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

Max Kapur

February 24, 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 $\succapprox_j$ over the set of admitted-student cohorts. Note that $\succsim_j$ 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 $\succapprox_j$ is considered exogenous to our model. This assumption is reasonable because in this model, students' preferences are deterministic.

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 $\succapprox_j$ as a random variable whose space is $2^{\mathcal{S}}!$. Specific realizations of $\succapprox_j$ are denoted $\succsim_j$. 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 $c_{j'}$, $\succapprox_j$ is statistically independent of $\succapprox_{j'}$.

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 $\succapprox_j$ 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 $\succsim_j$ of c_j's preference order. In our model, the distribution of $\succapprox$ 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

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

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)**.**

$$v_i(\mathcal{X}_i) = \mathrm{E}\left[\max_{j \in \mathcal{X}_i}\{t_{ij} \mid s_i \text{ admitted to } c_j\}\right] \tag{2}$$

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

$$u_i(x_{i\cdot}) = \sum_{l=1}^{\binom{m}{h}} v_i(\mathcal{X}_l) x_{il} \tag{3}$$

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

$$x_{i\cdot} \in \arg\max_{x_{i\cdot}} \left\{ u_i(x_{i\cdot}) : x_{il} \in [0,1], \sum_{l=1}^{\binom{m}{h}} x_{il} = 1 \right\}, \quad \forall i \in \mathcal{S} \tag{4}$$

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. Let c_0 denote Alma's outcome if she does not get into any college, with utility t_0 and $f_0 = 1$. Sort the schools so that $t_{j-1} \leq t_j$ for $j = 1 \dots m$.[2]

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 $x_j = 1 \iff j \in \mathcal{X}$ for $j = 1 \dots m$. The expected utility Alma receives from $\mathcal{X}$ is called the portfolio's *valuation*.

Definition 4 (Portfolio valuation function)**.** $v(\mathcal{X}) = \mathrm{E}\left[\max\{t_j Z_j : j \in \mathcal{X}\}\right]$.

It is helpful to define the random variable $X = \max\{t_j Z_j : j \in \mathcal{X}\}$ as the utility achieved by the portfolio, so that $v(\mathcal{X}) = \mathrm{E}[X]$.

[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. Once these factors are accounted for, the independence of the Z_j is a direct consequence of Assumption 1.

[2]It is without loss of generality to assume that $t_0 \leq t_1$ because schools for which $t_j < t_0$ can be trivially excluded from consideration.

Let $p_j(\mathcal{X})$ denote the probability that Alma attends c_j. Alma attends c_j if and only if she *applies* to c_j, is *admitted* to c_j, and is *rejected* from any school she prefers to c_j; that is, any school with higher index. Hence, for $j = 0 \dots m$,

$$p_j(\mathcal{X}) = \begin{cases} f_j \prod_{\substack{j' \in \mathcal{X}: \\ j' > j}} (1 - f_{j'}), & j \in \{0\} \cup \mathcal{X} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

$$\iff \quad p_j(x) = x_j f_j \prod_{j'=j+1}^{m} (1 - f_{j'} x_{j'}) \tag{6}$$

with the understanding that $x_0 = 1$ and the empty product equals one. The following proposition follows immediately.

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

$$v(\mathcal{X}) = \sum_{j \in \{0\} \cup \mathcal{X}} \Big(t_j f_j \prod_{\substack{j' \in \mathcal{X}: \\ j' > j}} (1 - f_{j'}) \Big), \quad \textit{or equivalently,} \tag{7}$$

$$v(x) = t_0 \prod_{j=1}^{m} (1 - f_j x_j) + \sum_{j=1}^{m} \Big(x_j t_j f_j \prod_{j'=j+1}^{m} (1 - f_{j'} x_{j'}) \Big) \tag{8}$$

Next, we show that without loss of generality, we may assume that $t_0 = 0$ (or any constant).

Theorem 1. *Let $\bar{t}_j = t_j - \gamma$ for $j = 0 \dots m$. Then $v(\mathcal{X}; \bar{t}_j) = v(\mathcal{X}; t_j) - \gamma$ regardless of $\mathcal{X}$.*

Proof. By definition, $\sum_{j=0}^{m} p_j(\mathcal{X}) = \sum_{j \in \{0\} \cup \mathcal{X}} p_j(\mathcal{X}) = 1$. Therefore

$$v(\mathcal{X}; \bar{t}_j) = \sum_{j \in \{0\} \cup \mathcal{X}} \bar{t}_j p_j(\mathcal{X}) = \sum_{j \in \{0\} \cup \mathcal{X}} (t_j - \gamma) p_j(\mathcal{X}) \tag{9}$$

$$= \sum_{j \in \{0\} \cup \mathcal{X}} t_j p_j(\mathcal{X}) - \gamma = v(\mathcal{X}; t_j) - \gamma \tag{10}$$

which completes the proof. □

Let us express the optimal college application problem as an INLP.

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

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

3.2 Naïve solution

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 (Naïve 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.

Example 1. Suppose $m = 3$, $q = 2$, and

$$\begin{aligned} t &= (70, 80, 90) \\ f &= (0.4, 0.4, 0.3) \\ \implies t * f &= (28, 32, 27) \end{aligned}$$

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

$$v(\tilde{x}) = 70(0.4)(1 - 0.4) + 80(0.4) = 48.8$$

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

$$v(x) = 80(0.4)(1 - 0.3) + 90(0.3) = 49.4$$

is the optimal solution. This shows that the naïve algorithm can produce a suboptimal solution.

In fact, the naïve algorithm is a $(1/h)$-approximation algorithm for Alma's problem, as expressed in the following theorem.

Theorem 2. *When the application limit is h, let $\mathcal{X}_h$ denote the optimal portfolio, and $\mathcal{T}_h$ the set of the h schools having the largest values of $t_j f_j$. Then $v(\mathcal{T}_h)/v(\mathcal{X}_h) \geq 1/h$.*

Proof. Because $\mathcal{T}_h$ maximizes the quantity $\mathrm{E}\big[\sum_{j\in\mathcal{T}_h}\{t_j Z_j\}\big]$, we have

$$\begin{aligned} v(\mathcal{X}_h) = \mathrm{E}\big[\max_{j\in\mathcal{X}_h}\{t_j Z_j\}\big] &\leq \mathrm{E}\big[\sum_{j\in\mathcal{X}_h}\{t_j Z_j\}\big] \leq \mathrm{E}\big[\sum_{j\in\mathcal{T}_h}\{t_j Z_j\}\big] \\ &= h\,\mathrm{E}\big[\tfrac{1}{h}\sum_{j\in\mathcal{T}_h}\{t_j Z_j\}\big] \leq h\,\mathrm{E}\big[\max_{j\in\mathcal{T}_h}\{t_j Z_j\}\big] = h\,v(\mathcal{T}_h) \end{aligned} \tag{12}$$

where the final inequality follows from the concavity of the $\max\{\}$ operator. □

The following example establishes the tightness of the approximation factor.

Example 2. Pick any h and let $m = 2h$. For a small constant $\varepsilon \in (0, 1)$, let

$$t = \Big(\underbrace{1, \ldots, 1}_{h}, \underbrace{\varepsilon^{-1}, \varepsilon^{-2}, \ldots, \varepsilon^{-(h-1)}, \varepsilon^{-h}}_{h}\Big)$$

$$\text{and} \quad f = \Big(\underbrace{1, \ldots, 1}_{h}, \underbrace{\varepsilon^{1}, \varepsilon^{2}, \ldots, \varepsilon^{h-1}, \varepsilon^{h}}_{h}\Big)$$

Since all $t_j f_j = 1$, the naïve algorithm can choose $\mathcal{T}_h = \{1, \dots, h\}$, with $v(\mathcal{T}_h) = 1$. But optimal solution is $\mathcal{X}_h = \{h+1, \dots, m\}$, with

$$v(\mathcal{X}_h) = \sum_{j=h+1}^{m} \Big(t_j f_j \prod_{j'=j+1}^{m} (1 - f_{j'}) \Big) = \sum_{j=1}^{h} (1-\varepsilon)^j \approx h$$

Thus, as ϵ approaches zero, we have $v(\mathcal{T}_h)/v(\mathcal{X}_h) \to 1/h$. (The optimality of $\mathcal{X}_h$ follows from the fact that it achieves the upper bound of Theorem 12.)

Hope is not lost. We can still find the optimal solution in time polynomial in h and m, as we will now show.

3.3 Solution

It turns out 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, then*

$$\mathcal{X}_1 \subset \mathcal{X}_2 \subset \cdots \subset \mathcal{X}_m \tag{13}$$

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

Proof. By induction on h. Applying Theorem 1, we assume that $t_0 = 0$.

(Base case.) First, we will show that $\mathcal{X}_1 \subset \mathcal{X}_2$. To get a contradiction, suppose that the optima are $\mathcal{X}_1 = \{j\}$ and $\mathcal{X}_2 = \{k, l\}$, where we may assume that $t_k \le t_l$. Optimality requires that

$$v(\mathcal{X}_1) = f_j t_j > v(\{k\}) = f_k t_k \tag{14}$$

and

$$\begin{aligned} v(\mathcal{X}_2) = f_k(1-f_l)t_k + f_l t_l > v(\{j,l\}) & \qquad (15)\\ = f_j(1-f_l)t_j + (1-f_j)f_l t_l + f_j f_l \max\{t_j, t_l\} & \qquad (16)\\ \ge f_j(1-f_l)t_j + (1-f_j)f_l t_l + f_j f_l t_l & \qquad (17)\\ = f_j(1-f_l)t_j + f_l t_l & \qquad (18)\\ \ge f_k(1-f_l)t_k + f_l t_l = v(\mathcal{X}_2) & \qquad (19) \end{aligned}$$

which is a contradiction.

(Inductive step.) Assume that $\mathcal{X}_1 \subset \cdots \subset \mathcal{X}_h$, and we will show $\mathcal{X}_h \subset \mathcal{X}_{h+1}$. Let $k = \arg\max\{t_k : k \in \mathcal{X}_{h+1}\}$ and write $\mathcal{X}_{h+1} = \mathcal{Y}_h \cup \{k\}$.

Suppose $k \notin \mathcal{X}_h$. To get a contradiction, assume that $v(\mathcal{Y}_h) < v(\mathcal{X}_h)$. Then

$$\begin{aligned} v(\mathcal{X}_{h+1}) &= v(\mathcal{Y}_h \cup \{k\}) \\ &= (1-f_k)v(\mathcal{Y}_h) + f_k t_k \\ &< (1-f_k)v(\mathcal{X}_h) + f_k \operatorname{E}\bigl[\max\{t_k, X_h\}\bigr] \\ &= v(\mathcal{X}_h \cup \{k\}) \end{aligned} \tag{20}$$

contradicts the optimality of $\mathcal{X}_{h+1}$.

Now suppose that $k \in \mathcal{X}_h$. We can write $\mathcal{X}_h = \mathcal{Y}_{h-1} \cup \{k\}$, where $\mathcal{Y}_{h-1}$ is some portfolio of size $h-1$. It suffices to show that $\mathcal{Y}_{h-1} \subset \mathcal{Y}_h$. By definition, $\mathcal{Y}_{h-1}$ (respectively, $\mathcal{Y}_h$) maximizes the function $v(\mathcal{Y} \cup \{k\})$ over portfolios of size $h-1$ (respectively, h) that do not include k. That is, $\mathcal{Y}_{h-1}$ and $\mathcal{Y}_h$ are the optimal *complements* to the singleton portfolio $\{k\}$.

We will use the function $w(\mathcal{Y})$ to grade portfolios $\mathcal{Y} \subseteq \mathcal{C} \setminus \{k\}$ according to how well they complement $\{k\}$. To construct $w(\mathcal{Y})$, let $\tilde{t}_j$ denote the expected utility Alma receives from school c_j *given* that she has been admitted to c_j and applied to c_k. For $j < k$, including $j = 0$, this is $\tilde{t}_j = t_j(1-f_k) + t_k f_k$; for $j > k$, this is $\tilde{t}_j = t_j$. This means that

$$v(\mathcal{Y} \cup \{k\}) = \sum_{j \in \{0\} \cup \mathcal{Y}} \tilde{t}_j p_j(\mathcal{Y}) \tag{21}$$

The transformation to $\tilde{t}$ does not change the order of the t_j-values. Therefore, the expression on the right side of (21) is itself a portfolio valuation function. In the corresponding market, t is replaced by $\tilde{t}$ and $\mathcal{C}$ is replaced by $\mathcal{C} \setminus \{k\}$. Now, we obtain $w(\mathcal{Y})$ through one more transformation: Define $\bar{t}_j = \tilde{t}_j - \tilde{t}_0$ so that $t_0 = 0$ and let

$$w(\mathcal{Y}) = \sum_{j \in \{0\} \cup \mathcal{Y}} \bar{t}_j p_j(\mathcal{Y}) = \sum_{j \in \{0\} \cup \mathcal{Y}} \tilde{t}_j p_j(\mathcal{Y}) - \tilde{t}_0 = v(\mathcal{Y} \cup \{k\}) - t_k f_k \tag{22}$$

where the second equality follows from Theorem 1. This identity says that the optimal complements to $\{k\}$, given by $\mathcal{Y}_{h-1}$ and $\mathcal{Y}_h$, are themselves optimal portfolios of size $h-1$ and h for the market whose objective function is $w(\mathcal{Y})$. Since $\bar{t}_0 = 0$ in the latter market, the inductive hypothesis implies that $\mathcal{Y}_{h-1} \subset \mathcal{Y}_h$, which completes the proof.[3] $\square$

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. Sorting t is $O(m \log m)$. At each of the h iterations, there are $O(m)$ candidates for k, and computing $v(\mathcal{X} \cup \{k\})$ is $O(h)$ using (7); therefore, the time complexity of this algorithm is $O(h^2 m + m \log m)$.

We reduce the computation time to $O(hm)$ by taking advantage of the transformation from the inductive step in the proof of Theorem 3. Once school k is added to $\mathcal{X}$, we remove it from the set $\mathcal{C} \setminus \mathcal{X}$ of candidates, and update the t_j-values of the remaining schools according to the

[3]We thank Yim Seho for discovering this critical transformation.

following transformation:

$$\bar{t}_j = \begin{cases} t_j(1 - f_k), & t_j \leq t_k \\ t_j - t_k f_k, & t_j > t_k \end{cases} \tag{23}$$

It is easy to verify that this is the composition of the two transformations (from t to $\tilde{t}$, and from $\tilde{t}$ to $\bar{t}$) given in the proof. Now, the *next* school added must be the optimal singleton portfolio in the modified market. But the optimal singleton portfolio consists simply of the school with the highest value of $f_j \bar{t}_j$. Therefore, by updating the t_j-values at each iteration according to (23), we eliminate the need to compute $v(\mathcal{X})$ entirely. Moreover, this algorithm does not require the schools to be indexed in ascending order by t_j, which removes the $O(m \log m)$ sorting cost.

The algorithm below outputs a list `X` of the h schools to which Alma should apply. The schools appear in the order of entry such that when the algorithm is run with $h = m$, the optimal portfolio of size h is given by $\mathcal{X}_h = \{\mathtt{X[1]}, \dots, \mathtt{X[h]}\}$. The entries of the list `V` give the valuation thereof.

Algorithm 1: Optimal portfolio algorithm for Alma's problem.

Data: Utility values $t \in [0, \infty)^m$, admissions probabilities $f \in [0, 1]^m$, application limit $h \leq m$.
$\mathcal{C} \leftarrow \{1 \dots m\}$;
$\mathtt{X}, \mathtt{V} \leftarrow$ empty lists;
for $i = 1 \dots h$ **do**
  $k \leftarrow \arg\max_{j \in \mathcal{C}} \{f_j t_j\}$;
  $\mathcal{C} \leftarrow \mathcal{C} \setminus \{k\}$;
  $\mathrm{append!}(\mathtt{X}, k)$;
  if $i = 1$ **then** $\mathrm{append!}(\mathtt{V}, f_k t_k)$ **else** $\mathrm{append!}(\mathtt{V}, \mathtt{V[i-1]} + f_k t_k)$;
  for $j \in \mathcal{C}$ **do**
    if $t_j \leq t_k$ **then** $t_j \leftarrow t_j(1 - f_k)$ **else** $t_j \leftarrow t_j - f_k t_k$;
  end
end
return $\mathtt{X}, \mathtt{V}$

Theorem 4 (Validity of Algorithm 1)**.** *Algorithm 1 produces an optimal application portfolio for Alma's problem in $O(hm)$ time.*

Proof. Optimality follows from the proof of Theorem 3. Suppose $\mathcal{C}$ is stored as a list. Then at each of the h iterations of the main loop, finding the top school costs $O(m)$, and the t_j-values of the remaining $O(m)$ schools are each updated in unit time. Therefore, the overall time complexity is $O(hm)$. □

In our numerical experiments, we found it effective to store $\mathcal{C}$ as a binary max heap rather than a list. The heap is ordered according to the criterion $i \geq j \iff f_i t_i \geq f_j t_j$. Nominally, using a heap increases the cost of the main loop from $O(hm)$ to $O(hm \log m)$ because the heap is rebalanced when each t_j-value is updated. However, typical problem instances do not achieve this upper bound because the order of the $f_j t_j$-values changes only slightly between iterations. The cost of updating each t_j-value can be reduced to unit time using a Fibonacci heap (Fredman and Tarjan 1987), yielding the same overall computation time.

3.4 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* $h = 2 \dots (m-1)$,

$$v(\mathcal{X}_h) - v(\mathcal{X}_{h-1}) \geq v(\mathcal{X}_{h+1}) - v(\mathcal{X}_h) \tag{24}$$

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

$$\begin{aligned} 2v(\mathcal{X}_h) &= v(\mathcal{X}_{h-1} \cup \{j\}) + v(\mathcal{X}_{h-1} \cup \{j\}) \\ &\geq v(\mathcal{X}_{h-1} \cup \{k\}) + v(\mathcal{X}_{h-1} \cup \{j\}) \\ &= v(\mathcal{X}_{h-1} \cup \{k\}) + (1-f_j)v(\mathcal{X}_{h-1}) + f_j \operatorname{E}[\max\{t_j, X_{h-1}\}] \\ &= v(\mathcal{X}_{h-1} \cup \{k\}) - f_j v(\mathcal{X}_{h-1}) + f_j \operatorname{E}[\max\{t_j, X_{h-1}\}] + v(\mathcal{X}_{h-1}) \\ &\geq v(\mathcal{X}_{h-1} \cup \{k\}) - f_j v(\mathcal{X}_{h-1} \cup \{k\}) + f_j \operatorname{E}[\max\{t_j, X_{h-1}\}] + v(\mathcal{X}_{h-1}) \\ &= (1-f_j) v(\mathcal{X}_{h-1} \cup \{k\}) + f_j \operatorname{E}[\max\{t_j, X_{h-1}\}] + v(\mathcal{X}_{h-1}) \\ &= v(\mathcal{X}_{h-1} \cup \{j,k\}) + v(\mathcal{X}_{h-1}) \\ &= v(\mathcal{X}_{h+1}) + v(\mathcal{X}_{h-1}) \end{aligned} \tag{25}$$

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:

$$\begin{aligned} 2v(\mathcal{X}_h) &= v(\mathcal{X}_{h-1} \cup \{j\}) + v(\mathcal{X}_{h-1} \cup \{j\}) \\ &\geq v(\mathcal{X}_{h-1} \cup \{k\}) + v(\mathcal{X}_{h-1} \cup \{j\}) \\ &= (1-f_k)v(\mathcal{X}_{h-1}) + f_k \operatorname{E}[\max\{t_k, X_{h-1}\}] + v(\mathcal{X}_{h-1} \cup \{j\}) \\ &= v(\mathcal{X}_{h-1}) - f_k v(\mathcal{X}_{h-1}) + f_k \operatorname{E}[\max\{t_k, X_{h-1}\}] + v(\mathcal{X}_{h-1} \cup \{j\}) \\ &\geq v(\mathcal{X}_{h-1}) - f_k v(\mathcal{X}_{h-1} \cup \{j\}) + f_k \operatorname{E}[\max\{t_k, X_{h-1}\}] + v(\mathcal{X}_{h-1} \cup \{j\}) \\ &= v(\mathcal{X}_{h-1}) + (1-f_k) v(\mathcal{X}_{h-1} \cup \{j\}) + f_k \operatorname{E}[\max\{t_k, X_{h-1}\}] \\ &= v(\mathcal{X}_{h-1}) + v(\mathcal{X}_{h-1} \cup \{j,k\}) \\ &= v(\mathcal{X}_{h-1}) + v(\mathcal{X}_{h+1}) \end{aligned} \tag{26}$$

□

It follows that when $\mathcal{X}_h$ is the optimal h-portfolio, for a given market, $v(\mathcal{X}_h)$ is $O(h)$. Example 2, in which $v(\mathcal{X}_h)$ can be made arbitrarily close to h, establishes the tightness of this bound.

To wrap up, we provide an example showing that if the entries of Z are dependent, then the optimal solution may violate the nestedness property of Theorem 3.

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

3.5 Extension to heterogeneous application costs

In this section, we consider a more general problem in which the constant g_j represents the *cost* of applying to c_j and the student, whom we now call Ellis, has a *budget* of H to spend on college applications. With minimal loss of generality, we assume that $g_j \in \mathbb{N}$ for $j = 1 \dots m$ and $H \in \mathbb{N}$. Applying Theorem 1, we assume $t_0 = 0$ and disregard c_0.

Definition 7 (Ellis's problem)**.** Ellis's optimal college application portfolio is given by the solution to the following integer nonlinear program:

$$\begin{aligned} \text{maximize} \quad & v(x) = \sum_{j=1}^{m} x_j t_j f_j \prod_{j'=j+1}^{m} (1 - f_{j'} x_{j'}) \\ \text{subject to} \quad & \sum_{j=1}^{m} g_j x_j \leq H \\ & x_j \in \{0, 1\}, \quad j = 1 \dots m \end{aligned} \tag{27}$$

The optima for Ellis's problem are not necessarily nested, nor is the number of schools in the optimal portfolio necessarily increasing in H. For example, if $f = (0.5, 0.5, 0.5)$, $t = (10, 3, 2020)$, and $g = (1, 1, 3)$, then it is evident that the optimal portfolio for $H = 2$ is $\{1, 2\}$ while that for $H = 3$ is $\{3\}$.

In fact, Ellis's problem is NP-complete, as we will show by a transformation from the binary knapsack problem, which is known to be NP-complete (Garey and Johnson 1979).

Definition 8 (Decision form of knapsack problem)**.** An *instance* consists of a set $\mathcal{B}$ of m objects; utility values $u_j \in \mathbb{N}$ and weight $w_j \in \mathbb{N}$ for each $j \in \mathcal{B}$; and target utility $U \in \mathbb{N}$ and knapsack capacity $W \in \mathbb{N}$. The instance is called a *yes-instance* if and only if there exists a set $\mathcal{B}' \subseteq \mathcal{B}$ having $\sum_{j \in \mathcal{B}'} u_j \geq U$ and $\sum_{j \in \mathcal{B}'} w_j \leq W$.

Definition 9 (Decision form of Ellis's problem)**.** An *instance* consists of an instance of Ellis's problem and a target valuation V. The instance is called a *yes-instance* if and only if there exists a portfolio $\mathcal{X} \subseteq \mathcal{C}$ having $v(\mathcal{X}) \geq V$ and $\sum_{j \in \mathcal{X}} g_j \leq H$.

Theorem 6. *The decision form of Ellis's problem is NP-complete.*

Proof. It is obvious that the problem is in NP.

Consider an instance of the knapsack problem, and we will construct an instance of Ellis's problem that is a yes-instance if and only if the corresponding knapsack instance is a yes-instance. Without loss of generality, we may assume that the the objects in $\mathcal{B}$ are indexed in increasing order of u_j, that each $u_j > 0$, and that the knapsack instance admits a feasible solution other than the empty set.

Let $U_{\max} = \sum_{j \in \mathcal{B}} u_j$ and $\delta = 1/mU_{\max} > 0$, and construct an instance of Ellis's problem with $\mathcal{C} = \mathcal{B}$, $H = W$, all $f_j = \delta$, and $t_j = u_j/\delta$ for all j. Clearly, $\mathcal{X} \subseteq \mathcal{C}$ is feasible for Ellis's problem if and only if it is feasible for the knapsack instance. Now, we observe that if $\mathcal{X}$ is nonempty,

$$\begin{aligned}
\sum_{j\in\mathcal{X}} u_j &= \sum_{j\in\mathcal{X}} t_j f_j > \sum_{j\in\mathcal{X}} \Big(t_j f_j \prod_{\substack{j'\in\mathcal{X}:\\ j'>j}} (1-f_{j'}) \Big) = v(\mathcal{X}) \\
&= \sum_{j\in\mathcal{X}} \Big(u_j \prod_{\substack{j'\in\mathcal{X}:\\ j'>j}} (1-\delta) \Big) \geq (1-\delta)^m \sum_{j\in\mathcal{X}} u_j \\
&\geq (1-m\delta) \sum_{j\in\mathcal{X}} u_j \geq \sum_{j\in\mathcal{X}} u_j - m\delta U_{\max} = \sum_{j\in\mathcal{X}} u_j - 1
\end{aligned} \tag{28}$$

This means that given a nonempty portfolio $\mathcal{X}$, its value for the corresponding knapsack instance is the smallest integer greater than $v(\mathcal{X})$. That is, $\sum_{j\in\mathcal{X}} u_j \geq U$ if and only if $v(\mathcal{X}) \geq U - 1$. Taking $V = U - 1$ completes the transformation and concludes the proof. □

We will provide an algorithmic solution to Ellis's problem that runs in $O(Hm + m\log m)$ time and $O(Hm)$ space. The algorithm resembles the dynamic programming algorithm for the binary knapsack problem. Because we cannot assume that $H \leq m$ (as was the case in Alma's problem), this represents a quasipolynomial-time solution (Dantzig 1957; Garey and Johnson 1979; *Wikipedia*, s.v. "Knapsack problem").

For $j = 0\dots m$ and $h = 0\dots H$, let $\mathcal{X}[j,h]$ denote the optimal portfolio using only the schools $\{1,\dots,j\}$ and costing no more than h, and let $V[j,h] = v(\mathcal{X}[j,h])$. It is clear that for $j = 0\dots m$, $\mathcal{X}[j,0] = \varnothing$ and $V[j,0] = 0$. For convenience, we also define $V[j,h] = -\infty$ for all $h < 0$.

For the remaining indices, $\mathcal{X}[j,h]$ either contains j or not. If it does not contain j, then $\mathcal{X}[j,h] = \mathcal{X}[j-1,h]$. On the other hand, if $\mathcal{X}[j,h]$ contains j, then its value is $(1-f_j)v(\mathcal{X}[j,h] \setminus \{j\}) + f_j t_j$. This requires that $\mathcal{X}[j,h] \setminus \{j\}$ make optimal use of the remaining budget over the remaining schools; that is, $\mathcal{X}[j,h] = \mathcal{X}[j-1,h-g_j] \cup \{j\}$. From these observations, we obtain the following Bellman equation for $j = 1\dots m$ and $h = 1\dots H$:

$$V[j,h] = \max\big\{V[j-1,h], (1-f_j)V[j-1,h-g_j] + f_j t_j\big\} \tag{29}$$

with the convention that $-\infty \cdot 0 = -\infty$. The corresponding optimal portfolios can be computed by observing that $\mathcal{X}[j,h]$ contains j if and only if $V[j,h] > V[j-1,h]$. The optimal solution is given by $\mathcal{X}[m,H]$. The algorithm below performs these computations and outputs the optimal portfolio $\mathcal{X}$.

Algorithm 2: Optimal portfolio algorithm for Ellis's problem.

Data: Utility values $t \in [0, \infty)^m$, admissions probabilities $f \in [0, 1]^m$, application costs $g \in \mathbb{N}^m$, budget $H \in \mathbb{N}$.

Index schools in ascending order by t;

function $V(j, h)$ **do**
  if $j = 0$ *or* $h = 0$ **then return** 0;
  else if $h < g_j$ **then return** $-\infty$;
  else return $\max\{V(j-1, h), (1-f_j)V(j-1, h-g_j) + f_j t_j\}$;
end
$h = H$;
$\mathcal{X} \leftarrow \emptyset$;
for $j = m, m-1, \ldots, 1$ **do**
  if $V(j-1, h) < V(j, h)$ **then**
    $\mathcal{X} \leftarrow \mathcal{X} \cup \{j\}$;
    $h \leftarrow h - g_j$;
  end
end
return $\mathcal{X}$

Theorem 7 (Validity of Algorithm 2)**.** *Algorithm 2 produces an optimal application portfolio for Ellis's problem in $O(Hm + m \log m)$ time and $O(Hm)$ space.*

Proof. Optimality follows from the foregoing discussion. Sorting by t_j is $O(m \log m)$. The number of indices at which $V(j, h)$ is nontrivially defined is $O(Hm)$. To prevent $V(j, h)$ from being evaluated more than once at a given index, its values can be stored in a dictionary as they are computed. □

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 $\gtrapprox$ whose space is $\{\mathcal{T} \subseteq \mathcal{S} : |\mathcal{T}| = q\}!$. Specific realizations of $\gtrapprox$ are denoted $\gtrsim$. 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 $\gtrapprox$ 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 $\gtrsim$ of c_1's preference order. In our model, the distribution of $\gtrapprox$ is regarded as an exogenous variable. Therefore, student i's *admissions probability* can be expressed as a function of her peers' application decisions.

Assumption 3 (Existence of $f(\cdot)$)**.** Let $x_{\setminus i}$ denote the application decisions of students other than i. Then

$$f_i(x_{\setminus i}) = \Pr \left[\begin{array}{l} i \text{ admitted to } c_1 \mid \\ i \text{ applies to } c_1 \text{ and others' application probabilities are } x_{\setminus i} \end{array} \right] \in [0, 1] \tag{30}$$

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 4 (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 $\gtrsim$ of c_1's preference order induces a realization y of Y.

Definition 10 (Expected ideal entering class)**.** Let

$$\mathcal{Y} = \underset{\mathcal{T} \subseteq \mathcal{S}}{\arg\max} \{\gtrapprox : |\mathcal{T}| = q\} \tag{31}$$

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

$$\bar{y} = f(\mathbf{1}) \tag{32}$$

and called the *expected ideal entering class*.

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

Theorem 8. *$x \in \{0,1\}^n$ 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

$$u_i(x) = x_i\big(f_i(x_{\setminus i})\big) + (1 - x_i)t_i \tag{33}$$

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 11 (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

$$x_i \in \arg\max_{x_i} \{u_i(x) : x_i \in [0,1]\}, \quad i = 1 \ldots n \tag{34}$$

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

$$x_i = 1 \quad \iff \quad f_i(x_{\setminus i}) \geq t_i, \quad i = 1 \dots n \tag{35}$$

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

Theorem 9 (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 13 (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 10 (Bounds on equilibrium size)**.** *The size of any pure-strategy equilibrium x is bound by*

$$q \leq k(x) \leq \frac{q}{\min(t_i)} \tag{36}$$

Proof. The lower bound is from Assumption 4. 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 14 (Stability index)**.** The statistic

$$\bar{S}(x) = \frac{x \cdot \bar{y}}{k(x)} \tag{37}$$

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 $\mathcal{B} = \{i : x_i = 0 \land y_i = 1\}$ 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.[4]

Let $\mathcal{A}_+ = \{i : x_i = 1 \land y_i = 1\}$ and $\mathcal{A}_- = \{i : x_i = 1 \land y_i = 0\}$. 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

$$S(x) = \frac{|\mathcal{A}_+|}{|\mathcal{A}_-| + |\mathcal{A}_+|} = \frac{x \cdot y}{k} \tag{38}$$

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 15 (Alignment index)**.** The statistic

$$\bar{T}(x) = \frac{x \cdot \bar{y}}{q} \tag{39}$$

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

[4] 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 8.

Definition 16 (Aggregate student welfare)**.** The sum of students' utility functions

$$\bar{U}(x) = \sum_{i=1}^{n} u_i(x) = q + (\mathbf{1} - x) \cdot t \tag{40}$$

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 (40) 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

Even in the two-school model, in which only c_1 has a nontrivial preference order, the high dimension of $\gtrapprox$ and the associated probability space presents a challenging modeling problem. In this chapter, we explore two parameterizations of $\gtrapprox$ that are tractable enough to enable the computation of equilibria and the associated efficiency measures but also conformable to realistic recruitment goals such as aggregate student talent and diversity.

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

Previous research in college course enrollment has used a quadratic objective function to model complementarities in students' preferences over combinations of courses. In the course allocation problem, the decision variable is a set of deterministic course schedules, and it is possible to obtain an equilibrium course schedule with several desirable fairness properties by having students bid on courses using an artificial currency and computing course prices using a tatônnment-like step rule (Othman et al. 2010; Budish 2011).

Our environment adds the new element of uncertainty, as well as being a sequential game.

In this section, we use a quadratic objective function with two terms to model combinatorial school preferences. We simulate the randomness in school preferences by solving the continuous relaxation of the quadratic program (QP) associated with c_1's optimal entering class.

5.2.1 Model

Let $z \in \{0,1\}^n$ denote the vector of students who attend c_1. We suppose that c_1's preference order is given by the concave objective function

$$w(z) = (1-\gamma)a \cdot z - \gamma \|b - z\|_2^2 \tag{41}$$

where $a, b \in \{c : \sum c_i = n, c \geq 0\}$ and $\gamma \in [0,1]$. c_1 then may determine its entering class by solving the following optimization problem.

$$\text{maximize} \;\; w(z) \quad \text{subject to} \;\; \sum z_i \leq q, \;\; z \in \{0,1\}^n \tag{42}$$

The first modification we make to this basic template is to model uncertainty in the preference order by allowing z to vary continuously zero and one. This yields the following convex QP:

$$\text{maximize } \ w(z) \quad \text{subject to } \ \sum z_i \leq q, \ \ \mathbf{0} \leq z \leq \mathbf{1} \tag{43}$$

Then the expected ideal entering class is $\bar{y} = z^*$, where z^* is the solution to (43).

We interpret $w(z)$ as follows. The first term, $a \cdot z$, is called the *substitution term*, as it treats students i and j as interchangeable in the ratio between the corresponding coefficients. For example, if $\gamma = 0$ and c_1 has a choice between two students s_i and s_j for whom $w_i > w_j$, it will always choose s_i. On the other hand, the quadratic term of $w(z)$ is called the *complementarity term* and measures the Euclidean distance between z and the point b. The effect of this term is to pull the admitted student vector toward the interior of the positive orthant. Thus, as γ increases, c_1 favors admitting s_i and s_j each to some extent in the continuous relaxation of the problem. Requiring the elements of b to sum to n ensures that the attracting point is on the outside of the constraint $\sum z_i \leq q$ so that this equation always holds with equality at the optimum.

Now, we need to construct the admissions probability function $f(x)$. We begin by modifying (43) to read $\mathbf{0} \leq z \leq x$. If x is binary, then this constraint simply says that c_1 can only admit students who apply. If s_i employs a mixed strategy, say $x_i = 0.5$, then this constraint says that c_1 can only admit her "to the extent" of $z = 0.5$. Next, recall that $f_i(x_{\setminus i})$ is a conditional probability, namely, the probability that s_i is admitted to c_1 if she applies with certainty. Therefore, its entries are determined by solving the counterfactual QP in which $x_i = 1$ and $x_{\setminus i}$ is unchanged. With these observations, we are ready to define $f(x)$ for the quadratic preference environment.

Definition 17 (Admissions probability function from QP)**.** Given an application vector x, the admissions probability function $f(x)$ is defined as follows: For $i = 1 \ldots n$, $f_i(x_i) = z^*(i)_i$, where $z^*(i)$ is the solution to the following QP.

$$\begin{aligned} \text{maximize} \quad & (1-\gamma) a \cdot z - \gamma \|b - z\|_2^2 \\ \text{subject to} \quad & \sum z_i \leq q \\ & 0 \leq z_{i'} \leq x_{i'}, \quad i' \neq i \\ & 0 \leq z_i \leq 1 \end{aligned} \tag{44}$$

The expected ideal entering class is defined as $\bar{y} = f(\mathbf{1})$.

Nominally, computing $f(x)$ requires solving n QPs. The computation time can be reduced by first solving

$$\text{maximize } \ w(z) \quad \text{subject to } \ \sum z_i \leq q, \ \ \mathbf{0} \leq z \leq x \tag{45}$$

and inspecting the dual variables λ associated the constraint $z \leq x$. If $\lambda_i = 0$, then relaxing this constraint yields no improvement in the optimal value of $w(z)$, and the optimal z-vectors for (44) and (45) are identical.

5.2.2 Equilibrium computation

To compute mixed-strategy equilibria in the quadratic preference environment, we start with an arbitrary strategy vector x and adjust it according to a decreasing step rule reminiscent of projected gradient descent.

Let $\mathrm{Clamp}(x)$ denote the projection of the vector x to the 1-hypercube:

$$\mathrm{Clamp}(x)_i = \max\big\{0, \min\{1, x_i\}\big\}, \tag{46}$$

Observe that a mixed strategy can be optimal for s_i only if $f_i = t_i$. If $f_i - t_i \gtrless 0$, then s_i can improve her utility by choosing $x_i = \frac{1}{0}$. Therefore, it is not difficult to see that x is an equilibrium if and only if

$$x = \mathrm{Clamp}\big[x + f(x) - t\big] \tag{47}$$

Thus, we use the following intuitive notion of approximate equilibrium as our convergence criterion.

Definition 18 (ε-approximate equilibrium)**.** For a given market, x is called an ε-approximate equilibrium if and only if

$$\Big\| x - \mathrm{Clamp}\big[x + f(x) - t\big] \Big\|_2 \leq \epsilon$$

We search for equilibria by generating an arbitrary mixed-strategy vector x and generate three kinds of neighbors, in a manner similar to that employed by Othman et al. (2010). updating in the direction of $f(x) - t$, as described in the following algorithm. The choice of step size ensures that, given infinite iterations, the algorithm can explore the entirety of the feasible space while refraining from cycling. Rather than iterating continuously from a single initial point, in our experiments, it proved more effective to terminate the search and choose a new seed point when the algorithm fails to converge within several iterations.

Algorithm 3: Equilibrium algorithm for quadratic school preferences.

Data: Market parameters $t \in [0,1]^n$, $a, b \in [0, \infty)^n$, $q \leq n$; step parameters $\alpha > 0, \beta \in (0, 1]$; tolerance $\epsilon > 0$.

for $j = 1 \dots J$ **do**
- $x^{(0)} \leftarrow$ an arbitrary vector in $[0,1]^n$;
- **for** $k = 1 \dots K$ **do**
  - $x^{(k)} \leftarrow \mathrm{Box}\Big[x^{(k-1)} + \frac{\alpha}{k^\beta}\big(f(x^{(k-1)}) - t\big)\Big]$;
  - **if** $\big|x_i^{(k)} - x_i^{(k-1)}\big| \leq \epsilon, i = 1 \dots n$ **then return** $x^{(k)}$;
- **end**

end

We were able to obtain a 10^{-5}-approximate equilibrium in 90 percent of our synthetic instances using the parameters $\alpha = 10$, $\beta = 0.01$, $J = 3$, and $K = 30$.

5.2.3 Computational experiments

We consider two experiments. In the heterogeneity experiment, we fix the value of γ to 0.5 and the *expected value* of t_i to 0.5. We vary the *variance* of t_i by drawing t_i independently from the $\mathrm{Beta}(B, B)$ distribution, where $1/B$ itself is drawn randomly from $\mathrm{Uniform}(0, 1)$.

In the complementarity experiment, we fix the distribution of t_i to $\mathrm{Beta}(10, 10)$ and vary γ by drawing it randomly from $\mathrm{Uniform}(0, 1)$.

In all cases, a, b, q, and n are ... and we create 500 instances.

Because the data used in these experiments is artificial, we elect to analyze the results graphically. T-tests show that convergence failures do not depend statistically on the experimental variables.

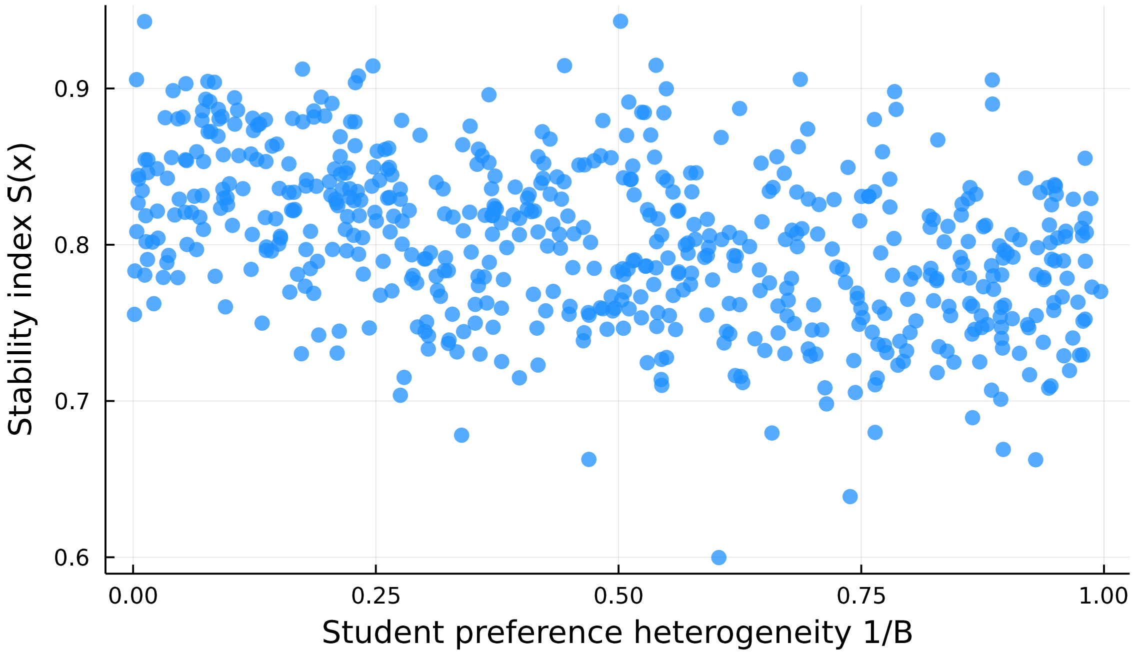

Figure 1: (Heterogeneity experiment.) Effect of heterogeneity in students' risk preferences on the stability index $\bar{S}(x)$, a measure of the equilibrium's robustness to envy. Risk parameters t_i were drawn from a $\mathrm{Beta}(B, B)$ distribution for 500 markets. 108 markets failed to converge to an equilibrium.

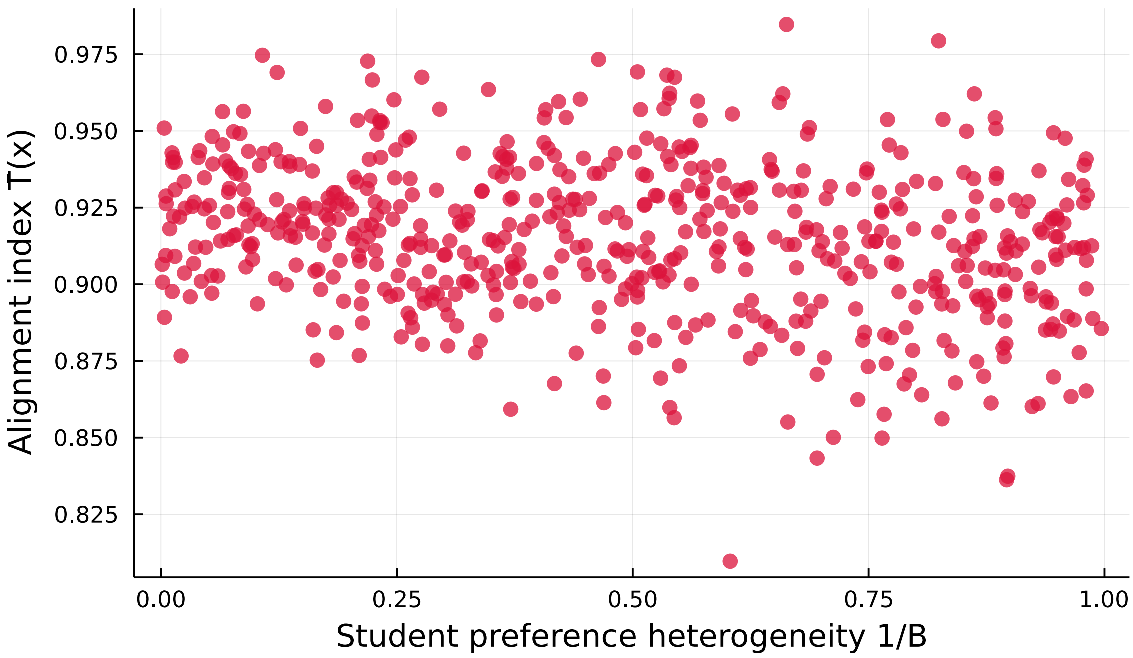

Figure 2: (Heterogeneity experiment.) Effect of heterogeneity in students' risk preferences on the alignment index $\bar{T}(x)$, a measure of the selective school's ability to approximate its ideal entering class using the subset of students who apply.

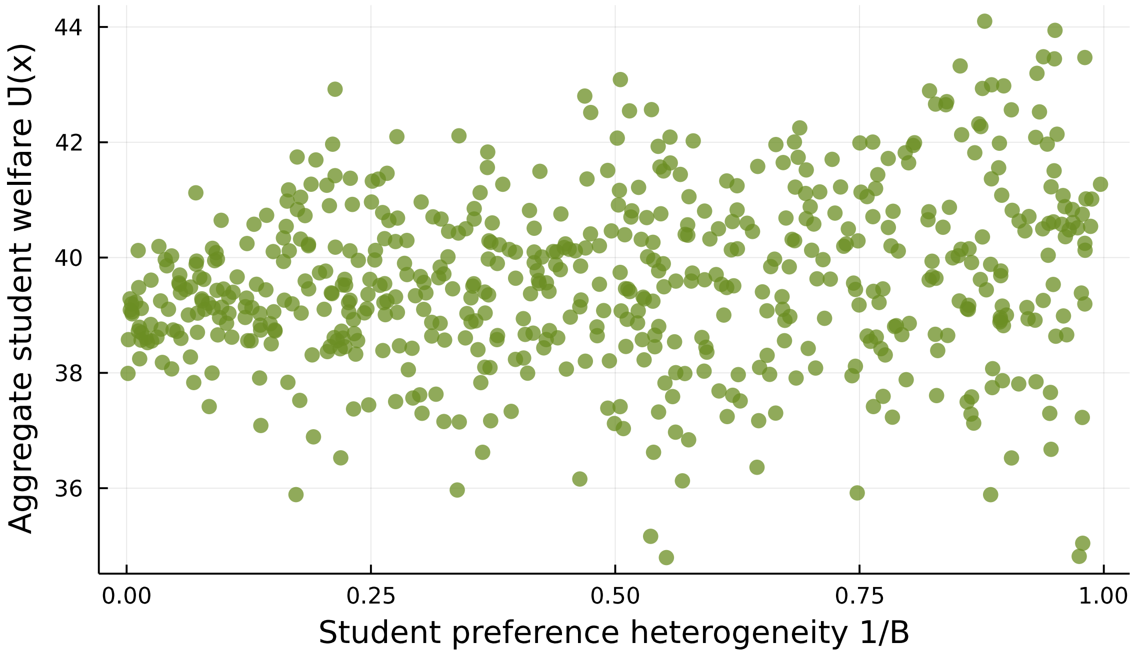

Figure 3: (Heterogeneity experiment.) Effect of heterogeneity in students' risk preferences on the aggregate student welfare $\bar{U}(x)$.

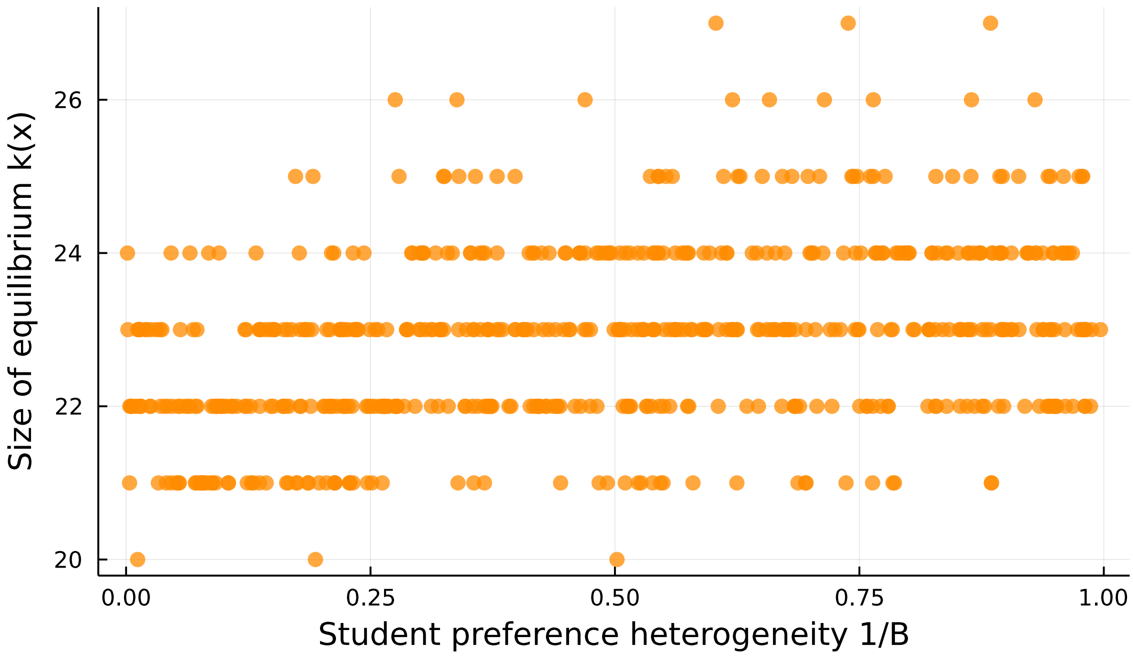

Figure 4: (Heterogeneity experiment.) Effect of heterogeneity in students' risk preferences on the size of the equilibrium $\bar{k}(x)$.

6 References

Budish, Eric. 2011. "The Combinatorial Assignment Problem: Approximate Competitive Equi-

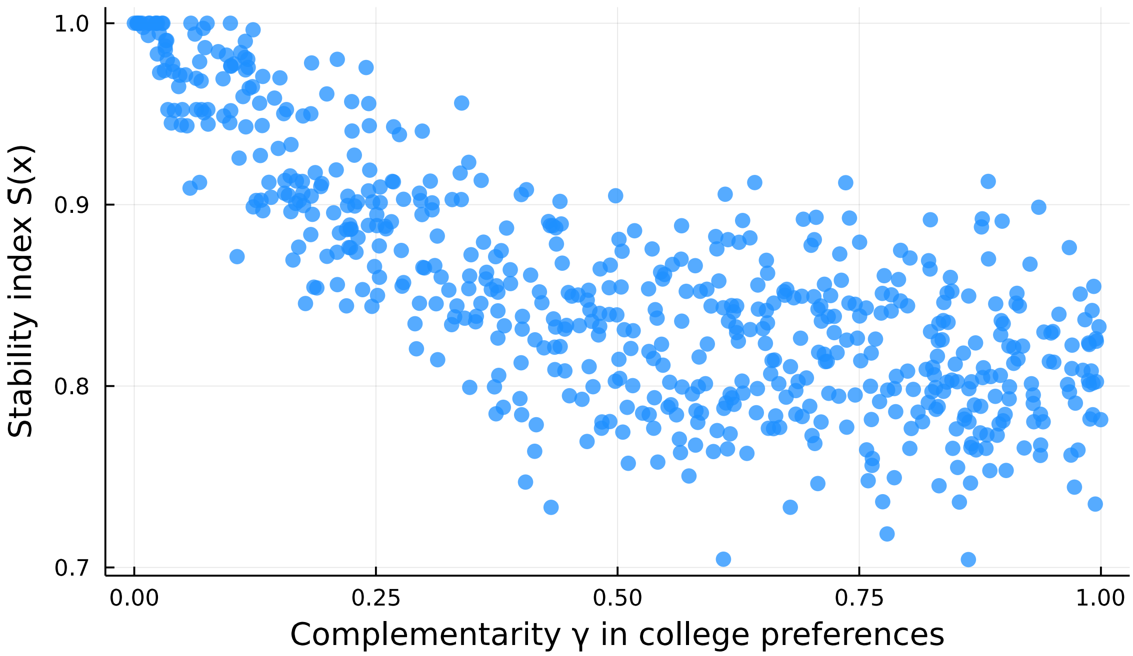

Figure 5: (Complementarity experiment.) Effect of nonlinearity in the selective school's objective function on the stability index $S(x)$, a measure of the equilibrium's robustness to envy. 500 random markets with $\gamma \sim \text{Uniform}(0, 1)$ were simulated. 142 markets failed to converge to an equilibrium.

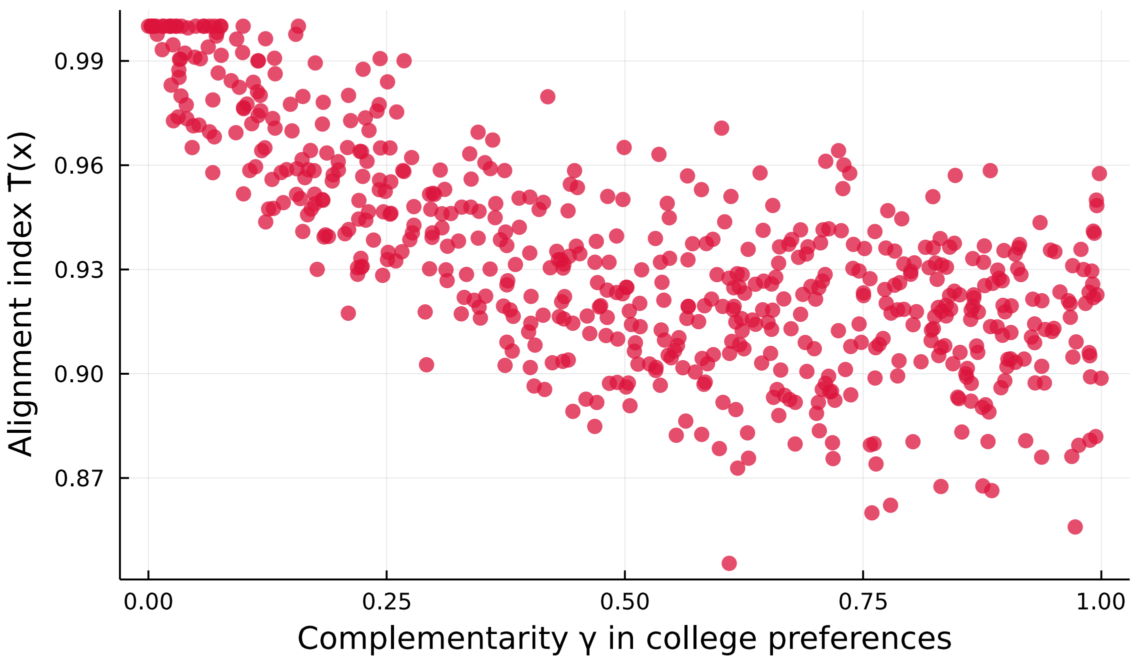

Figure 6: (Complementarity experiment.) Effect of nonlinearity in the school's objective function on the alignment index $\bar{T}(x)$, a measure of the selective school's ability to approximate its ideal entering class using the subset of students who apply.

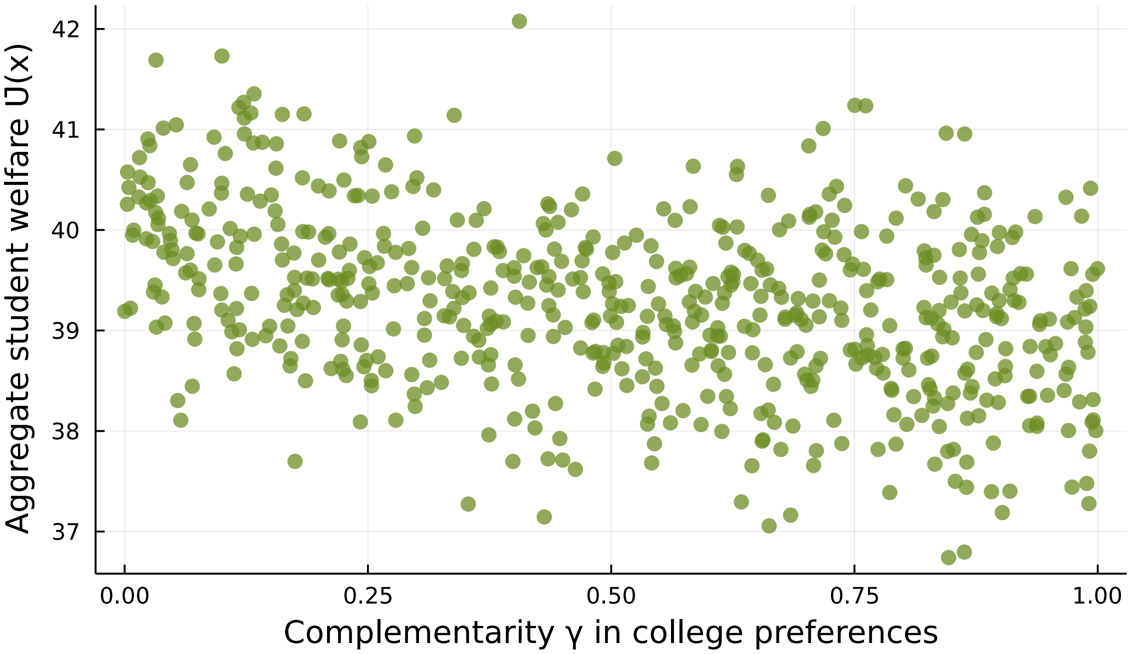

Figure 7: (Complementarity experiment.) Effect of nonlinearity in the school's objective function on the aggregate student welfare $\bar{U}(x)$.

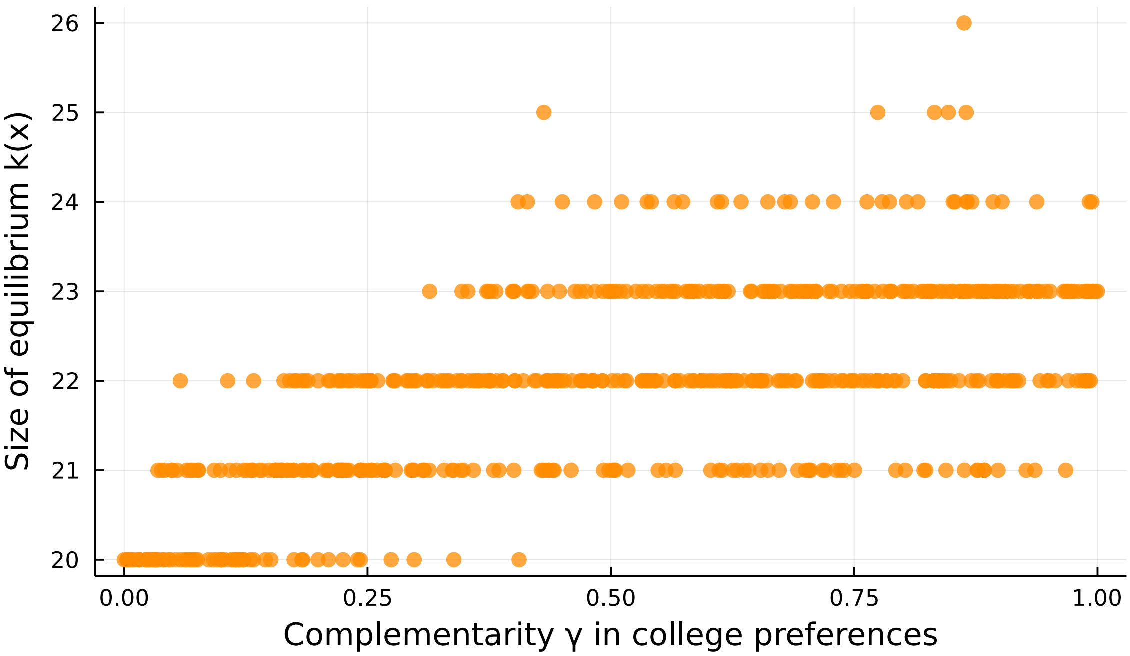

Figure 8: (Complementarity experiment.) Effect of nonlinearity in the school's objective function on the size of the equilibrium $\bar{k}(x)$.

librium from Equal Incomes." *Journal of Political Economy* 119 (6): 1061–1103. https://doi.org/10.1086/664613.

Dantzig, George B. 1957. "Discrete-Variable Extremum Problems." *Operations Research* 5 (2):

266–88.

Fisher, Marshall, George Nemhauser, and Laurence Wolsey. 1978. "An analysis of approximations for maximizing submodular set functions—I." *Mathematical Programming* 14: 265–94.

Fredman, Michael Lawrence and Robert Tarjan. 1987. "Fibonacci heaps and their uses in improved network optimization algorithms." *Journal of the Association for Computing Machinery* 34 (3): 596–615.

Fu, Chao. 2014. "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market." *Journal of Political Economy* 122 (2): 225–81. https://doi.org/10.1086/675503.

Garey, Michael and David Johnson. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* New York: W. H. Freeman and Company.

Othman, Abraham, Eric Budish, and Tuomas Sandholm. 2010. "Finding Approximate Competitive Equilibria: Efficient and Fair Course Allocation." In *Proceedings of 9th International Conference on Autonomous Agents and Multiagent Systems.* New York: ACM. https://dl.acm.org/doi/abs/10.5555/1838206.1838323.

Rozanov, Mark and Arie Tamir. 2020. "The nestedness property of the convex ordered median location problem on a tree." *Discrete Optimization* 36: 100581. https://doi.org/10.1016/j.disopt.2020.100581.

Vazirani, Vijay. 2001. *Approximation Algorithms.* Berlin: Springer.