Search and Information in Centralized School Choice Systems*

Juan F. Escobar† — University of Chile

Alfonso Montes‡ — University of Chile

February 9, 2024

Abstract

Centralized school assignment mechanisms play an important role in educational policy worldwide. In these systems, families face the non-trivial task of discovering and ranking schools. We evaluate the impact of information protocols on equilibrium search behavior and social welfare. We study a large market model in which students are assigned to schools using the deferred acceptance algorithm. We show that full transparency about the number of seats in the market is suboptimal. We also examine the effects of disclosing information about schools that are likely to be attractive to students, showing that transparency regarding top choices reduces congestion and increases welfare. Our analysis provides new insights for market designers as information interventions may subtly affect behavior and welfare.

1 Introduction

Centralized mechanisms for assigning students to schools are an essential tool for policymakers globally. The deferred acceptance algorithm proposed by Gale and Shapley (1962) is a common practice in this domain, providing a stable matching based on the

*Escobar acknowledges financial support from the Institute for Research in Market Imperfections and Public Policy, MIPP, IS130002. Escobar and Montes acknowledge financial support from the University of Chile, Proyecto Anillo ACT 210005.

†Email: juanescobar@uchile.cl

‡Email: amontes@dii.uchile.cl

preferences of schools and students. While this algorithm ensures truthfulness in reporting preferences (Roth, 1982), families face the daunting task of discovering and ranking schools.

The resultant uncertainty presents challenges and opportunities for authorities. Policymakers often institute protocols to assist families in their applications within matching markets. Some interventions, such as those detailed by Arteaga et al. (2022) and Elacqua et al. (2022), provide families with information about market congestion levels, encouraging them to apply to a broader range of schools. Others, as discussed by Hastings and Weinstein (2008), Andrabi et al. (2017), and Correa et al. (2022), offer information about schools' characteristics, aiding families in identifying attractive options based on location and academic performance.

While information protocols offer valuable guidance, it is well-established in the literature that information in markets and multi-person interactions can have subtle effects and, if mismanaged, may lead to reduced welfare.[1] This observation becomes particularly significant in matching markets utilizing the deferred acceptance algorithm. Here, the efforts of individual families to search and apply to schools may impact the admission chances of others and create excess demand for certain schools.

Our primary goal is to shed light on the impact of information interventions in matching markets. Focusing on a centralized school choice setting where students are assigned to schools using the deferred acceptance algorithm, we explore the policymaker's role in designing information protocols. Is transparency socially desirable? Should authorities refrain from disclosing some information? Is the impact of the intervention different when information is about admission probabilities than when information is about characteristics of schools?

To address these questions, we employ a large market model where a continuum of students applies to a finite number of schools (Azevedo and Leshno, 2016). Students face uncertainty about the available seats, and all schools have excess demand. Families discover schools through costly searching, deciding how many schools to inspect to form rank order lists. Families submit applications, schools rank applicants randomly, and the deferred acceptance algorithm outputs a stable matching (Gale and Shapley, 1962).

In our model, students face uncertainty about market congestion and do not know the schools they like. Students can inspect schools. The number of schools a student

[1]The literature includes Hirshleifer (1971), Kamien et al. (1990), Levin (2001) and recent works on information design in games such as Taneva (2019) and Bergemann and Morris (2016).

inspects depends on market fundamentals, including her belief about the supply of seats. Critically, search effort is determined by admission probabilities, and since these probabilities depend on the search intensity of all market participants, search behavior must be determined in equilibrium. Our results characterize equilibrium search patterns and the influence of information on equilibrium behavior and welfare.

Our first set of results characterizes equilibrium behavior. We show that by increasing the number of searched schools, a student reduces the chances of other students getting admission to their listed schools. Searching thus generates a negative externality. We also show that the model has strategic complementarities: the more a student searches, the larger the number of schools inspected by other students (Milgrom and Roberts, 1990). In equilibrium, students over-search and equilibria are Pareto-ranked.

The second set of results describes how information provision about the number of seats in the market changes equilibrium behavior and welfare. We show that full transparency is never optimal. When capacity turns out to be low, market congestion intensifies, and search behavior becomes overly intense under full transparency. Using information design tools, we show that a policymaker can avoid this outcome by withholding some information (Kamenica and Gentzkow, 2011).

Our final set of results explores the impact of disclosing information about schools. When the policymaker knows and discloses the name of her most preferred school to each student, the incremental value of learning is reduced for the student, resulting in reduced equilibrium congestion. By simplifying the decision problem of each family and reducing congestion, disclosing top schools results in welfare gains. In contrast, disclosing unattractive schools likely to be listed at the bottom of the rank order list increases equilibrium congestion and results in welfare losses.

Our analysis hinges on strategic complementarities in search behavior. This property depends on two critical restrictions in the model: valuations for schools follow independent distributions governed by a distribution that satisfies the monotone hazard rate property, and market congestion remains bounded. Exploiting these constraints, we demonstrate that as the search behavior of one student intensifies, so does that of others. This property allows us to leverage the theory of supermodular games to derive comparative statics results with respect to beliefs and to evaluate the impact of information interventions on market outcomes (Milgrom and Roberts, 1990).

Beyond its technical implications, the notion that a student's search intensity increases with the actions of others is empirically plausible. Arteaga et al. (2022) shows evidence from school choice programs in Chile and the US, where some families are informed about market congestion. They show that students engage in more active search behavior when they hold less optimistic views about their admission possibilities, substantiating the idea of strategic complementarities in search patterns.

Regarding practical implications, our theory provides important insights for the design of informational policies in matching markets.

Providing information that motivates increased search and applications is not guaranteed to result in Pareto improvements. Students who expand their rank order list after the information intervention would undoubtedly improve their admission chances, but that comes at the cost of leaving some other students unassigned or assigned to worse schools. Our model suggests that the evaluation of policies that disclose information to motivate search needs careful consideration of the winners and losers created by the intervention.

Targeted information interventions to motivate increased search could encourage some students to add schools with available seats, as in (Arteaga et al., 2022). While those seats would still be available in the aftermarket,[2] information interventions during the regular application process could facilitate search. As the interventions expand, the gains from additional applications to schools with slack capacity would eventually exhaust and the congestion effects captured by our model will become relevant.

Our model suggests that limiting the number of applications may be socially desirable. Imposing upper bounds on the number of schools students can apply to, a policy present in many school choice and college admission systems, could be an effective tool. Additionally, maintaining families' optimism about their admission chances may serve as another means to avoid an excessive number of applications.

Providing families with personalized information about schools they are likely to value has double benefits. It allows families to economize on search costs while simultaneously reducing market congestion and increasing overall welfare. Policies informing families about nearby schools or those with excellent academic performance could be particularly beneficial, but only when the disclosed schools align with the families' preferences.

Our theory provides a simple test to evaluate the welfare impact of a protocol disclosing personalized information about schools. The suggested schools' positions in the

[2]School choice systems typically have an aftermarket procedure in which students can get admission to uncongested schools.

rank order list serve as a clear indicator: when placed at the top, welfare is likely to increase due to reduced congestion; when situated towards the bottom, the protocol may contribute to congestion and reduce welfare.

Related Literature. Our paper relates to several theoretical works exploring informational aspects in matching markets.[3] Chen and He (2022) compares information acquisition in the Immediate and deferred acceptance algorithms. Most closely related, Immorlica et al. (2020) explores a college admissions model in which students form their preferences by acquiring costly information. To capture the information acquisition process, they explore the notion of regret-free stability and show how historical information can be used to induce efficient price discovery. In our model, students acquire information about schools before the matching. This is particularly relevant in school choice markets where the supply of seats and schools is uncertain and changes over time. Our results emphasize how students search under uncertainty and how information policies alter search behavior. We thus view our paper as complementary to Immorlica et al. (2020).

Our paper builds on the old idea in the matching literature that changing an agent's choices in a two-sided matching market has different consequences on both sides of the market. For instance, Kelso Jr and Crawford (1982) shows that adding workers to the market leaves firms better off, while increasing the number of firms leaves workers better off. Gale and Sotomayor (1985) shows that adding women to a marriage market can never make any of the men worse off.[4] Using choice rules as the tool to handle the agent's chosen alternatives, Echenique and Yenmez (2015) shows that expanding a college choice rule in a college-student matching market makes students better off. More generally, Chambers and Yenmez (2017) shows that expanding the choice rules of an agent on one side of a (many-to-many) matching market benefits the agents on the other side while worsening agents on the same side of the market. Thus, when a student adds a school to their application list in a centralized school admission system, all remaining students are weakly worse off. Our model captures such an effect as a negative externality that increases the system's congestion. We endogenize the length of the student's list through an individual search process, and we study how the system's congestion responds to

[3]For empirical work on informational policy interventions, see Arteaga et al. (2022), Andrabi et al. (2017), Hastings and Weinstein (2008).

[4]See also Blum et al. (1997), Kamada and Kojima (2014) and Konishi and Ünver (2006).

the fundamentals of the model. We show that the strategic interaction through the centralized school system results in an inefficient Nash equilibrium.

Finally, our paper relates to the literature on information design in congestion games. This literature studies how information can be used to reduce congestion in the context of traffic networks.[5] Our paper also uses tools from the Bayesian persuasion literature to study welfare-maximizing information policies in a congested two-sided matching market. Our optimal information intervention shares the intuition of the optimal policy in Das et al. (2017). Giving full information when a road has a lagged delay over crowds the non-congested road, and the authorities can do better by hiding information to reduce congestion. We apply similar tools to a different framework, so these results do not imply our conclusions.

The rest of the paper is organized as follows. Section 2 sets up the framework. Section 3 studies the strategic consequences of search and establishes the inefficiency of Nash equilibria. Section 4 characterizes the optimal public information disclosure policy regarding the system's capacities. Section 5 shows the effect of private information about families' valuation for schools on the equilibrium congestion. Finally, Section 6 concludes.

2 The Model

We consider a school choice model with a finite set of schools $S := \{1, \dots, N\}$, with $N \geq 2$, and a continuum of students in the interval $[0, 1]$. Students have types in $\{1, \dots, I\}$ and for each type, there is a fraction $\rho(i) \in]0, 1[$ of students having type i, with $\sum_{i=1}^{I} \rho(i) = 1$. All schools have the same capacity k, which is randomly realized from a finite set $K \subset]0, \infty[$ according to a distribution $\mu \in \Delta(K)$. We assume for all $k \in K$, $k < 1/N$.

Searching for schools. Students search for schools. Each student receives u_s from attending school s, which is distributed independently (across students and schools) according to a commonly known cumulative distribution function F over $\mathbb{R}_+$. We denote the derivative of F by f. We assume the hazard rate $\frac{f(u)}{1-F(u)}$ is non-decreasing and the reverse hazard rate $\frac{f(u)}{F(u)}$ is non-increasing. These conditions are satisfied when the

[5] See for instance Acemoglu et al. (2018), Wu et al. (2017) and Tavafoghi and Teneketzis (2017).

density is log-concave (that is, $\ln(f(u))$ is concave). Several commonly used distributions, including the uniform distribution, are log-concave. An unassigned student receives utility 0. We can also interpret unassigned students as students that are assigned to unattractive schools that have excess supply of seats.

Each student decides a number in $\{0, 1, \ldots, N\}$ of schools to be searched. After n is chosen by a student, a random (and uniformly chosen) list of schools $S' = \{s_1, \ldots, s_n\} \subseteq S$ of size n is realized. The student observes n realizations from F, the highest realization is the utility the student derives from school s_1, the second highest realization is the utility the student derives from school s_2, and so on. The student thus ranks the schools in S' and forms a rank order list containing the n schools in S'. In particular, the rank order list only contains schools the student inspects. Searching for n schools costs $c_i \cdot n$ to a student, where i is the type of the student and c_i is the type-dependent marginal cost of searching for schools. We assume that for all i, $0 < c_i < \mathbb{E}_F[u] = \int u dF(u)$ so that all students always search when admission to schools is ensured.[6] This cost captures the opportunity cost of time of inspecting and discovering schools.

By searching, a student discovers schools and learns how attractive the discovered schools are. By expanding the number of inspected schools, a student enlarges her rank order list and observes the values of additional schools. How attractive searching for schools is will depend on the procedure used to assign students to schools, which is what we describe now.

Centralized assignment. After searching for schools, students participate in a matching algorithm that assigns students to schools. Let $\bar{n}_i \in \{0, \ldots, N\}$ be the number of schools searched by type i students. Students submit rank order lists and schools rank applicants by drawing independent scores from the uniform distribution over $[0, 1]$.[7] After $k \in K$ is realized, the school district runs the (student-proposing) Gale-Shapley deferred acceptance algorithm (Gale and Shapley, 1962).

A profile $\bar{n} = (\bar{n}_i)_{i=1}^{I}$ induces a probability distribution $\pi^{\bar{n}} \in \Delta(\{1, \ldots, N\})$ over the number of searched schools. Formally, the measure of students searching for $l \in$

[6] Throughout the paper, for a random variable x distributed according to G, we use the term $\mathbb{E}_G[x]$ to denote the expectation of x with respect to the distribution G.

[7] We are thus assuming that ties are broken using multiple lotteries. We discuss the tie-breaking rule later in the paper.

$\{1, \ldots, N\}$ schools is

$$\pi_l^{\bar{n}} := \sum_{\{i: \bar{n}_i = l\}} \rho(i). \tag{2.1}$$

Stable matchings can be characterized by solutions to market clearing conditions (Azevedo and Leshno, 2016). Since schools are ex-ante identical, we can obtain a stable matching through a unique cutoff $p \in [0, 1]$ such that a student is assigned to the school she prefers the most among those where her scores exceed the cutoff p. Given a profile of search strategies $\bar{n} = (\bar{n}_i)_{i \in I}$, the cutoff p must satisfy the following market clearing condition:

$$\sum_{l=1}^{N} \pi_l^{\bar{n}} \sum_{\eta=1}^{l} (1-p) p^{\eta-1} \frac{1}{N} = k \tag{2.2}$$

To understand this condition, note that given that l schools have been ranked, the probability that a school is in the rank order list is l/N, and it will be ranked in position η with probability $1/l$. This explains the term $1/N$ on the left-hand side of (2.2). Finally, the term $(1-p)p^{\eta-1}$ is the probability that a student that ranks school in the η-position is accepted in the school. By rearranging equation (2.2), we obtain

$$\sum_{l=1}^{N} \pi_l^{\bar{n}} p^l = 1 - Nk. \tag{2.3}$$

When $\pi_0^{\bar{n}} < 1$ the market clearing condition (2.3) has a unique solution $p = p_{k,\bar{n}} \in [0, 1]$, where we emphasize the fact that the equilibrium cutoff depends on the realized capacity k and the profile $\bar{n} = (\bar{n}_i)_{i \in I}$ (through the distribution π^n).[8]

We are interested in equilibrium search patterns. The benefit that student obtains when exploring $n > 0$ schools is the expected utility from the school she is assigned to. Formally, given a profile $\bar{n} = (\bar{n}_i)_{i=1}^{I}$, the realized capacity k, and a cutoff $p = p_{k,\bar{n}}$, the expected benefit that a type i students obtains when searching for $n \in \{0, \ldots, N\}$ schools equals

$$b(n, p_{k,\bar{n}}, k) := \begin{cases} \mathbb{E}_{F_n} \left[\sum_{l=1}^{n} \tilde{u}_n^l (1 - p_{k,\bar{n}}) p_{k,\bar{n}}^{l-1} \right] & \text{if } n > 0 \\ 0 & \text{if } n = 0, \end{cases} \tag{2.4}$$

[8]When $\pi_0^n = 1$, schools receive no applications and we thus define $p_{n,k} = 0$. See Azevedo and Leshno (2016) for details.

where $\tilde{u}_n = (\tilde{u}_n^1, \ldots, \tilde{u}_n^n)$ is the vector of n ordered independent draws from the distribution F, whose (joint) distribution we denote by F_n. We assume that the draws are ordered from top to bottom, i.e. $\tilde{u}_n^1 \geq \cdots \geq \tilde{u}_n^n$. By defining

$$u_i(n, p_{k,\bar{n}}, \mu) := \mathbb{E}_\mu[b(n, p_{k,\bar{n}}, k)] - c_i \cdot n \tag{2.5}$$

we can characterize a Nash equilibrium of the search game G as a profile $\bar{n} = (\bar{n}_i)_{i \in I}$ such that for all i

$$\bar{n}_i \in \underset{n' \in \{0,\ldots,N\}}{\arg\max}\ u_i(n', p_{k,\bar{n}}, \mu). \tag{2.6}$$

This condition captures the idea that a student should search for a given number of schools to maximize the expected utility of the school where she will be assigned, taken the search behavior of all other students (including those of her own type) as given.[9] We denote by $NE(\mu)$ the set of all Nash equilibria given beliefs μ.

Since search costs are heterogeneous, in equilibrium some students search intensely and apply to several schools while others apply to few schools. Students are forward looking and Bayesian. In particular, students understand that the realized capacities will determine admission chances, but that their information about those admission chances is imperfect.

3 Congestion and Strategic Complementarities

This Section characterizes equilibrium search behaviour. We offer two main substantive results. First, we show that by increasing the number of searched schools, a student reduces the chances of other students getting admitted to their listed schools. Searching is thus a negative externality. Second, the more schools a student searches, the larger the number of schools searched by other students. In other words, our search game is supermodular and exhibits strategic complementarities (Milgrom and Roberts, 1990).

Define $B(n, \bar{n}, k) = b(n, p_{k,\bar{n}}, k)$ for $n \in \{1, \ldots, N\}$, $\bar{n} = (\bar{n}_i)$, $k \in K$, and probability measure μ over K. We also write $U_i(n, \bar{n}, \mu)$ for the expected utility of a type i student that searches for n schools when all other students search for $\bar{n}$ schools, given belief μ about k. Our first result describes the properties of the function $B(n, \bar{n}, k)$.

Lemma 1 (Negative externalities)**.** *The following hold:*

[9] Note that no player can change the probability distribution π when unilaterally changing her strategy.

a. $B(n, \bar{n}, k)$ *is increasing and concave in* n*.*

b. $B(n, \bar{n}, k)$ *is non-increasing in* $\bar{n}$*. Moreover, for* $n \geq 1$*,* $B(n, \bar{n}, k)$ *is decreasing in* $\bar{n}$*, that is, for* $\tilde{n} = (\tilde{n}_i)_i$ *and* $\bar{n} = (\bar{n}_i)_i$ *with* $\tilde{n}_i \geq \bar{n}_i$ *for all* i*,* $B(n, \bar{n}, k) \geq B(n, \tilde{n}, k)$

The fact that $B(n, \bar{n}, k)$ is increasing in the number of searched schools n captures the idea that searching for schools benefits the student. The concavity of the function shows that adding a school has a decreasing incremental value as more schools are already in the application. B is concave in n, and so is U. As a result, there exists a $N_0 \leq N$ such that a student searches at most N_0 schools.[10]

Since $B(n, \bar{n}, k)$ is decreasing in $\bar{n}$, a student is hurt when other students search more. Intuitively, when $\bar{n}$ increases, schools are more congested and therefore, the market clearing cutoff $p_{k,\bar{n}}$ rises. A higher cutoff reduces the admission chances of a student searching for $n \geq 1$ schools. In our model, searching for schools is a negative externality.

The negative externality result relies on the assumption that searched schools are oversubscribed. If students discover under-demanded schools, by searching, they could add schools with available seats and alleviate congestion in over-demanded schools. Our analysis is relevant in environments in which most search resources are spent to discover schools with more demand than available seats.

We now explore strategic feedback effects in our game. Fixing the prior belief μ and the profile of searched schools by all students in the market $\bar{n} \in \{0, \ldots, N_0\}$, consider the search incentives of a type i student:

$$BR_i(\bar{n}; \mu) = \underset{n \in \{0, \ldots, N_0\}}{\arg\max}\ U_i(n, \bar{n}, \mu).$$

We will show that the best response map BR_i is non-monotonic in $\bar{n}$. However, over an important range of parameters, $BR_i(\bar{n}; \mu)$ is non-decreasing in $\bar{n}$. Since BR_i is a set-valued map, we say that $BR_i(\bar{n}; \mu)$ is non-decreasing in $\bar{n}$ if $\max_n\{n \in BR_i(\bar{n}; \mu)\}$ and $\min_n\{n \in BR_i(\bar{n})\}$ are both non-decreasing in $\bar{n}$.

Let $u_n^l := \mathbb{E}[\tilde{u}_n^l]$ be the expected value of the school ranked l after searching for $n \geq l$ schools. Denote $u_n^{n+1} = 0$ and $\Delta_n^l := u_{n+1}^l - u_n^l$. Define $\tilde{p}$ by

[10] We define N_0 as:

$$N_0 := \max_{\{\bar{n}, \mu\}} \left\{ \underset{\{n\}}{\arg\max}\ U(n, \bar{n}, \mu) \right\}$$

$$\tilde{p} := \min_{n\in\{1,\dots,N_0\}} \left\{ 1 - \left(\frac{1}{n+1}\right) \frac{\sum_{l=1}^{n} \Delta_n^l}{u_{n+1}^{n+1}}, \min_{l\in\{1,\dots,n\}} \{\frac{\Delta_{n+1}^l}{\Delta_n^l}\} \right\}. \tag{3.1}$$

As shown in the Appendix, the decreasing hazard rate property implies $\tilde{p} \in]0,1[$.

Assumption 1. *For all $k \in K$, $k \geq \frac{1}{N}\left(1 - \tilde{p}^{N_0}\right) := \tilde{k}$.*

Assumption 1 says the school system has a non-negligible capacity. It implies that admission cutoffs are bounded away from 0, and therefore, students search knowing that there are non-trivial opportunities to get admission to over-demanded schools.

Proposition 1 (Strategic complementarities)**.** *Under Assumption 1, $BR_i(\bar{n};\mu)$ is non-decreasing in $\bar{n}$ for all i and all μ.*

This result is important for two reasons. First, it reveals that strategic feedback effects are positive: the incentives of a student to search are stronger when other students search more and admission chances are lower. This is an empirically plausible property. For example, in a school choice setup, Arteaga et al. (2022) show that students participating in centralized platforms under the deferred acceptance algorithm apply to more schools when admission chances are revealed lower. Second, the fact that the best response map is non-decreasing provides important insights about the basic economic forces behind search behavior and allows us to neatly characterize equilibria using the theory of supermodular games (Milgrom and Roberts, 1990).

Bounding cutoffs by above –as Assumption 1 does– is key for Proposition 1. When the assumption does not hold, the best a student can do when others search a lot is not to search.[11] In this case, $BR_i(\bar{n};\mu) = 0$ for $\bar{n}$ large enough; therefore, BR_i has a decreasing portion. In the proof of Proposition 1, we also establish that the expectation of the order statistic u_n^l is supermodular in (l,n), which is a new property that can be useful in other applications (see Lemma 5).

Proposition 2 (Properties of Nash equilibria)**.** *Under Assumption 1, the following hold.*

a. The game G has a nonempty set of Nash equilibria: $NE(\mu) \neq \emptyset$.

b. The set of Nash equilibria $NE(\mu)$ has smallest and a largest element.

[11]When Assumption 1 does not hold and others search a lot, the admission cutoffs are high, and therefore, any search effort is likely to be wasteful.

c. *Let $\bar{n}$ and $\tilde{n}$ be Nash equilibria with $\bar{n}_i \leq \tilde{n}_i$ for all i. Then, all students get higher payoffs under equilibrium $\bar{n}$ than under equilibrium $\tilde{n}$.*

d. *The smallest and largest Nash equilibria are non-increasing in the distribution $\mu \in \Delta(K)$, where we endow $\Delta(K)$ with the first order stochastic dominance.*

Proposition 2 establishes important properties of the set of Nash equilibria. It guarantees the existence of Nash equilibria. As in many games with strategic complements, there may be several equilibria.[12] Notably, there is a smallest and a largest Nash equilibrium. Equilibria can be Pareto-ranked: students prefer an equilibrium with low search. The result also shows that as μ increases and more weight is placed on high capacities, equilibrium search decreases.

Our main focus is on information interventions. Information interventions are common in practical applications of matching theory. They range from information about admission chances to suggestions of schools that are likely to be attractive to students. Surprisingly, little is known (theoretically or empirically) about the system-wide impacts of these interventions. The rest of the paper endeavours to fill this gap.

4 Disclosing Congestion

Several interventions provide information about congestion and admission chances to motivate students to apply to more schools. For example, Arteaga et al. (2022) report interventions in Chile and New Haven, where students receive information about their admission possibilities and are encouraged to add more schools to their applications. Elacqua et al. (2022) shows a similar intervention in Peru's centralized assignment of school teachers. This Section offers a counterpoint to those exercises by showing that fully disclosing how congested the market is has subtle equilibrium effects and results in welfare losses.

[12] Different school choice systems exhibit important differences in the lengths submitted applications. For example, in NYC high school matches, students rank seven schools on average, while in Chile, families rank only three schools (Abdulkadiroğlu et al., 2009; Correa et al., 2022). Of course, these differences could be explained by different market fundamentals. Our search game shows that identical markets may have different equilibrium application intensities.

4.1 Public Experiments

We consider game G and explore how information about the supply of schools k changes equilibrium outcomes. More specifically, we build on the information design literature and consider a social planner who can run a public experiment (Kamenica and Gentzkow, 2011). A public experiment consists of a finite realization space T and a family of probability distributions $\tau := (\tau_k)_{k \in K}$, where $\tau_k \in \Delta(T)$ for $k \in K$. We assume that students publicly observe the signals drawn from the experiment.

The timing of information provision is as follows. First, the planner commits to an experiment τ. Second, the state k is drawn according to μ, and the public signal is drawn according to τ_k. After observing $t \in T$, students form posterior beliefs $\gamma_t(k) = \mathbb{P}(k \mid t) = \frac{\tau_k(t)\mu(k)}{\sum_{k'} \tau_{k'}(t)\mu(k')}$. Finally, students play the game G given updated beliefs γ_t. We assume the planner can induce any equilibrium given beliefs γ_t.

An experiment τ publicly reveals information about capacities k. For example, when the realization space is $T = K$, and τ_k puts a weight of one on $t = k$, the realization of the experiment perfectly reveals the supply of seats in the system. In contrast, when τ_k does not depend on k, students learn nothing by observing the realized signal. More generally, an experiment τ may provide partial information, leaving students uncertain about the total supply of seats.

Experiments can also be implemented by suggesting incentive-compatible actions (numbers of schools to be applied to) to students (Bergemann and Morris, 2016). Suggesting the number of schools to be included in the application may be simpler to implement (and it is actually done in practical applications of matching theory), but working with experiments is analytically more convenient.

To state our information design problem, we abuse notation and let $\tau \in \Delta(\Delta(K))$ be the probability distribution over posterior beliefs induced by experiment τ.[13] Since students update beliefs using the Bayes rule, the distribution $\tau \in \Delta(\Delta(K))$ over posterior is *Bayes plausible*, i.e. the induced posterior beliefs average up to the prior. As discussed by Kamenica and Gentzkow (2011), Bayes plausibility is the only restriction imposed over posterior beliefs when students update using the Bayes rule.

Multiple equilibria may exist when belief γ is drawn from the experiment τ. Let $n^\gamma = (n_i^\gamma)$ be the smallest symmetric Nash equilibrium of the game G given belief γ.

[13] See Kamenica and Gentzkow (2011).

The social welfare given belief γ is[14]

$$V(\gamma) := \sum_{i=1}^{I} U_i(n_i^{\gamma}, n^{\gamma}, \gamma) w(i) \rho(i) \tag{4.1}$$

where $w(i) \geq 0$ is the Pareto weight that type i students receive in the welfare function. Our formulation allows for social welfare functions that put more weight on some groups. For example, the social planner could put more weight on groups with higher search costs and that therefore in equilibrium apply to fewer schools.

The social planner thus solves

$$\begin{aligned} \max_{\{\tau \in \Delta(\Delta(K))\}} \quad & \mathbb{E}_{\tau}[V(\gamma)] \\ s.t. \quad & \\ (i) \quad & \sum_{\gamma \in \text{supp}(\tau)} \gamma \tau(\gamma) = \mu \end{aligned} \tag{4.2}$$

where the expectation in the objective of (4.2) is taken over final beliefs γ distributed according to the experiment τ.

Define $\gamma_{\max K} \in \Delta(K)$ as the probability distribution that puts all the weight on the highest capacity in K. Clearly, for all γ, $n^{\gamma_{\max K}} \leq n^{\gamma}$. Define $\mathcal{M} = \{\gamma \mid n^{\gamma_{\max K}} = n^{\gamma}\}$. In words, $\mathcal{M}$ is the set of all beliefs that induce the lowest equilibrium search intensity.

Proposition 3 (Disclosing congestion)**.** *Under Assumption 1, suppose that $\mathcal{M} \subseteq \Delta(K)$ has nonempty interior and its convex hull is not $\Delta(K)$. The following hold.*

a. For all $\mu \in \Delta(K)$, perfectly revealing the state k is suboptimal.

b. For all $\mu \in \mathcal{M}$, an uninformative experiment $\tau_k = \tau_{k'}$ is optimal.

Proposition 3 characterizes optimal information provision in game G. The restriction over $\mathcal{M}$ is natural as it ensures equilibrium behavior reacts to information and also that for all beliefs close enough to $\gamma_{\max K}$ equilibrium behavior coincides with $n^{\gamma_{\max K}}$. The result shows that fully disclosing information is never optimal. Intuitively, full disclosure is not optimal because when k equals the lowest capacity, the system is congested, admission cutoffs will be high, and students will search a lot. By hiding some information, the planner can always avoid that outcome.

[14] As shown in Proposition 2, the set of Nash equilibria (given belief γ) has a smallest element. Since the equilibria can be Pareto-ranked, the smallest Nash equilibrium is the one the planner prefers the most. See the proof of Lemma 2 for detailed discussion.

The result also shows that when students put sufficiently high weight on the highest capacity in K, disclosing no information is optimal. In this case, students believe that lower capacities are so unlikely that there is no way to motivate them to apply to fewer schools.[15]

The proof of Proposition 3 employs ideas from information design by building the concave closure of the function $V(\gamma)$. We derive some properties of the function $V(\gamma)$ to characterize its concave closure and shape the planner's information design problem.

Lemma 2. *Under Assumption 1, the function $V(\gamma)$ is a piece-wise linear and upper semi-continuous function that is non-decreasing in γ.*

This lemma is key to establishing Proposition 3. In the proof, we repeatedly use the fact that, thanks to Assumption 1, our game G is supermodular and, as a result, its equilibrium set is monotone in the public belief γ. Figure 1 illustrates V and its concave closure.

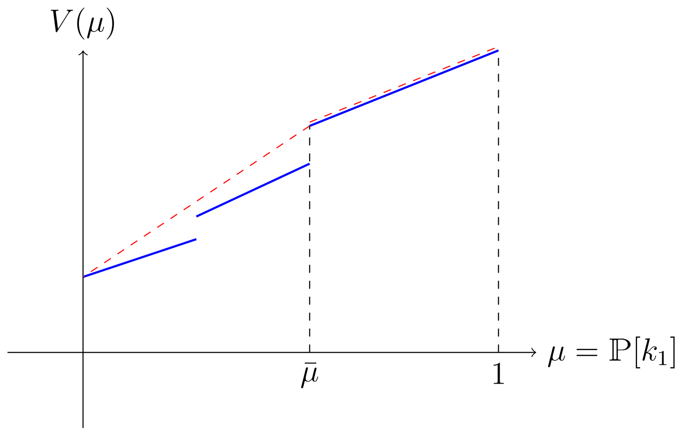

Figure 1: The welfare function V (in blue) and its concave closure (in dashed red) when $K = \{k_0, k_1\}$ and $k_0 < k_1$. The set $\mathcal{M}$ described in Proposition 3 is given by $[\bar{\mu}, 1]$.

Our focus is on public signals. Public information provides all students with the same details about the market and can therefore be desirable for normative reasons. As the following Subsection shows, private persuasion may be welfare enhancing

4.2 Private Experiments

We now explore private experiment. The general problem of information design in games is very hard to solve (Mathevet et al., 2020). We now show that when the welfare function

[15]In this case there are are several optimal experiments. Importantly, full disclosure is never optimal.

targets a small fraction of students, private signals may improve welfare over the optimal public experiment.

To see this, suppose that type i' students are a small fraction of the population ($\rho(i')$ is close to 0) but have a high welfare weight ($w(i') >> w(i)$ for $i \neq i'$). For example, type i' students may be disadvantaged, have a high search cost and therefore remain poorly informed about the schools they prefer. Students in the market are relatively optimistic so that μ puts almost all the weight on the highest capacity $\max K$. In this case, Proposition 3 shows that the experiment that maximizes (4.2) is uninformative. We assume that the Nash equilibrium given μ is strict.

Now, to build a private experiment, fix the equilibrium profile of students $\bar{n}_i$ for $i \neq i'$ under the uninformative experiment (that is, when their common belief is μ). Construct the game among type i' students in which, given a belief γ, they search for $\tilde{n}$ schools taken as fixed the search behavior of all other students in the market. The game among type i' students has a smallest Nash equilibrium $\underline{n}_{i'}^{\gamma}$ given their belief γ. Naturally, $\underline{n}_{i'}^{\mu} = \bar{n}_{i'}$. Let $\tau_{i'}$ be an experiment that reveals information to type i' students only. The planner may find optimal to reveal more detailed information to type i' students. By doing so, the planner ensures type i' students get better information and therefore their welfare is improved. Congestion in the market does not increase much because, by assumption, the fraction of type i' students is small. Other type of students $i \neq i'$ do not change their behavior as only a small fraction of the population of students is changing their behavior. While the information provided to type i' students is detrimental to $i \neq i'$, we are assuming that the authority puts most of the welfare weight on type i students. Private experiments may thus be welfare improving.

5 Disclosing Schools

In school choice systems, oftentimes platforms show some schools with much more prominence than others. In Chile, for example, students who access the centralized application platform are shown schools near their homes (Correa et al., 2022). Another common practice is to provide families with information about the academic effectiveness of schools, including report cards about school performance (Hastings and Weinstein, 2008; Andrabi et al., 2017; Elacqua et al., 2022). These policies simplify the search process faced by families and help them elucidate their preferences for schools. In this Section, we study the impact that information provision about schools has on equilibrium outcomes and welfare.

5.1 Disclosing Top Schools

In this subsection, we assume that the authority knows and fully discloses the school liked the most by each student out of the N schools in the market. This assumes that the authority knows the traits that are relevant for each family –academic performance, proximity, music programs– and is able to compute the most preferred school for each family. Each student receives the name of the school ranked top by her, searches for additional schools, and submits her application.

A student that applies to n schools will actually search for $n-1$ schools since her top school is revealed by the platform. Fixing the profile $\bar{n} = (\bar{n}_i)_{i=1}^{I}$ of schools students apply to, the expected utility a family gets when applying to n schools while receiving information about the name of the top school is denoted $U^I(n, \bar{n}, \mu)$. The utility function $U^I(n, \bar{n}, \mu)$ defines a game G^I similar to game G introduced in Section 2.[16]. Game G^I is also supermodular and has a smallest Nash equilibrium. We compare Nash equilibria of games G^I and G.

Proposition 4 (Disclosing top schools). *Under Assumption 1, the following hold:*

a. The smallest (resp. largest) Nash equilibrium of the game G^I (resp. G) is less than or equal (resp. greater than or equal) to any Nash equilibrium of the game G (resp. G^I).

b. Let n^I be the smallest Nash equilibrium of the game G^I. Then, each student gets strictly more welfare in the game G^I under equilibrium n^I than in any Nash equilibrium of the game G.

Proposition 4 shows that disclosing top schools reduces the equilibrium number of applications. The result also shows that the disclosure of top schools increases welfare. The following result is key to understanding the Proposition.

[16]Two observations are in order. First, in the game G^I, students receive information about top schools. Since the model is symmetric, different realizations of the signal disclosed (the name of the top school) do not change students' payoffs. So, a strategy in game G^I is effectively a number $n \in \{1, \ldots, N_0\}$. Second, in the game G^I, the market clearing condition is identical to the one discussed in Section 2

Lemma 3. *Under Assumption 1, for all $n \geq 1$,*

$$U^I(n+1, \bar{n}, \mu) - U^I(n, \bar{n}, \mu) \leq U(n+1, \bar{n}, \mu) - U(n, \bar{n}, \mu)$$

In other words, disclosing the top school reduces the incremental utility of adding a school to the rank order list. Intuitively, after the top school is disclosed, a student searches for schools that, on average, are less valuable and, as a result, her incentives to include more schools in her application list are reduced.

The reduction in search incentives when the top school is disclosed implies Proposition 4 part a.[17] Part b in the Proposition follows because in the lowest equilibrium of game G^I, each family solves a search problem with more information and has more chances to get admission than in any equilibrium of game G.

5.2 The Limits of Information About Schools

We have shown that disclosing the top school improves welfare. Is any kind of information about schools welfare improving? In this Subsection, we explore this question by assuming that each family receives the name of the least valued school. Disclosing the worst school is a rather stark way to model the idea that the information provided is about a school that is unlikely to be highly ranked by a student. Given the revealed school, each student searches for additional schools and submits her application. The student searching for $n-1$ schools will always add the school revealed by the authority at the bottom of her application.

Proposition 5 (Disclosing bottom schools)**.** *Under Assumption 1, suppose that F is uniform, and $N \geq N_0 + 1$. The following hold:*

a. *The smallest (resp. largest) Nash equilibrium of the game G (resp. G^{IB}) is less than or equal (resp. greater than or equal) to any Nash equilibrium of the game G^{IB} (resp. G).*
b. *Let n be the smallest Nash equilibrium of the game G and assume that n is strictly smaller than all Nash equilibria of game G^{IB}. Let $\bar{N}_0$ be the unique N_0 satisfying $N_0(1-Nk_1)^{N_0} = 1/15$. Then, if $N_0 \leq \bar{N}_0$, each student gets strictly more welfare in the game G under equilibrium n than in any Nash equilibrium of the game G^{IB}.*

[17] Milgrom and Roberts (1990) show comparative statics results for equilibria in supermodular games.

Proposition 5 shows that disclosing the bottom schools creates additional congestion in the market and reduces welfare. While Proposition 4 shows that disclosing top schools is socially desirable, Proposition 5 puts a limit to the idea that all types of information interventions about schools are desirable.

The following result is key to understanding Proposition 5.

Lemma 4. *Under Assumption 1, suppose that F is uniform and $N \geq N_0 + 1$. Then, for all $n \geq 2$, we have*

$$U^{IB}(n+1, \bar{n}, \mu) - U^{IB}(n, \bar{n}, \mu) \geq U(n+1, \bar{n}, \mu) - U(n, \bar{n}, \mu)$$

The lemma says that disclosing the least preferred school increases the incremental utility of adding a school to the rank order list. On the one hand, disclosing the least preferred school gives a free school to apply to a student. On the other hand, the disclosure of the least preferred school reduces the uncertainty about remaining schools, which reduces the search incentives. When $N \geq N_0 + 1$ is large enough, u_N^N is relatively small, and knowing that a school is the least attractive school in the market is not very informative. As a result, the main effect of disclosing the least preferred schools is to give students free applications.

Proposition 5 part a follows since students have stronger incentives to apply to schools when the bottom schools are disclosed. Part b in the Proposition follows because in the lowest equilibrium of game G, there is strictly less congestion than in the game G^{IB} and the additional information about bottom schools provided in G^{IB} has little value for each student.[18]

6 Conclusion

This paper explores information interventions in centralized school assignment mechanisms and reveals nuanced considerations for policymakers. Recognizing the challenges faced by families in the school selection process, our findings suggest a cautious approach to information disclosure. Full transparency about the system capacities and admission chances is never optimal. In contrast, if the authority has information about schools that

[18] Note that for information to reduce aggregate welfare, it has to induce more applications. If the equilibrium number of applications were the same, all agents would be better off with information because congestion would be the same.

are attractive to students, that information should be provided to the families. Our work contributes to ongoing discussion on optimizing centralized school assignment systems.

We now discuss some extensions and variations of our model.

Single tie breaking. We have assumed that schools rank students independently using multiple tie breaking, but in some systems a single tie breaking is used and scores are perfectly correlated (Abdulkadiroğlu et al., 2009).

In our model, under single tie breaking, all students are assigned to their top ranked school and cutoffs search intensities do not depend on the profile $\bar{n}$. In particular, each student' search problem does not depend on the search intensity of other students. As a result, under single tie breaking, full transparency about capacities and schools characteristics Pareto dominates any other information policy.

Several papers have shown that single tie breaking results in more students assigned to their top schools than multiple tie breaking. These results are sometimes used to favor single tie breaking in some practical applications. Our exercise provides an informational rational for the use of single tie breaking. Under single tie breaking, students search incentives are aligned with social goals and thus information can be fully disclosed and used.

Correlated preferences. We have assumed that students have independent preferences in the sense that knowing the top school of a student does not predict what are other attractive schools for the student. In practice, students are more likely to find attractive schools with similar characteristics. Our model and results can be accommodated to allow for correlated preferences.

Suppose that students live in one of two districts and each district has $N/2$ schools (N even). A student is more likely to find attractive schools in her own district. Concretely, if a student searches for $n \leq N/2$ schools, the student gets n draws from the distribution F but each school in her district is more likely to be part of the discovered schools. Schools are symmetric. A students will tend to apply more to schools in her own district, but will still place applications to schools in the other district. Our results extend immediately to this setup. Exploring our results in more general models with correlated preferences is an interesting venue for future research.

Appendix

A Proof of Lemma 1

Lemma 1 (Negative externalities). The following hold:

a. $B(n, \bar{n}, k)$ is increasing and concave in n.

b. $B(n, \bar{n}, k)$ is non-increasing in $\bar{n}$. Moreover, for $n \geq 1$, $B(n, \bar{n}, k)$ is decreasing in $\bar{n}$, that is, for $\tilde{n} = (\tilde{n}_i)_i$ and $\bar{n} = (\bar{n}_i)_i$ with $\tilde{n}_i \geq \bar{n}_i$ for all i, $B(n, \bar{n}, k) \geq B(n, \tilde{n}, k)$

Proof. We show that the function $b(n', p_{k,n}, k)$ satisfies [a] and [b], which implies the function $B(n, \bar{n}, k)$ does as well. We drop the sub-index k, n from p to simplify notation. We first show the part [a]. Let

$$\Delta b_n := b(n+1, p, k) - b(n, p, k) = \sum_{j=1}^{n} \Delta_n^j (1-p)p^{j-1} + u_{n+1}^{n+1}(1-p)p^n$$

By theorem 1 (b) in Watt (2021), when F satisfies the MHR property, $\Delta_n^j \geq 0$, and the function B increases in n. Furthermore, simple algebra shows that

$$\Delta b_n - \Delta b_{n-1} = \sum_{j=1}^{n-1} (\Delta_n^j - \Delta_{n-1}^j)(1-p)p^{j-1} + \Delta_n^n (1-p)^{n-1} + u_{n+1}^{n+1}(1-p)p^n - u_n^n (1-p)p^{n-1}$$

By Theorem 1 (b) in Watt (2021), when F satisfies the MHR property, $\Delta_n^j - \Delta_{n-1}^j \leq 0$. Hence,

$$\begin{aligned}
\Delta b_n - \Delta b_{n-1} &\leq (1-p)p^{n-2} \left(\sum_{j=1}^{n-1} (\Delta_n^j - \Delta_{n-1}^j) + p\left(\Delta_n^n + u_{n+1}^{n+1} p - u_n^n\right) \right) \\
&\leq (1-p)p^{n-2} \left(\sum_{j=1}^{n-1} (\Delta_n^j - \Delta_{n-1}^j) + p\left(\Delta_n^n + u_{n+1}^{n+1} - u_n^n\right) \right) \\
&\leq (1-p)p^{n-2} \left(\sum_{j=1}^{n-1} (\Delta_n^j - \Delta_{n-1}^j) + \Delta_n^n + u_{n+1}^{n+1} - u_n^n \right)
\end{aligned}$$

Where the last inequality follows because $\Delta_n^n + u_{n+1}^{n+1} - u_n^n \geq 0$ (see below). Hence, for b to be concave, it is enough to show that

$$\sum_{j=1}^{n-1} (\Delta_{n-1}^j - \Delta_{n-2}^j) + \Delta_n^n + u_{n+1}^{n+1} - u_n^n \leq 0 \tag{.1}$$

But our statistics u_n^j and u_{n-1}^j corresponds to the statistics μ_n^{n-j+1} and μ_{n-1}^{n-j} (respectively) when counting the statistics from bottom to top. Thus, from Lemma 2 (b) in David (1997), we have

$$\Delta_{n-1}^j = \mu_n^{n-j+1} - \mu_{n-1}^{n-j} \quad = \quad \binom{n-1}{n-j} \int_0^1 F(\tilde{u})^{n-j}(1-F(\tilde{u}))^j d\tilde{u} \tag{.2}$$

And

$$\Delta_n^j - \Delta_{n-1}^j \quad = \quad \binom{n-1}{n-j} \int_0^1 F(\tilde{u})^{n-j}(1-F(\tilde{u}))^j \left[\frac{n}{n-j+1} F(\tilde{u}) - 1\right] d\tilde{u} \tag{.3}$$

Also, from Lemma 2 (a) in David (1997)

$$u_n^n - u_{n+1}^{n+1} = \mu_n^1 - \mu_{n+1}^1 = \binom{n}{0} \int_0^1 F(\tilde{u})(1-F(\tilde{u}))^n d\tilde{u}$$

Thus,

$$\begin{aligned}
\Delta_n^n - (u_n^n - u_{n+1}^{n+1}) \quad &= \quad \binom{n}{1} \int_0^1 F(\tilde{u})(1-F(\tilde{u}))^n d\tilde{u} - \binom{n}{0} \int_0^1 F(\tilde{u})(1-F(\tilde{u}))^n d\tilde{u} \\
&= \quad \binom{n-1}{1} \int_0^1 F(\tilde{u})(1-F(\tilde{u}))^n d\tilde{u}
\end{aligned}$$

Now let

$$\begin{aligned}
A_0 \quad &:= \quad \binom{n-1}{1} \int_0^1 F(\tilde{u})(1-F(\tilde{u}))^n d\tilde{u} \\
A_1 \quad &:= \quad A_0 + \Delta_n^{n-1} - \Delta_{n-1}^{n-1} \\
&\vdots \\
A_m \quad &:= \quad A_{m-1} + \Delta_n^{n-m} - \Delta_{n-1}^{n-m}
\end{aligned}$$

defined for $m \leq n-1$. Now, we show that for all $m \leq n-2$,

$$A_m = \binom{n-1}{m+1} \int_0^1 F(\tilde{u})^{m+1}(1-F(\tilde{u}))^{n-m} d\tilde{u}$$

It is easy to see that it holds for $m = 0$. Now suppose it holds for $m \leq n-3$. Then,

$$\begin{aligned}
A_{m+1} &= \binom{n-1}{m+1} \int_0^1 F(\tilde{u})^{m+1}(1-F(\tilde{u}))^{n-m} d\tilde{u} + \Delta_n^{n-(m+1)} - \Delta_{n-1}^{n-(m+1)} \\
&= \binom{n-1}{m+1} \int_0^1 F(\tilde{u})^{m+1}(1-F(\tilde{u}))^{n-m} d\tilde{u} + ... \\
&\qquad ... \binom{n-1}{m+1} \int_0^1 F(\tilde{u})^{m+1}(1-F(\tilde{u}))^{n-m-1} \left[\frac{n}{m+2}F(\tilde{u}) - 1\right] d\tilde{u} \\
&= \binom{n-1}{m+1} \int_0^1 F(\tilde{u})^{m+1}(1-F(\tilde{u}))^{n-m-1} \left[(1-F(\tilde{u})) + \frac{n}{m+2}F(\tilde{u}) - 1\right] d\tilde{u} \\
&= \binom{n-1}{m+1} \int_0^1 F(\tilde{u})^{m+1}(1-F(\tilde{u}))^{n-m-1} \left[F\frac{n-m-2}{m+2}\right] d\tilde{u} \\
&= \binom{n-1}{m+2} \int_0^1 F(\tilde{u})^{m+2}(1-F(\tilde{u}))^{n-m-1} d\tilde{u}
\end{aligned}$$

Finally, it is easy to see that, when $m = n-2$, we obtain

$$A_{n-2} = \binom{n-1}{n-1} \int_0^1 F(\tilde{u})^{n-1}(1-F(\tilde{u}))^2 d\tilde{u}$$

And that

$$\begin{aligned}
\Delta_n^1 - \Delta_{n-1}^1 &= \binom{n-1}{n-1} \int_0^1 F(\tilde{u})^{n-1}(1-F(\tilde{u}))^1 \left[\frac{n}{n-1+1}F(\tilde{u}) - 1\right] d\tilde{u} \\
&= \binom{n-1}{n-1} \int_0^1 F(\tilde{u})^{n-1}(1-F(\tilde{u}))^2 d\tilde{u} \\
&= A_{n-2}
\end{aligned}$$

Thus, we conclude that equation (.1) is zero, because

$$\sum_{j=1}^{n-1}(\Delta_{n-1}^j - \Delta_{n-2}^j) + \Delta_n^n + u_{n+1}^{n+1} - u_n^n = \Delta_n^1 - \Delta_{n-1}^1 + A_{n-2} = 0$$

And so the function B is concave.

Now, we show part b. It is easy to see from equation (2.1) that for profiles $\bar{n} = (\bar{n}_i)$ and $\bar{n}' = (\bar{n}'_i)$, if $\bar{n}_i \geq \bar{n}'_i$, then $p_{\bar{n},k} \geq p_{\bar{n}',k}$ for all k. Thus, to show that $b(n, p_{k,\bar{n}}, k)$ is decreasing in $\bar{n}$, it is enough to show that it is decreasing in $p_{k,n}$. Thus, when $n \geq 1$

$$\begin{aligned}
\frac{\partial b}{\partial p} &= -u_n^1 + u_n^2(1-2p) + u_n^3(2p-3p^2) + ...u_n^n((n-1)p^{n-2} - np^{n-1}) \\
&= -(u_n^1 - u_n^2) - (u_n^2 - u_n^3)2p - (u_n^3 - u_n^4)3p^2... - (u_n^{n-1} - u_n^n)(n-1)p^{n-2} - np^{n-1} \\
&= -\sum_{l=1}^{n-1}(u_n^l - u_n^{l+1})lp^{l-1} - np^{n-1}
\end{aligned}$$

Which is less than zero because $u_n^l - u_n^{l+1} > 0$ for all l. Finally, when $n = 0$, $b(n, p, k) = 0$, and b is constant with respect to p. □

B Proof of Proposition 1

Proposition 1. Under Assumption 1, $BR(\bar{n}; \mu)$ is non-decreasing in $\bar{n} \in \{0, \ldots, N\}$ for all μ.

Proof. The proof of Propositions 1 and 2 rely on the following two Lemmas, whose proofs are below.

Lemma 5. *For all $n \leq N$ and $j \leq n$, Δ_{n-1}^j is non decreasing in j. That is,*

$$u_n^j - u_{n-1}^j \geq u_n^{j-1} - u_{n-1}^{j-1}$$

Lemma 6. *Define $\tilde{p}$ as:*

$$\tilde{p} := \min_{n\in\{1,\ldots,N_0\}} \left\{ 1 - \left(\frac{1}{n+1}\right) \frac{\sum_{j=1}^n \Delta_n^j}{u_{n+1}^{n+1}} \right\}. \tag{.4}$$

Then, $\tilde{p} \in]0,1[$. The function $b_i(n, p, k)$ has increasing differences in n and $p \leq \tilde{p}$.

Because of Lemma 6, under Assumption 1, properties $(A1) - (A4)$ in Milgrom and Roberts (1990) are satisfied, so the game is supermodular. Thus, by Topkis's Monotonicity Theorem (see Milgrom and Roberts (1990)), $BR_i(\bar{n}; \mu)$ is non-decreasing in $\bar{n} \in \{0, \ldots, N\}$ for all μ. □

B.1 Proof of Lemma 5

Proof. We first show that Δ^j_{n-1} is non-decreasing in j for $j \leq n-1$. Then, we show it is true for $j = n$. First, note that when counting from bottom to top, our statistic u^j_n corresponds to the statistic μ^{n-j+1}_n. Analogously, u^j_{n-1} corresponds to μ^{n-j}_{n-1}. From Watt (2021), we know that

$$\Delta^j_{n-1} := \mu^{n-j+1}_n - \mu^{n-j}_{n-1} \quad = \quad \binom{n-1}{n-j} \int_0^1 F(\tilde{u})^{n-j}(1-F(\tilde{u}))^j d\tilde{u} \tag{.5}$$

Thus, we want to show that $\Delta^j_{n-1} - \Delta^{j-1}_{n-1} \geq 0$. From equation (.2) in the proof of Lemma 1, we have

$$\begin{aligned}
\Delta^j_{n-1} - \Delta^{j-1}_{n-1} \quad &= \quad \binom{n-1}{n-j} \int_0^1 F(\tilde{u})^{n-j}(1-F(\tilde{u}))^{j-1} \left[1 - \frac{n}{n-j+1}F(\tilde{u})\right] d\tilde{u} &(.6)\\
&= \quad \binom{n-1}{n-j} \int_0^1 F(\tilde{u})^{n-j}(1-F(\tilde{u}))^{j-2} \left[1 - \frac{n}{n-j+1}F(\tilde{u})\right] f(\tilde{u})g(\tilde{u})d\tilde{u}
\end{aligned}$$

where $g(\tilde{u}) := \frac{1}{h(\tilde{u})}$ is decreasing by assumption, because $h(\tilde{u})$ is increasing. Let $\tilde{u}^*$ be such that

$$F(\tilde{u}^*) = \frac{n-j+1}{n}$$

So that for $\tilde{u} > \tilde{u}^*$, the integrand in (.6) is negative, for $\tilde{u} = \tilde{u}^*$ is zero, and for $\tilde{u} < \tilde{u}^*$ is positive. Consider now the integral

$$W(t) := \int_0^t F(\tilde{u})^{n-j}(1-F(\tilde{u}))^{j-2} \left[1 - \frac{n}{n-j+1}F(\tilde{u})\right] f(\tilde{u})d\tilde{u} \tag{.7}$$

It is clear that for $t \leq \tilde{u}^*$, $W(t) \geq 0$. For $t > \tilde{u}^*$, we have that $W(t) \geq W(t')$ for $t \leq t'$. Thus, if $W(1) \geq 0$, we can conclude that $W(t) \geq 0$ for all t. Making the change of variable $x = F(\tilde{u})$, we rewrite $W(1)$ as

$$W(1) := \int_0^1 x^{n-j}(1-x)^{j-2} \left[1 - \frac{n}{n-j+1}x\right] dx \tag{.8}$$

And using the fact that, for any m, n,

$$\int_0^1 x^m(1-x)^n dx = \frac{n!\ m!}{(n+m+1)!}$$

We obtain

$$W(1) \quad = \quad \frac{(n-j)!(j-2)!}{(n-1)!} - \frac{n}{n-j+1} \cdot \frac{(n-j+1)!(j-2)!}{n!} = 0$$

Thus, $W(t) \geq 0$ for all t. Finally, equation (.6) can be written as

$$\int_0^1 g(t)dW(t)$$

Integrating by parts with $u(t) = g(t)$ and $v(t) = W(t)$, obtain

$$\begin{aligned} \int_0^1 g(t)dW(t) \quad &= \quad g(t)W(t)\Big|_0^1 - \int_0^1 W(t)g'(t)dt \\ &= \quad -\int_0^t W(t)g'(t)dt \geq 0 \end{aligned}$$

where the second equality follows because $g(1) = 0$, $W(0) = 0$, $W(t) \geq 0$ for all t, and $g'(x) \leq 0$. Thus, we have shown $\Delta_n^j - \Delta_n^{j-1} \geq 0$. Now, we show it is also true for $j = n$. That is, we want to show that

$$\Delta_n^{n+1} - \Delta_n^n = u_{n+1}^{n+1} - (u_{n+1}^n - u_n^n) \geq 0$$

In terms of the statistics when counting from bottom to top, we want to show

$$\mu_{n+1}^1 - (\mu_{n+1}^2 - \mu_n^1) \geq 0$$

Let F_n^i denote the (cumulative) distribution of the i-th order statistic when counting from bottom to top. Thus,

$$\mu_n^i = \int_0^1 \tilde{u} dF_n^i(\tilde{u})$$

Integrating by parts, get

$$\mu_n^i = 1 - \int_0^1 F_n^i(\tilde{u})d\tilde{u}$$

Thus, obtain

$$\mu_{n+1}^1 - (\mu_{n+1}^2 - \mu_n^1) = 1 - \int_0^1 \left(F_{n+1}^1(\tilde{u}) + F_n^1(\tilde{u}) - F_{n+1}^2(\tilde{u})\right) d\tilde{u} \tag{.9}$$

But it is well known that[19]

$$F_n^i(\tilde{u}) = \sum_{k=i}^{n} \begin{pmatrix} n \\ k \end{pmatrix} F(\tilde{u})^k (1 - F(\tilde{u}))^{n-k}$$

Hence,

$$\begin{aligned} F_{n+1}^1(\tilde{u}) + F_n^1(\tilde{u}) - F_{n+1}^2(\tilde{u}) &= \begin{pmatrix} n+1 \\ 1 \end{pmatrix} F(\tilde{u})(1 - F(\tilde{u}))^n + \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} F(\tilde{u})^k (1 - F(\tilde{u}))^{n-k} - ... \\ & \quad ... \begin{pmatrix} n \\ 0 \end{pmatrix} (1 - F(\tilde{u}))^n \\ &= 1 - (n+1)(1 - F(\tilde{u}))^n \left[\frac{1}{n+1} - F\right] \end{aligned}$$

Hence, equation (.9) becomes

$$\begin{aligned} (.9) &= (n+1) \int_0^1 (1 - F(\tilde{u}))^n \left[\frac{1}{n+1} - F\right] d\tilde{u} \\ &= \int_0^1 (1 - F(\tilde{u}))^{n-1} \left[\frac{1}{n+1} - F\right] f(\tilde{u}) g(\tilde{u}) d\tilde{u} \end{aligned}$$

Where $g(\tilde{u}) = \frac{1 - F(\tilde{u})}{f(\tilde{u})}$ is the inverse hazard rate that is decreasing by assumption. Let

$$W(t) := \int_0^t (1 - F(\tilde{u}))^{n-1} \left[\frac{1}{n+1} - F\right] f(\tilde{u}) d\tilde{u}$$

Using the same argument as before, we need to show that $W(1) \geq 0$. Indeed,

$$W(1) := \frac{(n-1)!}{(n+1)n!} - \frac{(n-1)!}{(n+1)!} = 0$$

And we conclude that $\mu_{n+1}^1 - (\mu_{n+1}^2 - \mu_n^1) \geq 0$. $\square$

[19]See for instance Balakrishnan and Cohen (2014).

B.2 Proof of Lemma 6

Proof. As before, to show that $b(n', p_{k,n}, k)$ has decreasing differences in n', n, it is enough to show that it has decreasing differences $n', p_{k,n}$. Again, we drop the sub-index n, k from p to ease notation. From equation (2.4), it is easy to see that, for $n \geq 1$,

$$\begin{aligned}
\frac{\partial}{\partial p}\Big(b(n+1,p,k) - b(n,p,k)\Big) &= \sum_{j=1}^{n} \Delta_n^j\big((j-1)p^{j-2} - jp^{j-1}\big) + u_{n+1}^{n+1}(np^{n-1} - (n+1)p^n) \\
&= 1(\Delta_n^2 - \Delta_n^1) + 2p(\Delta_n^3 - \Delta_n^2) + \cdots + np^{n-1}(\Delta_n^{n+1} - \Delta_n^n) - .. \\
&\qquad ...u_{n+1}^{n+1}(n+1)p^n \\
&= \sum_{j=1}^{n} jp^{j-1}(\Delta_n^{j+1} - \Delta_n^j) - u_{n+1}^{n+1}(n+1)p^n
\end{aligned}$$

We know that Δ_n^j is non-decreasing in j for $j \leq n$. Then,

$$\begin{aligned}
\frac{\partial}{\partial p}\Big(b(n+1,p,k) - b(n,p,k)\Big) &\geq p^{n-1}\left(\sum_{j=1}^{n} j(\Delta_n^{j+1} - \Delta_n^j) - u_{n+1}^{n+1}(n+1)p\right) \\
&= p^{n-1}\left(\sum_{j=1}^{n}((j+1)\Delta_n^{j+1} - j\Delta_n^j) - u_{n+1}^{n+1}(n+1)p - \sum_{j=1}^{n}\Delta_n^{j+1}\right) \\
&= p^{n-1}\left((n+1)\Delta_n^{n+1} - \Delta_n^1 - u_{n+1}^{n+1}(n+1)p - \sum_{j=1}^{n}\Delta_n^{j+1}\right) \\
&= p^{n-1}\left((n+1)\Delta_n^{n+1} - u_{n+1}^{n+1}(n+1)p - \sum_{j=1}^{n}\Delta_n^{j}\right) \\
&= p^{n-1}\left((n+1)u_{n+1}^{n+1} - u_{n+1}^{n+1}(n+1)p - \sum_{j=1}^{n}\Delta_n^{j}\right)
\end{aligned}$$

So, a sufficient condition for $\frac{\partial}{\partial p}\Big(b(n+1,p,k) - b(n,p,k)\Big) \geq 0$ is

$$\begin{aligned}
0 &\leq (n+1)u_{n+1}^{n+1} - u_{n+1}^{n+1}(n+1)p - \sum_{j=1}^{n}\Delta_n^j \\
p &\leq 1 - \left(\frac{1}{n+1}\right)\frac{\sum_{j=1}^{n}\Delta_n^j}{u_{n+1}^{n+1}} = \tilde{p}
\end{aligned}$$

Note that $\tilde{p}$ is greater than zero because

$$\begin{aligned}
\Delta_n^{n+1} &\geq \Delta_n^j \\
(n+1)\Delta_n^{n+1} &> \sum_j \Delta_n^j \\
1 &> \left(\frac{1}{n+1}\right) \frac{\sum_{j=1}^n \Delta_n^j}{u_{n+1}^{n+1}}
\end{aligned}$$

Thus, whenever

$$p \leq \tilde{p} := \min_{\{n\}} \left\{ 1 - \left(\frac{1}{n+1}\right) \frac{\sum_{j=1}^n \Delta_n^j}{u_{n+1}^{n+1}} \right\}$$

the function $b(n,p)$ has increasing differences in n and p, which concludes the proof. □

C Proof of Proposition 2

Proposition 2. Under Assumption 1, the following hold.

a. The game G has a nonempty set of Nash equilibria: $NE(\mu) \neq \emptyset$.
b. The set of Nash equilibria $NE(\mu)$ has a smallest and a largest element.
c. Let $\bar{n}$ and $\tilde{n}$ be Nash equilibria with $\bar{n}_i \leq \tilde{n}_i$ for all i. Then, all students get higher payoffs under equilibrium $\bar{n}$ than under equilibrium $\tilde{n}$.
d. The smallest and largest Nash equilibria are non-increasing in the distribution $\mu \in \Delta(K)$, where we endow $\Delta(K)$ with the first order stochastic dominance.

Proof. Under Assumption 1, properties $(A1) - (A4)$ in Milgrom and Roberts (1990) are satisfied. Then, by Theorem 5 in Milgrom and Roberts (1990), given any $\mu \in \Delta(K)$, a pure strategy Nash equilibrium always exists. The largest and smallest Nash equilibrium profiles exist (see the proof of the corollaries of Theorem 5 in Milgrom and Roberts (1990)). Part [c.] follows directly from Theorem 7 in Milgrom and Roberts (1990). Moreover, it is easy to see that $u(n,p,\mu)$ has increasing differences in n and μ (for fixed p). That is, since $p_{n,k_0} \geq p_{n,k_1}$, and b is supermodular in $(n, p_{n,k})$ under Assumption 1, u has increasing differences in μ. Thus, by Theorem 6 in Milgrom and Roberts (1990), the largest and smallest Nash equilibrium profiles are (weakly) increasing in μ. □

D Proof of Proposition 3

Proposition 3. Under Assumption 1, suppose that $\mathcal{M} \subseteq \Delta(K)$ has nonempty interior and its convex hull is not $\Delta(K)$. The following hold.

a. For all $\mu \in \Delta(K)$, perfectly revealing the state k is suboptimal.

b. For all $\mu \in \mathcal{M}$, an uninformative experiment $\tau_k = \tau_{k'}$ is optimal.

Proof. Let $\textsc{cav}V$ denote the smallest concave function that is greater than or equal to V. Since $V(\gamma)$ is linear in $\gamma \in \mathcal{M}$ and $V(\gamma) > V(\gamma')$ for all $\gamma \in \mathcal{M}$ and $\gamma' \notin \mathcal{M}$, it follows that $\textsc{cav}V(\gamma) = V(\gamma)$ for all $\gamma \in \mathcal{M}$. In particular, this proves that when $\mu \in \mathcal{M}$, an uninformative experiment is optimal.

To prove that perfectly revealing the state is suboptimal, we will first argue that $\textsc{cav}V$ is not linear. Suppose $\textsc{cav}V$ is linear. Note that for all $\gamma \in \mathcal{M}$

$$\textsc{cav}V(\gamma) = V(\gamma) = \sum_{k \in K} \gamma(k) \sum_{i=1}^{I} U_i(\bar{n}, \bar{n}, k) w(i) \rho(i)$$

where $\bar{n} = n^\gamma$ for all $\gamma \in M$. Since $\mathcal{M}$ has nonempty interior and, by way of contradition, we are assuming that $\textsc{cav}V$ is linear over $\gamma \in \Delta(K)$, it follows that for all $\gamma \in \Delta(K)$,

$$\textsc{cav}V(\gamma) = \sum_{k \in K} \gamma(k) \sum_{i=1}^{I} U_i(\bar{n}, \bar{n}, k) w(i) \rho(i). \tag{.10}$$

This implies that for any γ, there exits an experiment that implements $\bar{n}$ for all signals. For γ' outside the convex hull of $\mathcal{M}$, this is a contradiction. Thus, $\textsc{cav}V$ is not linear over $\Delta(K)$. In particular, for $\gamma \notin \mathcal{M}$, a perfectly revealing experiment cannot maximize social welfare.

□

E Proof of Lemma 2

Lemma 2 Under Assumption 1, the function $V(\gamma)$ is a piece-wise linear and upper semi-continuous function that is non-decreasing in γ.

Proof. Let

$$V_i(\gamma) = \max_n U_i(n, n^\gamma, \gamma) = U_i(n^\gamma, n^\gamma, \gamma)$$

where n^γ is the smallest Nash equilibrium given beliefs γ. Clearly, $V_i(\gamma)$ is non-decreasing in γ because n^γ is non-increasing and the game has negative externalities. Moreover, since the set $\{n^\gamma \mid \gamma \in \Delta(K)\}$ is finite, it follows that $\Delta(K)$ can be written as a finite disjoint union

$$\Delta(K) = \bigcup_t E_t$$

so that for $\gamma', \gamma \in E_t$, $n^\gamma = n^{\gamma'}$. In particular, $V_i(\gamma)$ is linear in $\gamma \in E_t$. We now argue that V_i is upper semi-continuous. Take any sequence $\gamma_\nu \to \gamma$ as $\nu \to \infty$ and note that

$$n^\gamma = \min\{n \in NE(\gamma)\} \leq \liminf_{\nu\to\infty} n^{\gamma_\nu}$$

where the inequality (in the componentwise order) follows since $\liminf_\nu n^{\gamma_\nu} \in NE(\gamma)$. Now, we prove that

$$\limsup_\nu V_i(\gamma_\nu) \leq V_i(\gamma).$$

For simplicity, assume that the sequence γ_ν attains the limsup so

$$\limsup_\nu V_i(\gamma_\nu) = \lim_\nu V_i(\gamma_\nu).$$

It must be the case that for all ν large enough, $n^{\gamma_\nu} = \bar{n} \geq n^\gamma$ and therefore

$$\lim_\nu V_i(\gamma_\nu) = V_i(\bar{n}, \bar{n}, \gamma) \leq \max_{n'} V_i(n', \bar{n}, \gamma) \leq \max_{n'} V_i(n', n^\gamma, \gamma) = V_i(\gamma).$$

where the second inequality follows since $\bar{n} \geq n^\gamma$ and the game has negative externalities. The function $V(\gamma)$, being the weighted sum of piece-wise linear, non-decreasing and upper semi-continuous functions, satisfies the properties stated in the lemma. $\square$

F Proof of Proposition 4

Proposition 4. Under Assumption 1, the following hold:

a. The smallest (resp. largest) Nash equilibrium of game G^I (resp. G) is less than or equal (resp. greater than or equal) to any Nash equilibrium of game G (resp. G^I).

b. Let n^I be the smallest Nash equilibrium of the game G^I. Then, each student gets strictly more welfare in the game G^I under equilibrium n^I than in any Nash equilibrium of the game G.

Proof. Part [a.] of the proposition follows directly from Lemma 3. Part [b.] holds because when every player $j \neq i$ plays n^I, the cutoff is weakly lower than when every $j \neq i$ plays an equilibrium profile in the game G (because of Lemma 3). Then, each family i solves the same problem as in G, but with more information than before while facing a more favourable search pattern of the other families. Thus, family i is better off. □

G Proof of Lemma 3

Lemma 3 Under Assumption 1, for all $n \geq 1$,

$$U^I(n+1, \bar{n}, \mu) - U^I(n, \bar{n}, \mu) \leq U(n+1, \bar{n}, \mu) - U(n, \bar{n}, \mu)$$

Proof. When information about the top school is provided, the utility that a student gets when applying to n schools is

$$U^I(n, \bar{n}, \mu) = \mathbb{E}_\mu\Big((1 - p_{\bar{n},k})\mathbb{E}[u^1_N] + \sum_{l=1}^{n-1} \mathbb{E}[u^l_{n-1} \mid I](1 - p_{\bar{n},k})p^l_{\bar{n}}\Big) - c(n-1)$$

where $\mathbb{E}[u^l_{n-1} \mid I]$ denotes the expectation of the l-order statistics among $n-1$ schools after the information about the identity of the top schools is revealed. Thus, for $n \geq 2$,

$$U^I(n+1, \bar{n}, \mu) - U^I(n, \bar{n}, \mu) = \sum_{l=1}^{n-1} \Big(\mathbb{E}[u^l_n \mid I] - \mathbb{E}[u^l_{n-1} \mid I]\Big)(1 - p_{\bar{n}}, k)p^l_{\bar{n},k} + (1 - p_{\bar{n},k})p^{n-1}_{\bar{n},k}\mathbb{E}[u^n_{n+1} \mid I] - c$$

Thus, to prove the Lemma it is enough to show that

$$\begin{aligned} &\sum_{l=1}^{n} \Big(\mathbb{E}[u^l_{n`+1}] - \mathbb{E}[u^l_n]\Big)(1 - p_{\bar{n}})p^{l-1}_{\bar{n}} + (1 - p_{\bar{n}})p^n_{\bar{n}}\mathbb{E}[u^{n+1}_{n+1}] \\ &\quad \geq \sum_{l=1}^{n-1} \Big(\mathbb{E}[u^l_n \mid I] - \mathbb{E}[u^l_{n-1} \mid I]\Big)(1 - p_{\bar{n}})p^l_{\bar{n}} + (1 - p_{\bar{n}})p^{n-1}_{\bar{n}}\mathbb{E}[u^n_{n+1} \mid I]. \end{aligned}$$

Denoting $\tilde{\Delta}^l_n = \mathbb{E}[u^l_{n+1} \mid I] - \mathbb{E}[u^l_n \mid I]$, we show in Lemma 7 that $\tilde{\Delta}^l_n \leq \Delta^l_n$ (see below). Rerranging terms, it is enough to prove that

$$\mathbb{E}_\mu\Big((1 - p_{\bar{n},k})\big(\sum_{j=1}^{n-1}(\Delta^j_n - p_{\bar{n},k}\Delta^j_{n-1})p^{j-1}_{\bar{n},k} + \Delta^n_n p^{n-1}_{\bar{n},k} + (u^{n+1}_{n+1} - u^n_n)p^n_{\bar{n},k}\big)\Big) \geq 0$$

Thus, it is enough to show that, given profile $\bar{n}$, for all k, we have

$$(i) \qquad \Delta_n^j - p_{\bar{n},k}\Delta_{n-1}^j \geq 0 \quad \forall n \in \{2,\ldots,N_0\}\ \forall j \in \{1,\ldots,n-1\}$$
$$(ii) \qquad \Delta_n^n + p_{\bar{n},k}(u_{n+1}^{n+1} - u_n^n) \geq 0 \quad \forall n \in \{2,\ldots,N_0\}$$

Note that by construction, for all $\bar{n}$ and all k,

$$p_{\bar{n},k} \leq \tilde{p} \leq \min_{n\in\{1,\ldots,N_0\}} \min_{l\in\{1,\ldots,n\}} \{\frac{\Delta_{n+1}^l}{\Delta_n^l}\}$$

which established (i). Now, to show (ii) we note that

$$\Delta_n^n \;=\; n\int F(\tilde{u})(1-F(\tilde{u}))^n d\tilde{u}$$

and

$$\begin{aligned}
u_{n+1}^{n+1} &= \mu_{n+1}^1 = \int \tilde{u}dF_{n+1}^1(\tilde{u}) = 1 - \int F_{n+1}^1(\tilde{u})d\tilde{u} \\
&= 1 - \int \left\{ \sum_{k=1}^{n+1} \begin{pmatrix} n+1 \\ k \end{pmatrix} F(\tilde{u})^k(1-F(\tilde{u}))^{n+1-k}d\tilde{u} \right\} \\
&= 1 - \int \left\{ \sum_{k=0}^{n+1} \begin{pmatrix} n+1 \\ k \end{pmatrix} F(\tilde{u})^k(1-F(\tilde{u}))^{n+1-k}d\tilde{u} - (1-F(\tilde{u}))^{n+1} \right\} \\
&= \int (1-F(\tilde{u}))^{n+1}d\tilde{u}
\end{aligned}$$

Analogously, obtain $u_n^n = \int(1-F(\tilde{u}))^n d\tilde{u}$. Then, the left hand side of condition (ii) can be written as

$$\begin{aligned}
(b) \;&=\; n\int F(\tilde{u})(1-F(\tilde{u}))^n d\tilde{u} + p\left(\int (1-F(\tilde{u}))^{n+1}d\tilde{u} - \int (1-F(\tilde{u}))^n d\tilde{u}\right) \\
&=\; n\int (1-F(\tilde{u}))^n \left(F\left(1-\frac{p}{n}\right) d\tilde{u}\right. \\
&\geq\; 0
\end{aligned}$$

$\square$

Lemma 7. $\tilde{\Delta}_n^l \leq \Delta_n^l$

Proof. Note that

$$\tilde{\Delta}_n^l = \int \Big(\mathbb{E}[u_{n+1}^l \mid u_N^1] - \mathbb{E}[u_n^l \mid u_N^1]\Big) f_{u_N^1}(u_N^1) du_N^1.$$

Using the formula for the order statistics conditional on u_N^1, we obtain

$$\begin{aligned}\Big(\mathbb{E}[u_{n+1}^l \mid u_N^1] - \mathbb{E}[u_n^l \mid u_N^1]\Big) &= \binom{n}{n+1-l} \int_0^{u_N^1} \Big(\frac{F(u)}{F(u_N^1)}\Big)^{n+1-l}\Big(1-\frac{F(u)}{F(u_N^1)}\Big)^l du \\ &= \binom{n}{n+1-l} \int u^{n+1-l}(1-u)\frac{F(u_N^1)}{f(F^{-1}(uF(u_N^1)))} du.\end{aligned}$$

We claim that for each u, the function $u_N^1 \mapsto \frac{F(u_N^1)}{f(F^{-1}(uF^{-1}(u_N^1)))}$ is non-decreasing. To see this, define the increasing function $z(u_N^1) = F^{-1}(uF(u_N^1))$ and write

$$\frac{F(u_N^1)}{f(F^{-1}(uF^(u_N^1)))} = \frac{F(z(u_N^1))/u}{f(z(u_N^1))}.$$

Since $F(z)/f(z)$ is increasing, it follows that $\frac{F(u_N^1)}{f(F^{-1}(uF^{-1}(u_N^1)))}$ is non-decreasing in u_N^1. It thus follows that

$$\Big(\mathbb{E}[u_{n+1}^l \mid u_N^1] - \mathbb{E}[u_n^l \mid u_N^1]\Big) \leq \binom{n}{n+1-l} \int u^{n+1-l}(1-u)^l \frac{1}{f(F^{-1}(u))} du = \Delta_n^l$$

ans, as a result, $\tilde{\Delta}_n^l \leq \Delta_n^l$. □

H Proof of Proposition 5

Proposition 5. Under Assumption 1, suppose that F is uniform, and $N \geq N_0 + 1$. The following hold:

a. The smallest (resp. largest) Nash equilibrium of game G (resp. G^{IB}) is less than or equal (resp. greater than or equal) to any Nash equilibrium of game G^{IB} (resp. G).

b. Let n be the smallest Nash equilibrium of the game G and assume that n is strictly smaller than all Nash equilibria of game G^{IB}. Assume $N_0 \leq \bar{N}_0$, where $\bar{N}_0$ is the unique N_0 satisfying $N_0(1-Nk_1)^{N_0} = 1/15$. Then, each student gets strictly more welfare in the game G under equilibrium n than in any Nash equilibrium of game

G^{IB}.

Proof. Part (a) of the proposition follows directly from Lemma 4. Let n_{IB} be the smallest Nash equilibrium in the game G^{IB}. To proof part (b), it is enough to show that $U^I(n', p_{n_{IB},k}, \mu) \leq U(n', p_{n,k}, \mu)$ for all n'. To simplify notation and exposition, fix a state k and let $p_{IB} := p_{n_{IB},k}$ and $p_E := p_{n,k}$. Also, let

$$\begin{aligned} U^{IB}(n', p_{IB}, k) &= \sum_{l=1}^{n'-1} [(1-u_N^N)u_{n'-1}^l + u_N^N](1-p_{IB})p_{IB}^{l-1} + u_N^N(1-p_{IB})p_{IB}^{n'-1} - c(n'-1) \\ U(n', p_E, k) &= \sum_{l=1}^{n'} u_{n'}^l (1-p_E)p_E^{l-1} - cn' \end{aligned}$$

And note that

$$\begin{aligned} \frac{\partial U(n', p_E, k)}{\partial p} &= \sum_{l=1}^{n'-1} (u_{n'}^{l+1} - u_{n'}^l) l p_E^{l-1} - n' p_E^{n'-1} u_{n'}^{n'} \\ &= \frac{1}{n+1} \sum_{l=1}^{n'} l p_E^{l-1} \\ &\leq -\frac{n' p_E^{n'-1}}{2} \end{aligned}$$

and

$$\begin{aligned} U(n', p_{IB}, k) &= U(n', p_E, k) + \int_{p_E}^{p_{IB}} \frac{\partial U}{\partial p}(n', x, k) dx \\ &\leq U(n', p_E) - (p_{IB} - p_E) \frac{n' p_E^{n'-1}}{2} \end{aligned}$$

Then, we have

$$U^I(n', p_{IB}, k) - U(n', p_E, k) \leq U^I(n', p_{IB}, k) - U(n', p_{IB}) - (p_{IB} - p_E) \frac{n' p_E^{n'-1}}{2}$$

And we want to show that, for all k,

$$U^I(n', p_{IB}, k) - U(n', p_{IB}) \leq (p_{IB} - p_E) \frac{n' p_E^{n'-1}}{2}$$

$$\sum_{l=1}^{n'-1} [(1-u_N^N)u_{n'-1}^l + u_N^N - u_{n'}^l](1-p_{IB})p_{IB}^{l-1} + (u_N^N - u_{n'}^{n'})(1-p_{IB})p_{IB}^{n'-1} + c \quad \leq \quad (p_{IB} - p_E) \frac{n' p_E^{n'-1}}{2}$$

But $(1-u_N^N)u_{n'-1}^l + u_N^N - u_{n'}^l < 0$ and $u_N^N - u_{n'}^{n'} < 0$. Thus, it is enough to show that

$$c \leq \frac{1}{2} \leq (p_{IB} - p_E)\frac{n' p_E^{n'-1}}{2}$$

Noticing that,

$$\begin{aligned}
(p_{IB} - p_E) &= p_E((1-Nk)^{\frac{1}{n_{IB}}-\frac{1}{n}} - 1) \\
&\geq p_E\left(\left(\frac{1}{1-Nk}\right)^{\frac{1}{N_0^2}} - 1\right) \\
&\geq p_E\left(\left(\frac{1}{1-Nk_0}\right)^{\frac{1}{N_0^2}} - 1\right) \\
&= p_E\left(\frac{1}{\tilde{p}^2} - 1\right) \\
&\geq 15 p_E
\end{aligned}$$

We thus need

$$\frac{15 n' p_E^{n'}}{2} \geq \frac{1}{2}$$
$$n' p_E^{n'} \geq N_0(1-Nk_1)^{N_0} \geq \frac{1}{15}$$

Which concludes the proof. $\square$

I Proof of Lemma 4

Lemma 4 Under assumption 1, for all $n \geq 1$,

$$B^I(n+1, \bar{n}, \mu) - B^I(n, \bar{n}, \mu) \geq B(n+1, \bar{n}, \mu) - B(n, \bar{n}, \mu)$$

Proof. Let $\Delta B_n^I := B^I(n+1,\bar{n},\mu) - B^I(n,\bar{n},\mu)$ and $\Delta B_n := B(n+1,\bar{n},\mu) - B(n,\bar{n},\mu)$, and note that

$$\begin{aligned}
\Delta B_n^I &= \mathbb{E}_\mu\left[(1-u_N^N)\left((1-p_{\bar{n},k})\sum_{i=1}^{n-1}\Delta_{n-1}^i p_{\bar{n},k}^{i-1} + u_n^n(1-p_{\bar{n},k})p_{\bar{n},k}^{n-1}\right) + u_N^N(1-p_{\bar{n},k})p_{\bar{n},k}^n\right] \\
\Delta B_n &= \mathbb{E}_\mu\left[\left((1-p_{\bar{n},k})\sum_{i=1}^{n}\Delta_n^i p_{\bar{n},k}^{i-1} + u_{n+1}^{n+1}(1-p_{\bar{n},k})p_{\bar{n},k}^n\right)\right]
\end{aligned}$$

Thus, we want to show that

$$\mathbb{E}_\mu\left[(1-p_{\bar{n},k})\left((1-u_N^N)\sum_{i=1}^{n-1}\Delta_{n-1}^i p_{\bar{n},k}^{i-1}+(1-u_N^N)u_n^n p_{\bar{n},k}^{n-1}+u_N^N p_{\bar{n},k}^n\right.\right.$$
$$\left.\left.-\sum_{i=1}^{n}\Delta_n^i p_{\bar{n},k}^{i-1}-u_{n+1}^{n+1}p_{\bar{n},k}^n\right)\right]\geq 0$$

To simplify notation, we drop the subindex $\bar{n},k$ from p and show that, for all k

$$(1-u_N^N)\sum_{i=1}^{n-1}\Delta_{n-1}^i p^{i-1}+(1-u_N^N)u_n^n p^{n-1}+u_N^N p^n-\sum_{i=1}^{n}\Delta_n^i p^{i-1}-u_{n+1}^{n+1}p^n\geq 0$$

Since F is uniform, we have

$$u_n^j=\frac{n-j+1}{n+1}\qquad ;\qquad \Delta_n^j=\frac{j}{(n+1)(n+2)}$$

which implies that

$$\begin{aligned}\sum_{i=1}^{n-1}\left(\frac{N}{N+1}\Delta_{n-1}^i-\Delta_n^i\right)p^{i-1} &= \frac{2N-n}{n(n+1)(n+2)(N+1)}\sum_{i=1}^{n-1}ip^{i-1}\\ &\geq \frac{(2N-n)p^{n-2}}{n(n+1)(n+2)(N+1)}\sum_{i=1}^{n-1}i\\ &= \frac{(2N-n)(n-1)p^{n-2}}{2(n+1)(n+2)(N+1)}\end{aligned}$$

Hence, we need to show that

$$\frac{(2N-n)(n-1)}{2(n+1)(n+2)(N+1)}p^{n-2}+\frac{N}{N+1}\frac{1}{n+1}p^{n-1}+\frac{1}{N+1}p^n$$
$$-\frac{n}{(n+1)(n+2)}p^{n-1}-\frac{1}{n+2}p^n\geq 0$$

$$\Leftrightarrow \frac{(2N-n)(n-1)}{2(n+1)(n+2)(N+1)}p^{n-2}+\left(\frac{N}{(N+1)(n+1)}-\frac{n}{(n+1)(n+2)}\right)p^{n-1}$$
$$-\frac{N-(n+1)}{(N+1)(n+2)}p^n\geq 0$$

$$\Leftrightarrow (N-n/2)(n-1)+(2N-n)p-(n+1)(N-(n+1))p^2\geq 0$$

But assumption 1 holds, so $p \leq 1/4$, implying that

$$\begin{aligned}(N-n/2)(n-1)+(2N-n)p-(n+1)(N-(n+1))p^2 \geq \\ (N-n/2)(n-1)+(2N-n)p-(n+1)(N-(n+1))\frac{1}{16}\end{aligned}$$

Thus, it is enough to prove that

$$\begin{aligned}(N-n/2)(n-1)-(n+1)(N-(n+1))\frac{1}{16} &\geq 0 \\ \Leftrightarrow N &\geq \frac{(7n-3)(n-1)-4}{15n-17}\end{aligned}$$

This is always true because

$$N \geq N_0+1 \geq n+1 \geq \frac{(7n-3)(n-1)-4}{15n-17}$$

Which concludes the proof.

□

References

Abdulkadiroğlu, A., Pathak, P., and Roth, A. (2009). Strategy-proofness versus efficiency in matching with indifferences: Redesigning the nyc high school match. *American Economic Review*, 99:1974–1978.

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2018). Informational braess' paradox: The effect of information on traffic congestion. *Operations Research*, 66(4):893–917.

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

Arteaga, F., Kapor, A. J., Neilson, C. A., and Zimmerman, S. D. (2022). Smart matching platforms and heterogeneous beliefs in centralized school choice. *The Quarterly Journal of Economics*, 137(3):1791–1848.

Azevedo, E. M. and Leshno, J. D. (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy*, 124(5):1235–1268.

Balakrishnan, N. and Cohen, A. C. (2014). *Order statistics & inference: estimation methods*. Elsevier.

Bergemann, D. and Morris, S. (2016). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522.

Blum, Y., Roth, A. E., and Rothblum, U. G. (1997). Vacancy chains and equilibration in senior-level labor markets. *Journal of Economic theory*, 76(2):362–411.

Chambers, C. P. and Yenmez, M. B. (2017). Choice and matching. *American Economic Journal: Microeconomics*, 9(3):126–147.

Chen, Y. and He, Y. (2022). Information acquisition and provision in school choice: a theoretical investigation. *Economic Theory*, 74(1):293–327.

Correa, J., Epstein, N., Epstein, R., Escobar, J., Rios, I., Aramayo, N., Bahamondes, B., Bonet, C., Castillo, M., Cristi, A., Epstein, B., and Subiabre, F. (2022). School choice in chile. *Operations Research*, 70(2):1066–1087.

Das, S., Kamenica, E., and Mirka, R. (2017). Reducing congestion through information design. In *2017 55th annual allerton conference on communication, control, and computing (allerton)*, pages 1279–1284. IEEE.

David, H. (1997). Augmented order statistics and the biasing effect of outliers. *Statistics & probability letters*, 36(2):199–204.

Echenique, F. and Yenmez, M. B. (2015). How to control controlled school choice. *American Economic Review*, 105(8):2679–2694.

Elacqua, G., Gómez, L., Krussig, T., Méndez, C., and Neilson, C. (2022). The potential of smart matching platforms in teacher assignment: The case of ecuador.

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

Gale, D. and Sotomayor, M. (1985). Ms. machiavelli and the stable matching problem. *The American Mathematical Monthly*, 92(4):261–268.

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

Hirshleifer, J. (1971). The private and social value of information and the reward to inventive activity. *American Economic Review*, pages 541–556.

Immorlica, N., Leshno, J., Lo, I., and Lucier, B. (2020). Information acquisition in matching markets: The role of price discovery. *Available at SSRN 3705049*.

Kamada, Y. and Kojima, F. (2014). General theory of matching under distributional constraints. Technical report, mimeo.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kamien, M. I., Tauman, Y., and Zamir, S. (1990). On the value of information in a strategic conflict. *Games and Economic Behavior*, 2(2):129–153.

Kelso Jr, A. S. and Crawford, V. P. (1982). Job matching, coalition formation, and gross substitutes. *Econometrica: Journal of the Econometric Society*, pages 1483–1504.

Konishi, H. and Ünver, M. U. (2006). Games of capacity manipulation in hospital-intern markets. *Social choice and Welfare*, 27(1):3–24.

Levin, J. (2001). Information and the market for lemons. *RAND Journal of Economics*, pages 657–666.

Mathevet, L., Perego, J., and Taneva, I. (2020). On information design in games. *Journal of Political Economy*, 128(4):1370–1404.

Milgrom, P. and Roberts, J. (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica: Journal of the Econometric Society*, pages 1255–1277.

Roth, A. E. (1982). The economics of matching: Stability and incentives. *Mathematics of operations research*, 7(4):617–628.

Taneva, I. (2019). Information design. *American Economic Journal: Microeconomics*, 11(4):151–185.

Tavafoghi, H. and Teneketzis, D. (2017). Informational incentives for congestion games. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1285–1292. IEEE.

Watt, M. (2021). Concavity and convexity of order statistics in sample size. *arXiv preprint arXiv:2111.04702*.

Wu, M., Liu, J., and Amin, S. (2017). Informational aspects in a class of bayesian congestion games. In *2017 American Control Conference (ACC)*, pages 3650–3657. IEEE.