Congestion and Information Design in Matching Markets *

Juan F. Escobar†
University of Chile

Alfonso Montes‡
Diego Portales University

November, 2025

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, while policymakers implement information interventions to guide families' application strategies. We evaluate the impact of such information protocols on equilibrium 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 about top choices reduces congestion and increases welfare. Our analysis provides new insights for market design, as information interventions can subtly affect behavior, spillovers, and welfare in matching markets.

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 popular method that processes preference lists from students and schools

*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: alfonso.montes@udp.cl

to produce a stable matching. While this algorithm ensures students have incentives for truthful reporting (Roth, 1982), families still face the daunting task of discovering, evaluating, and ranking schools.

Policymakers often institute protocols to assist families in order to address informational frictions and make the systems more transparent. Some interventions provide families with information about market congestion, encouraging them to consider a wider range of schools in their application (Arteaga et al., 2022; Elacqua et al., 2022). Other interventions offer detailed information about school characteristics, aiding families in identifying attractive options based on factors like location and academic performance (Hastings and Weinstein, 2008; Andrabi et al., 2017; Correa et al., 2022; Cohodes et al., 2025).

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 (Hirshleifer, 1971; Kamien et al., 1990; Levin, 2001; Taneva, 2019; Bergemann and Morris, 2016). In particular, in matching markets that use the deferred acceptance algorithm, a student who places an additional application may reduce the admission chances of other students (Kelso Jr and Crawford, 1982; Blum et al., 1997; Chambers and Yenmez, 2017). As a result, information interventions that alter the search and application incentives of students in matching markets are not guaranteed to result in welfare improvements.

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 the information is about admission probabilities than when it 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. The number of schools a student 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. A student has lower chances of admission to her listed schools when all other students in the market increase the number of searched schools. Searching thus generates a negative externality. Under some conditions, the model has strategic complementarities: a student searches more when so do other students (Milgrom and Roberts, 1990). In equilibrium, students over-search and equilibria are Pareto-ranked. We also show that placing limits on the number of schools students can apply to can improve students' welfare.

The second set of results describes how information provision about the number of seats in the market changes equilibrium behavior and welfare. In our model, information about the number of seats is a simple way to provide information about congestion in the market.[1] We show that full transparency is never optimal. When capacity turns out to be low, market congestion intensifies and search behavior becomes too intense under full transparency. We show that a policymaker can avoid this outcome by withholding some information. As a result, transparency about market congestion or admission cutoffs is a suboptimal policy.

Our final set of results explores the impact of disclosing information about schools. When the policymaker knows and discloses the name of the most preferred school to each student, the incremental value of learning is reduced for the student, which in turn induces all students to search less and reduces 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

[1]Policymakers frequently release information about market congestion, like the acceptance probability, or the total number of applications from the previous year; see Arteaga et al. (2022), Elacqua et al. (2022).

independent distributions that satisfy the monotone hazard rate property, and the system capacity is non-negligible. 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. Idoux (2023) shows that students in the NYC school choice program who faced decreases in admission odds lengthened their applications. Arteaga et al. (2022) provide evidence from the school choice program in Chile and show that students engage in more active search and apply to more schools when they hold less optimistic views about their admission possibilities. These studies show that as perceived competition heats up –implying others are also searching actively– a student increases her own search effort. This real-world behavior supports our model's premise of strategic complementarities in the school search process.

Our theory delivers 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. 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.[2]

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 used in many school choice and college admission systems– reduces market congestion and could be an effective tool to reduce market congestion.

Providing families with personalized information about schools they will likely 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 theoretical framework can be used to assess the welfare consequences of a proto-

[2]Targeted information interventions to motivate increased search could encourage some students to add schools with available seats, as in (Arteaga et al., 2022). 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.

col that discloses personalized school information. The positions of the disclosed schools within students' rank ordered lists provides insights on the protocol's welfare effect. Specifically, when the schools disclosed are highly rated for students, the intervention is likely to improve the welfare by alleviating market congestion. Conversely, when the schools disclosed rank poorly, the protocol may be reducing welfare by exacerbating congestion.

Related Literature. This work relates to recent research exploring information acquisition in matching markets.[3] Artemov (2021) and Chen and He (2022) explore the incentives to acquire information under different school choice mechanisms. Maxey (2024) observes that search externalities may impact welfare in matching markets under common values. Immorlica et al. (2020) and Hakimov et al. (2023) argue that information about cutoffs may facilitate the search process for each student but do not analyze welfare. Bucher and Caplin (2021) emphasizes the heterogeneous impact of information costs on welfare in a school choice model. We contribute to this growing literature by providing a tractable framework to explore the role of information interventions on students' beliefs, equilibrium search behavior, and overall welfare.

This paper contributes to the literature on information design, which studies how the provision of information influences the beliefs and behavior of individual users. This research has found applications in various domains, such as the design of traffic information for drivers (Acemoglu et al., 2018; Das et al., 2017), or product recommendations for users of digital platforms (Che and Hörner, 2018). Our work explores information design questions that are specific to matching markets. We analyze whether a central planner should disclose congestion measures (e.g., admission cutoffs, available seats) or provide specific information related to agents' preferences. These questions have no analogue in the information design literature and require new arguments and formal results.

Our analysis connects with some findings in the search theory literature, particularly the insight that the equilibrium search is excessively inefficient (Rogerson et al., 2005; Wright et al., 2021). Arnosti et al. (2021) examine congestion in decentralized matching markets and show that externalities can be alleviated by restricting application lists.[4] We depart from the literature by evaluating how *information interventions* shape congestion,

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

[4] He and Magnac (2022) empirically study how application costs affect congestion in matching markets under the deferred acceptance algorithm.

search and application behavior, and welfare.

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. Section 6 discusses extensions. Finally, Section 7 concludes. All omitted proofs are in the Appendix.

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 uniformly distributed in the interval $[0, 1]$. Each student is characterized by $i \in [0, 1]$. All schools have the same capacity k, which is randomly realized from a finite set $K \subset]0, \infty[$ according to a distribution μ which is in the interior of $\Delta(K)$. We assume for all $k \in K$, $k \leq 1/N$, with strict inequality for some $k \in K$. We denote the smallest and the largest capacity $k \in K$ as k_m and k_M, respectively.

Searching 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 the interval $[0, 1]$. 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 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 who are assigned to unattractive schools that have an excess supply of seats.

A student searches for schools to determine her consideration sets and to learn about her private valuations. Similar to Stigler (1961), each student i decides a number in $\{0, 1, \dots, N_0\}$ of schools to search and inspect, with $1 \leq N_0 \leq N$. After n is chosen by the student i, an ordered list of schools $S' = (s_1, \dots, s_n)$, with $s_m \in S$, is realized according to a uniform distribution. The student observes n realizations from F, the highest realization is the utility the student derives from her top school in S', the second highest realization is the utility the student derives from the school ranked second in

S', 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 $c \cdot n$. This cost captures the opportunity cost of time of discovering, inspecting, and applying to schools. We assume that $0 < c < N \cdot k_m \cdot \mathbb{E}_F[u]$.[5] This condition means that discovering and inspecting schools is costly and, as we will see later in the paper, guarantees that in equilibrium students search for at least one school.

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 student $i \in [0,1]$. Students submit rank order lists and schools rank applicants by drawing independent scores from the uniform distribution over $[0,1]$.[6] After $k \in K$ is realized, the school district runs the (student-proposing) Gale-Shapley deferred acceptance algorithm (Gale and Shapley, 1962).

A profile $\bar{n} = (\bar{n}_i)_{i\in[0,1]}$ 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 $\ell \in \{1, \ldots, N\}$ schools is

$$\pi_\ell^{\bar{n}} := \int_{\{i:\bar{n}_i=\ell\}} di. \tag{2.1}$$

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 $\bar{n} = (\bar{n}_i)_{i\in[0,1]}$, the cutoff p must satisfy the following market clearing condition:

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

To understand this condition, note that given that ℓ schools have been ranked, the

[5]Throughout the paper, $\mathbb{E}_G[x]$ denotes the expectation of a random variable x with distribution G.

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

probability that a school is in the rank order list is ℓ/N, and it will be ranked in position η with probability $1/\ell$. 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 who ranks school in the η-position is accepted in the school. By rearranging equation (2.2), we obtain

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

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 $\bar{n} = (\bar{n}_i)_{i\in[0,1]}$ (through the distribution π^n).[7]

Equilibrium. We are interested in equilibrium search patterns. The benefit that a 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

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

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. $\tilde{u}_n^1 \geq \cdots \geq \tilde{u}_n^n$, and we refer to $\tilde{u}_n^\ell$ as the ℓ^{th} top order statistic. By defining

$$U(n, \bar{n}, \mu) := \mathbb{E}_\mu[B(n, \bar{n}, k)] - c \cdot n \tag{2.5}$$

we can characterize a Nash equilibrium of the search game G as a profile $\bar{n} = (\bar{n}_i)_{i\in[0,1]}$ such that for all i

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

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

This condition captures the idea that a student searches for a given number of schools to maximize her expected utility, taking the search behavior of all other students as given.[8]

Our game is symmetric and thus we also introduce symmetric Nash equilibria. In a symmetric Nash equilibrium, all agents search for the same number of schools. Note that when all students search for the same number of schools $\bar{n} \in \{0, \dots, N\}$, the stable matching cutoff assumes the form $p_{k,\bar{n}} = (1 - Nk)^{1/\bar{n}}$. In this case, we abuse notation and write $B(n, \bar{n}, k)$ and $U(n, \bar{n}, \mu)$ as the expected benefit and expected utility of an agent who searches for n schools when all other agents in the market search for $\bar{n}$ schools. We denote by $\mathrm{NE}(\mu)$ the set of all symmetric Nash equilibria given beliefs μ, defined as

$$\mathrm{NE}(\mu) = \Big\{ n \in \{0, \dots, N\} \mid n \in \arg \max_{n' \in \{0, \dots, N\}} \mathbb{E}_{\mu}[U(n', n, k)] \Big\}.$$

Discussion. The model is designed to capture some key features of centralized school choice procedures, where assignments result from the deferred acceptance algorithm but students face uncertainty about market fundamentals. In the model, the total number of seats is unknown to students. This assumption captures the difficulty faced by a family to assess admission probabilities. Although we have modeled this as a random supply of seats, our analysis also applies to environments in which the mass of students entering the market is unknown.

Through search, a student discovers schools and learns their attractiveness, thereby enlarging her consideration set. A critical feature of the model is that the decision to search, and thus the size of the consideration set, is determined endogenously by equilibrium congestion. This captures an important aspect highlighted by recent empirical research showing that the sizes of rank order lists in school choice programs are endogenous to beliefs about market congestion (Ajayi and Sidibe, 2020; Arteaga et al., 2022; Idoux, 2023).

We assume that a student applies only to schools that have been searched for. In practice, students may be familiar with some schools and apply to them without incurring any search costs. This possibility can be incorporated by assuming that students are endowed with a set of familiar schools they can rank at no cost. Our analysis applies inasmuch as students can still expand their rank order list by searching for additional schools in response to market conditions.

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

3 Congestion and Strategic Complementarities

This section characterizes equilibrium search behavior. We offer three main 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 imposes a negative externality. Second, a student has stronger incentives to search for schools when other students in the market search more. In other words, our search game is supermodular and exhibits strategic complementarities (Milgrom and Roberts, 1990). Third, placing limits on the number of schools students can apply to can improve students' welfare.

Our first result describes the properties of the function $B(n, \bar{n}, k)$.

Proposition 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* $\bar{n}$. *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.

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 creates a negative externality.

The negative externality result relies on the assumption that searched schools are oversubscribed. If, by searching, students discovered under-demanded schools, 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 a strategy profile $\bar{n} = (\bar{n}_i)_{i \in [0,1]}$, we consider the search incentives of a student and characterize her best response as:

$$BR(\bar{n};\mu) = \underset{n\in\{0,\dots,N_0\}}{\arg\max}\, U(n,\bar{n},\mu).$$

We will show that over an important set of parameters, the best response map BR is non-decreasing in $\bar{n}$. We say that $BR(\bar{n};\mu)$ is *non-decreasing* in $\bar{n}$ if $\max_n\{n \in BR(\bar{n};\mu)\}$ and $\min_n\{n \in BR(\bar{n})\}$ are both non-decreasing in $\bar{n}$.

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

$$\tilde{k} := \frac{1}{N}\left(1 - \min_{\{n\leq N_0-1,\ell\leq n\}} \min\left\{\frac{n(u_n^n - \delta_{n+1}^n)}{(n+1)u_{n+1}^{n+1}}, \frac{\delta_{n+1}^\ell}{\delta_n^\ell}\left(1+\frac{\delta_{n+1}^{n+1}}{n\delta_n^\ell}\left(\frac{\delta_{n+1}^\ell}{\delta_n^\ell}\right)^{\frac{(n+1)(n-\ell)}{\ell}}\right)\right\}^{N_0}\right)$$

As shown in the Appendix, $\tilde{k} \in]0, \frac{1}{N}[$.

Assumption 1. *For all $k \in K$, $k \geq \tilde{k}$.*

Assumption 1 says the school system has 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. We maintain Assumption 1 throughout the paper. A key implication from Assumption 1 is established in the following result.

Theorem 1 (Strategic complementarities)**.** *For all μ, $BR(\bar{n};\mu) \in \{1,\dots,N_0\}$ is non-decreasing in $\bar{n} = (\bar{n}_i)_{i\in[0,1]}$, where $\bar{n}_i \in \{1,\dots,N_0\}$.*

Theorem 1 shows 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. Arteaga et al. (2022) and None Idoux (2023) show that students participating in centralized platforms under the deferred acceptance algorithm apply to more schools and expand their consideration sets when their admission chances are revealed to be lower. Since admission chances are lower when students in the market search for more schools, the property best responses are non-decreasing, is empirically plausible.

Bounding capacity from below –as Assumption 1 does– is key for Theorem 1. When the assumption does not hold, the best a student can do when others search a lot is not to search. To see this, note that when Assumption 1 does not hold and other students in the market search for a high number of schools, admission cutoffs are arbitrarily high, and therefore, any search effort is wasteful for a student. As a result, the best a student can do is not to search for any school. In this case, $BR(\bar{n};\mu) = 0$ for $\bar{n}$ large enough and BR has a decreasing portion.[9] Assumption 1 thus ensures best responses are non-decreasing, which not only is a key property for the technical analysis of the model but is also an important regularity established for several markets by the empirical market design literature.

The proof of Theorem 1 derives explicit bounds on the capacity of the market to ensure that best responses are non-decreasing. In the proof, we also show that the expectation of the ℓ^{th} top order statistic u_n^ℓ has increasing differences in (ℓ, n). This is a technical contribution that can be useful in other models and applications. See Theorem 4 in the Appendix.

Theorem 1 allows us to neatly characterize Nash equilibria using the theory of supermodular games (Milgrom and Roberts, 1990).

Proposition 2 (Properties of Nash equilibria)**.** *The following hold:*

a. The set of Nash equilibria is nonempty.

b. The set of Nash equilibria has a smallest and a largest element. The smallest and largest Nash equilibria are symmetric.

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 order.[10]

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. Notably, there is a smallest and a largest Nash equilibrium. Equilibria can be Pareto-ranked: students prefer an equilibrium with low search. In particular, the smallest Nash equilibrium not only is symmetric but also Pareto dominates all other Nash equilibria. The result also shows that as μ increases and more weight is placed on high capacities, equilibrium search decreases.

In our search game students ignore search externalities. It is thus natural to explore

[9] Theorem 1 need not hold when $\bar{n} \equiv 0$ is allowed. This is not relevant for our analysis as, under our working assumptions, in equilibrium all students search.

[10] We endow K with the linear order on the reals.

protocols to moderate these externalities. Many matching platforms place limits on the number of schools students can apply to. Naturally, a limit on the number of schools students can apply to does not bound the number of schools students search for. We now show that such limit on the number of applications can be welfare improving. Take a socially optimal number of schools n^{SP} such that

$$n^{SP} \in \arg\max \{U(n', n', \mu) \mid n' \in \{0, 1, \ldots, N\}\}.$$

Consider the game G_L in which students can search any number of schools but can apply to at most n^{SP} schools.[11]

Proposition 3 (Limiting the number of applications)**.** *The following hold:*

a. For any Nash equilibrium $n = (n_i)_{i \in [0,1]}$ of game G, $n^{SP} \leq n_i$ for all i.

b. All Nash equilibria of game G_L result in higher payoffs than any Nash equilibrium of game G.

This result shows that in equilibrium, students over-search: a social planner would like students to search less than they do in any equilibrium Nash equilibrium of the original game G. The intuition is immediate from Proposition 1 as students ignore the congestion impact of their search decisions in equilibrium. More importantly, the Proposition shows that by limiting the number of schools a student can apply to, we can ensure congestion does not blow up, and thus, students have more chances to get to the schools they apply to.[12]

The rest of the paper explores the impact of information interventions on equilibrium and welfare. Information interventions are common in practical applications of matching theory. They range from information about market congestion to suggestions of schools that are likely to be attractive to students. Surprisingly little is known about the system-wide impacts of these interventions.

[11] In game G_L, given equilibrium cutoffs p, the payoff function of a student searching for n schools is:

$$U_L(n,p) = \begin{cases} \sum_{\ell=1}^{n_{SP}} u_n^\ell (1-p) p^{\ell-1} - c \cdot n & \text{if } n \geq n_{SP} \\ \sum_{\ell=1}^{n} u_n^\ell (1-p) p^{\ell-1} - c \cdot n & \text{if } n < n_{SP}. \end{cases}$$

[12] The deferred acceptance algorithm with a restricted application list is not strategy-proof. See Haeringer and Klijn (2009) and Calsamiglia et al. (2010). In our model all schools have the same admission cutoffs, it is thus incentive compatible for a student who searches for $n > n_{SP}$ schools to report her top n_{SP} schools truthfully.

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.

We consider game G and explore how information about the supply of schools k changes equilibrium outcomes. In practice, instead of revealing the supply of seats, authorities often reveal admission chances or suggest the number of schools students should apply to. We will discuss the connection between all of these implementation protocols. More generally, the disclosure of information about the number of seats in the market is an analytically tractable way to model how information about the level of competition for schools shapes families' search strategies.

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 $\gamma_t(k) = \mathbb{P}(k \mid t)$ according to Bayes' rule. Finally, students play the game G given updated beliefs γ_t. We assume the planner can induce any equilibrium given beliefs γ_t (hence, selects the smallest, Pareto-dominant equilibrium).[13]

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

[13]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.

about the total supply of seats.

To state our problem, we abuse notation and let $\tau \in \Delta(\Delta(K))$ be the probability distribution over posterior beliefs induced by experiment τ.[14] 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. We denote by $\mathcal{B}(\mu)$ the set of all Bayes plausible distributions:

$$\mathcal{B}(\mu) = \Big\{\tau \in \Delta(\Delta(K)) : \sum_{\gamma \in \text{supp}(\tau)} \gamma\tau(\gamma) = \mu\Big\}.$$

As discussed by Kamenica and Gentzkow (2011), Bayes plausibility is the only restriction imposed over posterior beliefs when students update using the Bayes rule after observing a public signal.

Let $n^\gamma \in \{1, \ldots, N\}$ be the smallest symmetric Nash equilibrium of the game G given belief γ (characterized in Proposition 2). The social welfare given belief $\gamma \in \Delta(K)$ is[15]

$$V(\gamma) := U(n^\gamma, n^\gamma, \gamma) \tag{4.1}$$

The social planner thus solves

$$\max_{\{\tau \in \mathcal{B}(\mu)\}} \qquad \mathbb{E}_\tau[V(\gamma)] \tag{4.2}$$

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

Before characterizing the solution to (4.2), we observe that experiments can also be implemented by suggesting Nash equilibrium actions n^γ to students, where γ is distributed according to τ and is the realized public belief (Bergemann and Morris, 2016). Suggesting the number of schools to be included in the application is simpler to implement (and it is actually done in some implementations). In contrast, information about admission cutoffs can hardly be used to achieve the outcome induced by τ. To see why, note that given the realized public belief γ, a player's best response depends on the full distribution of cutoffs $(p_{n^\gamma,k})_{k\in K}$—not just its expected value $\mathbb{E}_\gamma[p_{n^\gamma,k}] = \sum_k p_{n^\gamma,k}\gamma(k)$. Thus, unless τ perfectly reveals the capacity (which, as we will show, is suboptimal),

[14] See Kamenica and Gentzkow (2011).

[15] 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.

implementing τ using cutoff information is overly complicated.

Let γ_M be the belief that put all the weight in state k_M, and let $\mathcal{M} := \{\gamma \in \Delta(K) : n^{\gamma} = n^{\gamma_M}\}$.

Theorem 2 (Disclosing market congestion)**.** *Suppose that $\mathcal{M} \subsetneq \Delta(K)$ and has nonempty interior. The following holds:*

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

b. For all $\mu \in \mathcal{M}$, an uninformative experiment is optimal.

Theorem 2 characterizes the optimal information disclosure policy for a planner in game G. The assumption on $\mathcal{M}$—that it has a nonempty interior and is not $\Delta(K)$—is natural and ensures that equilibrium search intensity varies with beliefs, making information persuasion possible. Theorem 2 shows that perfectly revealing the state of the market is never optimal. Furthermore, if students' prior beliefs already induce the minimum possible search intensity (i.e., $\mu \in \mathcal{M}$), then the best a planner can do is provide no additional information.[16]

To understand Theorem 2, note that information has two conflicting effects on welfare. On one hand, there is a private benefit to information: students can make better search decisions when they know the true capacity in the market. On the other hand, information has a strategic effect that creates a negative externality: news suggesting high congestion can trigger a "rat race" where everyone applies to more schools, making the market more congested.

We prove that this negative strategic externality outweighs the private benefit of information. To see how these two forces interact, consider a belief μ that places probability 1 on some capacity, under which students search n^{μ}. Suppose the planner could induce an alternative belief μ', resulting in a lower search intensity $n^{\mu'} < n^{\mu}$. This change is welfare-improving (from the perspective of the original state being true) provided that $U(n^{\mu'}, n^{\mu'}, \mu) > U(n^{\mu}, n^{\mu}, \mu)$, which can be rewritten as:

$$\underbrace{U(n^{\mu'}, n^{\mu'}, \mu) - U(n^{\mu'}, n^{\mu}, \mu)}_{\text{Gain from reduced congestion}} > \underbrace{U(n^{\mu}, n^{\mu}, \mu) - U(n^{\mu'}, n^{\mu}, \mu)}_{\text{Value of tailoring search to belief } \mu} \tag{4.3}$$

The left-hand side of this inequality measures the positive impact on a student that a reduction in others' search decisions has. The right-hand side represents the utility

[16] In this case, any uninformative experiment is optimal. For instance, an experiment that sends the same signal regardless of the state, thereby leaving the posterior belief equal to the prior, is optimal.

gain a student gets from being able to tailor the search decision (n^μ) to the precise information revealed by μ.

The proof of Theorem 2 has two key steps. First, we establish that the benefits from mitigating the search externality are strong enough to outweigh the private value of tailoring search decisions to the true state. Second, we then develop a new variational argument that allows us to perturb the perfectly revealing experiment to improve welfare. As a result, perfectly revealing the market capacity is suboptimal.

5 Disclosing Schools

In school choice systems, oftentimes platforms show some schools with much more prominence than others. In several systems, 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; Cohodes et al., 2025). These policies aim to 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.

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. Each student receives the name of the school ranked top by her, searches for additional schools, and submits her application. This information protocol is an idealized benchmark where the authority can perfectly identify and disclose each student's top school. This assumes that the authority knows the traits that are relevant for each family –proximity, academic performance, art programs– and is able to compute the most preferred school for each family. This assumption (which is relaxed later) captures the goal of personalized recommendation systems that aim to identify 'best fit' schools to each family.

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.[17]. Game G^I is also supermodular and has a smallest Nash equilibrium. We compare Nash equilibria of games G^I and G.

Theorem 3 (Disclosing top schools)**.** *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.*

Theorem 3 establishes that disclosing each family's top-ranked school reduces equilibrium application intensity and improves student welfare. Thus, Theorem 3 provides support to the practice of suggesting schools that are likely to be highly ranked by students and families.

Theorem 3 explores an information intervention that provides private information to each student. We establish that once a student knows her best possible match, the incremental benefit of an application decreases (Lemma 5, Appendix A.4). Intuitively, information disclosure has two competing effects. On the one hand, learning about the top school reduces uncertainty about the quality of the remaining schools. Moreover, the uncertainty about the most attractive school in the list (the top one) is fully resolved. The reduced uncertainty allows for a more accurate search, which is a force that reduces search and application incentives. On the other hand, the free school also changes the incremental value of applications which, as we prove in the Appendix, is a relatively weak force to apply to more schools. Thus, we establish that the reduction in uncertainty effect always dominates the free school effect. As students' decisions are strategic complements, they submit shorter rank-order lists in equilibrium (part a). The resulting decrease in application volume reduces market-wide congestion. Students thus face better admission chances while receiving a free school and additional information, improving overall welfare (part b).

[17] 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

We have shown that disclosing the top school to each family improves welfare. Is any kind of information about schools welfare improving? We now 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. For instance, an authority might recommend a school based on its high performance in mathematics, but for a particular family, that school may be a poor fit due to its location.

Formally, we denote by G^{IB} the game in which each family is revealed the worst school. Given the school revealed, 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 the application.

Proposition 4 (Disclosing bottom schools)**.** *Suppose that F is uniform. 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 $\bar{n}$ be the smallest Nash equilibrium of the game G and assume that $\bar{n}$ is strictly smaller than all Nash equilibria of game G^{IB}. Then, there exists a $\bar{c} > 0$ such that, when $c \leq \bar{c}$, each student gets strictly more welfare in the game G under equilibrium $\bar{n}$ than in any Nash equilibrium of the game G^{IB}.

Proposition 4 shows that disclosing the bottom-ranked schools may create additional congestion in the market, resulting in reduced welfare. As in Theorem 3, the disclosure has two opposing effects. On the one hand, information reduces the uncertainty about the quality of the remaining schools. However, unlike the disclosure of the top school, the uncertainty of the most valuable element on the list does not vanish. On the other hand, the free application also increases the incremental value of applications. The free application effect is dominant, and the disclosure of the bottom school raises the incremental utility of lengthening the application list (Lemma 6, Appendix A.4). Thus, information increases congestion (part a). When c is small, seach is cheap, the platform's information is not very valuable, and the bottom school is of little value to the student who adds it. However, the additional application imposes a negative externality, displacing some students from schools they valued more highly, leading to an overall reduction in welfare (part b).

Taken together, Theorem 3 and Proposition 4 show that the welfare consequences of information disclosure about schools are subtle. While Theorem 3 shows that disclosing top ranked schools is socially desirable, Proposition 4 qualifies the idea that all information interventions are beneficial, showing that disclosing bottom ranked schools can be detrimental. Information may have different impacts on search incentives and market congestion. Disclosing a high-value school satisfies a student's primary search objective, thereby reducing their need to generate a long list of alternatives and mitigating market wide competition. In contrast, disclosing a low-value school facilitates a marginal, low-value application. While this is harmless for the individual student, it generates negative externalities for other students.

The two policies introduced in this section are polar cases that isolate some important mechanisms at play. The main insights are robust to a more realistic setting where the disclosure is noisy. Formally, our results do not require that the disclosed school is the absolute top or bottom with certainty. For instance, a policy that discloses a school with a high probability of being a student's top choice will still lower the expected incremental benefit of search, thereby pushing the equilibrium towards shorter lists and reduced congestion. Thus, the mechanisms identified in Theorem 3 and Proposition 4 remain operative when the information provided by the authority is noisy.

In practice, the positions of the disclosed schools within students' rank ordered lists serve as a sufficient statistic for the protocol's welfare effect. When the revealed schools rank highly for students, the intervention is welfare enhancing, as shown by Theorem 3. In contrast, when the disclosed schools rank poorly, the protocol is welfare reducing, as shown by Proposition 4. These observations provide a clear criterion for evaluating disclosure practices that are already in place, such as recommending nearby or high performing schools. Such policies are likely to be beneficial only if the schools they highlight are placed at the top of families' ranked order lists.

6 Extensions

We now discuss some extensions and variations of our model.

Search technology. We have assumed that by searching n schools, a student observes n random draws from F. Suppose now that a student searching for n schools observes the top n schools among all N schools on the market. In other words, a student observes $u_N^1, u_N^2, \ldots, u_N^n$. This, given the admission cutoff p, a student that searches for $n \geq 1$ schools has utility function given by

$$\sum_{\ell=1}^{n} \mathbb{E}[u_N^\ell](1-p)p^{\ell-1} - cn.$$

This is a model in which the search technology allows families to direct search and learn about their best schools. All our results apply to this model, but Assumption 1 takes a slightly different form. Concretely, we assume that for all $k \in K$,

$$k \geq \frac{1}{N}\left(1 - \left(\min_{n=1,\ldots,N_0-1}\left(\frac{n}{n+1}\right)\right)^{N_0}\right)$$

Under this condition, best responses are non-decreasing and Theorem 1 holds. We can also establish a condition similar to (4.3) to deduce Theorem 2. The results about disclosure of schools in Section 5 are immediate. In this model, the disclosure of top schools does not change the equilibrium number of applications but allows agents to economize on seach costs. Thus, Theorem 3 holds in this model.

Correlated scores. 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 ex-ante symmetric model, under single tie breaking, all students are assigned to their top ranked school and cutoffs 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 rationale 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.

Geographic preferences. We have assumed that students have independent preferences in the sense that knowing the top school of a student does not predict what other attractive schools are for the student. In practice, preferences are often correlated, for instance by geography. Our model and results can be accommodated to allow for such a correlation.

Assume that $N \geq 2$ schools are located equidistantly on the circumference of a circle normalized to 1. A continuum of students is also uniformly distributed along this circle. A student's preference for a school is determined by both an idiosyncratic match component and a penalty for distance. Concretely, a student derives utility $u = \tilde{u} - \psi d$ from a school at distance d (the shortest arc), where $\tilde{u}$ is a match value drawn from a distribution F and $\psi \geq 0$ is a travel cost. A student that searches for n schools observes the utilities from a uniformly and randomly chosen set of n schools, and then submits a rank-order list based on these utilities.

When $\psi = 0$, we recover the original model with independent preferences. For any $\psi > 0$, a student's ranking over the discovered schools is biased towards proximity. However, the random match component $\tilde{u}$ ensures that students will still consider and potentially apply to schools far from their own location. The aggregate demand for each school is symmetric due to the equidistant school placement and the uniform student distribution. Our results on equilibrium search patterns and the impact of information interventions extend to this model.

7 Concluding Comments

This paper explores information interventions in centralized school assignment mechanisms. Our findings suggest a cautious approach to information disclosure. While transparency is often lauded, our analysis of the strategic interactions in student search reveals that not all information is beneficial. Our work contributes to the ongoing discussion on optimizing centralized school assignment systems by providing a clear framework for evaluating such policies. The analysis yields important takeaways.

Providing full transparency about market congestion or admission chances can be detrimental. We show that when families learn that competition is high, their perceived odds of admission fall. This incentivizes them to search more aggressively, leading to excessive search. Policymakers should be cautious about releasing data about market congestion (such as admission cutoffs), as this information can increase the very congestion it describes and harm students.

Disclosing personalized information about schools that students are likely to find attractive is welfare-improving. By providing a family with a school they are likely to rank highly, the intervention reduces their need to engage in costly and extensive search to find a good match. This reduces their individual search effort, and through strategic complementarity, eases aggregate congestion for all participants. As a result, interventions should focus on helping families discover good matches, not on broadcasting market-wide competition levels. Personalized recommendations, such as information on high-performing or nearby schools that align with students' preferences, can both simplify the family's task and improve the overall efficiency of the market. On the other hand, disclosing schools that students are unlikely to value can increase congestion and reduce welfare.

The model's tractable and parsimonious structure allows for a clear analysis of the interplay between policy-relevant information interventions, the matching algorithm, student search patterns, and equilibrium welfare. Our main findings are robust to variations and extensions of the model, including alternative search technologies and geographical preferences. Exploring the model under a dynamic search technology–in which students can adjust their search effort based on past outcomes–is an extension left for future research.

Appendix

This Appendix consists of four parts. Appendix A.1 shows that the expectation of the order statistics has increasing differences. Appendix A.2 provides proofs for Section 3. Appendix A.3 provides proofs for Section 4. Appendix A.4 provides proofs for Section 5.

A.1 Increasing Differences and Order Statistics

In this section, we derive new properties for order statistics. Specifically, we show that the ℓ^{th} top order statistic out of a sample of size n has increasing differences in (n, ℓ). These results are key for the analysis of our search model, and could have other applications in mechanism design and auction theory.

Theorem 4. *Let u_n^ℓ be the expected value of the ℓ^{th} top order statistic, obtained from n independent draws from the distribution F. Suppose that the hazard rate, $f(u)/1-F(u)$, is non-decreasing. Then, u_n^ℓ has increasing differences in (n, ℓ). That is, for all $n \leq N$ and $\ell \leq n$, we have*

$$u_n^\ell - u_{n-1}^\ell \quad \geq \quad u_n^{\ell-1} - u_{n-1}^{\ell-1} \tag{A1}$$

Proof. The proof of the Theorem uses the following two technical lemmas. The first lemma, from David (1997)), gives an expression for the incremental expected value of the ℓ^{th} top order statistic when increasing the number of draws. The second lemma establishes specific properties of the product between a measure and a nonnegative function.

Lemma 1. *(David (1997))*

$$u_n^\ell - u_{n-1}^\ell \quad = \quad \binom{n-1}{n-\ell} \int_0^1 F(\tilde{u})^{n-\ell}(1-F(\tilde{u}))^\ell d\tilde{u}$$

Lemma 2. *Let W be a measure on the interval (a, b) and g a nonnegative function defined on the same interval.*

i. Suppose that g is non-increasing and $\int_a^t dW(x) \geq 0$ for all $t \in (a, b)$. Then, $\int_a^b g(x)dW(x) \geq 0$.

For the proof of Lemma 1, see David (1997). The proof of Lemma 2 follows at the end of this subsection. Let $u^n_{n-1} = 0$, and define $\Delta^\ell_{n-1} := u^\ell_n - u^\ell_{n-1}$ all $\ell \leq n$. We first show that Δ^ℓ_{n-1} is non-decreasing in ℓ for $\ell \leq n-1$. Then, we show it is true for $\ell = n$. Using the definition in Lemma 1, straightforward algebra shows that

$$\begin{aligned}
\Delta^\ell_{n-1} - \Delta^{\ell-1}_{n-1} &= \binom{n-1}{n-\ell} \int_0^1 F(\tilde{u})^{n-\ell}(1-F(\tilde{u}))^{\ell-1} \left[1 - \frac{n}{n-\ell+1} F(\tilde{u})\right] d\tilde{u} \qquad \text{(A2)} \\
&= \binom{n-1}{n-\ell} \int_0^1 F(\tilde{u})^{n-\ell}(1-F(\tilde{u}))^{\ell-2} \left[1 - \frac{n}{n-\ell+1} F(\tilde{u})\right] f(\tilde{u})g(\tilde{u})d\tilde{u}
\end{aligned}$$

where $g(\tilde{u}) := \frac{1-F(\tilde{u})}{f(\tilde{u})}$ is decreasing by assumption. Let $\tilde{u}^*$ be such that

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

So that for $\tilde{u} > \tilde{u}^*$, the integrand in (A2) is negative, for $\tilde{u} = \tilde{u}^*$ is zero, and for $\tilde{u} < \tilde{u}^*$ is positive. Consider now the integral

$$W(t) := \int_0^t F(\tilde{u})^{n-\ell}(1-F(\tilde{u}))^{\ell-2} \left[1 - \frac{n}{n-\ell+1} F(\tilde{u})\right] f(\tilde{u})d\tilde{u} \qquad \text{(A3)}$$

So that the right hand side of equation (A2) can be written as

$$\int_0^1 g(\tilde{u})dW(\tilde{u})$$

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

$$W(1) := \int_0^1 x^{n-\ell}(1-x)^{\ell-2} \left[1 - \frac{n}{n-\ell+1} x\right] dx \qquad \text{(A4)}$$

And using the fact that, for any m, n non negative integers,

$$\int_0^1 x^m(1-x)^n dx = \frac{n!\ m!}{(n+m+1)!} \qquad \text{(A5)}$$

We obtain

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

Thus, $W(t) \geq 0$ for all t. Finally, by Lemma 2, the right-hand side of A2 is positive, concluding that $\Delta_n^\ell - \Delta_n^{\ell-1} \geq 0$ for all $\ell \leq n-1$. Now, we show it is also true for $\ell = n$. That is, we want to show that

$$\Delta_{n-1}^n - \Delta_{n-1}^{n-1} = u_n^n - \Delta_{n-1}^{n-1} \geq 0 \tag{A6}$$

Recall that our ℓ^{th} top order statistic, u_n^ℓ, corresponds to the $(n-\ell+1)^{th}$ *bottom* order statistic, $\mu_n^{n-\ell+1}$, so the distribution of u_n^ℓ is given by[18]

$$\begin{aligned} F_n^\ell(\tilde{u}) &= \sum_{k=n-\ell+1}^{n} \binom{n}{k} F(\tilde{u})^k (1-F(\tilde{u}))^{n-k} \\ &= \sum_{k=0}^{\ell-1} \binom{n}{k} F(\tilde{u})^{n-k} (1-F(\tilde{u}))^k \end{aligned} \tag{A7}$$

Thus, using equation A7, we can write u_n^n as

$$\begin{aligned} u_n^n &= \int \tilde{u} dF_n^n(\tilde{u}) \\ &= 1 - \int F_n^n(\tilde{u}) d\tilde{u} \\ &= 1 - \int \left\{ \sum_{k=0}^{n-1} \binom{n}{k} F(\tilde{u})^{n-k} (1-F(\tilde{u}))^k d\tilde{u} \right\} \\ &= 1 - \int \left\{ \sum_{k=0}^{n} \binom{n}{k} F(\tilde{u})^{n-k} (1-F(\tilde{u}))^k d\tilde{u} - (1-F(\tilde{u}))^n \right\} \\ &= \int (1-F(\tilde{u}))^n d\tilde{u} \end{aligned} \tag{A8}$$

where the first equality follows from the definition of u_n^n, the second follows through integration by parts, and the last equality follows because

$$1 = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \tag{A9}$$

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

Finally, using the expression for Δ_{n-1}^{n-1} given in Lemma 1, we get

$$\begin{aligned} u_n^n - \Delta_{n-1}^{n-1} &= n\int (1-F(\tilde{u}))^{n-1}\left\{\frac{1}{n} - F(\tilde{u})\right\} d\tilde{u} \\ &= n\int (1-F(\tilde{u}))^{n-2}\left\{\frac{1}{n} - F(\tilde{u})\right\} f(\tilde{u})g(\tilde{u})d\tilde{u} \end{aligned}$$

where $g(\tilde{u}) := \frac{1-F(\tilde{u})}{f(\tilde{u})}$ is decreasing by assumption. Let

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

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

$$W(1) := \frac{(n-2)!}{n(n-1)!} - \frac{(n-2)!}{n!} = 0$$

And we conclude that (A6) holds, completing the proof of the theorem.

Proof of Lemma 2. Integrating by parts using $u(t) = g(t)$ and $v(t) = W(t)$, obtain

$$\begin{aligned} \int_a^b g(x)dW(x) &= g(t)W(t)\Big|_a^b - \int_a^b W(t)g'(t)dt \\ &= g(b)W(b) - \int_a^b W(t)g'(t)dt \geq 0 \end{aligned}$$

Since g is non negative, $g(b)W(b) \geq 0$. Furthermore, since $W(a) = 0$, $g(a)\cdot W(a) = 0$. Finally, since $W(t) \geq 0$ for all t and $g'(t) \leq 0$, we have $-\int_a^b W(t)g'(t)dt \geq 0$. □

A.2 Proofs for Section 3

A.2.1 Proof of Proposition 1

Proof. We first show part b. It is easy to see from equations (2.1) and (2.2) that, for profiles $\bar{n} = (\bar{n}_i)$ and $\bar{n}' = (\bar{n}'_i)$ such that $\bar{n}_i \geq \bar{n}'_i$, then $p_{\bar{n},k} \geq p_{\bar{n}',k}$ for all k. Thus, to show that $B(n, \bar{n}, k)$ is decreasing in $\bar{n}$, it is enough to show that it is decreasing in $p_{k,\bar{n}}$.

We drop the sub-index $k, \bar{n}$ from p to simplify notation, and write, for $n \geq 1$

$$\begin{aligned}
\frac{\partial B}{\partial p} &= -u_n^1 + u_n^2(1-2p) + u_n^3(2p-3p^2) + ...u_n^n((n-1)p^{n-2} - np^{n-1}) \\
&= -(u_n^1 - u_n^2) - (u_n^2 - u_n^3)2p - (u_n^3 - u_n^4)3p^2 ... - (u_n^{n-1} - u_n^n)(n-1)p^{n-2} - nu_n^n p^{n-1} \\
&= -\sum_{\ell=1}^{n-1}(u_n^\ell - u_n^{\ell+1})\ell p^{\ell-1} - nu_n^n p^{n-1} \qquad \text{(A10)}
\end{aligned}$$

Which is less than zero because $u_n^\ell - u_n^{\ell+1} > 0$ for all ℓ. When $n = 0$, $B(n, \bar{n}, k) = 0$, and B is constant with respect to p, concluding the proof of part b. To show part a., we use the following technical result due to Li (2005). The lemma says that when the distribution F has a non-decreasing hazard rate, the expected value of the ℓ^{th} top order statistic is increasing and concave in the sample size n.

Lemma 3. *(Li (2005)) Let u_n^ℓ be the expected value of the ℓ^{th} top order statistic, obtained from n independent draws from the distribution F. Suppose that the hazard rate, $f(u)/1-F(u)$, is non-decreasing. Then, u_n^ℓ is increasing and concave in the sample size n. That is, for all $\ell \leq n$, we have $u_n^\ell - u_{n-1}^\ell \geq 0$ and*

$$u_n^\ell - u_{n-1}^\ell \geq u_{n+1}^\ell - u_n^\ell$$

Now define

$$\Delta B_n := B(n+1, \bar{n}, k) - B(n, \bar{n}, k) = \sum_{\ell=1}^{n} \Delta_n^\ell (1-p)p^{\ell-1} + u_{n+1}^{n+1}(1-p)p^n$$

Where, we use $\Delta_n^\ell = u_{n+1}^\ell - u_n^\ell$. By Lemma 3, $\Delta_n^\ell \geq 0$, and the function B increases in n ($\Delta B_n \geq 0$). Furthermore, simple algebra shows that

$$\Delta B_n - \Delta B_{n-1} = (1-p)\left(\sum_{\ell=1}^{n-1}(\Delta_n^\ell - \Delta_{n-1}^\ell)p^{\ell-1} + \Delta_n^n p^{n-1} + u_{n+1}^{n+1}p^n - u_n^n p^{n-1}\right)$$

By Lemma 3, $\Delta_n^\ell - \Delta_{n-1}^\ell \leq 0$. Hence,

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

Then, it suffices for concavity to have

$$\sum_{\ell=1}^{n-1}(\Delta_n^\ell - \Delta_{n-1}^\ell) + p\left(\Delta_n^n + u_{n+1}^{n+1} - u_n^n\right) \leq 0$$

Using the definition of Δ_n^ℓ in Lemma 1 and the expression for u_n^n in equation A8, we obtain

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

Using equation A18 and the expression for u_n^n given in A8, we obtain

$$\begin{aligned} \sum_{\ell=1}^{n} \Delta_n^\ell &= \int (1-F(\tilde{u}))\left\{1-(1-F(u))^n\right\} du \\ &= u_1^1 - u_{n+1}^{n+1} \end{aligned}$$

Thus, it is direct to observe that

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

So the function B is concave.

□

A.2.2 Proof of Theorem 1

The proof relies on the following lemma.

Lemma 4. *Define $\hat{k}$ as:*

$$\hat{k} := \frac{1}{N}\left(1 - \left(\min_{n\in\{1,\dots,N_0-1\}} \left\{\frac{n\left(u_n^n - \delta_{n+1}^n\right)}{u_{n+1}^{n+1}(n+1)}\right\}\right)^{N_0}\right) \tag{A11}$$

Then, $\hat{k}$ $\in]0,1[$, and the function $B(n,\bar{n},k)$ has increasing differences in $n,\bar{n}$ for all $k \geq \hat{k}$.

Proof of Lemma 4 . To show that $B(n',\bar{n},k)$ has increasing differences in $n,\bar{n}$, it is enough to show that it has increasing differences $n', p_{k,\bar{n}}$. Again, we drop the sub-index

n, k from p to ease notation. From equation (A10)

$$\frac{\partial B_n}{\partial p} = -\sum_{\ell=1}^{n-1} \delta_n^\ell \ell p^{\ell-1} - n u_n^n p^{n-1}$$

where $\delta_n^\ell := u_n^\ell - u_n^{\ell+1}$. Then,

$$\frac{\partial B_n}{\partial p} - \frac{\partial B_{n+1}}{\partial p} = \sum_{\ell=1}^{n-1} \left(\delta_{n+1}^\ell - \delta_n^\ell\right) \ell p^{\ell-1} + \delta_{n+1}^n n p^{n-1} + (n+1) u_{n+1}^{n+1} p^n - n u_n^n p^{n-1}$$

So it is enough to have

$$\begin{aligned} \delta_{n+1}^n n p^{n-1} + (n+1) u_{n+1}^{n+1} p^n &\leq n u_n^n p^{n-1} \\ p &\leq \frac{n(u_n^n - \delta_{n+1}^n)}{u_{n+1}^{n+1}(n+1)} \end{aligned}$$

Let

$$\hat{p} := \min_{n \in \{1,\ldots,N_0-1\}} \left\{ \frac{n(u_n^n - \delta_{n+1}^n)}{u_{n+1}^{n+1}(n+1)} \right\}$$

And note that $\hat{p} > 0$ because

$$\begin{aligned} u_n^n - \delta_{n+1}^n &= \int (1 - F(\tilde{u}))^n d\tilde{u} - (n+1) \int F(1 - F(\tilde{u}))^n d\tilde{u} \\ &= (n+1) \int (1 - F(\tilde{u}))^n \left(\frac{1}{n+1} - F(\tilde{u}) \right) d\tilde{u} \end{aligned}$$

So by letting $W(t) = \int_0^t (1 - F(\tilde{u}))^{n-1} \left(\frac{1}{n+1} - F(\tilde{u})\right) f(\tilde{u}) d\tilde{u}$, we can follow the argument in the proof of Theorem 4 and show that $u_n^n - \delta_{n+1}^n > 0$ by showing that $W(1) \geq 0$, which always holds because

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

where we have used once again equation A5. Then, from equation 2.3, it is easy to see that $B(n', \bar{n}, k)$ has increasing differences in $n, \bar{n}$ whenever $\hat{p}^{N_0} \leq (1 - Nk)$, or when

$$k \geq \frac{1}{N} \left(1 - \hat{p}^{N_0}\right) \geq \hat{k} \geq \tilde{k}$$

where $\tilde{k}$ is given in Assumption 1. Since $\tilde{p} \in (0, 1)$, we have $\tilde{k} \in (0, 1)$. Thus, the result follows. $\square$

To deduce the Theorem, note that because of Lemma 4, properties $(A1)-(A4)$ in Milgrom and Roberts (1990) are satisfied. Thus the game is supermodular. 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.

A.2.3 Proof of Proposition 2

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 decreasing 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 decreasing differences. Thus, by Theorem 6 in Milgrom and Roberts (1990), the largest and smallest Nash equilibrium profiles are non-increasing in μ. $\square$

A.2.4 Proof of Proposition 3

Proof. To prove Part [a.], take n and suppose $n < n^{SP}$. Then,

$$\begin{aligned} U(n^{SP},n^{SP}) - U(n,n) &\geq 0 \\ U(n^{SP},n^{SP}) - U(n^{SP},n) + U(n^{SP},n) - U(n,n) &\geq 0 \end{aligned}$$

But $U(n^{SP},n^{SP})-U(n^{SP},n) < 0$ because of the negative externality, which implies that $U(n^{SP},n)-U(n,n) > 0$. Thus, for $n < n^{SP}$, $n \notin BR(n)$. It follows that any symmetric Nash equilibrium of game G must be greater than or equal than n^{SP}. Since the smallest Nash equilibrium of game G is symmetric, the result follows.

Part [b] follows from two observations. First, in any symmetric Nash equilibrium of game G_L, students search $\tilde{n} \geq n_{SP}$ (and apply to n_{SP} schools). To see this, suppose that $\tilde{n} < n_{SP}$. Since $\tilde{n}$ is a Nash equilibrium of G_L, $\tilde{n} \in \arg\max_{n' \leq n_{SP}} \sum_{\ell=1}^{n'} \mathbb{E}_{\mu}[u^{\ell}_{n'}(1-$

$p_{\tilde{n},k})p_{\tilde{n},k}^{\ell-1}] - cn'$. So $\tilde{n}$ is a local maximum of the concave function

$$n \in \{0,\dots,N\} \mapsto \sum_{\ell=1}^{n'} \mathbb{E}_\mu[u_{n'}^{\ell}(1-p_{\tilde{n},k})p_{\tilde{n},k}^{\ell-1}] - cn' = U(n',\tilde{n},\mu).$$

It thus follows that $\tilde{n} < n^{SP}$ is a Nash equilibrium of game G. This contradicts the fact that n^{SP} is less than or equal to any Nash equilibrium of game G.

The second observation is that students searching for $\tilde{n}$ schools and applying to n_{SP} schools get higher payoffs than in any symmetric Nash equilibrium of game G. To see that, note that

$$\begin{aligned}\sum_{\ell=1}^{n_{SP}} \mathbb{E}_\mu[u_{\tilde{n}}^{\ell}(1-p_{n_{SP},k})p_{n_{SP},k}^{\ell-1}] - c\tilde{n} &\geq \sum_{\ell=1}^{n_{SP}} \mathbb{E}_\mu[u_{n_{SP}}^{\ell}(1-p_{n_{SP},k})p_{n_{SP},k}^{\ell-1}] - cn_{SP} \\ &= U(n_{SP},n_{SP},\mu) \\ &\geq U(n,n,\mu)\end{aligned}$$

where n is any symmetric Nash equilibrium of game G. The first inequality follows since $\tilde{n} \geq n_{SP}$ is the equilibrium search number of schools in the game G_L. The equality is by definition of U. The last inequality holds by definition of n_{SP}. This completes the proof. □

A.3 Proofs for Section 4

A.3.1 Proof of Theorem 2

Proof. Let $\gamma_j \in \Delta(K)$ be the belief that puts all the weight in state $k_j \in K$, and let $n^\gamma = (n_i^\gamma)_i$ be the smallest Nash equilibrium profile under common belief γ.[19] First, we show that when

$$U(n^{\gamma_M}, n^{\gamma_M}, \gamma_j) - U(n^{\gamma}, n^{\gamma}, \gamma_j) > 0 \qquad \forall\ n^\gamma > n^{\gamma_M}, \quad \forall\ j \tag{A12}$$

deviations from the fully informative experiment are strictly profitable for all $\mu \in \Delta(K)$. Then, we show that when (A12) holds, an uninformative experiment is optimal for all $\mu \in \mathcal{M}$. Finally, we show that under Assumption 1, condition (A12) is always satisfied.

The fully informative experiment, τ^F, which has support $\sup\{\tau^F\} = (\gamma_j)_{j\in K}$, and

[19] Recall that k_m and k_M denote the smallest and largest capacities in K, respectively.

probability $\mu_j := \mu(k_j)$ of observing posterior γ_j, i.e., $\tau^F(\gamma_j) = \mu_j$, gives a payoff of

$$V_{\tau^F} := \sum_{j \in K} U(n^{\gamma_j}, n^{\gamma_j}, \gamma_j)\tau^F(\gamma_j)$$

Since the prior is interior, $\tau^F(\gamma_j) > 0$ for all j. Now consider a perturbation $\tilde{\tau}$ of τ^F such that the posteriors in the support of $\tilde{\tau}$ are $\tilde{\gamma}_j = \gamma_j$ for all $j \neq M$, and $\tilde{\gamma}_M = (\epsilon, 0, ..., 0, 1-\epsilon)$.

It is easy to see that for $\tilde{\tau}$ to be Bayes plausible, $\tilde{\tau}(\gamma_j) = \mu_j$ for $j \notin \{m, M\}$, $\tilde{\tau}(\tilde{\gamma}_M) = \frac{\mu_M}{1-\epsilon}$, and $\tilde{\tau}(\gamma_m) = \mu_m - \frac{\epsilon}{1-\epsilon}\mu_M$. Thus, the payoff of experiment $\tilde{\tau}$ is

$$V_{\tilde{\tau}} := \sum_{j \in K} U(n^{\tilde{\gamma}_j}, n^{\tilde{\gamma}_j}, \tilde{\gamma}_j)\tilde{\tau}(\tilde{\gamma}_j)$$

Taking the difference between $V_{\tilde{\tau}}$ and V_{τ^F}, obtain

$$\begin{aligned} V_{\tilde{\tau}} - V_{\tau^F} &= \sum_j U(n^{\tilde{\gamma}_j}, n^{\tilde{\gamma}_j}, \tilde{\gamma}_j)\tilde{\tau}(\tilde{\gamma}_j) - \sum_j U(n^{\gamma_j}, n^{\gamma_j}, \gamma_j)\tau^F(\gamma_j) \\ &= U(n^{\gamma_m}, \gamma_m)\tilde{\tau}(\gamma_m) + U(n^{\tilde{\gamma}_M}, \tilde{\gamma}_M)\tilde{\tau}(\gamma_M)... \\ &\quad -U(n^{\gamma_m}, \gamma_m)\tau^F(\gamma_m) - U(n^{\gamma_M}, \gamma_M)\tau^F(\gamma_M) \\ &= \frac{\mu_M}{1-\epsilon}\left\{U(n^{\tilde{\gamma}_M}, \tilde{\gamma}_M) - \epsilon U(n^{\gamma_m}, \gamma_m) - (1-\epsilon)U(n^{\gamma_M}, \gamma_M)\right\} \qquad \text{(A13)} \end{aligned}$$

Where we have shortened the notation using $U(n^\gamma, \gamma) := U(n^\gamma, n^\gamma, \gamma)$. Since $\mathcal{M}$ has non-empty interior, for ϵ small, $n^{\tilde{\gamma}_M} = n^{\gamma_M}$. Then,

$$\begin{aligned} U(n^{\gamma_M}, \tilde{\gamma}_M) &= \epsilon\left\{B(n^{\gamma_M}, n^{\gamma_M}, k_m) - cn^{\gamma_M}\right\} + (1-\epsilon)\left\{B(n^{\gamma_M}, n^{\gamma_M}, k_M) - cn^{\gamma_M}\right\} \\ &= \epsilon U(n^{\gamma_M}, \gamma_m) + (1-\epsilon)U(n^{\gamma_M}, \gamma_M) \end{aligned}$$

Replacing back into (A13), obtain $V_{\tilde{\tau}} - V_{\tau^F} \geq 0$ whenever $U(n^{\gamma_M}, \gamma_m) - U(n^{\gamma_m}, \gamma_m) \geq 0$, which is true by (A12).

Now suppose $\mu \in \mathcal{M}$, and consider an experiment $\hat{\tau}$ inducing a posterior $\gamma \notin \mathcal{M}$. Since $\mathcal{M} \subsetneq \Delta(K)$, $n^\gamma > n^{\gamma_M}$. We can rewrite γ as a convex combination of extreme points of the simplex, obtaining

$$
\begin{aligned}
U(n^\gamma, n^\gamma, \gamma) &= \sum_j \alpha_j U(n^\gamma, n^\gamma, \gamma_j) \\
&< \sum_j \alpha_j U(n^{\gamma_M}, n^{\gamma_M}, \gamma_j) \\
&= U(n^{\gamma_M}, n^{\gamma_M}, \gamma) \qquad \text{(A14)}
\end{aligned}
$$

Where the inequality holds by (A12). Thus, the payoff of an uninformative experiment is

$$
\begin{aligned}
V_{\tau^N} &= U(n^{\gamma_M}, \mu) \\
&= \sum_{\sup\{\hat{\tau}\}} U(n^{\gamma_M}, \gamma)\hat{\tau}(\gamma) \\
&> \sum_{\sup\{\hat{\tau}\}} U(n^{\gamma}, \gamma)\hat{\tau}(\gamma) \\
&= V_{\hat{\tau}}
\end{aligned}
$$

where the second line follows because $\hat{\tau}$ is Bayes plausible, and the inequality follows by (A14). We conclude the proof by showing that, under assumption 1, condition (A12) always hold. Note that a sufficient condition for (A12) to hold is, for all $n^\gamma > n^{\gamma_M}$ and all γ_j

$$
U(n^{\gamma_M}, n^{\gamma_M}, \gamma_j) - U(n^\gamma, n^\gamma, \gamma_j) > U(n^{\gamma_M}, n^{\gamma_M}, \gamma_M) - U(n^\gamma, n^\gamma, \gamma_M) \qquad \text{(A15)}
$$

where the right-hand side of (A15) is always greater than zero because

$$
U(n^{\gamma_M}, n^{\gamma_M}, \gamma_M) \geq U(n^\gamma, n^{\gamma_M}, \gamma_M) > U(n^\gamma, n^\gamma, \gamma_M)
$$

The first inequality comes from the Nash equilibrium condition, and the second inequality comes from the negative externality. Thus, (A15) implies (A12). Letting $B(n^\gamma, k) := B(n^\gamma, n^\gamma, k)$, we can rewrite (A15) as

$$
B(n^\gamma, k_M) - B(n^\gamma, k_j) > B(n^{\gamma_M}, k_M) - B(n^{\gamma_M}, k_j)
$$

for all $n^{\gamma} > n^{\gamma_M}$ and all k_j. Thus, it is enough to show that for any $k \geq \tilde{k}$,

$$B(n+1, k_M) - B(n+1, k) > B(n, k_M) - B(n, k) \tag{A16}$$

Using the definition of B in equation 2.4 and $\delta_n^\ell = u_n^\ell - u_n^{\ell+1}$, we obtain

$$B(n+1,k) - B(n,k) \quad = \quad u_{n+1}^1 - u_n^1 + \sum_{\ell=1}^{n} \left\{ p_n^\ell \delta_n^\ell - p_{n+1}^\ell \delta_{n+1}^\ell \right\} - p_{n+1}^{n+1} \delta_{n+1}^{n+1}$$

From equation 2.3, it is easy to see that $p_n = (1-Nk)^{\frac{1}{n}}$, so condition A16 holds whenever

$$\begin{aligned} \sum_{\ell=1}^{n} \left\{ (1-Nk)^{\frac{\ell}{n}} \delta_n^\ell - (1-Nk)^{\frac{\ell}{n+1}} \delta_{n+1}^\ell \right\} &< (1-Nk)\delta_{n+1}^{n+1} \\ \Leftrightarrow \sum_{\ell=1}^{n} \left\{ y^{\frac{n(n+1)}{\ell}} \frac{u_{n+1}^{n+1}}{n} + y^n \delta_{n+1}^\ell - y^{n+1} \delta_n^\ell \right\} &> 0 \end{aligned}$$

where we have defined $y := (1-Nk)^{\frac{\ell}{n(n+1)}}$ and use the fact that $\delta_{n+1}^{n+1} = u_{n+1}^{n+1}$. Imposing the inequality for each ℓ, we have

$$\begin{aligned} 0 &< y^{\frac{n(n+1)}{\ell}} \cdot \frac{u_{n+1}^{n+1}}{n} + \delta_{n+1}^\ell y^n - \delta_n^\ell y^{n+1} \\ y &< \frac{\delta_{n+1}^\ell}{\delta_n^\ell} + y^{\frac{n(n+1)}{\ell} - n} \cdot \frac{u_{n+1}^{n+1}}{n \delta_n^\ell} \qquad (=: \Phi(y)) \end{aligned}$$

But for $y_0 := \frac{\delta_{n+1}^\ell}{\delta_n^\ell}$, $y_0 < \Phi(y_0) < \Phi(\Phi(y_0))...$, so condition (A16) holds for all y satisfying

$$\begin{aligned} y &\leq \frac{\delta_{n+1}^\ell}{\delta_n^\ell} + \frac{u_{n+1}^{n+1}}{n\delta_n^\ell} \left(\frac{\delta_{n+1}^\ell}{\delta_n^\ell} \right)^{\frac{n(n+1)}{\ell} - n} \\ 1 - Nk &\leq \min_{\{n \leq N_0 - 1\}} \min_{\{\ell \leq n\}} \left\{ \frac{\delta_{n+1}^\ell}{\delta_n^\ell} + \frac{u_{n+1}^{n+1}}{n\delta_n^\ell} \left(\frac{\delta_{n+1}^\ell}{\delta_n^\ell} \right)^{\frac{n(n+1)}{\ell} - n} \right\}^{\frac{n(n+1)}{\ell}} \\ k &\geq \frac{1}{N} \left(1 - \min_{\{n \leq N_0 - 1\}} \min_{\{\ell \leq n\}} \left\{ \frac{\delta_{n+1}^\ell}{\delta_n^\ell} + \frac{u_{n+1}^{n+1}}{n\delta_n^\ell} \left(\frac{\delta_{n+1}^\ell}{\delta_n^\ell} \right)^{\frac{n(n+1)}{\ell} - n} \right\}^{\frac{n(n+1)}{\ell}} \right) \\ k &\geq \frac{1}{N} \left(1 - \tilde{p}^{N_0} \right) \end{aligned}$$

Where $\tilde{p} = \min_{\{n \le N_0-1\}} \min_{\{\ell \le n\}} \left\{ \frac{\delta^\ell_{n+1}}{\delta^\ell_n} + \frac{u^{n+1}_{n+1}}{n\delta^\ell_n} \left(\frac{\delta^\ell_{n+1}}{\delta^\ell_n} \right)^{\frac{n(n+1)}{\ell} - n} \right\}^{\frac{n(n+1)}{\ell \cdot N_0}}$. Thus, under Assumption 1, (A12) always hold. □

A.4 Proofs for Section 5

A.4.1 Proof of Theorem 3

Proof. Part [a.] of the Theorem follows directly from Theorem 6 (and its corollary) in Milgrom and Roberts (1990) and the following Lemma

Lemma 5. *For all $n \ge 1$,*

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

Part [b.] holds because when every player $j \ne i$ plays n^I, the cutoff is weakly lower than when every $j \ne i$ plays an equilibrium profile in the game G (because of Lemma 5). 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.

□

A.4.2 Proof of Lemma 5

Proof. When the authority fully discloses the family's top school, the utility that a student gets when applying to n schools is

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

where $\mathbb{E}[u^\ell_{n-1} \mid I]$ denotes the expectation of the ℓ^{th} top order statistics among $n-1$ schools after the information about the identity of the top schools is revealed. Thus, for $n \ge 2$, we let $\Delta U^I_n := U^I(n+1, \bar{n}, \mu) - U^I(n, \bar{n}, \mu)$ and obtain

$$\Delta U^I_n = \sum_{\ell=1}^{n-1} \Big(\mathbb{E}[u^\ell_n \mid I] - \mathbb{E}[u^\ell_{n-1} \mid I] \Big)(1 - p_{\bar{n}}, k) p^\ell_{\bar{n},k} + (1 - p_{\bar{n},k}) p^n_{\bar{n},k} \mathbb{E}[u^n_n \mid I] - c$$

Letting $\Delta U_n := U(n+1, \bar{n}, \mu) - U(n, \bar{n}, \mu)$, to prove the Lemma it is enough to show that $\Delta U_n \geq \Delta U_n^I$. That is,

$$\sum_{\ell=1}^{n} \Big(u_{n+1}^{\ell} - u_n^{\ell}\Big)(1 - p_{\bar{n},k}) p_{\bar{n},k}^{\ell-1} + (1 - p_{\bar{n},k}) p_{\bar{n},k}^{n} u_{n+1}^{n+1}$$
$$\geq \sum_{\ell=1}^{n-1} \Big(\mathbb{E}[u_n^{\ell} \mid I] - \mathbb{E}[u_{n-1}^{\ell} \mid I]\Big)(1 - p_{\bar{n},k}) p_{\bar{n},k}^{\ell} + (1 - p_{\bar{n}}) p_{\bar{n},k}^{n} \mathbb{E}[u_n^n \mid I].$$

We denote $\tilde{\Delta}_n^{\ell} = \mathbb{E}[u_{n+1}^{\ell} \mid I] - \mathbb{E}[u_n^{\ell} \mid I]$, and show below $\tilde{\Delta}_n^{\ell} \leq \Delta_n^{\ell}$. We also note that $u_n^n > \mathbb{E}[u_{n+1}^{\ell} \mid I]$. Thus, rearranging terms, it is enough to prove that

$$\mathbb{E}_{\mu}\Big((1 - p_{\bar{n},k})\big(\sum_{\ell=1}^{n-1} (\Delta_n^{\ell} - p_{\bar{n},k} \Delta_{n-1}^{\ell}) p_{\bar{n},k}^{\ell-1} + \Delta_n^n p_{\bar{n},k}^{n-1} + (u_{n+1}^{n+1} - u_n^n) p_{\bar{n},k}^{n}\big)\Big) \geq 0$$

which will always hold if, for all p,

$$\sum_{\ell=1}^{n} \Delta_n^{\ell} p^{\ell-1} \geq \sum_{\ell=1}^{n-1} \Delta_{n-1}^{\ell} p^{\ell} + (u_n^n - u_{n+1}^{n+1}) p^n \tag{A17}$$

We let $\bar{F}(\tilde{u}) := (1 - F(\tilde{u}))$, and observe that

$$\begin{aligned}
\sum_{\ell=1}^{n} \Delta_n^{\ell} p^{\ell-1} &= \frac{1}{p} \sum_{\ell=1}^{n} \begin{pmatrix} n \\ n-l+1 \end{pmatrix} \int F(\tilde{u})^{n-\ell+1} (p\bar{F}(\tilde{u}))^{\ell} d\tilde{u} \\
&= \frac{1}{p} \int \sum_{\ell=0}^{n-1} \begin{pmatrix} n \\ n-\ell \end{pmatrix} F(\tilde{u})^{n-\ell} (p\bar{F}(\tilde{u}))^{\ell+1} d\tilde{u} \\
&= \int \bar{F}(\tilde{u}) \sum_{\ell=0}^{n-1} \begin{pmatrix} n \\ n-\ell \end{pmatrix} F(\tilde{u})^{n-\ell} (p\bar{F}(\tilde{u}))^{\ell} d\tilde{u} \\
&= \int \bar{F}(\tilde{u}) \left\{ \sum_{\ell=0}^{n} \begin{pmatrix} n \\ n-\ell \end{pmatrix} F(\tilde{u})^{n-\ell} (p\bar{F}(\tilde{u}))^{\ell} - (p\bar{F}(\tilde{u}))^n \right\} d\tilde{u} \\
&= \int \bar{F}(\tilde{u}) \left\{ (F(\tilde{u}) + p\bar{F}(\tilde{u}))^n - (p\bar{F}(\tilde{u}))^n \right\} d\tilde{u} \qquad \text{(A18)}
\end{aligned}$$

Where the last equality follows because

$$(x+y)^n = \sum_{\ell=0}^{n} \begin{pmatrix} n \\ n-l \end{pmatrix} x^k y^{n-k} \tag{A19}$$

Analogously, it is easy to check that

$$\sum_{\ell=1}^{n-1} \Delta_{n-1}^{\ell} p^{\ell} = p \int \bar{F}(\tilde{u}) \left\{ (F(\tilde{u}) + p\bar{F}(\tilde{u}))^{n-1} - (p\bar{F}(\tilde{u}))^{n-1} \right\} d\tilde{u}$$

Using the expressions for u_n^n and u_{n+1}^{n+1} provided in equation A8, we can see that $(u_n^n - u_{n+1}^{n+1})p^n = \int F(\tilde{u})(p\bar{F}(\tilde{u}))^n d\tilde{u}$, and our condition A17 becomes

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

Concluding that condition A17 always hold. To conclude the proof of the lemma, it remains to show that $\tilde{\Delta}_n^{\ell} \leq \Delta_n^{\ell}$. When the authority fully discloses $\tilde{u}_N^1 = u_N^1$, the families learn that the random variables $\tilde{u}$ are drawn from the distribution F truncated on the right at u_N^1. Thus,

$$\mathbb{E}[u_{n+1}^{\ell} \mid u_N^1] - \mathbb{E}[u_n^{\ell} \mid u_N^1] = \binom{n}{n+1-\ell} \int_0^{u_N^1} \left(\frac{F(u)}{F(u_N^1)} \right)^{n+1-\ell} \left(1 - \frac{F(u)}{F(u_N^1)} \right)^{\ell} du$$

Changing variables, $x = \frac{F(u)}{F(u_N^1)}$ and $dx = \frac{dF(u)}{F(u_N^1)} du$, we obtain

$$\mathbb{E}[u_{n+1}^{\ell} \mid u_N^1] - \mathbb{E}[u_n^{\ell} \mid u_N^1] \quad = \binom{n}{n+1-\ell} \int_0^1 x^{n+1-\ell}(1-x)^{\ell} \frac{F(u_N^1)}{f(F^{-1}(xF(u_N^1)))} dx.$$

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

$$\frac{F(u_N^1)}{f(F^{-1}(xF(u_N^1)))} = \frac{F(z(u_N^1))/x}{f(z(u_N^1))}$$

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. Thus, taking $u_N^1 = 1$, obtain $F(u_N^1) = 1$ and $x = F(u)$. It thus follows that

$$\begin{aligned}
\mathbb{E}[u_{n+1}^{\ell} \mid u_N^1] - \mathbb{E}[u_n^{\ell} \mid u_N^1] &\leq \begin{pmatrix} n \\ n+1-\ell \end{pmatrix} \int_0^1 x^{n+1-\ell}(1-x)^{\ell} \frac{1}{f(F^{-1}(x))} dx \\
&= \begin{pmatrix} n \\ n+1-\ell \end{pmatrix} \int_0^1 F(u)^{n+1-\ell}(1-F(u))^{\ell} du \\
&= \Delta_n^{\ell}
\end{aligned}$$

and, as a result, $\tilde{\Delta}_n^{\ell} \leq \Delta_n^{\ell}$.

□

A.4.3 Proof of Proposition 4

Proof. The proof of part [a.] follows directly from Theorem 6 (and its corollary) in Milgrom and Roberts (1990) and the following result.

Lemma 6. *Suppose that F is uniform. Then, for all $n \geq 2$, we have*

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

Let n_{IB} be the smallest Nash equilibrium in the game G^{IB}, $p_{IB} := (p_{n_{IB},k})_{k \in K}$ and $p_E := (p_{\bar{n},k})_{k \in K}$, where , $p_{n_{IB},k} = (1 - Nk)^{\frac{1}{n_{IB}}}$ and $p_{\bar{n},k} = (1 - Nk)^{\frac{1}{\bar{n}}}$. To prove part (b), it is enough to show that $U^{IB}(n, p_{n_{IB}}, \mu) \leq U(n, p_E, \mu)$ for all n.

$$\begin{aligned}
U^{IB}(n, p_{IB}, \mu) &= \mathbb{E}_{\mu}\left[\sum_{\ell=1}^{n-1} \left[(1-u_N^N)u_{n-1}^{\ell} + u_N^N\right](1-p_{IB,k})p_{IB,k}^{\ell-1} + u_N^N(1-p_{IB,k})p_{IB,k}^{n-1} - c(n-1)\right] \\
U(n, p_E, \mu) &= \mathbb{E}_{\mu}\left[\sum_{\ell=1}^{n} u_n^{\ell}(1-p_{E,k})p_{E,k}^{\ell-1} - cn\right]
\end{aligned}$$

And note that

$$\begin{aligned}
\frac{\partial U(n,p_E,\mu)}{\partial p} &= -\mathbb{E}_\mu\left[\sum_{\ell=1}^{n-1}(u_n^\ell - u_n^{\ell+1})lp_{E,k}^{\ell-1} - np_{E,k}^{n-1}u_n^n\right] \\
&= -\mathbb{E}_\mu\left[\frac{1}{n+1}\sum_{\ell=1}^{n}\ell p_{E,k}^{\ell-1}\right] \\
&\leq -\mathbb{E}_\mu\left[\frac{np_{E,k}^{n-1}}{2}\right]
\end{aligned}$$

and

$$\begin{aligned}
U(n,p_{IB},\mu) &= U(n,p_E,\mu) + \int_{p_E}^{p_{IB}}\frac{\partial U}{\partial p}(n,x,\mu)dx \\
&\leq U(n,p_E,\mu) - \mathbb{E}_\mu\left[(p_{IB,k}-p_{E,k})\frac{np_{E,k}^{n-1}}{2}\right]
\end{aligned}$$

Then, we have

$$U^{IB}(n,p_{IB},\mu) - U(n,p_E,\mu) \leq U^{IB}(n,p_{IB},\mu) - U(n,p_{IB},\mu) - \mathbb{E}_\mu\left[(p_{IB,k}-p_{E,k})\frac{np_{E,k}^{n-1}}{2}\right]$$

And it is enough to show that,

$$U^{IB}(n,p_{IB},\mu) - U(n,p_{IB},\mu) \leq \mathbb{E}_\mu\left[(p_{IB,k}-p_{E,k})\frac{np_{E,k}^{n-1}}{2}\right] \tag{A20}$$

By noticing that

$$\begin{aligned}
U^{IB}(n,p_{IB},\mu) - U(n,p_{IB},\mu) &= \mathbb{E}_\mu\Big[\sum_{\ell=1}^{n-1}[(1-u_N^N)u_{n-1}^\ell + u_N^N - u_n^\ell](1-p_{IB,k})p_{IB,k}^{\ell-1}... \\
&\quad ... + (u_N^N - u_n^n)(1-p_{IB,k})p_{IB,k}^{n-1}\Big] + c
\end{aligned}$$

We can define

$$Z(n,p_{IB},\mu) \;:=\; -\mathbb{E}_\mu\left[\sum_{\ell=1}^{n-1}[(1-u_N^N)u_{n-1}^\ell + u_N^N - u_n^\ell](1-p_{IB,k})p_{IB,k}^{\ell-1} + (u_N^N - u_n^n)(1-p_{IB,k})p_{IB,k}^{n-1}\right]$$

And condition A20 holds whenever

$$c \leq \bar{c} := Z(n, p_{IB}, \mu) + \mathbb{E}_\mu \left[(p_{IB,k} - p_{E,k}) \frac{n p_{E,k}^{n-1}}{2} \right]$$

It is easy to see that $(1 - u_N^N) u_{n-1}^\ell + u_N^N - u_n^\ell < 0$ and $u_N^N - u_n^n < 0$. Thus, since μ is interior, $\bar{c} > 0$. This concludes the proof. $\square$

A.4.4 Proof of Lemma 6

Proof. When the authority fully discloses the family's bottom school, the utility that a student gets when applying to n schools is

$$U^{IB}(n, \bar{n}, \mu) = \mathbb{E}_\mu \Big(\sum_{\ell=1}^{n-1} \mathbb{E}[u_{n-1}^\ell \mid IB](1 - p_{\bar{n},k}) p_{\bar{n}}^{\ell-1} + (1 - p_{\bar{n},k}) p_{\bar{n},k}^{n-1} u_N^N \Big) - c(n-1)$$

It is well known (see equation 2.4.4 in Balakrishnan and Cohen (2014)) that the density of the ℓ^{th} top order statistic, when the underlying distribution is F, is given by

$$f_n^\ell(u) = \frac{n!}{(n-\ell)!(\ell-1)} F(u)^{n-\ell} (1 - F(\tilde{u}))^{\ell-1} f(u)$$

Thus, for F uniform over $[0, 1]$, we get

$$\begin{aligned} u_n^\ell &= \int_0^1 \tilde{u} f_n^\ell(\tilde{u}) d\tilde{u} \\ &= \frac{n!}{(n-\ell)!(\ell-1)!} \int_0^1 \tilde{u} \cdot \tilde{u}^{n-\ell} (1-\tilde{u})^{\ell-1} d\tilde{u} \\ &= \frac{n!}{(n-\ell)!(\ell-1)!} \int_0^1 \tilde{u}^{n-\ell+1} (1-\tilde{u})^{\ell-1} d\tilde{u} \\ &= \frac{n!}{(n-\ell)!(\ell-1)!} \frac{(n-\ell+1)!(\ell-1)!}{(n+1)!} \\ &= \frac{n-\ell+1}{n+1} \end{aligned} \tag{A21}$$

Where we have used the definition of u_n^ℓ and equation A5. It is direct to see that

$$\Delta_n^\ell = \frac{\ell}{(n+1)(n+2)} \tag{A22}$$

When the authority fully discloses $\tilde{u}_N^N$, the families learn that the random variables $\tilde{u}$ are drawn from the distribution F truncated on the left at $\tilde{u}_N^N$. Thus, when F is uniform, we obtain

$$\begin{aligned}
\mathbb{E}[u_n^\ell|IB] &= \mathbb{E}[\mathbb{E}[u_n^\ell|\tilde{u}_N^N]] \\
&= \mathbb{E}\left[\frac{n!}{(n-\ell)!(\ell-1)}\int_{\tilde{u}_N^N}^1 u\left(\frac{u-\tilde{u}_N^N}{1-\tilde{u}_N^N}\right)^{n-\ell}\left(1-\frac{u-\tilde{u}_N^N}{1-\tilde{u}_N^N}\right)^{\ell-1}\frac{1}{1-\tilde{u}_N^N}du\right] \\
&= \mathbb{E}\left[\frac{n!}{(n-\ell)!(\ell-1)}\int_0^1 \left(x(1-\tilde{u}_N^N)+\tilde{u}_N^N\right)x^{n-\ell}(1-x)^{\ell-1}dx\right] \\
&= \mathbb{E}\left[\tilde{u}_N^N+(1-\tilde{u}_N^N)u_n^\ell\right] \\
&= u_N^N+(1-u_N^N)u_n^\ell
\end{aligned}$$

Where we have changed variables $x=\frac{u-\tilde{u}_N^N}{1-\tilde{u}_N^N}$, and used again equation A5. Therefore, we let $\Delta U_n^{IB} := U^{IB}(n+1,\bar{n},\mu)-U^{IB}(n,\bar{n},\mu)$ to obtain

$$\Delta U_n^{IB} = \mathbb{E}_\mu\left[(1-u_N^N)\left((1-p_{\bar{n},k})\sum_{\ell=1}^{n-1}\Delta_{n-1}^\ell p_{\bar{n},k}^{\ell-1}+u_n^n(1-p_{\bar{n},k})p_{\bar{n},k}^{n-1}\right)+u_N^N(1-p_{\bar{n},k})p_{\bar{n},k}^n\right]$$

Hence, letting $\Delta U_n := U(n+1,\bar{n},\mu)-U(n,\bar{n},\mu)$, it is direct to see that $\Delta U_n^{IB}\geq\Delta U_n$ if and only if

$$\begin{aligned}
\mathbb{E}_\mu\Bigg[(1-p_{\bar{n},k})\Bigg((1-u_N^N)\sum_{\ell=1}^{n-1}\Delta_{n-1}^\ell p_{\bar{n},k}^{\ell-1}+(1-u_N^N)u_n^n p_{\bar{n},k}^{n-1}+u_N^N p_{\bar{n},k}^n \\
-\sum_{\ell=1}^{n}\Delta_n^\ell p_{\bar{n},k}^{\ell-1}-u_{n+1}^{n+1}p_{\bar{n},k}^n\Bigg)\Bigg]\geq 0 \qquad \text{(A23)}
\end{aligned}$$

To simplify notation, we drop the subindex $\bar{n},k$ from p and note that right hand side of A23 is positive if, for all p

$$(1-u_N^N)\sum_{\ell=1}^{n-1}\Delta_{n-1}^\ell p^{\ell-1}+(1-u_N^N)u_n^n p^{n-1}+u_N^N p^n-\sum_{\ell=1}^{n}\Delta_n^\ell p^{\ell-1}-u_{n+1}^{n+1}p^n\geq 0 \quad \text{(A24)}$$

However,

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

Hence, replacing the order statistics u_n^{ℓ} for the uniform case, condition A24 becomes

$$\begin{aligned}&\frac{(2N-n)(n-1)}{2(n+1)(n+2)(N+1)}p^{n-2}+\frac{N}{N+1}\frac{1}{n+1}p^{n-1}+\frac{1}{N+1}p^{n}\ldots\\ &\qquad\ldots-\frac{n}{(n+1)(n+2)}p^{n-1}-\frac{1}{n+2}p^{n} \quad\geq\quad 0\end{aligned}$$

Multiplying both sides of the above inequality by $\frac{(N+1)(n+2)}{p^{n-2}}$ and arrenging terms, we obtain

$$\frac{(2N-n)(n-1)}{2(n+1)}+p\frac{(2N-n)}{n+1}-p^2(N-(n+1)) \quad\geq\quad 0 \tag{A25}$$

The left-hand side of A25 is a concave function of p, which is positive at $p=0$, and it is increasing for p close to zero. Thus, it is enough to guarantee that condition A25 holds for $p=1$. That is,

$$\begin{aligned}\frac{(2N-n)(n-1)}{2(n+1)}+\frac{(2N-n)}{n+1}-(N-(n+1)) &\geq 0\\ \frac{(2N-n)}{n+1}\left(\frac{n-1}{2}+1\right)-(N-n-1) &\geq 0\\ \frac{1}{2}(2N-n)-(N-n-1)) &\geq 0\\ \frac{n+2}{2} &\geq 0\end{aligned}$$

This concludes the proof.

$\square$

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.

Ajayi, K. and Sidibe, M. (2020). School choice under imperfect information. *Working Paper*.

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.

Arnosti, N., Johari, R., and Kanoria, Y. (2021). Managing congestion in matching markets. *Manufacturing & Service Operations Management*, 23(3):620–636.

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.

Artemov, G. (2021). Assignment mechanisms: Common preferences and information acquisition. *Journal of Economic Theory*, 198:105370.

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.

Bucher, S. and Caplin, A. (2021). Inattention and inequity in school matching. Technical report, National Bureau of Economic Research.

Calsamiglia, C., Haeringer, G., and Klijn, F. (2010). Constrained school choice: An experimental study. *American Economic Review*, 100(4):1860–1874.

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

Che, Y.-K. and Hörner, J. (2018). Recommender systems as mechanisms for social learning. *The Quarterly Journal of Economics*, 133(2):871–925.

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

Cohodes, S., Corcoran, S., Jennings, J., and Sattin-Bajaj, C. (2025). When do informational interventions work? experimental evidence from new york city high school choice. *Educational Evaluation and Policy Analysis*, 47(1):208–236.

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.

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.

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

Hakimov, R., Kübler, D., and Pan, S. (2023). Costly information acquisition in centralized matching markets. *Quantitative Economics*, 14(4):1447–1490.

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.

He, Y. and Magnac, T. (2022). Application costs and congestion in matching markets. *The Economic Journal*, 132(648):2918–2950.

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

Idoux, C. (2023). Integrating new york city schools: The role of admission criteria and family preferences. *Working paper*.

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

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.

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

Li, X. (2005). A note on expected rent in auction theory. *Operations Research Letters*, 33(5):531–534.

Maxey, T. (2024). School choice with costly information acquisition. *Games and Economic Behavior*, 143:248–268.

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

Rogerson, R., Shimer, R., and Wright, R. (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4):959–988.

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

Stigler, G. J. (1961). The economics of information. *Journal of political economy*, 69(3):213–225.

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

Wright, R., Kircher, P., Julien, B., and Guerrieri, V. (2021). Directed search and competitive search equilibrium: A guided tour. *Journal of Economic Literature*, 59(1):90–148.