows that in Valparaíso, the Duncan index can be reduced from 0.316 to 0.247, while the proportion of unassigned students increases from 9.46% to 9.81% and the proportion of students in Pareto improving pairs reduces from 7.81% to 6.68%. How society weights all these outcomes will determine the optimal minority reserve.[31] More generally, our simulations show that the minority reserve is a significant policy decision that has sizable impact on important market outcomes and its size should be carefully evaluated in practical implementations.

[31]The social welfare function $W = W(D, T, P, U)$ is likely to be decreasing in the proportion of students in Pareto improving pairs $P$, and in the proportion of unassigned students $U$. The function $W$ will be increasing in the proportion of students assigned to their top schools $T$. The dependence of $W$ on $D$ will capture how educational outcomes depend on peer effects and how society values integration. For discussion on peer effects and educational outcomes, see Hoxby (2000), Hanushek, Kain, Markman, and Rivkin (2003), Angrist and Lang (2004), and Rao (2019). At least in some nontrivial range, $W$ should be decreasing in $D$.

| **Valparaíso** | $f = 0\%$ | $f = 15\%$ | $f = 44\%$ | $f = 75\%$ | $f = 100\%$ |
|---|---|---|---|---|---|
| Duncan index (Proposition 1) | 0.316 (0.005) | 0.312 (0.005) | 0.247 (0.003) | 0.311 (0.003) | 0.317 (0.002) |
| Minority students assigned to their top choice (Lemma 1) | 71.73 (0.52) | 72.05 (0.4) | 78.89 (0.38) | 88.69 (0.29) | 89.54 (0.25) |
| Regular students assigned to their top choice (Lemma 1) | 63.67 (0.37) | 63.45 (0.48) | 59.88 (0.35) | 56.25 (0.35) | 55.96 (0.27) |
| Students assigned to their top choice (Proposition 2 a) | 67.21 (0.23) | 67.23 (0.33) | 68.23 (0.25) | 70.49 (0.22) | 70.7 (0.21) |
| Students assigned to their fourth choice or worst (Proposition 2 c) | 7.32 (0.21) | 7.29 (0.17) | 7.6 (0.17) | 8.24 (0.16) | 8.31 (0.16) |
| Students unassigned (Proposition 2 c) | 9.46 (0.16) | 9.48 (0.19) | 9.81 (0.14) | 10.56 (0.13) | 10.63 (0.13) |
| Students in Pareto improving pairs (Proposition 3) | 7.81 (0.53) | 7.9 (0.51) | 6.68 (0.57) | 3.66 (0.26) | 3.3 (0.33) |
| **Concepción** | $f = 0\%$ | $f = 15\%$ | $f = 43\%$ | $f = 75\%$ | $f = 100\%$ |
| Duncan index (Proposition 1) | 0.353 (0.004) | 0.342 (0.005) | 0.264 (0.003) | 0.355 (0.003) | 0.358 (0.003) |
| Minority students assigned to their top choice (Lemma 1) | 67.77 (0.43) | 68.57 (0.44) | 76.74 (0.37) | 88.01 (0.2) | 88.46 (0.22) |
| Regular students assigned to their top choice (Lemma 1) | 58.06 (0.47) | 57.4 (0.46) | 54.11 (0.39) | 50.68 (0.33) | 50.5 (0.27) |
| Students assigned to their top choice (Proposition 2 a) | 62.23 (0.29) | 62.2 (0.33) | 63.84 (0.26) | 66.73 (0.24) | 66.81 (0.17) |
| Students assigned to their fourth choice or worst (Proposition 2 c) | 11.42 (0.21) | 11.41 (0.25) | 11.46 (0.21) | 11.95 (0.15) | 11.97 (0.16) |
| Students unassigned (Proposition 2 c) | 12.65 (0.16) | 12.75 (0.13) | 13.08 (0.15) | 13.61 (0.1) | 13.63 (0.1) |
| Students in Pareto improving pairs (Proposition 3) | 12.31 (0.41) | 12.34 (0.54) | 10.25 (0.44) | 5.64 (0.37) | 5.35 (0.3) |
| **Santiago** | $f = 0\%$ | $f = 15\%$ | $f = 37\%$ | $f = 75\%$ | $f = 100\%$ |
| Duncan index (Proposition 1) | 0.312 (0.002) | 0.303 (0.002) | 0.246 (0.001) | 0.328 (0.001) | 0.331 (0.001) |
| Minority students assigned to their top choice (Lemma 1) | 70.86 (0.19) | 71.58 (0.17) | 77.81 (0.16) | 90.79 (0.08) | 91.35 (0.07) |
| Regular students assigned to their top choice (Lemma 1) | 56.48 (0.17) | 56.15 (0.19) | 54.05 (0.12) | 50.29 (0.11) | 50.1 (0.1) |
| Students assigned to their top choice (Proposition 2 a) | 61.87 (0.12) | 61.93 (0.13) | 62.95 (0.09) | 65.46 (0.08) | 65.55 (0.07) |
| Students assigned to their fourth choice or worst (Proposition 2 c) | 12.3 (0.09) | 12.32 (0.1) | 12.49 (0.09) | 13.14 (0.07) | 13.19 (0.06) |
| Students unassigned (Proposition 2 c) | 12.49 (0.06) | 12.49 (0.05) | 12.67 (0.06) | 13.1 (0.05) | 13.13 (0.04) |
| Students in Pareto improving pairs (Proposition 3) | 9.5 (0.2) | 9.43 (0.19) | 8.06 (0.16) | 4.52 (0.13) | 4.39 (0.12) |

Table 3: Average impact of minority reserves on market outcomes. Excluding the Duncan index, all values are percentages. Simulation standard deviations inside parentheses.

## 5.1 Other algorithmic decisions

To put the design of the minority reserves in perspective, we now discuss the impact of other algorithmic decisions on market outcomes. We only report results for Santiago (similar results are obtained for Valparaíso and Conception).

### 5.1.1 Double reserve policy

We now simulate the double reserve policy (Section 4.4.3). We compare minority reserves to a double reserve policy, where we reserve a fraction of seat equals to the proportion of the group in the market.

| | Minority reserve | Double reserve |
|---|---|---|
| **Santiago** | | |
| Duncan index | 0.246 (0.001) | 0.232 (0.001) |
| Students assigned to their top choice | 62.95 (0.09) | 62.89 (0.12) |
| Students assigned to their fourth choice or worst | 12.49 (0.09) | 12.37 (0.08) |
| Students unassigned | 12.67 (0.06) | 12.66 (0.05) |
| Students in Pareto improving pairs | 8.06 (0.16) | 8.06 (0.17) |

Table 4: Average impact of single and double reserves on market outcomes. Excluding the Duncan index, all values are percentages. Simulation standard deviations inside parentheses.

Table 4 shows that moving from (single) minority reserves to a double reserve policy has a much smaller impact than moving from no reserve to minority reserve. This is precisely what our theory predicts. Note that the introduction of the ideal point policy reduces the number of students assigned to their top schools, similar to the model with polarized preferences in Example 3.

#### 5.1.2 Set asides

We also simulated each of the markets using the set aside affirmative action policy. Consistent with Proposition 5, increasing the magnitude of the affirmative action policy has similar impacts under minority reserves and set asides. Tables 3 and 5 also confirm the prediction of Proposition 4 that fixing the reserves $\rho$, changing the interpretation of the affirmative action policy from minority reserves to set asides increases the number of students assigned to top schools and reduces the number of Pareto improving pairs. Under set asides, segregation is minimized for a reserve below the proportion of minority students in the population.

| **Santiago** | $f = 0\%$ | $f = 15\%$ | $f = 37\%$ | $f = 75\%$ | $f = 100\%$ |
|---|---|---|---|---|---|
| Duncan index | 0.312 (0.002) | 0.279 (0.001) | 0.308 (0.001) | 0.331 (0.001) | 0.331 (0.001) |
| Minority students assigned to their top choice | 70.91 (0.19) | 80.22 (0.18) | 88.35 (0.12) | 91.33 (0.08) | 91.34 (0.06) |
| Regular students assigned to their top choice | 56.5 (0.16) | 52.75 (0.14) | 50.78 (0.12) | 50.12 (0.12) | 50.13 (0.1) |
| Students assigned to their top choice | 61.9 (0.11) | 63.05 (0.1) | 64.85 (0.09) | 65.56 (0.08) | 65.57 (0.07) |
| Students assigned to their fourth choice or worst | 12.3 (0.08) | 12.7 (0.08) | 13.04 (0.07) | 13.2 (0.08) | 13.19 (0.06) |
| Students unassigned | 12.49 (0.05) | 12.71 (0.05) | 13.01 (0.05) | 13.13 (0.04) | 13.13 (0.05) |
| Students in Pareto improving pairs | 9.46 (0.2) | 8.09 (0.18) | 5.53 (0.15) | 4.38 (0.13) | 4.35 (0.14) |

Table 5: Average impact of set asides on market outcomes. Excluding the Duncan index, all values are percentages. Simulation standard deviations inside parentheses.

### 5.1.3 Single tie-breaking

When considering lotteries over students using the deferred acceptance algorithm, whether the lottery is run school-by-school (multiple tie-breaking) or system-wide (single tie-breaking) makes a difference in outcomes (Abdulkadiroğlu and Sönmez 2003, Abdulkadiroğlu, Pathak, and Roth 2009, De Haan, Gautier, Oosterbeek, and Van der Klaauw 2015, Ashlagi and Nikzad 2020). Theoretical results and practical experience show that moving from multiple to single tie-breaking increases the number of students assigned to their first preferences, but more students are unassigned. These results are also confirmed by the simulations reported in Table 6.

In terms of the cumulative rank distribution, our theoretical results and simulations show that increasing reserves has an impact similar to moving from multiple to single lottery. Notably, moving from multiple to single tie-breaking has almost no impact on segregation. Minortity reserves are thus an important and distinctive policy decision.

| **Santiago** | $f = 15\%$ |
|---|---|
| Duncan index | 0.302 (0.002) |
| Minority students assigned to their top choice | 75.69 (0.19) |
| Regular students assigned to their top choice | 61.8 (0.14) |
| Students assigned to their top choice | 67 (0.08) |
| Students assigned to their fourth choice or worst | 12.04 (0.06) |
| Students unassigned | 13.12 (0.05) |
| Students in Pareto improving pairs | 1.06 (0.09) |

Table 6: Impact of single-tie breaking on market outcomes. Excluding the Duncan index, all values are percentages. Simulation standard deviations inside parentheses.

## 6 Conclusions

In this paper, we have examined the influence of minority reserves on segregation and efficiency within school choice programs. Our findings highlight the significant role of minority reserves in mitigating segregation in schools. While minority reserves improve overall efficiency, they also lead to more students assigned to unattractive schools. This paper contributes to the field of market design by describing the impacts of minority reserves on important market outcomes.

Our model illustrates basic forces that govern the effect of minority reserves on the final assignments. Specifically, our theoretical analysis reveals that the qualitative impact of minority reserves on matching efficiency hinges on the dissimilarity of preferences among distinct groups and the tightness of the market. Our analysis has practical relevance. The simulations using data from a large-scale implementation show that minority reserves represent an important and distinct policy choice.

An important driver of our results is the observation that low-income groups tend to apply less frequently to high-demand institutions. Similar patterns have been observed in various contexts, including school choice programs in Boston and Amsterdam (Laverde 2020, Oosterbeek, Sóvágó, and van der Klaauw 2021) and college admission in the United States (Hoxby and Avery 2013). We hope our findings hold relevance for discussions on strategies to reduce segregation in different settings.

As the literature shows, it is hard to obtain general comparative statics results in matching models (Hafalir, Yenmez, and Yildirim 2013, Dur, Kominers, Pathak, and Sönmez 2018). Our main comparative statics result, Proposition 2, leverage the large market assumption to quantify the winners and losers created by minority reserves. Extending our results to more general setups is an interesting theoretical question.

In our theoretical framework, the DA algorithm is strategy-proof. We therefore assume –as do Abdulkadiroğlu, Pathak, and Roth (2009) and Che and Tercieux (2019), among others– that students do not alter their applications following a change in a strategy-proof mechanism. However, minority reserves could change application patterns as a result of location decisions or in reaction to school compositions (Epple and Romano 2003, Baum-Snow and Lutz 2011, Avery and Pathak 2021, Calsamiglia, Martínez-Mora, and Miralles 2021, Akbarpour, Kapor, Neilson, van Dijk, and Zimmerman 2022). Application patterns may also be influenced by behavioral economics forces (Rees-Jones and Shorrer 2023). These questions warrant exploration in future research.[32]

# References


Abdulkadiroğlu, A. (2005): "College Admissions with Affirmative Action," *International Journal of Game Theory*, 33(4), 535–549.

Abdulkadiroglu, A., Y.-K. Che, and Y. Yasuda (2015): "Expanding" Choice" in School Choice," *American Economic Journal: Microeconomics*, 7(1), 1.

Abdulkadiroglu, A., P. Pathak, A. Roth, and T. Sönmez (2006): "Changing the Boston School Choice Mechanism: Strategy-proofness as Equal Access," *MIT*.


[32]Note that the reserve policy used in Chile has not changed since the introduction of the centralized school choice system in 2016. Thus, it is challenging to empirically explore the impact of minority reserves on applications using the Chilean context.

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2009): "Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*, 99(5), 1954–78.

Abdulkadiroğlu, A., and T. Sönmez (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93(3), 729–747.

Akbarpour, M., A. Kapor, C. Neilson, W. van Dijk, and S. Zimmerman (2022): "Centralized School choice with unequal outside options," *Journal of Public Economics*, 210, 104644.

Angrist, J., and K. Lang (2004): "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," *American Economic Review*, 94(5), 1613–1634.

Ashlagi, I., and A. Nikzad (2020): "What Matters in School Choice Tie Breakings? How Competition Guides Design," *Journal of Economic Theory*, 190.

Avery, C., and P. Pathak (2021): "The Distributional Consequences of Public School Choice," *American Economic Review*, 111(1), 129–52.

Aygun, O., and I. Bó (2021): "College Admission with Multidimensional Privileges: The Brazilian Affirmative Action Case," *American Economic Journal: Microeconomics*.

Aygün, O., and B. Turhan (2020): "Dynamic reserves in matching markets," *Journal of Economic Theory*, 188, 105069.

Azevedo, E., and J. Leshno (2016): "A Supply and Demand Framework for Two-Sided Matching Markets," *Journal of Political Economy*, 124(5), 1235–1268.

Baum-Snow, N., and B. Lutz (2011): "School Desegregation, School Choice, and Changes in Residential Location Patterns by Race," *American Economic Review*, 101(7), 3019–46.

Calsamiglia, C., F. Martínez-Mora, and A. Miralles (2021): "School choice design, risk aversion and cardinal segregation," *The Economic Journal*, 131(635), 1081–1104.

Che, Y.-K., and O. Tercieux (2019): "Efficiency and Stability in Large Matching Markets," *Journal of Political Economy*, 127(5), 2301–2342.

Correa, J., R. Epstein, J. Escobar, I. Rios, B. Bahamondes, C. Bonet, N. Epstein, N. Aramayo, M. Castillo, and A. Cristi (2022): "School Choice in Chile," *Operations Research*, 70(2).

De Haan, M., P. Gautier, H. Oosterbeek, and B. Van der Klaauw (2015): "The Performance of School Assignment Mechanisms in Practice," .

Duncan, O., and B. Duncan (1955): "A Methodological Analysis of Segregation Indexes," *American Sociological Review*, 20(2), 210–217.

Dur, U., S. Kominers, P. Pathak, and T. Sönmez (2018): "Reserve Design: Unintended Consequences and the Demise of Boston's Walk Zones," *Journal of Political Economy*, 126(6), 2457–2479.

Dur, U., P. Pathak, and T. Sönmez (2020): "Explicit vs. Statistical Targeting in Affirmative Action: Theory and Evidence from Chicago's Exam Schools," *Journal of Economic Theory*, 187, 104996.

Echenique, F., and M. Yenmez (2015): "How to Control Controlled School Choice," *American Economic Review*, 105(8), 2679–94.

Ehlers, L., I. Hafalir, M. Yenmez, and M. Yildirim (2014): "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds," *Journal of Economic Theory*, 153, 648–683.

Epple, D., and R. Romano (2003): "Neighborhood Schools, Choice, and the Distribution of Educational Benefits," in *The Economics of School Choice*, pp. 227–286. University of Chicago Press.

Featherstone, C. (2020): "Rank Efficiency: Modeling a Common Policymaker Objective," Discussion paper, The Wharton School, University of Pennsylvania.

Fragiadakis, D., and P. Troyan (2017): "Improving Matching under Hard Distributional Constraints," *Theoretical Economics*, 12(2), 863–908.

Frankel, D., and O. Volij (2011): "Measuring School Segregation," *Journal of Economic Theory*, 146(1), 1–38.

Gale, D., and L. Shapley (1962): "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 69(1), 9–15.

Hafalir, I., M. Yenmez, and M. Yildirim (2013): "Effective Affirmative Action in School Choice," *Theoretical Economics*, 8(2), 325–363.

Hanushek, E., J. Kain, J. Markman, and S. Rivkin (2003): "Does Peer Ability Affect Student Achievement?," *Journal of Applied Econometrics*, 18(5), 527–544.

Hastings, J., T. Kane, and D. Staiger (2009): "Heterogeneous Preferences and the Efficacy of Public School Choice," *NBER*.

Honey, N., and A. Carrasco (2022): "A New Admission System in Chile and Its Foreseen Moderate Impact on Access for Low-Income Students," *Educational Evaluation and Policy Analysis*, p. 01623737221093374.

Hoxby, C., and C. Avery (2013): "Missing'One-Offs': The Hidden Supply of High-Achieving, Low-Income Students," *Brookings Papers on Economic Activity*, (1), 1–46.

Hoxby, C. M. (2000): "Peer Effects in the Classroom: Learning from Gender and Race Variation," Discussion paper, National Bureau of Economic Research.

Hutchens, R. (2004): "One Measure of Segregation," *International Economic Review*, 45(2), 555–578.

Karsten, S. (2010): "School segregation," in *Equal Opportunities The Labour Market Integration of the Children of Immigrants.*

Kojima, F. (2012): "School Choice: Impossibilities for Affirmative Action," *Games and Economic Behavior*, 75(2), 685–693.

Kominers, S., and T. Sönmez (2016): "Matching with Slot-Specific Priorities: Theory," *Theoretical Economics*, 11(2), 683–710.

Kutscher, M., S. Nath, and S. Urzua (2020): "Centralized Admission Systems and School Segregation: Evidence from a National Reform," Working paper, University of Maryland.

Laverde, M. (2020): "Unequal Assignments to Public Schools and the Limits of School Choice," Working paper, Chicago.

Leshno, J., and I. Lo (2021): "The Cutoff Structure of Top Trading Cycles in School Choice," *The Review of Economic Studies*, 88(4), 1582–1623.

Nguyen, T., and R. Vohra (2019): "Stable Matching with Proportionality Constraints," *Operations Research*, 67(6), 1503–1519.

Oosterbeek, H., S. Sóvágó, and B. van der Klaauw (2021): "Preference Heterogeneity and School Segregation," *Journal of Public Economics*, 197, 104400.

Pathak, P., T. Sönmez, M. Ünver, and M. Yenmez (2020): "Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Health Care Rationing," Discussion paper, Boston College.

Rao, G. (2019): "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools," *American Economic Review*, 109(3), 774–809.

Rees-Jones, A., and R. Shorrer (2023): "Behavioral economics in education market design: a forward-looking review," *Journal of Political Economy Microeconomics*, 1(3), 557–613.

Rios, I., T. Larroucau, G. Parra, and R. Cominetti (2021): "Improving the Chilean college admissions system," *Operations Research*, 69(4), 1186–1205.

Roth, A., and M. Sotomayor (1990): *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* Cambridge Univ, Press.

Son, S. (2020): "Distributional Impacts of Centralized School Choice," Discussion paper, Yale.

SÖNMEZ, T., AND M. YENMEZ (2022): "Affirmative Action in India via Vertical, Horizontal, and Overlapping Reservations," *Econometrica*, 90(3), 1143–1176.

WESTKAMP, A. (2013): "An Analysis of the German University Admissions System," *Economic Theory*, 53(3), 561–589.

# APPENDIX

Appendix A contains proofs and complementary results. Appendix B provides details about our field data. Appendix C desagregates the impact of minority reserves on segregation. Appendix D explores simulations with abundant capacity and shows the empirical relevance of Proposition 2 part b. The Online Appendix provides additional details.

## A Proofs and complementary results

*Proof of Proposition 1.* Note that for $\rho_1 \in R$, each tier 1 school has $k_1 - \rho_1$ regular students and $\rho_1$ minority students, while a tier 2 school has $(1 - n(k_1 - \rho_1))/N$ regular students and $(\beta - n\rho_1)/N$ minority students.[33] Therefore,

$$D(\rho) = \frac{1}{2}\Big\{n|\frac{k_1 - \rho_1}{1} - \frac{\rho_1}{\beta}| + N|\frac{(1-(k_1-\rho_1)n)/N}{1} - \frac{(\beta - n\rho_1)/N}{\beta}|\Big\}.$$

The first (resp. second) term inside the bracket captures the summation defining $D(\rho)$ over tier 1 schools (resp. tier 2 schools). Thus, for $\rho_1 \in R$,

$$D(\rho) = \frac{1}{2}\Big\{n|k_1 - \rho_1(1+\frac{1}{\beta})| + n|k_1 - \rho_1(1+\frac{1}{\beta})|\Big\}.$$

Note that for $\rho_1 \notin R$, $D(\rho)$ is flat. The result follows. $\square$

*Proof of Proposition 2.* (a) We first note that for $q \leq L_t$, $F_t(q) = 1 - G_t(p_t)^q(1 - \Pi_t(q-1))$. Thus,

$$\begin{aligned}(1+\beta)&\frac{\partial F}{\partial \rho_1}(q,\rho_1)\\ &= -\, qG_r(p_r)^{q-1}g_r(p_r)(1-M_r(q-1))\frac{\partial p_r}{\partial \rho_1} - \beta q G_m(p_m)^{q-1}g_m(p_m)(1-M_m(q-1))\frac{\partial p_m}{\partial \rho_1}.\end{aligned}$$

[33] To see the distribution of students in tier 2 schools, note that $1 - n(k_1 - \rho_1)$ regular students are not assigned to tier 1 schools. Regular students that are not assigned to tier 1 schools demand tier 2 schools uniformly.

Now, recall that $\mathbb{E}_r[1-G_r(p_r)^{l_r}] = (k_1-\rho_1)n \quad \beta\mathbb{E}_m[1-G_m(p_m)^{l_m}] = \rho_1 n$. Taking derivatives

$$\beta g_m(p_m)\frac{\partial p_m}{\partial \rho_1} = \frac{-n}{\mathbb{E}_m[l_m G_m(p_m)^{l_m-1}]} \quad g_r(p_r)\frac{\partial p_r}{\partial \rho_1} = \frac{n}{\mathbb{E}_r[l_r G_r(p_r)^{l_r-1}]}$$

We deduce that for $q \leq L_m$

$$\frac{\partial F}{\partial \rho_1}(q) < 0 \text{ (resp. } > 0) \quad \text{iff} \quad \frac{G_m(p_m)^{q-1}(1-\Pi_m(q-1))}{\mathbb{E}_m[l_m G_m(p_m)^{l_m-1}]} < \frac{G_r(p_r)^{q-1}(1-\Pi_r(q-1))}{\mathbb{E}_r[l_r G_r(p_r)^{l_r-1}]} \text{ (resp. } >). \tag{A.1}$$

In particular,

$$\frac{\partial F}{\partial \rho_1}(1,\rho_1) > 0 \quad \text{iff} \quad \frac{\mathbb{E}_r[l_r G_r(p_r)^{l_r-1}]}{1-\Pi_r(0)} > \frac{\mathbb{E}_m[l_m G_m(p_m)^{l_m-1}]}{1-\Pi_m(0)}. \tag{A.2}$$

To derive a sufficient condition for the condition on the right in (A.2), we note that

$$\frac{\mathbb{E}_m[l_m G_m(p_m)^{l_m-1}]}{1-\Pi_m(0)} < \frac{\mathbb{E}_m[l_m]}{1-\Pi_m(0)} = \mathbb{E}_m[l_m \mid l_m \geq 1]. \tag{A.3}$$

Also,

$$\pi_r(0) + (1-\pi_r(0))G_r(p_r) > \mathbb{E}_r[G_r(p_r)^{l_r}] = 1 - n(k_1-\rho_1) > 1 - nk_1$$

and therefore

$$G_r(p_r) > 1 - \frac{nk_1}{1-\pi_r(0)}$$

As a result,

$$\mathbb{E}_r[l_r G_r(p_r)^{l_r-1}] > \mathbb{E}_r[l_r\big(1-\frac{nk_1}{1-\pi_r(0)})^{l_r-1}\big)] = \mathbb{E}_r[l_r\big(1-\frac{nk_1}{1-\pi_r(0)})^{l_r-1}\big) \mid l_r \geq 1](1-\Pi_r(0)). \tag{A.4}$$

Using (A.2), (A.3) and (A.4), it follows that $\frac{\partial F(1,\rho_1)}{\partial \rho_1} > 0$ when

$$\mathbb{E}_r[l_r\big(1-\frac{nk_1}{1-\pi_r(0)})^{l_r-1}\big) \mid l_r \geq 1] \geq \mathbb{E}_m[l_m \mid l_m \geq 1].$$

By assumption, $\mathbb{E}_r[l_r \mid l_r \geq 1] > \mathbb{E}_m[l_m \mid l_m \geq 1]$. As a result, we can define

$$\bar{K} = \max\Big\{K \in [0,1] \mid \mathbb{E}_r[l_r(1-\frac{K}{1-\pi_r(0)})^{l_r} \mid l_r \geq 1] \geq \mathbb{E}_m[l_m \mid l_m \geq 1]\Big\} \in ]0, 1-\pi_r(0)[$$

and for all $nk_1 \leq \bar{K}$, $\frac{\partial F}{\partial \rho_1}(1) > 0$ for all $\rho_1$.

(b) Following the proof of part (a), it follows that

$$\frac{\partial F}{\partial \rho_1}(1,\rho_1) < 0 \quad \text{iff} \quad \mathbb{E}_r[l_r G_r(p_r)^{l_r-1} \mid l_r \geq 1] < \mathbb{E}_m[l_m G_m(p_m)^{l_m-1} \mid l_m \geq 1].$$

For $x \in [0,1]$ and $t \in \{r, m\}$, define $\Phi_t(x) = \mathbb{E}_t[l_r x^{l_t - 1} \mid l_t \geq 1]$. Since $\Phi_t(0) = \frac{\pi_t(1)}{1-\Pi_t(0)}$ and $\Pi_r$ dominates $\Pi_m$, it follows that $\Phi_r(0) < \Phi_m(0)$. Take $\bar{x} > 0$ such that $\Phi_r(x) < \Phi_m(0)$ for all $x < \bar{x}$.

Now, the market clearing condition for regular students is given by

$$(1 - \pi_r(0)) - \sum_{l_r=1}^{L_r} \pi_r(l_r) G_r(p_r)^{l_r} = n(k_1 - \rho_1).$$

In particular, there exists $\bar{y} = \bar{y}(\Pi_r, \Pi_m) > \pi_r(0)$ such that for all $n(k_1 - \rho_1) > 1 - \bar{y}$, $G_r(p_r) < \bar{x}$. Therefore, for $nk_1 > 1 - \bar{y} + \beta(1 - \pi_m(0))$ and $\rho_1 \in R$, $G_r(p_r) < \bar{x}$. Thus

$$\Phi_r(G_r(p_r)) < \Phi_m(0) < \Phi_m(G_m(p_m)).$$

Defining $\hat{K} = 1 - \bar{y} + \beta(1 - \pi_m(0)) < 1 - \pi_r(0) + \beta(1 - \pi_m(0))$, the result follows.

(c) From Lemma 1, for $q \in \{L_m + 1, \ldots, L_r\}$,

$$\frac{\partial F}{\partial \rho_1}(q) = \frac{\partial}{\partial \rho_1}\Big(\frac{\beta + F_r(q)}{1 + \beta}\Big) < 0.$$

We thus deduce Proposition 2. $\square$

*Proof of Proposition 3.* Consider any student $s$ who is assigned to a tier 1 school $c = \mu_\rho(s)$ that is not her top choice . Let $\bar{c}$ be the top choice of student $s$. Consider the (positive measure) set $\bar{S} \subset S$ of all students such that they rank school $c$ first, and school $\bar{c}$ second. Define $\hat{S} \subseteq \bar{S}$ by $\hat{S} = \{s' \in \bar{S}, \omega_c^{s'} < p_\rho < \omega_{\bar{c}}^{s'}\}$. By construction, $\hat{S}$ has positive measure. For any $s' \in \hat{S}$, $c \succ_{s'} \mu_\rho(s') = \bar{c}$. As a result, $s$ can Pareto improve by switching school with $s' \in \hat{S}$.

If $s$ is assigned to a tier 1 school that is her top choice, then it is clear that $s$ cannot Pareto improve by switching school.

If $s$ is assigned to a tier 2 school, then $s$ is either assigned to her top choice or $s$ would prefer a tier 1 school. If $s$ is assigned to her top choice, then $s$ cannot Pareo improve by switching school. If $s$ would like to move to some tier 1 school, then all students assigned to that tier 1 school prefer their current school to the tier 2 school $s$ is assigned to. So, $s$ cannot Pareto improve by switching school.

It thus follows that a student $s$ can Pareto improve by switching school iff $s$ is assigned to a tier 1 school that is not her top choice. It thus follows that

$$\begin{aligned} P(\rho_1) &= 1 - F(1, \rho_1) - \frac{\sum_{l_r=1}^{L_r} \pi_r(l) G_r(p_r)^{l_r} + \beta \sum_{l_m=1}^{L_m} \pi_m(l) G_m(p_m)^{l_m}}{1 + \beta} \\ &= 1 - F(1, \rho_1) - \frac{(1 - \pi_r(0)) + \beta(1 - \pi_m) - nk_1}{1 + \beta} \end{aligned}$$

which is decreasing in $\rho_1$ under the conditions of Proposition 2. $\square$

*Proof of Proposition 5.* The proof is identical to the proof of Proposition 2. $\square$

*Proof of Proposition 4.* The cutoffs $p_m^{SA}$ and $p_r^{SA}$ are entirely determined by the intersection of the market clearing (4.1) and set aside (4.4) conditions. The set aside condition is to the left of the minority reserve condition (see also Figure 6) and therefore $p_r^{SA} > p_r$ and $p_m^{SA} < p_m$. By increasing $\rho_1$, the minority reserve condition (4.2) moves to the left. As a result, we can find $\rho_1' > \rho_1$ such that the cutoffs $p_m'$ and $p_r'$ under minority reserves $\rho_1'$ satisfy $p_m^{SA} = p_m'$ and $p_r^{SA} = p_r'$. Proposition 2 implies that for a fixed $\rho_1$ more students are assigned to their top school under set asides than under minority reserves.

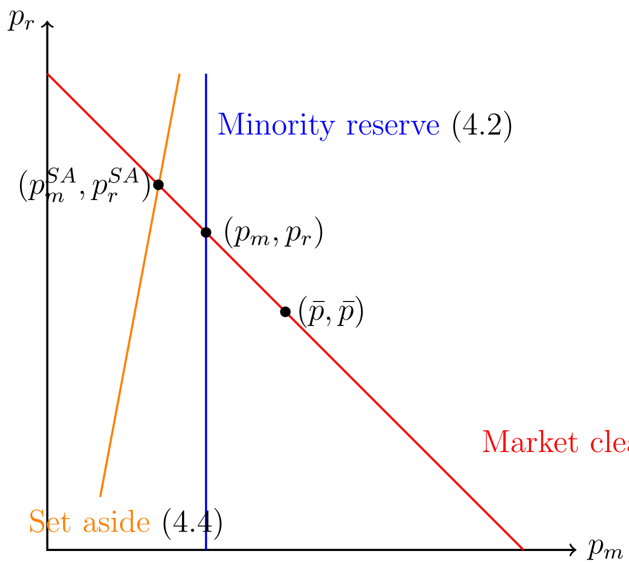



Figure 6: The market clearing condition and the set aside condition determine cutoffs $p_r^{SA}$ and $p_m^{SA}$. For a given $\rho_1$, the set aside condition is to the left of the minority reserve condition.

$\square$

**Proposition 6.** *Assume that* $L_r = L_m = L$. *If*

$$\sum_{l=1}^{L} \pi_r(l) l \big(1 - \frac{n(k_1 - \rho_1)}{1 - \pi_r(0)}\big)^{l-L} \leq \frac{\pi_r(L)}{\pi_m(L)} (1 - \pi_m(0))$$

*then* $\frac{\partial F}{\partial \rho_1}(L, \rho_1) < 0$.

*Proof.* Following the proof of Proposition 2, it follows that $\frac{\partial}{\partial \rho_1} F(q = L) < 0$ iff

$$\frac{\sum_{l=1}^{L} \pi_r(l) l G_r^{l-L}}{\sum_{l=1}^{L} \pi_m(l) l G_m^{l-L}} < \frac{\pi_r(L)}{\pi_m(L)}.$$

Since $G_r \geq 1 - \frac{n(k_1 - \rho_1)}{1 - \pi_r(0)}$,

$$\sum_{l=1}^{L} \pi_r(l) l G_r^{l-L} \leq \sum_{l=1}^{L} \pi_r(l) l \big(1 - \frac{n(k_1 - \rho_1)}{1 - \pi_r(0)}\big)^{l-L}$$

As $G_m \leq 1$,

$$\sum_{l=1}^{L} \pi_m(l) l G_m^{l-L} \geq (1 - \pi_m(0)).$$

Therefore,

$$\frac{\sum_{l=1}^{L} \pi_r(l) l G_r^{l-L}}{\sum_{l=1}^{L} \pi_m(l) l G_m^{l-L}} \leq \frac{\sum_{l=1}^{L} \pi_r(l) l \big(1 - \frac{n(k_1 - \rho_1)}{1 - \pi_r(0)}\big)^{l-L}}{1 - \pi_m(0)} \leq \frac{\pi_r(L)}{\pi_m(L)}.$$

□

# B Field data

## B.1 Markets

The Chilean centralized system runs nationwide. While any student could apply to any school in the country, virtually all students apply exclusively within their provinces or districts. The system is thus composed of several isolated markets. We show that each of our marketsis indeed isolated and virtually independent from the rest of the markets in the country.

We first define our markets. The Valparaiso market includes each school located in the provincial department of Valparaiso. The Concepcion market includes each school located in the provincial department of Concepcion. The Santiago market includes each school located in the Metropolitan Region of Santiago.[34] So, the boundary of each of our markets follows administrative definitions.

For each market, we consider all students that apply exclusively within the market. Thus a student with a rank order list including some schools in Valparaiso and others outside Valparaiso is excluded from our exercise. This set of students is small as big urban centers heavily concentrate applications. In our database, 99.76% of all nationwide applications listing some school in the Santiago market list

[34]As Santiago is the main urban center in Chile the country, the provincial department of Santiago excludes several towns close to Santiago whose students apply to schools in the city. We thus work with the Metropolitan Region of Santiago.

exclusively schools in Santiago. The numbers for Valparaiso and Concepcion are 98.85% and 99.66%, respectively. The following table shows the characterization for each market:

Table 7: Valparaiso, Concepcion and Santiago markets

| | Valparaíso | Concepción | Santiago |
|---|---|---|---|
| Number of provincial departments | 1 | 1 | 7 |
| Number of counties | 10 | 12 | 52 |
| Number of schools | 275 | 250 | 1,214 |
| Applicants to the market applying exclusively inside the market | 98.85% | 99.66% | 99.76% |

Thus, in practical terms, each of our markets is isolated and independent from all other markets in the country.

## B.2 Alternative popularity definitions

We explore an alternative definition of popularity. A school is $\nu$-popular if it is oversubscribed after running the DA algorithm (with no reserves) with probability at least $\nu$ (since schools define priorities using random inputs, a school may be oversuscribed for some but not all realizations of the lottery numbers). Clearly, a school $c$ with $pop(c) \geq 1$ will be $\nu$-popular for all $\nu \in [0, 1]$. For concreteness, set $\nu = 0.9$.

We construct the empirical distribution of application intensities for $\nu$-popular schools. As shown below, we obtain the exact same results reported in Section 2.3.2. This shows the robustness of our claim that minority students apply with less intensity to high demand schools.

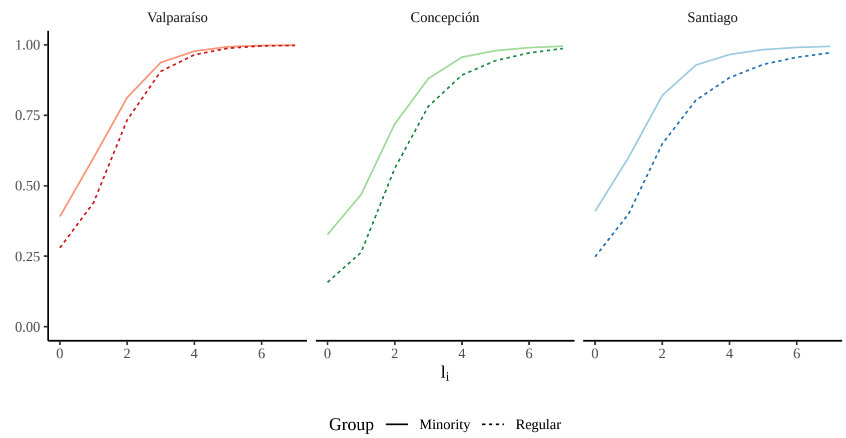



Figure 7: Empirical distributions of application intensities to $\nu$-popular schools for each group $t$.

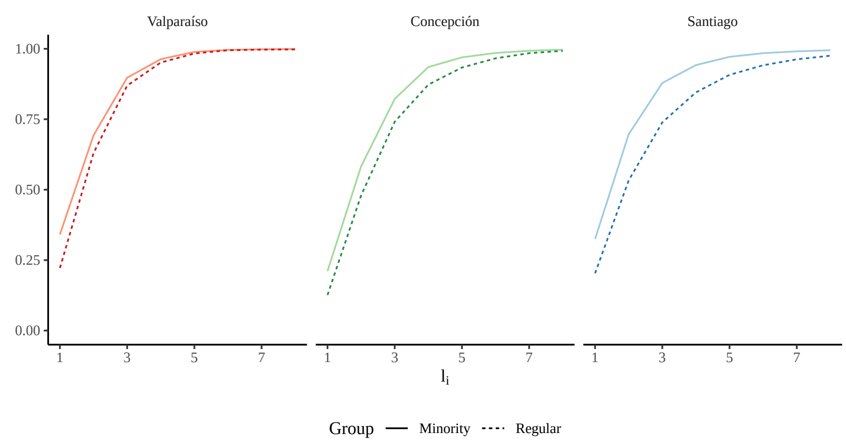



Figure 8: Empirical distributions of application intensities to $\nu$-popular schools, conditional on applying to at least one $\nu$-popular school.

# C Segregation in schools

In the main body of the paper, we have explored how an aggregate segregation index (the Duncan index) changes as minority reserves increase. Figure 9 shows how segregation in each school is determined by its popularity and by the minority reserve. Each school is an observation. As can be seen popular schools tend to have a lower fraction of minority students. The upper graphs are derived without any minority reserve. The upper graphs show that few popular schools have overrepresented minority students. The lower graphs are derived with minority reserves equal to the fraction of minority students in the population.

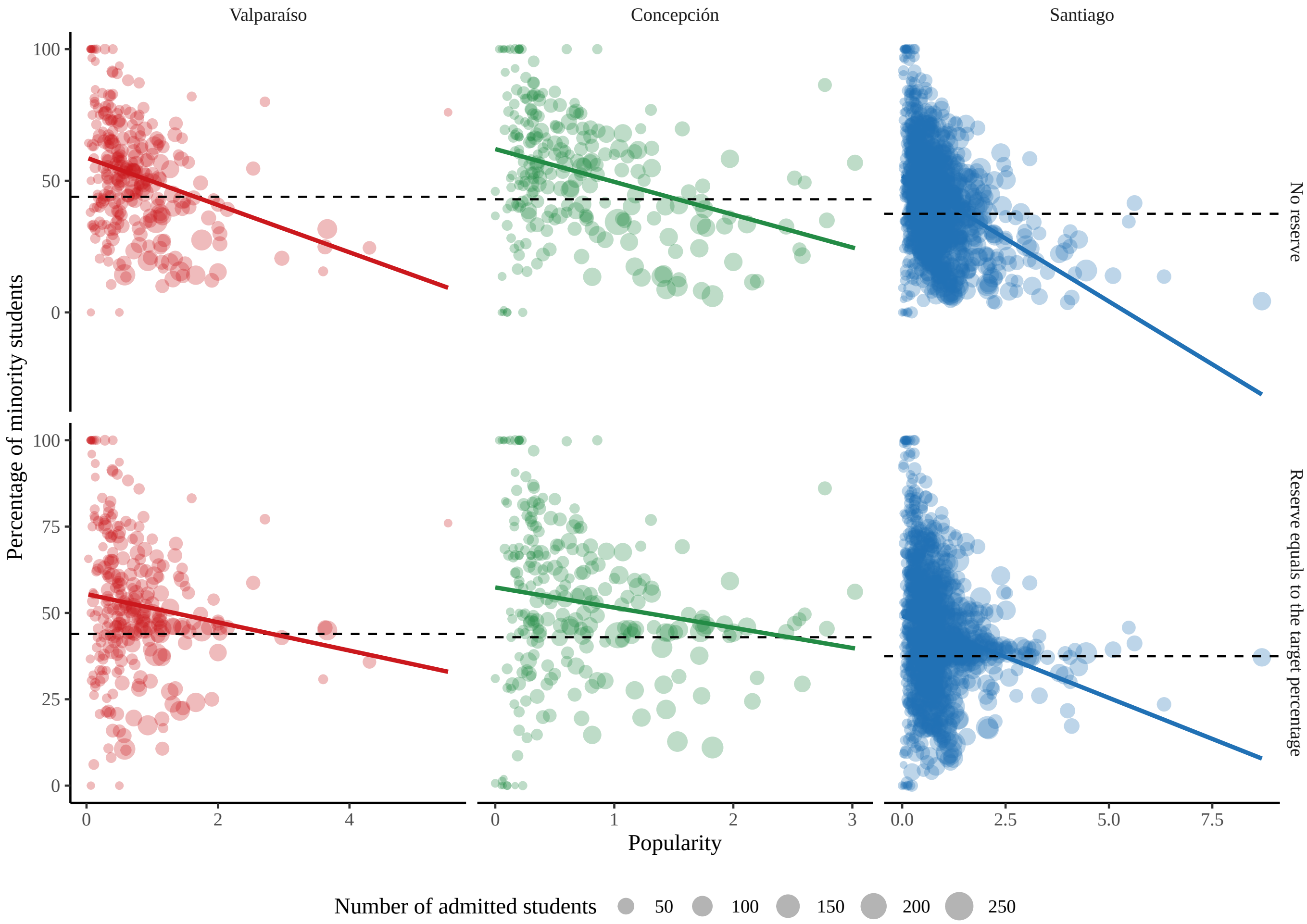



Figure 9: Schools composition. Minority students are under-represented in popular schools.

# D Other markets

## D.1 A slack market

Proposition 2 part b shows that when the market is not tight, increasing reserves reduces the fraction of students assigned to their top schools. To validate this claim empirically, we have built a slack market using the Santiago data as follows. First, we have dropped all regular students with application intensity less than or equal to 1 (that is, we dropped all regular students applying first to a non-popular school or applying first to a popular schools followed by a non-popular school). We recompute the popularity of schools and obtain a smaller set of popular schools. We again dropped students who place first a non popular school (this does not change the number of popular schools). The resulting market has 16,814 students (10,376 regular and 6,438 minority) and the percentage of minority students is

38.3%.

For each $x \in [0,1]$, we expand capacities as follows:

– If $pop(c) < 1$, then the number of seats of the school equals the number of students in the market.

– If $pop(c) \geq 1$, then the number of seats of the school equals $n_r(c) + xn_m(c)$ (approximated to the lowest integer), where $n_t(x)$ is the number of type $t$ students applying first to $c$.

For any given $x \in [0,1]$, the resulting market is more slack than the original Santiago market. As shown in Table 8, when $x = 0.6$, the market is slack and the fraction of students assigned to top schools decreases as reserves increase (when $x = 0.2$, the market is not slack enough). This is consistent with our theoretical results.

Table 8: Percentage of students assigned to their top choices in a fictitious slack market

| Value of $x$ | $f = 0\%$ | $f = 15\%$ | $f = 40\%$ | $f = 50\%$ | $f = 60\%$ | $f = 75\%$ | $f = 100\%$ |
|---|---|---|---|---|---|---|---|
| 0.2 | 66.28 (0.29) | 66.31 (0.26) | 66.21 (0.23) | 66.27 (0.24) | 66.62 (0.26) | 67.27 (0.24) | 68.02 (0.21) |
| 0.6 | 88.61 (0.25) | 88.63 (0.24) | 88.57 (0.23) | 88.44 (0.26) | 88.26 (0.25) | 87.98 (0.21) | 88.04 (0.22) |

## D.2 Simulations for 9th grade

The centralized system in Chile operates across all school levels, as documented in (Correa, Epstein, Escobar, Rios, Bahamondes, Bonet, Epstein, Aramayo, Castillo, and Cristi 2022). This implies that, except for PreK, a significant portion of students already have their assignments determined and do not engage with the centralized platform. We report simulations for the 9th grade in Santiago. While the 9th grade has a smaller student population compared to PreK, it boasts a substantially larger enrollment compared to all other levels.

Table 9: Santiago 9th grade

| Category | Value |
|---|---|
| Number of schools | 808 |
| Total capacity | 91,119 |
| Available seats | 49,116 |
| Number of students in the centralized system | 37,548 |
| Regular | 18,821 (50.13%) |
| Minority | 18,727 (49.87%) |
| Mean number of submited preferences | 4.1 |

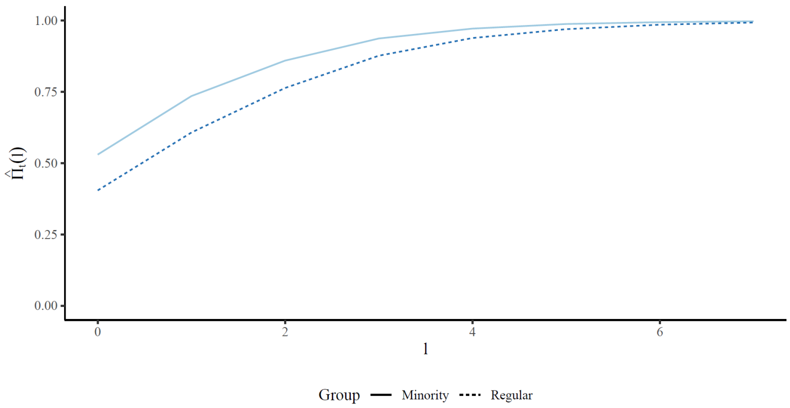



Figure 10: Application intensities. Santiago 9th grade.

Figure 10 shows that regular students apply to popular schools with higher intensity. As a result, Table 10 also show that our main simulation results hold for this grade: As reserves increase, the Duncan index is U-shaped and the number of students assigned to top schools increases.[35]

Table 10: Results summary

| Reserve | $f = 0\%$ | $f = 15\%$ | $f = 35\%$ | $f = 40\%$ | $f = 75\%$ | $f = 100\%$ |
|---|---|---|---|---|---|---|
| Duncan index | 0.234 (0.002) | 0.23 (0.002) | 0.218 (0.002) | 0.22 (0.002) | 0.309 (0.001) | 0.316 (0.001) |
| Minority students assigned to their top choice | 61.02 (0.17) | 61.4 (0.17) | 63.91 (0.19) | 64.97 (0.16) | 73.73 (0.1) | 74.21 (0.1) |
| Regular students assigned to their top choice | 50.28 (0.18) | 49.84 (0.18) | 47.8 (0.16) | 47.09 (0.15) | 41.75 (0.13) | 41.54 (0.13) |
| Students assigned to their top choice | 55.63 (0.11) | 55.61 (0.11) | 55.83 (0.11) | 56 (0.1) | 57.7 (0.08) | 57.83 (0.09) |
| Students assigned to their fourth choice or worst | 12.5 (0.08) | 12.5 (0.08) | 12.46 (0.08) | 12.46 (0.08) | 12.68 (0.07) | 12.68 (0.07) |
| Students unassigned | 7.89 (0.07) | 7.9 (0.07) | 7.98 (0.06) | 8.04 (0.07) | 8.43 (0.06) | 8.44 (0.05) |
| Students in Pareto improving pairs | 8.35 (0.23) | 8.39 (0.24) | 8.13 (0.23) | 7.94 (0.22) | 5.63 (0.18) | 5.39 (0.17) |

[35]We report the Duncan index for the final assignment, which includes students that did not participate in the centralized system.

# ONLINE APPENDIX

This Online Appendix contains further supportive results. Online Appendix E provides a version of Proposition 1 for alternative segregation indexes. Online Appendix F details application patterns in schools. Online Appendix G provides some evidence about the impact of the centralized platform in Chile on segregation.

# E Other segregation indexes

We adapt Proposition 1 for the Hutchens index (Hutchens 2004):

$$H_\mu = 1 - \sum_{c \in C} \sqrt{\eta^r_\mu(c) \cdot \frac{\eta^m_\mu(c)}{\beta}}$$

Note first that $H_\mu$ does not depend on $\rho_1$ when $\rho_1 \notin [\frac{\alpha_m \beta}{n}(1 - G_m(\bar{p}_m))^l_m, \min\{\frac{\alpha_m \beta}{n}, k_1\}]$.

Recall that for $\rho_1 \in [\frac{\alpha_m \beta}{n}(1 - G_m(\bar{p}_m))^l_m, \min\{\frac{\alpha_m \beta}{n}, k_1\}]$, each tier 1 school has $k_1 - \rho_1$ regular students and $\rho_1$ minority ones. Each tier 2 school has $\frac{1-n(k_1-r_1)}{N}$ regular students and $\frac{\beta - nr_1}{N}$ minority ones. Thus, for $\rho_1$ in this range, the H-index is computed as:

$$H_\mu = 1 - \underbrace{n\sqrt{\frac{\rho_1(k_1-\rho_1)}{\beta}}}_{H_1} - \underbrace{m\sqrt{\frac{(1-n(k_1-\rho_1))(\beta - n\rho_1)}{m^2\beta}}}_{H_2}$$

where the terms $H_1$ and $H_2$ correspond to the sum across tier 1 and tier 2 schools respectively.

Taking derivatives we get that:

$$\frac{\partial H_\mu}{\partial \rho_1} = -\frac{n}{2\sqrt{\beta}} \left( \frac{k_1 - 2\rho_1}{\sqrt{\rho_1(k_1-\rho_1)}} + \frac{\beta - 1 + n(k_1 - 2\rho_1)}{\sqrt{(1-n(k_1-\rho_1))(\beta - n\rho_1)}} \right)$$

And also that:

$$\frac{\partial^2 H_\mu}{\partial \rho_1^2} = \frac{n}{4\sqrt{\beta}} \left( \frac{k_1^2}{[\rho_1(k_1-\rho_1)]^{3/2}} + \frac{n(\beta + 1 - nk_1)^2}{[(1-n(k_1-\rho_1))(\beta - n\rho_1)]^{3/2}} \right) > 0$$

So we deduce that $H_\mu$ is a strictly convex function. Since $\frac{\partial H_\mu}{\partial \rho_1} = 0$ when $\rho_1 = \frac{\beta}{1+\beta} k_1$, the result follows.

The Atkinson index (Frankel and Volij 2011) can be defined in our setup as:

$$A_\mu = 1 - \left[ \sum_{c \in C} \eta^r_\mu(c)^\delta \cdot \left( \frac{\eta^m_\mu(c)}{\beta} \right)^{1-\delta} \right]^{\frac{1}{1-\delta}}$$

Where $\delta \in (0,1)$ is a fixed weight. In the symmetric case in which both types are treated equally in the segregation index, $\delta = \frac{1}{2}$ and thus the Atkinson index is obtained by an increasing transformation of the Hutchens index. The result follows.

# F Application patterns, standardized tests, and location

Understanding why minority students apply less to popular schools is beyond the scope of this paper. We observe that distance may be playing a role because minority students tend to live farther away from popular schools. To see this, in each market, we restrict our set of students to those that are market as *properly georeferenced* by the Chilean Ministry of Education[36]. For these set of students, we compute the distance to the closest popular school ($pop(c) > 1$) using the Vincenty (ellipsoid) method provided by the GEOSPHERE package from the R Statistical Software. The resulting distributions are presented below:

| | Valparaíso | | Concepción | | Santiago | |
|---|---|---|---|---|---|---|
| | Regular | Minority | Regular | Minority | Regular | Minority |
| Sample (number of students) | 2494 | 1833 | 2907 | 2029 | 22508 | 13089 |
| First quartile | 0.47 | 0.52 | 0.42 | 0.48 | 0.38 | 0.42 |
| Median | 0.81 | 0.92 | 0.70 | 0.82 | 0.63 | 0.69 |
| Mean | 2.86 | 1.37 | 1.13 | 1.27 | 0.97 | 1.19 |
| Third quartile | 1.36 | 1.59 | 1.16 | 1.38 | 1.02 | 1.07 |

Table 11: Distance (Km.) to the closest popular school

[36]Students that shared their location when applying on the platform or those whose location held a unique response and was marked as "rooftop" or "range_interpolated" in the "location_type" variable of Google's Geocoding API.

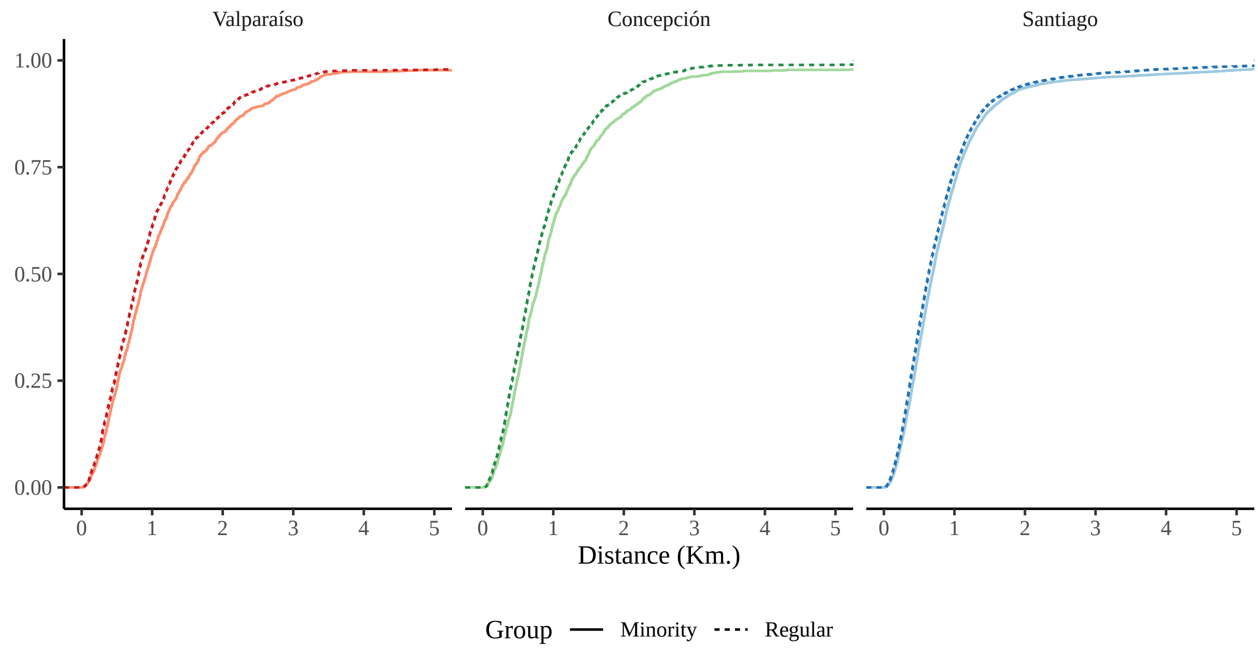



Figure 11: Distance to closest popular school. Minority students live farther away from popular schools than regular students.

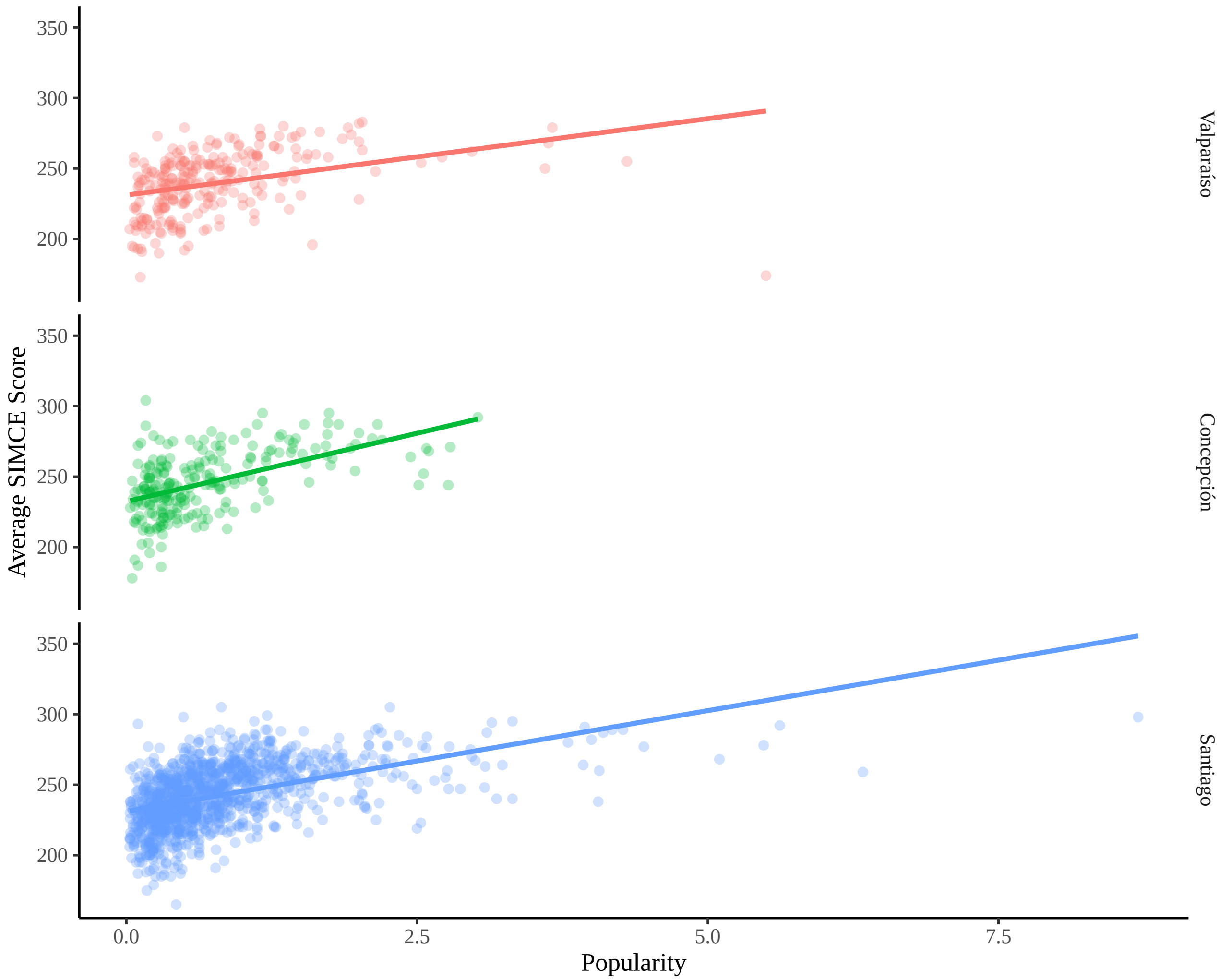
As discussed in the text, popular schools tend to perform better in standardized tests. For each market, we restrict our set of schools to those such that: (1) took part in SIMCE 2015 test[37] (2) reported valid SIMCE scores. This slightly decreases the set of schools we considered (so Table 12 has fewer schools than Table 7). We only use data from the Language test of second degree students in 2015. Popular schools are those such that $pop(c) > 1$.

Table 12: SIMCE scores

| | Valparaíso | | Concepción | | Santiago | |
|---|---|---|---|---|---|---|
| | Not popular | Popular | Not popular | Popular | Not popular | Popular |
| Sample (number of schools) | 200 | 61 | 185 | 53 | 884 | 302 |
| First quartile | 222 | 244 | 224 | 259 | 223.75 | 248.00 |
| Median | 239 | 259 | 239 | 269 | 236 | 260 |
| Mean | 234.84 | 254.21 | 239.21 | 267.47 | 236.87 | 259.10 |
| Third quartile | 250 | 269 | 252 | 277 | 251 | 271 |

[37]SIMCE is a standardized test taken to all students in the country

Figure 12: Popular schools tend to have higher SIMCE scores




These results show that popular schools have better performance in standardized tests. Obviously, this exercise is just illustrative and we are not claiming any causal effect.

# G Centralized platform and its impact on segregation

We now present some evidence about the impact of the centralized system on segregation in Chilean cities. Our main data set is built using 3 sources of information:

- *Student enrollment*, available at:
  http://datos.mineduc.cl/dashboards/19776/descarga-bases-de-datos-de-matricula-por-estudiante/
- *Disadvantaged students*, available at:
  http://datos.mineduc.cl/dashboards/19939/bases-de-datos-alumnos-prioritarios/

- *SAE's supply 2020*, available at:

  http://datos.mineduc.cl/dashboards/20514/descarga-bases-de-datos-de-los-proceso-de-admision-escolar-anos-2016-y-2017/

We consider data from 2013 onward. For each year and each of the 16 regions in Chile, we define a market by selecting every school located in the *main provincial department* of the region. We consider every student enrolled in Pre Kinder in schools that were part of SAE in 2020 (this excludes private schools). We build a dummy variable ($\text{HAS\_SAE}_{i,t}$) if the assignment in region $i$ and year $t$ was centralized using SAE. As discussed in the main body of the paper, SAE was gradually introduced in the country. For each region $i$ and year $t$, we also compute the fraction of disadvantaged students ($\text{DISAD\_FRAC}_{i,t}$) and the Duncan index ($\text{SEG\_INDEX}_{i,t}$).

The following are the main regression results. For different specifications, the introduction of SAE results in a relatively modest reductions of the Duncan index. This shows that the impact that minority reserves have on segregation is relatively important.

Table 13: Regression results

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | seg_index | | | | |
| | *OLS* | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| has_sae1 | −0.012 (0.011) | −0.005 (0.012) | −0.012 (0.009) | −0.034 (0.023) | −0.017* (0.010) |
| disad_frac | | 0.075 (0.057) | −0.002 (0.083) | | |
| Region fixed effect | | | Yes | | Yes |
| Year fixed effect | | | | Yes | Yes |
| Constant | 0.313*** (0.006) | 0.269*** (0.034) | 0.241*** (0.049) | 0.292*** (0.013) | 0.221*** (0.009) |
| Observations | 122 | 122 | 122 | 122 | 122 |
| $R^2$ | 0.009 | 0.024 | 0.717 | 0.191 | 0.893 |
| Adjusted $R^2$ | 0.001 | 0.007 | 0.671 | 0.134 | 0.868 |
| Residual Std. Error | 0.056 (df = 120) | 0.056 (df = 119) | 0.032 (df = 104) | 0.052 (df = 113) | 0.020 (df = 98) |
| F Statistic | 1.112 (df = 1; 120) | 1.438 (df = 2; 119) | 15.535*** (df = 17; 104) | 3.339*** (df = 8; 113) | 35.700*** (df = 23; 98) |
| *Note:* | | | | | *p<0.1; **p<0.05; ***p<0.01 |