A greedy algorithm for the cardinality-constrained college application problem

Max Kapur[1], Seho Yim[2], and Sung-Pil Hong[3]

[1,2,3]Department of Industrial Engineering, Seoul National University

March 26, 2023

Abstract

This paper considers a novel portfolio optimization problem called the college application problem. We show that the objective function is a nondecreasing submodular set function and provide a quadratic-time algorithm for its maximization over a cardinality constraint. The algorithm's validity rests on the nested structure of the optimal solution, a property that is of independent interest. Further analysis of the optimal application strategy helps explain perplexing findings from regression studies of college applicant behavior: In our model, systematic differences in the risk allocations of low- and high-income students emerge spontaneously from differences in applicants' budgets, even in the absence of risk aversion. Also, our analysis suggests that the underwhelming effectiveness of behavioral nudges (such as application fee waivers) designed to encourage talented low-income students to apply to more colleges can be attributed to the fact that these interventions have typically targeted students with small shadow costs.

Keywords: combinatorial optimization, college application, submodular maximization, portfolio optimization, submodular maximization with cardinality constraint

Correspondence may be addressed to Max Kapur.

Email: maxkapur@gmail.com

Address: 39-411, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

Contents

1 Introduction

In the college admissions market, the desirability of a school correlates with its competitiveness, meaning that college application represents a challenging risk allocation problem: Applicants must weigh competitive, attractive "reach schools" against less attractive "safety schools" where admission is a safer bet. The difficulty of the college application problem, along with the perceived importance of college admissions to future career outcomes, drives demand for a lucrative industry known as admissions consulting. In the US, students pay private consultants an average of $200 per hour for assistance in preparing application materials, estimating their admissions odds, and deciding which schools to apply to [25]. However, when admissions consultants advise their clients on where to apply, they typically rely on qualitative heuristics, such as balancing applications evenly between safety, target, and reach schools [15, 22]. Can a formal statement of the college application problem help us improve upon these heuristics?

Regression studies of college application behavior have revealed intriguing systematic differences in students' application strategies: High-achieving students from wealthy backgrounds tend to apply to more, and more competitive, colleges than their low-income peers [10, 13]. Does this phenomenon reflect underlying psychological differences in students' risk aversion, or rational decision-making under different resource constraints?

The present study addresses the questions above by formulating the student's college application strategy as the following combinatorial optimization problem:

$$\begin{aligned}
\text{maximize} \quad & v(\mathcal{X}) = \mathrm{E}\Big[\max\{t_0, \max\{t_j Z_j : j \in \mathcal{X}\}\}\Big] \\
\text{subject to} \quad & \mathcal{X} \subseteq \mathcal{C}, \quad |\mathcal{X}| \leq h
\end{aligned} \tag{1}$$

Here $\mathcal{C} = \{1 \ldots m\}$ represents the set of colleges in the market and $h > 0$ is the number of colleges the student can apply to. For $j = 1 \ldots m$, let f_j denote the student's probability of being admitted to school j if she applies and Z_j a random, independent Bernoulli variable with probability f_j, so that Z_j equals one if she is admitted and zero if not. We assume that each Z_j is statistically independent.[1] For $j = 0 \ldots m$, $t_j \geq 0$ indicates the utility she associates with attending school j. School 0, with $f_0 = 1$ and utility t_0, represents the outside option of not attending college. Then the student's objective is to maximize the expected utility associated with the best school she is admitted to. Therefore, her optimal college application strategy is given by the solution $\mathcal{X}$ to the problem above, where $\mathcal{X}$ represents the set of schools to which she applies, which we call her application *portfolio.*

The problem can also be applied to other competitive matching games such as job application or marriage proposals. Here the budget constraint may arise due to the time required to communicate with each applicant or a legal limit on the number of applications allowed.

We show that the optimal portfolios for the cardinality-constrained college are *nested* in the budget constraint, meaning that when $\mathcal{X}_h$ is the optimal portfolio of size h, $\mathcal{X}_h \subset \mathcal{X}_{h+1}$. The nestedness property is equivalent to the optimality of the *greedy* algorithm that iteratively adds to $\mathcal{X}$ the school that elicits the greatest increase in the objective value. The proof of the nestedness property yields insight into the properties of the optimal college application strategy: Students who cannot afford to apply to many schools allocate a greater proportion of their application budget to safety schools than students with a large application budget. This phenomenon, which we call optimal risk elasticity, reveals a shortcoming of the reach/safety school heuristics used in the admissions consulting industry: Whereas admissions consultants recommend a *static* risk allocation that allocates a fixed proportion of the portfolio to reach, target, and safety schools, we find that the optimal application strategy gradually favors reach schools over safety schools as h increases. Optimal risk elasticity therefore provides a rational basis for the divergent risk behavior observed in the social science research cited above: It may not be the case that low-income students are inherently more risk-averse than their wealthy peers, but rather, their constrained budget induces a more risk-averse optimal policy.

Our analysis also shows that the valuation of the college application problem is sublinear in the budget h, and that the shadow cost of the budget constraint is highest for students with low admissions probability at the schools currently in their portfolio. This finding suggests an economic explanation for the underwhelming efficacy of behavioral nudges, such as application fee waivers, designed to encourage talented low-income students to apply to selective universities [9].

1.1   Related optimization problems

The objective function of (1) is a nondecreasing, submodular set function in the sense first described by [21]. They showed that the greedy algorithm is asymptotically $(1 - 1/e)$-optimal for optimizing a monotonic submodular function over a cardinality constraint, and a subsequent result of [20] proved that this approximation ratio is the best achievable by a polynomial-time algorithm. For some submodular optimization problems, however, the same greedy algorithm offers a much tighter approximation guarantee [6]. We show that college application represents an easy instance of submodular maximization: The greedy algorithm is exact.

The proof of the greedy algorithm's optimality rests on the nested structure of the optimal

---

[1]Thus, f should give admissions probabilities estimated specifically for this student, as opposed to generic acceptance rates; in the latter case, admissions outcomes correlate between schools.

portfolios, a property that is of independent interest. As [24] articulate, the knowledge that the optima are nested aids not only in computing the optimal solution, but in the implementation thereof under uncertain information. For example, in the United States, many colleges have application deadlines at the beginning of November, and students typically begin work on their applications during the prior summer because colleges favor students who tailor their essays to the target school. However, students may not know how many schools they can afford to apply to until late October. The nestedness property implies that even in the absence of complete budget information, students can begin to carry out the optimal application strategy by writing essays for schools in the order that they enter the optimal portfolio.

In a narrow sense, the college application problem is a static variant of the Pandora's Box problem [27, 5]. In a Pandora's Box formulation of college application, the student applies to schools one by one, each time paying an application fee and observing her admissions outcome after a certain time delay. The problem is to determine an optimal *halting rule* for when Pandora should stop applying to new schools and enroll at the best school she has been admitted to so far. [27] showed that the optimal policy is to end the search when the value of the current best offer exceeds the maximum *reservation price,* a statistic that represents the marginal value of an additional application.

The static model considered in this study is arguable more hostile to students than the Pandora's Box problem. If, for example, an unlucky Pandora is rejected from a safety school at an early round of application, then she can compensate for the unexpected loss by pivoting to a more risk-averse application strategy. By contrast, the decision maker in our college application problem must commit at the outset to applying to every school in her portfolio. The admissions process used in the United States can be viewed as the concatenation of both problems: In the fall, students solve (1) and send out a batch of applications. Then, upon observing their admissions outcomes in March, they use the Pandora's Box strategy to pursue additional offers by applying to schools that offer rolling admissions.

1.2 Admissions markets

Most quantitative research on college admissions to date has been descriptive in nature. A standard benchmark for logistic regression techniques, for example, involves estimating students' admissions probabilities from information about their past academic performance [1, 18]. And the social science literature offers a trove of studies that compare students' application strategies across demographic backgrounds and levels of academic ability [9, 10, 11, 12].

Some economists have constructed equilibrium models of college admissions that feature the student's application decision as a prominent subproblem. For example, a problem reducible to ours arose in a sequential game model of the United States college market by [7]. (We describe the reduction in section 3.2.) However, to ensure the robustness of her parameter estimates, Fu clustered the colleges in the market into $m = 8$ broad categories, yielding a small instance that could be solved by enumerating all 2^m possible portfolios. Our solution algorithm computes the optimal portfolio in time quadratic in m.

The literature on admissions markets also includes a vast body of research on the deferred acceptance algorithm, an algorithm for matching students and schools that generalizes the Gale–Shapley algorithm for stable marriage. The algorithm takes students' ordinal preferences over schools and schools' ordinal preferences over students as input, and produces an assignment that possesses a number of desirable properties: It is a stable matching, meaning that no student–school pair is incentivized to deviate from the assignment. And the preference-reporting mecha-

nism is incentive compatible, meaning that no student can improve her outcome by lying about her preferences [8, 23]. In design, deferred acceptance is a centralized assignment algorithm; however, stable assignment can also be interpreted as the equilibrium of a decentralized admissions game in which agents possess perfect information about one another's preferences [3].

The utility model implied by stable assignment differs from that of the present study in two key ways: First, under stable assignment, students' preferences are ordinal, rather than cardinal (although it is possible to estimate cardinal utility values from stable matchings in certain special cases, such as when colleges' preferences are identical and students' preferences are determined by the multinomial logit choice model [16]). Second, stable assignment is deterministic: The typical student-proposing deferred acceptance algorithm always produces the unique, student-optimal stable assignment. There exist variants of the deferred acceptance algorithm that introduce randomization in order to break ties in schools' preference lists or optimize for distributional goals such as gender parity, but because the randomization occurs after students submit their preferences, it has no effect on the optimal application strategy, which remains to report one's preferences honestly [2, 4]. In the admissions process considered in this study, students face a tougher strategic challenge.

To the best of our knowledge, the first systematic study of the college application problem was Kapur's master's thesis [17], which considered (1) as well as the more general case of college application with a knapsack constraint. The present study extends Kapur's results with an analysis of heuristics used in the admissions consulting industry and the implications of the nestedness and submodularity properties for student welfare.

1.3 Outline

This paper has five sections. In section 2, we derive a closed-form expression for the objective function of (1), show that it is a monotonic submodular function, and state a technical lemma. In section 3, we prove that the optimal portfolios are nested in the budget h, which implies the validity of the greedy algorithm. Further refinement reduces the algorithm's computation time from $O(h^2 m + m \log m)$ to $O(hm)$. In section 4, we examine heuristic solution mechanisms used in the admissions consulting industry, assess their performance relative to the optimal solution, and offer a few conjectures concerning policy interventions to improve student equity. A brief conclusion follows in section 5.

2 Preliminaries

For the remainder of the paper, unless otherwise noted, we assume with minimal loss of generality that $0 < h \leq m$, $f_j \in (0, 1]$, and $t_0 < t_1 \leq \cdots \leq t_m$. Unless otherwise noted, we assume $t_0 = 0$, an assumption justified presently.

2.1 The objective function

First we derive a closed-form expression for the objective function of (1).

We refer to the set $\mathcal{X} \subseteq \mathcal{C}$ of schools to which a student applies as her *application portfolio*. The expected utility the student receives from $\mathcal{X}$ is called its *valuation*. Given an application portfolio, let $p_j(\mathcal{X})$ denote the probability that the student attends school j. This occurs if and only if she *applies* to school j, is *admitted* to school j, and is *rejected* from any school she prefers

to j; that is, any school with higher index. Hence, for $j = 0 \ldots m$,

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

The following proposition follows by computing $v(\mathcal{X}) = \sum_{j=0}^{m} t_j p_j(\mathcal{X})$.

Proposition 1 (Closed form of portfolio valuation function).

$$
v(\mathcal{X}) = \sum_{j=0}^{m} t_j p_j(\mathcal{X}) = \sum_{j \in \{0\} \cup \mathcal{X}} \left(f_j t_j \prod_{\substack{i \in \mathcal{X}: \\ i > j}} (1 - f_i) \right)
$$

Next, we show that without loss of generality, we may assume that $t_0 = 0$.

Lemma 1. *For some $\gamma \leq t_0$, let $\bar{t}_j = t_j - \gamma$ for $j = 0 \ldots m$. Then $v(\mathcal{X}; \bar{t}_j) = v(\mathcal{X}; t_j) - \gamma$ for any $\mathcal{X} \subseteq \mathcal{C}$.*

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

$$
\begin{aligned}
v(\mathcal{X}; \bar{t}_j) &= \sum_{j \in \{0\} \cup \mathcal{X}} \bar{t}_j p_j(\mathcal{X}) = \sum_{j \in \{0\} \cup \mathcal{X}} (t_j - \gamma) p_j(\mathcal{X}) \\
&= \sum_{j \in \{0\} \cup \mathcal{X}} t_j p_j(\mathcal{X}) - \gamma = v(\mathcal{X}; t_j) - \gamma
\end{aligned}
$$

which completes the proof. $\qquad \square$

2.2 An elimination technique

Now, we present a variable elimination technique that will prove useful throughout the paper. Suppose that the student has already resolved to apply to school k, and the remainder of her decision consists of determining which *other* schools to apply to. Writing her total application portfolio as $\mathcal{X} = \mathcal{Y} \cup \{k\}$, we devise a function $w(\mathcal{Y})$ that orders portfolios according to how well they "complement" the singleton portfolio $\{k\}$. Specifically, the difference between $v(\mathcal{Y} \cup \{k\})$ and $w(\mathcal{Y})$ is the constant $f_k t_k$.

To construct $w(\mathcal{Y})$, let $\tilde{t}_j$ denote the expected utility the student receives from school j *given* that she has been admitted to school j and applied to school k. For $j < k$ (including $j = 0$), this is $\tilde{t}_j = (1 - f_k) t_j + f_k t_k$; for $j > k$, this is $\tilde{t}_j = t_j$. This means that

$$
v(\mathcal{Y} \cup \{k\}) = \sum_{j \in \{0\} \cup \mathcal{Y}} \tilde{t}_j p_j(\mathcal{Y}). \tag{2}
$$

The transformation to $\tilde{t}$ does not change the order of the t_j-values. Therefore, the expression on the right side of (2) is itself a portfolio valuation function. In the corresponding market, t is replaced by $\tilde{t}$ and $\mathcal{C}$ is replaced by $\mathcal{C} \setminus \{k\}$. To restore our convention that $t_0 = 0$, we obtain $w(\mathcal{Y})$ by taking $\bar{t}_j = \tilde{t}_j - \tilde{t}_0$ for all $j \neq k$ and letting

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

where the second equality follows from Lemma 1. The validity of this transformation is summarized in the following theorem, where we write $v(\mathcal{X}; \bar{t})$ instead of $w(\mathcal{Y})$ to emphasize that $w(\mathcal{Y})$ is, in form, a portfolio valuation function.

Lemma 2 (Eliminate school k). *For* $\mathcal{X} \subseteq \mathcal{C} \setminus \{k\}$, $v(\mathcal{X} \cup \{k\}; t) = v(\mathcal{X}; \bar{t}) + f_k t_k$, *where*

$$\bar{t}_j = \begin{cases} (1 - f_k)t_j, & t_j \leq t_k \\ t_j - f_k t_k, & t_j > t_k. \end{cases} \tag{3}$$

Proof. It is easy to verify that (3) is the composition of the two transformations (from t to $\tilde{t}$, and from $\tilde{t}$ to $\bar{t}$) discussed above. $\square$

2.3 Submodularity of the objective

Now, we show that the portfolio valuation function is submodular. This result is primarily of taxonomical interest and may be safely skipped, as our subsequent results do not rely on submodular analysis.

Definition 1 (Submodular set function)**.** Given a ground set $\mathcal{C}$ and function $v : 2^{\mathcal{C}} \mapsto \mathbb{R}$, $v(\mathcal{X})$ is called a *submodular set function* if and only if $v(\mathcal{X}) + v(\mathcal{Y}) \geq v(\mathcal{X} \cup \mathcal{Y}) + v(\mathcal{X} \cap \mathcal{Y})$ for all $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{C}$. Furthermore, if $v(\mathcal{X} \cup \{k\}) - v(\mathcal{X}) \geq 0$ for all $\mathcal{X} \subset \mathcal{C}$ and $k \in \mathcal{C} \setminus \mathcal{X}$, $v(\mathcal{X})$ is said to be a *nondecreasing* submodular set function.

Theorem 1. *The college application portfolio valuation function* $v(\mathcal{X})$ *is a nondecreasing submodular set function.*

Proof. It is self-evident that $v(\mathcal{X})$ is nondecreasing. To establish its submodularity, we apply proposition 2.1.iii of [20] and show that

$$v(\mathcal{X} \cup \{j\}) - v(\mathcal{X}) \geq v(\mathcal{X} \cup \{j, k\}) - v(\mathcal{X} \cup \{k\}) \tag{4}$$

for $\mathcal{X} \subset \mathcal{C}$ and $j \neq k \in \mathcal{C} \setminus \mathcal{X}$. By Lemma 2, we can repeatedly eliminate the schools in $\mathcal{X}$ according to (3) to obtain a portfolio valuation function $w(\mathcal{Y})$ with parameter $\bar{t}$ such that $w(\mathcal{Y}) = v(\mathcal{X} \cup \mathcal{Y}) + \text{const.}$ for any $\mathcal{Y} \subseteq \mathcal{C} \setminus \mathcal{X}$. Therefore, (4) is equivalent to

$$w(\{j\}) - w(\varnothing) \geq w(\{j, k\}) - w(\{k\})$$
$$\iff \quad w(\{j\}) + w(\{k\}) \geq w(\{j, k\})$$
$$\iff \quad \mathrm{E}[\,\bar{t}_j Z_j\,] + \mathrm{E}[\,\bar{t}_k Z_k\,] \geq \mathrm{E}\big[\max\{\bar{t}_j Z_j, \bar{t}_k Z_k\}\big]$$

which is true because t_j and t_k are nonnegative. $\square$

3 The greedy algorithm

3.1 The nestedness property

The optimality of the greedy algorithm for the college application problem rests on the fact the the optimal solution possesses a special structure: An optimal portfolio of size $h + 1$ includes an optimal portfolio of size h as a subset.

Theorem 2 (Nestedness of optimal application portfolios). *There exists a sequence of portfolios* $\{\mathcal{X}_h\}_{h=1}^m$ *satisfying the nestedness relation*

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

such that each $\mathcal{X}_h$ is an optimal application portfolio when the application limit is h.

Proof. By induction on h. Applying Lemma 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 $t_k \leq t_l$. Optimality requires that

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

and

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

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, suppose that $v(\mathcal{Y}_h) < v(\mathcal{X}_h)$ and $v(\mathcal{X}_{h+1}) > v(\mathcal{X}_h \cup \{k\})$. Then

$$
\begin{aligned}
v(\mathcal{X}_{h+1}) &= v(\mathcal{Y}_h \cup \{k\}) \\
&= (1 - f_k)v(\mathcal{Y}_h) + f_k t_k \\
&\leq (1 - f_k)v(\mathcal{X}_h) + f_k \mathrm{E}\big[\max\{t_k, X_h\}\big] \\
&= v(\mathcal{X}_h \cup \{k\})
\end{aligned}
$$

is a contradiction.

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 $\{k\}$.

Applying Lemma 2, we eliminate school k, transform the remaining t_j-values to $\bar{t}_j$ according to (3), and obtain a function $w(\mathcal{Y}) = v(\mathcal{Y} \cup \{k\}) - f_k t_k$ that grades portfolios $\mathcal{Y} \subseteq \mathcal{C} \setminus \{k\}$ according to how well they complement $\{k\}$. Since $w(\mathcal{Y})$ is itself a portfolio valuation function and $\bar{t}_0 = 0$, the inductive hypothesis implies that $\mathcal{Y}_{h-1} \subset \mathcal{Y}_h$, which completes the proof. $\qquad\square$

3.2 Quadratic-time solution

Applying Theorem 2 yields an efficient greedy 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 $O(h)$; 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 Lemma 2. Once school k is added to $\mathcal{X}$, we eliminate it from the set $\mathcal{C} \setminus \mathcal{X}$ of candidates, and update the t_j-values of the remaining schools according to (3). 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 (3), 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.

Algorithm 1: Optimal portfolio algorithm for the college application problem.

Input: Utility values $t \in (0, \infty)^m$, admissions probabilities $f \in (0, 1]^m$, application limit $h \leq m$.

1   $\mathcal{C} \leftarrow \{1 \ldots m\}$;
2   $\mathtt{X}, \mathtt{V} \leftarrow$ empty h-arrays;
3   **for** $i = 1 \ldots h$ **do**
4      $k \leftarrow \arg\max_{j \in \mathcal{C}} \{f_j t_j\}$;
5      $\mathcal{C} \leftarrow \mathcal{C} \setminus \{k\}$;
6      $\mathtt{X}[i] \leftarrow k$;
7   **if** $i = 1$ **then** $\mathtt{V}[i] \leftarrow f_k t_k$ **else** $\mathtt{V}[i] \leftarrow \mathtt{V}[i-1] + f_k t_k$;
8   **for** $j \in \mathcal{C}$ **do**
9   **if** $t_j \leq t_k$ **then** $t_j \leftarrow (1 - f_k) t_j$ **else** $t_j \leftarrow t_j - f_k t_k$;
10   **end**
11 **end**
12 **return** $\mathtt{X}, \mathtt{V}$

Theorem 3 (Validity of Algorithm 1). *Algorithm 1 produces an optimal application portfolio for the cardinality-constrained college application problem in $O(hm)$ time.*

Proof. Optimality follows from the proof of Theorem 2. 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 $O(hm)$. $\qquad\square$

By running Algorithm 1 with $h = m$, one can obtain the optimal portfolios for *all* h by taking $\mathcal{X}_h = \{\mathtt{X}[1], \ldots, \mathtt{X}[h]\}$. Thus, a trivial modification of the algorithm yields $O(m^2)$-time solution for Fu (2014)'s variant of the college application problem, stated as

$$\text{maximize} \quad v(\mathcal{X}) - c(|\mathcal{X}|)$$

where $c(\cdot)$ is an increasing cost function. By separability, the optimal portfolio in this problem must be optimal over portfolios of the same size. Hence, it suffices to compute

$$h^* = \underset{h=1\ldots m}{\arg\max} \big\{ \mathtt{V}[h] - c(h) \big\}$$

and return $\mathcal{X}_{h^*}$.

[17] describes two implementations of Algorithm 1 in the Julia language and provides performance benchmarks.

4 Discussion: Heuristics and student welfare

The college application problem possesses a somewhat unusual property that we call *optimal risk elasticity:* Students with large budgets allocate a larger proportion of their portfolio to competitive reach schools than those without. This property implies the inoptimality of heuristic solutions that target a fixed distribution of reach and safety schools in the application portfolio. It also suggests a rational, rather than risk-averse, explanation for why low-income students are relatively unlikely to apply to reach schools. Next, we analyze the shadow cost of the budget constraint and profile the kind of student who stands to benefit most from an intervention, such as an application fee waiver, to eliminate barriers to application.

Throughout the following discussion, reach, target, and safety schools should be understood as relative concepts: A "reach school" is not necessarily an elite university, but rather a school with high utility and low admissions odds relative to the others in a *given* student's input data.

4.1 The distributional heuristic and optimal risk elasticity

The college application problem is relatively easy when the applicant can identify schools that have both high admissions probability f_j and high utility t_j. In typical admissions markets, however, f_j and t_j correlate inversely, because desirable schools attract greater numbers of applicants. Students must make tough dichotomous decisions between schools that have similar expected utility but different risk profiles. Many admissions consultants therefore advise applicants to allocate applications evenly across reach, target, and safety schools [15, 22]. How does this *distributional heuristic* perform compared to the optimal application strategy?

The distributional heuristic does not specify the boundary between target, safety, and reach schools in numerical terms. However, regardless of how this boundary is specified, we argue that the distributional heuristic is limited by the fact that it does not respond adaptively to the application budget h. In a market where f_j and t_j are inversely correlated, the first schools to enter the optimal portfolio in the greedy algorithm are safety schools, and the allocation shifts to incorporate more target and reach schools as the application budget h grows. This phenomenon, which we call *optimal risk elasticity,* implies that any heuristic that targets a *static* risk distribution becomes excessively risk averse for students with a sufficiently capacious application budget. Students with a large budget can tolerate, and indeed benefit, from taking bigger risks than those without.

To see why large values of h favor a riskier portfolio allocation, consider two schools i and j having $t_i < t_j$ and $f_i t_i = f_j t_j = 1$. That is, the expected utility of applying *individually* to either school is identical, and a risk-neutral student is indifferent between the safety school i and the reach school j. Now suppose that the student has resolved to apply to another school, k, and the question is whether to apply to $\mathcal{X} = \{i, k\}$ or $\{j, k\}$. There are three cases:

- If $t_k \leq t_i < t_j$, then

$$v(\{i, k\}) - v(\{j, k\}) = (1 - f_i)f_k t_k + f_i t_i - (1 - f_j)f_k t_k - f_j t_j$$
$$= -f_i f_k t_k + f_j f_k t_k \qquad\qquad < 0$$

because $f_i > f_j$.

- If $t_i \leq t_k \leq t_j$, then

$$
\begin{aligned}
v(\{i,k\}) - v(\{j,k\}) &= (1-f_k)f_i t_i + f_k t_k - (1-f_j)f_k t_k - f_j t_j \\
&= (1-f_k) + f_k t_k - \left(1 - \frac{1}{t_j}\right)f_k t_k - 1 \\
&= -f_k + \frac{1}{t_j}f_k t_k \qquad\qquad\qquad\qquad\qquad \leq 0
\end{aligned}
$$

because $t_k \leq t_j$.

- If $t_i < t_j \leq t_k$, then

$$
v(\{i,k\}) - v(\{j,k\}) = (1-f_k)f_i t_i + f_k t_k - (1-f_k)f_j t_j - f_k t_k \qquad = 0.
$$

In every case, it is (weakly) optimal to choose the reach school j over the safety school i. Thus, however one defines a "target school," it is always better to pair a target school with a reach school than with a safety school as long as the choices are equivalent in their actuarial value. By the elimination argument, one can show that the same result holds with respect to adding a reach or safety school to a portfolio that is otherwise balanced.

4.2 Implications of optimal risk elasticity for econometric research

Our formulation of the college application problem has presumed a risk-neutral student who maximizes her portfolio's expected value without regard to its variance. Is this model complex enough to capture the psychology of real college applicants? Perhaps not: Regression studies of college applicant behavior have shown that low-income students tend to apply to a few safety schools, whereas their wealthy peers tend to apply to more schools and take bigger risks, even when comparing students of similar academic ability [12]. One interpretation of these findings is as part of a larger pattern of greater risk aversion in those from disadvantaged backgrounds, a finding that has been noted in other areas of psychological research [10, 26, 13].

Risk aversion almost certainly factors into real students' decision-making; however, optimal risk elasticity implies that students may *appear* to differ systematically in their risk preferences as an artifact of underlying differences in application resources. For a student with no risk aversion but a small application budget, applying to a few safety schools is often the *rational* strategy. Thus, to the extent that low-income students have smaller application budgets than their wealthy peers, it may be incorrect to attribute variance in students' allocation across reach, target, and safety schools solely or primarily to psychological risk aversion. We propose that future regression studies of applicant behavior control for the number of schools in the application portfolio when comparing students' risk allocation across demographic categories.

Optimal risk elasticity distinguishes the college application problem from classical portfolio optimization models in the [19] tradition, which feature an exogenous risk-aversion parameter. In the Markowitz model, the optimal portfolio is expressed as a proportion of the total budget, and this allocation is constant for budgets of all sizes. Therefore, if an investment adviser knows a client's risk parameter, then she can recommend an optimal portfolio allocation without asking the client's total investment budget. In college application, on the other hand, even among students with *no* risk aversion and identical market parameters f and t, the proportion of the portfolio allocated to safe and risky assets depends on the budget h. Allowing for nonzero risk aversion and variation in the input data only complicates this picture further.

4.3 Diminishing returns to application and equity interventions

Diminishing returns to the size of $\mathcal{X}$ are an intrinsic feature of submodular set functions. The college application problem is no exception, as the following theorem states.

Theorem 4. *For $h = 2 \ldots (m-1)$,*

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

Proof. Applying Theorem 2, we write $\mathcal{X}_h = \mathcal{X}_{h-1} \cup \{j\}$ and $\mathcal{X}_{h+1} = \mathcal{X}_{h-1} \cup \{j, k\}$. By optimality, $v(\mathcal{X}_h) - v(\mathcal{X}_{h-1}) \geq v(\mathcal{X}_{h-1} \cup \{k\}) - v(\mathcal{X}_{h-1})$. By submodularity, $v(\mathcal{X}_{h-1} \cup \{k\}) - v(\mathcal{X}_{h-1}) \geq v(\mathcal{X}_h \cup \{k\}) - v(\mathcal{X}_h) = v(\mathcal{X}_{h+1}) - v(\mathcal{X}_h)$. $\square$

(An elementary proof appears in [17].) It follows that when $\mathcal{X}_h$ is the optimal h-portfolio for fixed values of f and t, $v(\mathcal{X}_h) = t_0 + O(h)$. The following example establishes the tightness of this bound.

Example 1. For a small constant $\varepsilon \in (0,1)$, define the market as follows.

j	1	2	$\cdots$	$m-1$	m
f_j	ε^1	ε^2	$\cdots$	ε^{m-1}	ε^m
t_j	ε^{-1}	ε^{-2}	$\cdots$	$\varepsilon^{-(m-1)}$	ε^{-m}

The optimal portfolios are given by $\mathcal{X}_h = \{m - h + 1, \ldots, m\}$, for which

$$v(\mathcal{X}_h) = \sum_{j=m-h+1}^{m} \left(f_j t_j \prod_{i=j+1}^{m} (1 - f_i) \right) = \sum_{j=m-h+1}^{m} \left(\prod_{i=j+1}^{m} (1 - \varepsilon^i) \right).$$

This expression is bound from above by $\sum_{j=m-h+1}^{m} 1 = h$, and from below by

$$\sum_{j=m-h+1}^{m} \left(\prod_{i=j+1}^{m} (1 - \varepsilon) \right) = \sum_{j=m-h+1}^{m} \left(\prod_{i=j+1}^{m} (1 - \varepsilon)^{m-j} \right) \geq \sum_{j=m-h+1}^{m} (1 - \varepsilon)^m = h(1 - \varepsilon)^m.$$

Thus, as ε approaches zero, we have $v(\mathcal{X}_h) \to h$.

In Example 1, the ability of the student to achieve a valuation linear in h hinges on the fact that the admissions probabilities f_j are very small; when some schools have f_j close to one, the valuation function tends to "flatten out" as these schools are incorporated into the optimal portfolio. This means that, ceteris paribus, students with many reach schools in their portfolio benefit more from the opportunity to apply to an additional college than students with many safety schools.

As mentioned above, high-achieving, low-income students tend to self-select out of applying to competitive universities, and many scholars have proposed interventions such as application fee waivers to encourage these so-called "one-off" students to expand their application portfolio [12, 11]. However, experiments on nudges such as fee waivers and deadline reminders have elicited somewhat underwhelming results: Students who receive nudges did apply to more colleges, but their enrollment patterns did not change relative to a control group [9]. Our analysis of the college application problem suggests that these interventions may limit their efficacy by narrowly targeting high-achieving, low-income students. While the opportunity to apply to more schools benefits all students, the students who stand to gain the most are not necessarily high achievers,

but rather students of middling ability who are nonetheless qualified for college. For these students, every school is (in a relative sense) a "reach school," and each additional application substantially increases the chance of getting into *somewhere*. Therefore, we speculate that broad interventions to decrease the costs of application, such as direct admission (in which colleges offer proactive admission to high school students who meet certain minimal requirements; see [14]), may yield better equity outcomes than nudges targeted to low-income, high-achieving students.

5   Conclusion

This paper has analyzed a novel submodular optimization problem called the college application problem. When optimizing over a cardinality constraint, we have shown that the optimal portfolios satisfy a nestedness property that implies the optimality of a simple greedy algorithm. We found that the problem exhibits optimal risk elasticity: Students with large budgets should apply to more reach schools than students without. This finding may provide a rational explanation for the apparent risk aversion of low-income students in the college application process. Additionally, the sublinearity of the valuation function suggests that broad interventions to remove barriers to college application may promote greater student equity than programs narrowly targeted to low-income, high-achieving students.

There is ample room to extend the college application problem to more complex constraint structures, such as matroids, knapsacks, and their intersections. For example, a knapsack constraint can model the scenario in which colleges have different application fees and the student has a finite monetary budget to spend on applications. [17] shows that the college application problem with a knapsack constraint is NP-complete and provides a fully polynomial-time approximation scheme. Matroidal constraints also appear in real-world admissions markets: In the Korean admissions process, for example, the online application form contains three multiple-choice fields, labeled *ga, na,* and *da* for the first three letters of the Korean alphabet, which students use to indicate the colleges to which they wish to apply. Most schools appear only in one or two of the three fields. Therefore, students are restricted not only in the number of applications they can submit, but in the combinations of schools that may coincide in a feasible application portfolio. From a portfolio optimization perspective, this is a *diversification constraint,* because its primary purpose is to prevent students from applying only to top-tier schools. When the fields *ga, na,* and *da* partition $\mathcal{C}$, is not difficult to see that the set of possible portfolios forms a matroid.

A potential application of our model is an econometric study that estimates students' perceptions of college quality (t_j) from their observed application behavior $(\mathcal{X})$ under the assumption of utility maximization. As discussed in section 4.2, real-world applicant behavior draws on a combination of utility-maximizing and risk-averse strategies. Therefore, adding a variance penalty of the following form to the objective function could improve our model's ability to capture real student behavior:

$$\tilde{v}(\mathcal{X}) = \mathrm{E}\Big[\max\{t_0, \max\{t_j Z_j : j \in \mathcal{X}\}\}\Big] - \gamma \operatorname{Var}\Big[\max\{t_0, \max\{t_j Z_j : j \in \mathcal{X}\}\}\Big]$$

($\gamma > 0$ is the risk-aversion parameter.) Our study has not considered the portfolio optimization problem associated with this more complicated objective, let alone the econometric task of estimating its parameters from applicant behavior.

References

[1]   Mohan S. Acharya, Asfia Armaan, and Aneeta S. Antony. "A Comparison of Regression Models for Prediction of Graduate Admissions". In: *Second International Conference on Computational Intelligence in Data Science*. 2019. DOI: 10.1109/ICCIDS.2019.8862140.

[2]   Itai Ashlagi and Afshin Nikzad. "What Matters in School Choice Tie-Breaking? How Competition Guides Design". In: *Journal of Economic Theory* 190 (2020), article no. 105120. DOI: 10.1016/j.jet.2020.105120.

[3]   Eduardo Azevedo and Jacob Leshno. "A Supply and Demand Framework for Two-Sided Matching Markets". In: *Journal of Political Economy* 124.5 (2016), pp. 1235–68. DOI: 10.1086/687476.

[4]   Aaron Bodoh-Creed. "Optimizing for Distributional Goals in School Choice Problems". In: *Management Science* 66.8 (2020), pp. 3657–76. DOI: 10.1287/mnsc.2019.3376.

[5]   Shant Boodaghians et al. "Pandora's Box Problem with Order Constraints". In: *Mathematics of Operations Research Articles in Advance* 0.0 (2022), null. DOI: 10.1287/moor.2022.1271.

[6]   Ivan Contreras and Elena Fernández. "Hub Location as the Minimization of a Supermodular Set Function". In: *Operations Research* 62.3 (2014), pp. 557–70. DOI: 10.1287/opre.2014.1263.

[7]   Chao Fu. "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market". In: *Journal of Political Economy* 122.2 (2014), pp. 225–81. DOI: 10.1086/675503.

[8]   David Gale and Lloyd Shapley. "College Admissions and the Stability of Marriage". In: *American Mathematics Monthly* 69.1 (1962), pp. 9–15. DOI: 10.2307/2312726.

[9]   Oded Gurantz et al. "A National-Level Informational Experiment to Promote Enrollment in Selective Colleges". In: *Journal of Policy Analysis and Management* 40.2 (2021), pp. 453–79. DOI: 10.1002/pam.22262.

[10]   Vanessa Hartlaub and Thorsten Schneider. "Educational Choice and Risk Aversion: How Important Is Structural vs. Individual Risk Aversion?" In: *SOEPpapers on Multidisciplinary Panel Data Research*. 433. 2012. URL: https://www.diw.de/documents/publikationen/73/diw_01.c.394455.de/diw_sp0433.pdf.

[11]   Caroline Hoxby and Sarah Turner. *Expanding college opportunities for high-achieving, low income students*. Tech. rep. 014. Stanford Institute for Economic Policy Research, 2013, p. 7.

[12]   Caroline M. Hoxby and Christopher Avery. *The missing "one-offs": The hidden supply of high-achieving, low income students*. Tech. rep. National Bureau of Economic Research, 2012. DOI: 10.3386/w18586.

[13]   Thomas van Huizen and Rob Alessie. "Risk Aversion and Job Mobility". In: *Journal of Economic Behavior & Organization* 164 (2019), pp. 91–106. DOI: 10.1016/j.jebo.2019.01.021.

[14]   Scott Jaschik. "Direct Admissions Takes Off". In: *Inside Higher Ed* (July 18, 2022). URL: https://www.insidehighered.com/admissions/article/2022/07/18/direct-admissions-takes-individual-colleges.

[15] Minhee Jeon. "College application strategy: Six chances total. . . divide applications across reach, target, and safety schools". In: *Jungang Ilbo* (Aug. 26, 2015). URL: https://www.joongang.co.kr/article/18524069.

[16] Max Kapur. "Characterizing Nonatomic Admissions Markets". In: *ArXiv* (July 3, 2021). URL: https://arxiv.org/abs/2107.01340.

[17] Max Kapur. "The College Application Problem". MA thesis. Seoul National University, 2022. URL: https://github.com/maxkapur/CollegeApplication.

[18] Daniel Kibum Lim. "A Simulation Approach to Predicting College Admissions". MA thesis. University of California Los Angeles, 2013. URL: https://escholarship.org/uc/item/8r2695n5.

[19] Harry Markowitz. "Portfolio Selection". In: *The Journal of Finance* 7.1 (1952), pp. 77–91. URL: https://www.jstor.org/stable/2975974.

[20] George Nemhauser and Laurence Wolsey. "Best Algorithms for Approximating the Maximum of a Submodular Set Function". In: *Mathematics of Operations Research* 3.3 (1978), pp. 177–88. DOI: 10.1287/moor.3.3.177.

[21] George Nemhauser, Laurence Wolsey, and Marshall Fisher. "An Analysis of Approximations for Maximizing Submodular Set Functions—I". In: *Mathematical Programming* 14 (1978), pp. 265–94. DOI: 10.1007/BF01588971.

[22] Timothy Peck. *Safety, Target, & Reach Schools: How to Find the Right Ones.* 2021. URL: https://blog.collegevine.com/the-college-list-decoded-safeties-targets-and-reaches/.

[23] Alvin E. Roth. "The Economics of Matching: Stability and Incentives". In: *Mathematics of Operations Research* 7.4 (1982), pp. 617–28. URL: https://www.jstor.org/stable/3689483.

[24] Mark Rozanov and Arie Tamir. "The Nestedness Property of the Convex Ordered Median Location Problem on a Tree". In: *Discrete Optimization* 36 (2020), article no. 100581. DOI: 10.1016/j.disopt.2020.100581.

[25] Mark Sklarow. *State of the Profession 2018: The 10 Trends Reshaping Independent Educational Consulting.* Tech. rep. Independent Educational Consultants Association, 2018. URL: https://www.iecaonline.com/wp-content/uploads/2020/02/IECA-Current-Trends-2018.pdf.

[26] Tyler Watts, Greg J. Duncan, and Haonan Quan. "Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Delay of Gratification and Later Outcomes". In: *Psychological Science* 29.7 (2018), pp. 1159–77. DOI: 10.1177/0956797618761661.

[27] Martin Weitzman. "Optimal Search for the Best Alternative". In: *Econometrica* 47.3 (1979), pp. 641–54. DOI: 10.2307/1910412.