College admissions markets with combinatorial preferences, constrained applications, and uncertainty

Max Kapur

January 28, 2022

Contents

1 Introduction

2 Equilibrium model

This study considers the efficiency of a college admissions market in which the following three features coincide:

1. Colleges have *combinatorial* preferences over the composition of their entering class.

2. There is *uncertainty* in colleges' preferences; that is, students have only probabilistic information about their admissions prospects at the time of application.

3. Students are *constrained* in the number of schools to which they can apply.

The model contains finite sets $\mathcal{S} = \{1 \dots n\}$ of students and $\mathcal{C} = \{1 \dots m\}$ of colleges. The name of student i is s_i, and the name of college j is c_j.

Feature 1 says that each school has a preference order $\gtrapprox_j$ over the set of admitted-student cohorts. Note that $\gtrsim_j$ orders sets of *admitted* students, with the understanding that only a subset of admits will actually enroll at the college. The college-specific optimization problem that produces $\gtrapprox_j$ is considered exogenous to our model.

Feature 2 arises in real admissions markets because there is asymmetry in the information that students have about themselves and the information that colleges collect about students in the application process. For example, students may have a rough idea of the quality of their personal essay, but the college's evaluation of the same will depend on the biases and mood of the reader. We model this asymmetry by regarding each $\gtrapprox_j$ as a random variable whose space is $2^{\mathcal{S}}!$. Specific realizations of $\gtrapprox_j$ are denoted $\gtrsim_j$. Notice that this definition alone does not impose any constraint on whether it is students who possess complete information about their qualifications and colleges who add measurement noise, or vice-versa. However, the following assumption (which greatly simplifies the subsequent analysis of student utility) implies that the noise is added by the colleges:

Assumption 1. For two colleges c_j and $c_{j'}$, $\gtrapprox_j$ is statistically independent of $\gtrapprox_{j'}$.

Now we can specify the admissions procedure. First, students submit applications to a subset of colleges. Let $\mathcal{X}_i$ denote the set of schools to which s_i applies. Next, each school draws a realization of its preference order $\gtrapprox_j$ and applies it to the set of applicants to determine which students to admit. Thus, when student i applies to c_j, her admissions outcome is determined entirely by two variables: first, the set of applicants with whom she must compete; and second, the realization $\gtrsim_j$ of c_j's preference order. In our model, the distribution of $\gtrapprox$ is exogenized. Therefore, student i's *admissions probability* can be expressed as a function of her peers' application decisions.

Assumption 2 (Existence of $f(\cdot)$)**.** Let $\mathcal{X}_{\setminus i}$ denote the application decisions of students other than s_i. Then

$$f_{ij}(\mathcal{X}_{\setminus i}) = \Pr \begin{bmatrix} s_i \text{ admitted to } c_j \mid \\ s_i \text{ applies to } c_j \text{ and others' application decisions are } \mathcal{X}_{\setminus i} \end{bmatrix} \in [0, 1] \qquad (1)$$

Assumption 3 (c_1 fills capacity)**.**

$$|\{i' \neq i : j \in \mathcal{X}_{i'}\}| < q \implies f_{ij}(\mathcal{X}_{\setminus i}) = 1 \tag{2}$$

Going forward, we interact with the distributions of the random variables $\gtrapprox_j$ primarily via the function $f(\cdot)$. However, it is worth bearing in mind that $f_{\cdot j}$ is a low-dimension projection of c_j's true preference distribution. For example, suppose that $n = 4$, $q = 2$, and when all four students apply to c_j, $f_{\cdot j} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. This could mean that c_j prefers all admitted-student cohorts with equal probability, or it could mean that c_j's preferred cohort is $\{s_1, s_2\}$ with probability $1/2$ and $\{s_3, s_4\}$ with probability $1/2$.

We model feature 3 by allowing each student to apply to only h colleges. Let t_{ij} denote the utility that s_i receives from attending c_j. Assume that students receive zero utility if they do not attend college. Then we regard the expected utility associated with the application portfolio $\mathcal{X}_i$ as the value it provides the student.

Definition 1 (Portfolio valuation function)**.**

$$v_i(\mathcal{X}_i) = \mathrm{E}\left[\max_{j \in \mathcal{X}_i}\{t_{ij} \mid s_i \text{ admitted to } c_j\}\right] \tag{3}$$

The optimal application portfolio for s_i is a set of h schools that maximizes $v_i(\mathcal{X}_i)$. A polynomial-time algorithm for this combinatorial optimization problem is provided in §3.

If there are multiple optimal portfolios, then s_i may elect to choose randomly among them. In fact, s_i may find it strategically advantageous to do so. Thus, in the broadest conception, s_i's decision variable is a probability vector $x_{i\cdot}$ over the $\binom{m}{h}$ possible portfolios. We will refer to $x_{i\cdot}$ as s_i's *application probability vector*, and regard each student's expected portfolio valuation as her utility function.

Definition 2 (Student utility function)**.** The function

$$u_i(x_{i\cdot}) = \sum_{l=1}^{\binom{m}{h}} v_i(\mathcal{X}_l) x_{il} \tag{4}$$

is called student s_i's *utility function*, where l is an index of the possible h-school application portfolios and x_{il} is the probability that s_i applies to the schools in $\mathcal{X}_l$,

Now we are ready to define the market equilibrium.

Definition 3 (Nash equilibrium)**.** The matrix of application probability vectors x, where x_{il} represents the probability that s_i applies to the h-school subset indicated by l, is said to be a (mixed-strategy) *Nash equilibrium* if

$$x_{i\cdot} \in \arg\max_{x_{i\cdot}} \left\{ u_i(x_{i\cdot}) : x_{il} \in [0,1], \sum_{l=1}^{\binom{m}{h}} x_{il} = 1 \right\}, \quad \forall i \in \mathcal{S} \tag{5}$$

If, furthermore, $x_{il} \in \{0, 1\}$ for all i and l, then x is called a *pure-strategy* (Nash) equilibrium.

3 The optimal college application strategy

In this section, we consider the optimal college application strategy for a single student. As Chao (2014) remarked, this represents a somewhat subtle portfolio optimization problem. The traditional Markowitz model trades off the expected value across all assets with a risk term, obtaining a concave maximization problem with linear constraints. But college applicants maximize the observed value of their *best* asset: If a student is admitted to her jth choice, then she is indifferent to whether or not she gets into her $(j+1)$th choice. As a result, the valuation function that students maximize is *convex* in the expected utility associated with individual applications. Risk management is implicit in the college application problem because, in a typical admissions market, college preferability is negatively correlated with competitiveness. Thus, students must negotiate a tradeoff between highly attractive, selective "reach schools" and less preferable "safety schools" where admission is a safer bet. Finally, the combinatorial nature of the college application problem makes it difficult to solve using the gradient-based techniques used in continuous portfolio optimization.

Chao estimated her model (which considers application as a *cost* rather than a constraint) by clustering the schools so that $m = 8$, a scale at which enumeration is possible. However, subject to certain assumptions on the quality of the data available to students in their decision-making process, an optimal application portfolio for a single student can be computed in time polynomial in h and m, as we show presently.

As this section considers a single student's optimization problem, we drop subscripts where appropriate.

3.1 Problem formulation

Consider a college admissions market with m schools, $\mathcal{C} = \{1 \dots m\}$. The jth school is named c_j. By government regulation, students are allowed to apply to no more than h schools. (In the Korean case, $m = 202$ and $h = 3$.) We consider the optimal application strategy for a single student, whom we will call Alma.

For $j = 1 \dots m$, let $t_j > 0$ denote the utility that Alma receives from attending c_j, and let f_j denote the probability that she is admitted if she applies. Let the random variable Z_j equal one if Alma gets into c_j and zero otherwise. We assume that Alma's admissions outcome at each school is independent of her outcome at the other schools.[1] Thus Z is a vector of independent Bernoulli variables with probabilities given by f. To ease notation, let c_0 denote Alma's outcome if she does not get into any college, where $t_0 = 0$ and $f_0 = 1$. Sort the schools so that $t_{j-1} \leq t_j$ for $j = 1 \dots m$.

Let $\mathcal{X}$ denote the set of schools to which Alma applies, called her *application portfolio*, and let x denote the same encoded as a binary vector, where $x_j = \mathbf{1}[j \in \mathcal{X}]$ for $j = 1 \dots m$. The expected utility Alma receives from $\mathcal{X}$ is called the portfolio's *valuation*.

[1]This assumption is appropriate when f gives the admissions probabilities *specifically* for Alma. Recall that in the equilibrium model, the entries of $f = f_i(\mathcal{X}_{\setminus i})$ depend on Alma's index i and the application decisions of students other than Alma.

Definition 4 (Portfolio valuation function)**.**

$$v(\mathcal{X}) = \mathrm{E}\left[\max\{t_j Z_j : j \in \mathcal{X}\}\right] \tag{6}$$

Theorem 1 (Closed form of portfolio valuation function)**.**

$$v(\mathcal{X}) = \sum_{j \in \mathcal{X}} t_j f_j \prod_{\substack{j' \in \mathcal{X}: \\ j' > j}} (1 - f_{j'}) = \sum_{j=1}^{m} x_j t_j f_j \prod_{j'=j+1}^{m} (1 - f_{j'} x_{j'}) \tag{7}$$

where the empty product equals one.

Proof. First consider the case where Alma applies to every school. Let B denote the index of the best school that Alma gets into, where $B = 0$ if she gets in nowhere. Then $B = j$ if and only if Alma is *admitted* to j and *rejected* from any school she prefers to j; that is, any school with higher index.

$$\Pr[B = j \mid \text{apply everywhere}] = f_j(1 - f_{j+1})(1 - f_{j+2})(\cdots)(1 - f_m), \quad j = 1 \ldots m \tag{8}$$

When Alma applies to only a subset of schools, then the product should run over the set of schools that Alma prefers to j *and* actually applies to. Since $(1 - f_j x_j) = 1$ if $x_j = 0$, we can express this by writing

$$\Pr[B = j] = x_j f_j (1 - f_{j+1} x_{j+1})(1 - f_{j+2} x_{j+2})(\cdots)(1 - f_m x_m), \quad j = 1 \ldots m \tag{9}$$

Now computing the expectation

$$v(\mathcal{X}) = \sum_{j=0}^{m} t_j \Pr[B = j] \tag{10}$$

yields the result. □

Let us express the optimization problem as an INLP.

Definition 5 (Alma's problem)**.** The optimal college application portfolio for Alma is given by the solution to the following integer nonlinear program:

$$\begin{aligned} \text{maximize} \quad & \sum_{j=1}^{m} x_j t_j f_j \prod_{j'=j+1}^{m} (1 - f_{j'} x_{j'}) \\ \text{subject to} \quad & \sum_{j=1}^{m} x_j \leq h \\ & x_j \in \{0, 1\}, \quad j = 1 \ldots m \end{aligned} \tag{11}$$

Notice that for a given school c_j, the expected utility associated with applying to c_j is simply $\mathrm{E}[t_j Z_j] = t_j f_j$. It is therefore tempting to adopt the following greedy strategy, which turns out to be inoptimal.

Definition 6 (Greedy algorithm for Alma's problem)**.** Apply to the h schools having the highest expected utility $t_j f_j$.

The basic error of this algorithm is that it maximizes $\mathrm{E}\left[\sum t_j Z_j\right]$ instead of $\mathrm{E}\left[\max\{t_j Z_j\}\right]$. The latter is what Alma is truly concerned with, since in the end she can attend only one school.

Theorem 2. *The greedy algorithm can produce a suboptimal solution.*

Proof. Suppose $m = 3$, $q = 2$, and

$$\begin{aligned} t &= (0.7, 0.8, 0.9) \\ f &= (0.4, 0.4, 0.3) \\ \implies t * f &= (0.28, 0.32, 0.27) \end{aligned}$$

The greedy algorithm picks $\tilde{x} = (1, 1, 0)$ with

$$v(\tilde{x}) = 0.7(0.4)(1 - 0.4) + 0.8(0.4) = 0.488$$

But $x = (0, 1, 1)$ with

$$v(x) = 0.8(0.4)(1 - 0.3) + 0.9(0.3) = 0.494$$

is the optimal solution. □

Hope is not lost. We can still find the optimal solution in $O(h^2 m + m \log m)$ time, as we will now show.

3.2 Solution

The following theorem states that the solution to Alma's problem possesses a special structure: An optimal portfolio of size $h + 1$ includes an optimal portfolio of size h as a subset.

Theorem 3 (Nestedness of optimal application portfolios)**.** *Let $\mathcal{X}_h$ denote Alma's optimal application portfolio when the application limit is h. If each $\mathcal{X}_h$ is unique,*

$$\mathcal{X}_1 \subset \mathcal{X}_2 \subset \cdots \subset \mathcal{X}_m \tag{12}$$

If the optimal portfolios are not unique, then there is a sequence of optimal portfolios satisfying the above.

Proof. The proof is a nested induction. First, we show that $\mathcal{X}_1 \subset \mathcal{X}_2$ (step 1) and $\mathcal{X}_1 \subset \mathcal{X}_{h-1} \implies \mathcal{X}_1 \subset \mathcal{X}_h$ for $h = 3 \dots m$ (step 2). We then assume that $\mathcal{X}_1 \subset \cdots \subset \mathcal{X}_{h-1}$, and show that $\mathcal{X}_{h-n} \subset \mathcal{X}_h \implies \mathcal{X}_{h-n+1} \subset \mathcal{X}_h$ for $h = 3 \dots m$ and $n = 2 \dots (h-1)$ (step 3). It follows that $\mathcal{X}_{h-1} \subset \mathcal{X}_h$ for $h = 2 \dots m$, which completes the proof.

(Step 1.) To get a contradiction, suppose that the unique optima are $\mathcal{X}_1 = \{j\}$ and $\mathcal{X}_2 = \{k, l\}$, where we may assume that $t_k \leq t_l$. Optimality requires that

$$v(\mathcal{X}_1) = f_j t_j > v(\{k\}) = f_k t_k \tag{13}$$

and

$$
\begin{aligned}
v(\mathcal{X}_2) = f_k(1-f_l)t_k + f_l t_l > v(\{j,l\}) && (14)\\
= f_j(1-f_l)t_j + (1-f_j)f_l t_l + f_j f_l \max\{t_j, t_l\} && (15)\\
\geq f_j(1-f_l)t_j + (1-f_j)f_l t_l + f_j f_l t_l && (16)\\
= f_j(1-f_l)t_j + f_l t_l && (17)\\
> f_k(1-f_l)t_k + f_l t_l = v(\mathcal{X}_2) && (18)
\end{aligned}
$$

a contradiction.

(Step 2.) Label the schools in $\mathcal{X}_h$ as $[1]\dots[h]$, where $t_{[1]} \leq \cdots \leq t_{[h]}$. Then

$$v(\mathcal{X}_h) = (1 - f_{[h]})v(\mathcal{X}_h \setminus [h]) + t_{[h]}f_{[h]} \tag{19}$$

If $[h] = j$, then this step is complete. If not, then $\{[1]\dots[h-1]\}$ must form an optimal $(h-1)$-portfolio over $\mathcal{C} \setminus [h]$. By assumption, this portfolio includes the optimal 1-portfolio over $\mathcal{C} \setminus [h]$, namely $\mathcal{X}_1 = \{j\}$.

(Step 3.) Label the schools in $\mathcal{X}_h \setminus \mathcal{X}_{h-n}$ as $(1)\dots(n)$, where $t_{(1)} \leq \cdots \leq t_{(n)}$. If $(n) \in \mathcal{X}_{h-n+1}$, then the inductive hypothesis implies $\mathcal{X}_{h-n+1} = \mathcal{X}_{h-n} \cup \{(n)\}$ and there is nothing left to show. If not, then write

$$v(\mathcal{X}_h) = (1 - f_{(n)})v\big(\mathcal{X}_{h-n} \cup \{(1)\dots(n-1)\}\big) + f_{(n)}\,\mathrm{E}\big[\max\{X_{h-n}, t_{(n)}\}\big] \tag{20}$$

where the random variable X_{h-n} denotes the value achieved by the portfolio $\mathcal{X}_{h-n}$. Once $\mathcal{X}_{h-n}$ and (n) are fixed, the remaining schools must maximize $v(\mathcal{X}_{h-n} \cup \{(1)\dots(n-1)\})$. That is, they form an optimal $(h-1)$-portfolio, call it $\mathcal{Y}_{h-1}$, over $\mathcal{C} \setminus (n)$. By the inductive hypothesis, $\mathcal{Y}_{h-1}$ is a superset of the optimal $(h-n+1)$-portfolio, call it $\mathcal{Y}_{h-n+1}$, over $\mathcal{C} \setminus (n)$. But $\mathcal{X}_{h-n+1}$ is optimal over *all* of $\mathcal{C}$ and does not include (n), so we can take $\mathcal{Y}_{h-n+1} = \mathcal{X}_{h-n+1}$, which completes this step. □

Applying the result above yields an efficient algorithm for the optimal portfolio: Start with the empty set and add schools one at a time, maximizing $v(\mathcal{X} \cup \{k\})$ at each addition.

Algorithm 1: Dynamic programming algorithm for Alma's problem.

Data: Utility values $t \in [0,\infty)^m$, admissions probabilities $f \in [0,1]^m$.
Reindex schools in ascending order of t;
$\mathcal{X} \leftarrow \emptyset$;
for $i = 1 \dots h$ **do**
  $j \leftarrow \arg\max_k \big\{v(\mathcal{X} \cup \{k\}) : k \notin \mathcal{X}\big\}$;
  $\mathcal{X} \leftarrow \mathcal{X} \cup \{j\})$;
end
return $\mathcal{X}$

Theorem 4 (Validity of Algorithm 1)**.** *Algorithm 1 produces an optimal application portfolio in $O(h^2 m + m \log m)$ time.*

Proof. Optimality follows from Theorem 3. At each of the h iterations, there are $m - i +$

1 candidates for k which must be inspected. Computing $v(\cdot)$ is $\Theta(h)$ using (7). Adding the $O(m \log m)$ time required to sort the entries of t yields the result. $\square$

The computation time can be further reduced by observing that the school added at each iteration must be maximal in the tuple (f_j, t_j). That is, if $f_j \geq f_k \wedge t_j \geq t_k$ holds for two candidate schools j and k, then k cannot be the uniquely optimal addition, since admission to j is both weakly preferable and weakly more probable. Therefore, we may restrict our attention to the subset of schools *not* dominated by another in this way.

Proposition 1 (Efficient frontier for Alma's problem)**.** Given a set of schools $\mathcal{D}$, define the set

$$\mathrm{EF}(\mathcal{D}) = \{j \in \mathcal{D} : f_j \geq f_k \vee t_j \geq t_k, \forall k \in \mathcal{D}\} \tag{21}$$

as the *efficient frontier* of $\mathcal{D}$. Then

$$l \in \mathrm{EF}(\mathcal{C} \setminus \mathcal{X}_{h-1}) \tag{22}$$

where l is the school added at iteration h of Algorithm 1.

3.3 Additional theoretical results

The nestedness property implies that Alma's expected utility is a concave function of h.

Theorem 5 (Optimal portfolio valuation concave in h)**.** *For* $h = 2 \dots (m-1)$,

$$v(\mathcal{X}_h) - v(\mathcal{X}_{h-1}) \geq v(\mathcal{X}_{h+1}) - v(\mathcal{X}_h) \tag{23}$$

Proof. We will prove the equivalent expression $2v(\mathcal{X}_h) \geq v(\mathcal{X}_{h+1}) + v(\mathcal{X}_{h-1})$. Applying Theorem 3, we write $\mathcal{X}_h = \mathcal{X}_{h-1} \cup \{j\}$ and $\mathcal{X}_{h+1} = \mathcal{X}_{h-1} \cup \{j, k\}$. Define the random variables X_i as above. If $t_k \leq t_j$, then

$$\begin{aligned}
2v(\mathcal{X}_h) &= v(\mathcal{X}_{h-1} \cup \{j\}) + v(\mathcal{X}_{h-1} \cup \{j\}) && (24)\\
&\geq v(\mathcal{X}_{h-1} \cup \{k\}) + v(\mathcal{X}_{h-1} \cup \{j\}) && (25)\\
&= v(\mathcal{X}_{h-1} \cup \{k\}) + (1 - f_j) v(\mathcal{X}_{h-1}) + f_j \, \mathrm{E}[\max\{t_j, X_{h-1}\}] && (26)\\
&= v(\mathcal{X}_{h-1} \cup \{k\}) - f_j v(\mathcal{X}_{h-1}) + f_j \, \mathrm{E}[\max\{t_j, X_{h-1}\}] + v(\mathcal{X}_{h-1}) && (27)\\
&\geq v(\mathcal{X}_{h-1} \cup \{k\}) - f_j v(\mathcal{X}_{h-1} \cup \{k\}) + f_j \, \mathrm{E}[\max\{t_j, X_{h-1}\}] + v(\mathcal{X}_{h-1}) && (28)\\
&= (1 - f_j) v(\mathcal{X}_{h-1} \cup \{k\}) + f_j \, \mathrm{E}[\max\{t_j, X_{h-1}\}] + v(\mathcal{X}_{h-1}) && (29)\\
&= v(\mathcal{X}_{h-1} \cup \{j, k\}) + v(\mathcal{X}_{h-1}) && (30)\\
&= v(\mathcal{X}_{h+1}) + v(\mathcal{X}_{h-1}) && (31)
\end{aligned}$$

The first inequality follows from the optimality of $\mathcal{X}_h$, while the second follows from the fact that adding k to $\mathcal{X}_{h-1}$ can only increase its valuation.

If $t_k \geq t_j$, then the steps are analogous:

$$\begin{aligned}
2v(\mathcal{X}_h) &= v(\mathcal{X}_{h-1} \cup \{j\}) + v(\mathcal{X}_{h-1} \cup \{j\}) && (32)\\
&\geq v(\mathcal{X}_{h-1} \cup \{k\}) + v(\mathcal{X}_{h-1} \cup \{j\}) && (33)\\
&= (1 - f_k) v(\mathcal{X}_{h-1}) + f_k \operatorname{E}[\max\{t_k, X_{h-1}\}] + v(\mathcal{X}_{h-1} \cup \{j\}) && (34)\\
&= v(\mathcal{X}_{h-1}) - f_k v(\mathcal{X}_{h-1}) + f_k \operatorname{E}[\max\{t_k, X_{h-1}\}] + v(\mathcal{X}_{h-1} \cup \{j\}) && (35)\\
&\geq v(\mathcal{X}_{h-1}) - f_k v(\mathcal{X}_{h-1} \cup \{j\}) + f_k \operatorname{E}[\max\{t_k, X_{h-1}\}] + v(\mathcal{X}_{h-1} \cup \{j\}) && (36)\\
&= v(\mathcal{X}_{h-1}) + (1 - f_k) v(\mathcal{X}_{h-1} \cup \{j\}) + f_k \operatorname{E}[\max\{t_k, X_{h-1}\}] && (37)\\
&= v(\mathcal{X}_{h-1}) + v(\mathcal{X}_{h-1} \cup \{j, k\}) && (38)\\
&= v(\mathcal{X}_{h-1}) + v(\mathcal{X}_{h+1}) && \square
\end{aligned}$$

To wrap up, we show that the nestedness property depends on the assumption of independent admissions by providing a simple counterexample.

Theorem 6. *If the entries of Z are dependent, then the optimal solution may violate the nestedness property of Theorem 3.*

Proof. Let $t = (3, 3, 4)$, $Z_1 \sim \operatorname{Bernoulli}(0.5)$, $Z_2 = 1 - Z_1$, and $Z_3 \sim \operatorname{Bernoulli}(0.5)$. Then it is easy to verify that the unique optimal portfolios are $\mathcal{X}_1 = \{3\}$ and $\mathcal{X}_2 = \{1, 2\}$. □

4 Two-school model

The market described above is intractably complex. However, we argue that we can maintain its most important features while restricting our attention to a tractable, stylized market with $m = 2$ and $h = 1$.

The coexistence of an application limit and uncertainty in college preferences requires students to strategize in selecting the set of schools to which they apply. We can think of the student's application decision of consisting of two parts: First, she must rank the colleges by preferability and identify her admissions probability at each. Second, she must allocate her limited applications across the set of schools in the market.

For a typical student, college preferability and admissions probability are negatively correlated. Thus, the second stage of the application decision boils down to trading off schools that are desirable but hard to get into (reach schools) with schools that are less desirable but easy to get into (safety schools). The optimal allocation between reach schools and safety schools depends on the individual student's tolerance for risk.

4.1 Market participants

To highlight the essential nature of the tradeoff between reach schools and safety schools, we consider a stylized admissions market with two schools, c_1 and c_2. c_1 is a competitive university, whereas c_2 represents the safety school, which admits any applicant. Students are allowed to apply to only one school. Every student prefers attending c_1 to c_2, and c_2 to nonattendance, but students differ in the strength of these preferences as well as in their admissions probabilities at c_1, as detailed below.

Let $\mathcal{S} = \{1 \dots n\}$ denote the set of *students*, and let the natural number $q < n$ denote c_1's *capacity*. c_1 has an ordinal preference order over the set of possible entering classes comprised of q students. This preference order is a random variable $\underset{\approx}{\succsim}$ whose space is $\{\mathcal{T} \subseteq \mathcal{S} : |\mathcal{T}| = q\}!$. Specific realizations of $\underset{\approx}{\succsim}$ are denoted $\underset{\sim}{\succ}$. The safety school, c_2, admits every applicant.

4.2 The competitive admissions process

Let $x \in [0, 1]^n$ denote the *application probability vector*, where x_i is the probability that student i applies to c_1 and $1 - x_i$ is the probability that i applies to c_2. If $x \in \{0, 1\}^n$, then it is called a (deterministic) *application vector*.

If more than q students apply to c_1, then c_1 draws a realization of its preference order $\underset{\approx}{\succsim}$ and applies it to x to determine the entering class. If q or fewer students apply to c_1, all are admitted. Thus, when student i applies to c_1, her admissions outcome is determined entirely by two parameters: first, the set of applicants with whom she must compete; and second, the realization $\underset{\sim}{\succ}$ of c_1's preference order. In our model, the distribution of $\underset{\approx}{\succsim}$ is regarded as an exogenous variable. Therefore, student i's *admissions probability* can be expressed as a function of her peers' application decisions.

Assumption 4 (Existence of $f(\cdot)$)**.** Let $x_{\setminus i}$ denote the application decisions of students other than i. Then

$$f_i(x_{\setminus i}) = \Pr \left[\begin{array}{l} i \text{ admitted to } c_1 \mid \\ i \text{ applies to } c_1 \text{ and others' application probabilities are } x_{\setminus i} \end{array} \right] \in [0, 1] \tag{39}$$

We will use the vector $f(x)$ to denote the concatenation of these probabilities, with the understanding that the ith entry of $f(x)$ does not depend on x_i.

Assumption 5 (c_1 fills capacity)**.** $x \in \{0, 1\}^n$ and $\sum_{j \neq i} x_j < q \implies f_i(x_{\setminus i}) = 1$.

This assumption need not hold for mixed strategies: If each student applies with near-zero probability, say ϵ, then there is an ϵ^n chance that *every* student applies, and some rejections must occur.

It is convenient to define the random variable Y as the admissions vector that arises when *every* student applies to c_1; that is, c_1's "global optimum." A given realization $\underset{\sim}{\succ}$ of c_1's preference order induces a realization y of Y.

Definition 7 (Expected ideal entering class)**.** Let

$$\mathcal{Y} = \underset{\mathcal{T} \subseteq \mathcal{S}}{\arg\max} \{\underset{\approx}{\succsim} : |\mathcal{T}| = q\} \tag{40}$$

denote the set of students c_1 admits when all students apply, and let $Y_i = \mathbf{1}[i \in \mathcal{Y}]$ denote the same encoded as a binary vector. Then the expectation of Y is denoted

$$\bar{y} = f(\mathbf{1}) \tag{41}$$

and called the *expected ideal entering class*.

Going forward, we will interact with $\gtrapprox$ primarily via the function $f(\cdot)$ and the statistic $\bar{y}$. The following fact is helpful.

Theorem 7. *$x \in \{0,1\}^n$ and $\sum x_i \geq q \implies x \cdot f(x) = q$.*

Proof. When x is fixed, the quantity $x_i f_i(x_{\setminus i})$ represents the probability that student i will attend c_1. Thus, summing over i yields the expected size of c_1's entering class. By assumption, this is always q. □

As above, this result may not hold for mixed strategies.

Corollary 1. *$\sum \bar{y}_i = q$.*

4.3 Student preferences

Each student receives one unit of utility if she attends c_1, $t_i \in (0,1)$ units of utility if she attends c_2, and zero units of utility if she is unable to enroll in college this season (that is, if she applies to c_1 and is rejected). We regard student i's expected utility

$$u_i(x) = x_i\big(f_i(x_{\setminus i})\big) + (1 - x_i)t_i \tag{42}$$

as her utility function. t_i is called student i's *risk aversion parameter*, as explained below.

4.4 Notion of equilibrium

Each student seeks to maximize her utility. The market reaches equilibrium when no student, acting alone, can increase her expected payoff by changing her application strategy.

Definition 8 (Nash equilibrium)**.** The application probability vector x, where $x_i \in [0,1]$ represents the probability that student i applies to c_1 instead of c_2, is said to be a (mixed-strategy) *Nash equilibrium* if

$$x_i \in \arg\max_{x_i} \{u_i(x) : x_i \in [0,1]\}, \quad i = 1 \ldots n \tag{43}$$

If, furthermore, $x_i \in \{0,1\}$ for all i, then x is called a *pure-strategy* (Nash) equilibrium.

Behavioral economics research tells us that humans often make decisions in terms of risk-mitigating heuristics rather than explicit payoff functions. The notion of equilibrium defined above admits an alternative interpretation in which t_i is a parameter that represents student i's risk aversion. In particular, suppose that each student resolves to apply to c_1 only if her probability of admission is at least t_i. Then we can define an equilibrium as an admissions vector in which students' stated risk preferences accord with their actual application decisions.

Definition 9 (Risk equilibrium)**.** The application vector x is said to be a *risk equilibrium* if and only if $x_i \in \{0,1\}$ and

$$x_i = 1 \quad \iff \quad f_i(x_{\setminus i}) \geq t_i, \quad i = 1 \ldots n \tag{44}$$

Risk equilibria and Nash equilibria are related by the following theorem.

Theorem 8 (Risk equilibria and Nash equilibria)**.** *Let $\mathcal{X}_r$, $\mathcal{X}_p$, and $\mathcal{X}_n$ denote the sets of risk, pure-strategy, and Nash equilibria for a given market. Then $\mathcal{X}_r \subseteq \mathcal{X}_p \subseteq \mathcal{X}_n$.*

Proof. The result follows immediately from the fact that $u_i(x)$ is linear in x_i. □

Corollary 2. *If x is a Nash equilibrium and $f_i(x_{\setminus i}) \neq t_i$ for all i, then $x \in \{0,1\}^n$, and x is also a pure-strategy equilibrium and a risk equilibrium.*

This study will concern itself primarily with risk equilibria. In a mixed-strategy equilibrium, there is a chance that zero students apply to c_1, which makes some of our efficiency measures undefined.

4.5 Size of equilibrium

Definition 10 (Size of x)**.** The number of applicants $k(x) = \sum x_i$ is referred to as the *size* of the application vector x.

We will write simply k when the value of x is clear from context.

Theorem 9 (Bounds on equilibrium size)**.** *The size of any pure-strategy equilibrium x is bound by*

$$q \leq k(x) \leq \frac{q}{\min(t_i)} \tag{45}$$

Proof. The lower bound is from Assumption 5. For the upper bound, fix x and notice that among the set of applicants, the *average* admissions probability is q/k. If this value is less than $\min(t_i)$, then there must be at least applicant whose admissions probability is below t_i; thus x is not an equilibrium. □

Neither bound is necessary for mixed-strategy equilibria. Consider a market in which the $\gtrapprox$ is distributed uniformly on its support; that is, c_1 picks a random subset of applicants with uniform probability. For all i, let $x_i = \chi \in (0,1)$; then $f_i(x_{\setminus i})$ is some common constant $\varphi \in (0,1)$, and picking $t_i = \phi$ makes x an equilibrium. χ, and therefore $k(x)$, can be made arbitrarily small. To violate the upper bound,

4.6 Measures of efficiency

In our model, the level of utility experienced by the students is incommensurate with the utility experienced by the schools. Thus, there is no single index akin to market surplus that can capture the overall efficiency of the admissions process. Instead, we propose separate measures which serve as indices of fairness, school utility, and student welfare. We will first define these measures under the assumption that x is a binary vector, then discuss the extension to mixed strategies in §4.6.4.

4.6.1 Fairness

We consider two notions of fairness, both derived from the dot product $x \cdot \bar{y}$.

Definition 11 (Stability index)**.** The statistic

$$\bar{S}(x) = \frac{x \cdot \bar{y}}{k(x)} \tag{46}$$

is called the *stability index* of the application vector x.

For any equilibrium, since $x \in [0,1]^n$, $\sum \bar{y}_i = q$, and $k \geq q$, we have $\bar{S}(x) \in [0,1]$.

We interpret the stability index as follows: The uncertain nature of the admissions process means that even in equilibrium, there are some market realizations in which the students who attend c_2 are more qualified than those who attend c_1. The stability index captures the equilibrium's robustness to the envy that arises from these mismatches.

To see this, consider a realization $\succsim$ of c_1's preference order, and let y denote the associated realization of Y. Let $\mathcal{B} = \{i : x_i = 0 \wedge y_i = 1\}$ denote the set of students who applied to c_2 but appear in y. These students form a blocking coalition for the outcome induced by x and $\succsim$. That is, if the students in $\mathcal{B}$ collectively decide to apply to c_1 alongside the students already in x, they will surely be admitted.[2]

Let $\mathcal{A}_+ = \{i : x_i = 1 \wedge y_i = 1\}$ and $\mathcal{A}_- = \{i : x_i = 1 \wedge y_i = 0\}$. When the coalition $\mathcal{B}$ mobilizes, any student in $\mathcal{A}_+$ receives a weakly better outcome: Either she goes from being rejected from c_1 to being admitted, or she is admitted in both cases. On the other hand, any student in $\mathcal{A}_-$ receives a weakly worse outcome. The quantity

$$S(x) = \frac{|\mathcal{A}_+|}{|\mathcal{A}_-| + |\mathcal{A}_+|} = \frac{x \cdot y}{k} \tag{47}$$

represents the proportion of applicants who are "safe" from disruption by $\mathcal{B}$, and $\bar{S}(x)$ represents the expectation thereof.

The second measure of fairness is as follows.

Definition 12 (Alignment index)**.** The statistic

$$\bar{T}(x) = \frac{x \cdot \bar{y}}{q} \tag{48}$$

is called the *alignment index* of the application vector x.

By Corollary 1 and the fact that $x \in [0,1]^n$, we have $\bar{T}(x) \in [0,1]$.

The alignment index captures the intuitive notion that "the best students go to the best school." The dot product $x \cdot \bar{y}$ represents, in expectation, the degree of overlap between c_1's set of applicants x and its ideal entering class $\bar{y}$. Since the entries of $\bar{y}$ are nonnegative and $x \in [0,1]^n$, this quantity is maximized when $x = \mathbf{1}$, yielding $\mathbf{1} \cdot \bar{y} = q$. Thus, $\bar{T}(x)$ represents

[2] Note that for the students in $\mathcal{B}$, coordination is not a significant challenge. Once the students in $\mathcal{B}$ have agreed to mobilize, any single student who fails to comply does so at her own expense. One may nonetheless regard the expected number of students in $\mathcal{B}$ as a meaningful efficiency property. This quantity is $(\mathbf{1} - x) \cdot \bar{y} = q - x \cdot \bar{y}$, which is perfectly inversely correlated with the alignment index $\bar{T}(x)$ defined below.

the extent to which c_1 is able to approximate its ideal entering class using only the students in x.

The stability and alignment indices differ only slightly in form, but they capture distinct notions of fairness. As our computational results will show, the two measures are sometimes in tension.

4.6.2 School utility

The discussion above implies that the alignment index is a heuristic indicator of c_1's utility. In the computational experiments, we will construct more precise indicators by using a utility function to induce the distribution of $\gtrapprox$.

As for c_2, the assumption that c_2 admits every applicant implies that c_2's utility is increasing in the number of students in its entering class, namely $n - k$.

4.6.3 Student welfare

The following definition of student welfare is simply the sum of student utility after applying Theorem 7.

Definition 13 (Aggregate student welfare)**.** The sum of students' utility functions

$$\bar{U}(x) = \sum_{i=1}^{n} u_i(x) = q + (\mathbf{1} - x) \cdot t \tag{49}$$

is called the *aggregate student welfare* of the application vector x.

If we choose to interpret t_i as a risk parameter rather than a utility valuation, then the measure above is inappropriate. An alternative measure of student disutility is the number of students who fail to enroll in either school, that is, $k - q$, which depends only on the size of the equilibrium. In a given market, if students have equal risk aversion t_i, then (49) depends only on the size of the equilibrium as well. Thus, both $\bar{U}(x)$ and the size criterion order the market's equilibria in the same way.

4.6.4 Summary

Given a market and one of its equilibria x, we regard the stability index $\bar{S}(x)$, the alignment index $\bar{T}(x)$, the aggregate student welfare $\bar{U}(x)$, and the size of the equilibrium k as relevant economic indicators.

If x is a binary vector—that is, a pure-strategy equilibrium—then these are deterministic measures. On the other hand, if x is a vector of mixed strategies, then each student's application decision is the random Bernoulli variable X_i which equals 1 with probability x_i. We may assume that the entries of X_i are independent of one another and of $\gtrapprox$ (and therefore of Y). It follows from the linearity of expectations that $\mathrm{E}[\bar{T}(X)] = \bar{T}(x)$, $\mathrm{E}[\bar{U}(X)] = \bar{U}(x)$, and $\mathrm{E}[k(X)] = k(x)$. $\bar{S}(x)$ is the exception: When students play mixed strategies, there is a small chance that $k(X) = 0$, rendering the expectation of $\bar{S}(X)$ undefined. One option is to redefine $\bar{S}(x)$ to equal zero in the case that $k < q$, with the understanding that an application vector smaller than

q is maximally vulnerable to the blocking coalition, then compute the expectation. Another is to simply report $\bar{S}(x)$ according to the function's definition. To avoid imposing this interpretive choice, in the computational results that follow, we consider only pure strategies.

5 Computational results

5.1 Additive school utility with homogeneous risk aversion

5.1.1 Existence of sorted equilibrium

5.1.2 Bounds on the efficiency measures

5.1.3 Sorted equilibrium algorithm

5.1.4 Computational experiments

5.2 Combinatorial admissions with uncertainty from QP

5.2.1 Model

5.2.2 Computational experiments

6 References

Fu, Chao. 2014. "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market." *Journal of Political Economy* 122, no. 2: 225–81. https://doi.org/10.1086/675503.
Fu, Chao. 2014. "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market." *Journal of Political Economy* 122, no. 2: 225–81. https://doi.org/10.1086/675503.