Simultaneous Search[*]

Hector Chade[†]

Arizona State University
Department of Economics
Tempe, AZ 85287-3806

Lones Smith[‡]

University of Michigan
Department of Economics
Ann Arbor, MI 48109-1220

this version: June 2004[§]

Abstract

We introduce and solve a new class of static portfolio choice problems, where only the best realized alternative matters. A decision maker must simultaneously choose among independent ranked options, and the better alternatives have a lower chance of panning out. Each choice is costly, and just one option may be exercised. This often emerges in practice:

- A student must make a costly and simultaneous application to many colleges, and is accepted with smaller chances by the better schools.
- An economics department must decide which of several PhD job candidates to fly out, and the better recruits will be available with smaller probability.

We show that such portfolio choice problems quite generally entail maximizing a submodular function of finite sets — which is NP hard in general. Still, we develop a marginal improvement algorithm that produces the optimal set for our binary option structure in a quadratic number of steps. Applying it, we then show that the optimal choices are less risky than the sequentially optimal ones in Weitzman (1979), but riskier than the best singleton college choices. We also give practical rules of thumb, such as: (*i*) don't insure, choosing a safety school; instead, take risks — unless success rates are positively correlated; (*ii*) apply to an upwardly diverse portfolio of schools. We also provide comparative statics on the chosen set.

---

[*]The usage of the term 'search' rather than 'choice' here reflects a precedent set in Weitzman (1979), and in the directed search literature. We have benefited from seminars at the Midwest Economic Theory Meetings, ITAM, LBS, Penn, Duke, Michigan, and the Toronto Matching Conference. We are very grateful for research assistance and nontrivial insights of Kan Takeuchi, and the feedback of Miles Kimball and Steve Salant. Lones is grateful for the financial support of the National Science Foundation.

[†]Email: `hchade@asu.edu`

[‡]Email: `lones@umich.edu`; web: `www.umich.edu/~lones`.

[§]The first widely circulated version was August 2003.

1 Introduction

We introduce and solve a new class of static portfolio choice problems, where only the best realized alternative matters. A decision maker (DM) must choose from among many ranked options, better alternatives have a lower chance of panning out, and only one option may be exercised. We completely solve two search problems where a portfolio of options is chosen *simultaneously* once-and-for-all: (*i*) the DM can choose sets of any size, and each selection is costly, and (*ii*) the DM is restricted to choice sets of a given size.

Our paper generalizes the first 1961 search theory paper by Stigler who describes the optimal fixed size wage search. But unlike Stigler, we do not assume *a priori* identical prizes, and thus characterize not only the optimal search size but also the optimal *choice composition*. As such, we provide a foundation for problems in *directed search*.

Weitzman (1979) also explored a problem with *a priori* distinct prizes — but in the sequential world. Weitzman's method is a nice application of Gittins' solution of the bandit problem. Each option can be assigned an index in isolation of all others; the selection rule is simply to choose the unexplored option with the highest index. In our paper, no such simple index rule presents itself. Instead, we find ourselves faced with the maximization of a submodular function of sets of alternatives; formally equivalent to maximizing a convex function, this problem is known to be NP hard, in general. Nevertheless, we introduce an economically natural algorithm that produces the optimal set in quadratically many steps. Since submodular maximization is intractable, this algorithm alone allows us to derive all of our characterization results.

We explore basic properties of the optimal set. A key question we ask is how much of a risk taker one should be. It is frequently mused, for instance, that one should include an 'insurance option' or 'safety school'. We show that unless the options have correlated success rates, this conventional wisdom is false. The optimal choice set is riskier than the one consisting solely of the best options evaluated in isolation. For instance, if a tier two school is your best solo possibility, your optimal application portfolio cannot include tier three schools. On the other hand, we show that the optimal static choices are less risky than the sequential choices dictated in Weitzman (1979).

We also ask how varied should the choices be. We argue in favor of an upwardly diverse portfolio: For a rich enough array of possible colleges, a connected 'interval' of similarly-risky prospects is not optimal. With low enough costs, individuals should eventually add a distinctly riskier option to their choice set. We also provide some nonstandard comparative statics, showing how the choice set improves when acceptance chances rise, and the acceptance chances of better alternatives rise proportionately more.

We believe that our problem is not without substantive practical value.

Example 1. A student must make a costly and simultaneously application to several colleges, and is accepted with smaller chances by the better schools.

Example 2. An economics department must fly out several new PhD job candidates; the fly-outs are costly. Each school ranks the job candidates, and schools have correlated preferences. Thus, better candidates are harder to hire. To map the decision problem confronting each school into our simple framework, assume that (*i*) fly-outs do not inform the hiring decision; (*ii*) each department simply wishes to hire a single job candidate; (*iii*) after the fly-out stage, the market clears top to bottom (the best schools choose first), so that the better recruits are available with smaller chance to any school below the top.

Example 3. Akin to Weitzman's inspirational application, a research department of a large firm wishes to choose a technology; several are available, and all are costly to explore; some will work out, and others will not. Finally, it is in a hurry (e.g., it is in a race with other firms), and must simultaneously choose which technologies to explore.

Our algorithm yields the following practical advice to aspiring college applicants: Identify your best singleton school, and apply to it. Next figure out which school adds most value to this sample size one coalition, and ask "Does the marginal increment pay for itself?" Continue, until the answer is 'no'. We think that this economically natural algorithm is new, and is the first recipe that identifies the optimal choice set for a class of submodular maximization problems other than linear programming.

While our paper is very much of a sequel to Weitzman's work, it may more topically be viewed by search economists as a foundation work towards the recent literature on directed search.[1] This line of work seeks to improve upon purely random matching, by allowing, for instance, employees to choose which jobs to apply to. Our paper solves this decision problem for multiple applications and heterogeneous jobs.[2]

[1] A recent salient contribution here is Burdett, Shi, and Wright (2001).

[2] See Albrecht, Gautier, and Vroman (2002). Perhaps the first equilibrium paper with multiple simultaneous searches is Burdett and Judd (1983).

Finally, we offer a key theoretical perspective on our contribution. Optimization of submodular set functions has appeared in the economics literature from time to time. Our contribution, however, turns on an absolutely critical difference. We study the open problem of *maximization* of a submodular set-function. The existing literature, for instance, Gul and Stacchetti (1999) and Kelso and Crawford (1982), relies on the now understood problem of the *minimization* of submodular set functions (Fujishige 1991). We argue that our problem is formally the maximization of a convex function in $\mathbb{R}^N$, whose complexity should be manifest. Absent a single tool to grasp onto — for instance, methods developed in monotone comparative statics are worthless for a submodular maximization — our proofs are mostly via induction or application of our algorithm. We defer a discussion of **?**, who studies a related problem, until §4.2-C.

While this paper will proceed for definiteness using the example and language of the college admissions problem, the applications to other problems must be kept in mind.[3] We first describe the search problem and give its sequential solution. We then provide our static algorithm and prove its optimality. We then explore the properties of the optimal set: Do students insure themselves or gamble? Are their choices similar, or disparate? What if success rates are correlated? The appendix collects the longer proofs.

2 The Model and Two Competing Approaches

A. The College Problem. A student must choose once and for all a portfolio S of colleges from $\{1, 2, \ldots, N\}$ to which she wants to apply for admission. Here, N is a natural number, but in an abuse of notation, we denote this set by N too, and its subsets by 2^N. We assume either a fixed sample size, or a fixed application cost $c > 0$.

The best college is 1, the second best 2, and so on. The student's cardinal utility from going to college i is u_i, where $u_1 > u_2 > \cdots > u_N$. His chance of being admitted to college i is $\alpha_i \in (0, 1]$.[4] We shall often posit the rather intuitive inverse ordering $\alpha_1 < \alpha_2 < \cdots < \alpha_N$, but this is inessential for most of our results. The acceptance

[3] To be precise, we hit this problem when trying to solve an equilibrium version of the College Admissions problem, with uncertainty. That paper (Chade, Lewis, and Smith 2003) is a work in progress, and uses the general results of this paper.

[4] In Chade, Lewis, and Smith (2003), these chances α_i arise naturally as the probabilities that the signal of the quality of a student lies above the standard of college i.

decisions by any group of colleges are independent. In the college example, one can rationalize this as follows: Colleges perceive noisy conditionally iid signals of a student's type, and a student fully knows his true type. We sometimes explore the problem with correlated acceptance chances, where one does not fully know one's type.

Denote by $z_i = \alpha_i u_i$ the *unconditional expected payoff* of applying *just* to college i. Given an application cost, we assume $z_i > c$ for all i, pruning weakly dominated colleges.

B. The Sequential Choice Problem. Consider first the case where a student can apply to the colleges sequentially, observing whether one accepts her before she applies to the next. In which order should she apply, given the cost $c > 0$? The optimal policy in Weitzman (1979) — which coincides with the Gittins index for bandits (Gittins and Jones 1974) — is derived as follows for our context. To each college i, associate an intrinsic *index* or reservation value I_i; this leaves the student indifferent between a terminal payoff I_i, and applying to college i and then earning payoff I_i if rejected:

$$I_i = z_i - c + (1-\alpha_i)I_i \qquad \Rightarrow \qquad I_i = (z_i - c)/\alpha_i = u_i - c/\alpha_i \tag{1}$$

The optimal policy orders the colleges by their indices I_i; the student starts at the colleges with the maximal I_i, moving down the list until one accepts him.

C. The Objective Function for the Static Optimization. We are concerned in this paper with a one-shot decision. First of all, let $f(S)$ be the expected *gross value of school subset* S, ignoring application costs. Working recursively, one either gets into the best college, or one does not; if rejected, one either gets into the next best, or not, etc. Let $\rho(S) \equiv \Pi_{i\in S}(1-\alpha_i)$ be the probability of being rejected by all colleges in the set S. Thus, we compute a few decompositions for the *gross payoff* as follows:

$$\begin{aligned} f(S) &= \textstyle\sum_{i=1}^{|S|} z_{(i)} \prod_{\ell=0}^{i-1}(1-\alpha_{(\ell)}) = \sum_{i=1}^{|S|} z_{(i)}\rho_{(i-1)}(S) && (2)\\ \Rightarrow f(U+L) &= f(U) + \rho(U)f(L) \qquad \text{for all sets } U \text{ above } L \text{ in } N && (3) \end{aligned}$$

where (i) is the i-th best ranked college in the set S, so that $z_{(i)} \equiv \alpha_{(i)}u_{(i)}$, and $\rho_{(i-1)}(S) = \prod_{\ell=0}^{i-1}(1-\alpha_{(\ell)})$ is the chance of being rejected by the top-ranked $i-1$ schools in set S.

The (net) *value of applying to the set S of schools* is $v(S) \equiv f(S) - |S|\, c$. Applying nowhere is worthless: $v(\varnothing) = 0$. We consider two problems. First, for fixed $c > 0$, maximize the net value:

$$\max_{S \subseteq N} v(S) \equiv f(S) - |S|\, c \tag{4}$$

Let $S^*(c) = \arg\max_{S \subseteq N} v(S)$ be the optimal set. Second, for fixed sample size $n \in N$, maximize the gross payoff:

$$\max_{S \subseteq N \& |S| = n} f(S) \tag{5}$$

Denote the optimal set for (5) by $S_n^* = \arg\max_{S \subseteq N} f(S)$. Since all applications are equally costly, for any $c > 0$, if $n = |S^*(c)|$, then $S_n^* = S^*(c)$. Call this set S^*.

C. Consistency Checks on the Optimal Set. Computing the optimal set turns out to be very hard, but we are able now to provide two useful tests that it must obey.

Lemma 1 *Let the optimal portfolio be $S_n^* = \{j_1, \ldots, j_n\}$, where $j_1 < \cdots < j_n$. Then:*
(a) For all k, $\{j_k, \ldots, j_n\}$ is an optimal $(n - k + 1)$ selection from[5] $[j_k, N]$.
(b) Assume $z_i > z_j$ and $\alpha_i < \alpha_j$. If $j \in S^*$, then $i \in S^*$.

Proof Part (a) follows from the 'downward recursive' structure of (2), as seen in (3).

Part (b) is intuitive: College i not only has a higher expected payoff but also yields a higher rejection externality for lesser schools. Formally, since $i < j$, we may write $S^* = U + j + M + L$ and $S^* - j + i = U + M + i + L$, where (upper set) $U = [1, j) \cap S^*$, (middle set) $M = (i, j) \cap S^*$, and (lower set) $L = (i, N] \cap S^*$. Consider the following *suboptimal* implementation policy for $S^* - j + i$: Accept the best available option, unless it is i, in which case accept the best opportunity in M (if available) over i. So by (3),

$$\begin{aligned} f(S^* - j + i) &\geq f(U) + \rho(U)\,(f(M) + \rho(M)[z_i + (1 - \alpha_i) f(L)]) \\ &> f(U) + \rho(U)\,(f(M) + \rho(M)[z_j + (1 - \alpha_j) f(L)]) = f(S^*) \end{aligned}$$

where we use the facts that $\rho(U)\rho(M) > 0$ and $z_i + (1 - \alpha_i) f(L) > z_j + (1 - \alpha_j) f(L)$. Hence, we have shown that if $i \notin S^*$ then $j \notin S^*$. □

[5] We hereby introduce the suggestive notation (i, j) for the set of colleges ranked strictly between i and j; similarly, define $(i, j]$. Thus, $[1, j) \cap S$ is the set of all colleges in S that are better than j.

3 Portfolio Choice Problems are Submodular

Let's consider a generalization of the problem (4), for a stronger result. Assume N options, each characterized by a c.d.f. G_i, $i = 1, 2, \ldots, N$, over the set of prizes that i can deliver. Let $[0, \bar{u}]$ contain the union of the supports of all G_i. The decision maker chooses a subset $S \subseteq N$ at a cost c per option included in S; after the choice is made, he draws a prize from each G_i, $i \in S$; and chooses the maximum of the realized prizes:

$$\max_{S \subseteq N} v(S) \equiv f(S) - |S|c = \int_0^{\bar{u}} \left(1 - \prod_{i \in S} G_i(u)\right) du - c|S|. \tag{6}$$

This formula owes to the well-known fact that the expected value of a nonnegative random variable equals the integral of the survivor distribution.

Lemma 2 (Value Submodularity) *The function* $v : 2^N \mapsto \mathbb{R}$ *in* (6) *is submodular.*

Proof We need to show that for any two subsets S and T of N, $v(S \cap T) + v(S \cup T) \leq v(S) + v(T)$. A simpler characterization (Proposition 1.1 in Lovász (1982)) is *diminishing returns*: for any $j \notin S \subset N$, the function $v(S + j) - v(S)$ is decreasing in S.[6, 7]

First, $-c|S|$ is a modular, and thus submodular, function. Thus, it suffices to show that $f(S + j) - f(S)$ is decreasing in S. But this is clear, as $\prod_{i \in S} G_i(u)$ falls in S, and

$$f(S+j) - f(S) = \int_0^{\bar{u}} (1 - G_j(u)) \prod_{i \in S} G_i(u) du. \qquad \Box$$

The intuition is simply that with more options, any addition to the current portfolio is less valuable. As noted, this 'diminishing returns' suffices to establish submodularity.

By Lemma 2, problem (6) entails the maximization of a submodular function over the set of subsets of N. For pedagogical reasons, we provide a brief simple intuition from §4 of Lovász (1983) for the inherent difficulty of maximizing submodular functions. Associate to the set function f another function $\hat{f}$ on 0-1 vectors in $\mathbb{R}^N$ as follows: For each $S \subset N$, let a^S be a vector of 0's and 1's, with $a_i^S = 1$ if and only if school i belongs to S. First

[6] We will use the shorter and suggestive notation $S + j = S \cup \{j\}$ and $S - j = S \setminus \{j\}$.

[7] This property has most recently appeared in the economics literature in Gul and Stacchetti (1999). See also related work by Kelso and Crawford (1982) on the gross substitutes condition.

set $\hat{f}(a^S) = f(S)$. Next, given 0-1 vectors $a_1 \geq a_2 \geq \cdots \geq a_N$, uniquely express any $b \in \mathbb{R}^N$ as $b = \sum_{i=1}^N \lambda_i a_i$, for scalars $\lambda_i > 0$. Finally, define $\hat{f}(b) = \sum_{i=1}^N \lambda_i \hat{f}(a_i)$.

Proposition *The extension $\hat{f}$ is convex on $\mathbb{R}^N$ if and only if f is submodular on 2^N.*

In other words, our problem is tantamount to the *maximization of a convex function* on the space $[0,1]^N$. The difficulty of our exercise now should be apparent. The argmin of a convex function is a convex set, and more generally, the argmin of a submodular function is a well-behaved sublattice. However, the argmax of a convex function has no known regularity properties, nor does the argmax of a submodular function. Intuitively, one expects a corner solution — but which corner? (There are 2^N corners.) By exploiting the special functional form of our objective function v, we next provide an algorithm that quickly finds the optimal portfolio S^*.

4 An Optimal Marginal Improvement Algorithm

4.1 The College Problem Solved

The *Marginal Improvement Algorithm* (MIA) identifies a choice set S^* via a simple inductive procedure as follows:

Step 1 *Let $i_1 = \arg\max_{i \in N} \; f(\{i\})$. Then $S_1 = \{i_1\} \subseteq S^*$.*

Step n *Let $i_k = \arg\max_{i \in N - S_{k-1}} \; f(S_{k-1} + i) - f(S_{k-1})$, for $k = 1, \ldots, n-1$. Fix $S_{n-1} = \{i_1, \ldots, i_{n-1}\} \subseteq S^*$. Define $i_n = \arg\max_{i \in N \setminus S_{n-1}} \; f(S_{n-1} + i) - f(S_{n-1})$.*

Stopping Rule for (4) *If $f(S_{n-1} + i_n) - f(S_{n-1}) > c$, then $i_n \in S^*$. Otherwise, stop.*

In other words, one first identifies the college $i = i_1$ whose unconditional expected payoff z_i is largest. In the induction step, one finds the college i_n that affords the largest marginal benefit over the college set constructed so far. For problem (4) with an application cost, one stops if the *net* marginal benefit turns negative.

Theorem 1 *The MIA identifies the optimal set for problem* (5). *With the stopping rule, it identifies the optimal set for problem* (4).

As we mentioned, this problem is, in general, NP-hard.[8] One must in principle calculate the values of all 2^N college application patterns. Yet our algorithm works in polynomial time. Initially, the student examines N application patterns and then finds the best college to apply first. At the second round, she examines $N-1$ patterns and finds the second college, etc. At most, she will examine a total of $N + (N-1) + \cdots + 2 + 1 = N(N+1)/2$ patterns. We thus reduce the difficulty of the problem from $O(2^N)$ to $O(N^2)$.

The optimality of the MIA is shown in the appendix, but since it is a key result, we will summarize our proof, which proceeds by policy improvement. Specifically, for Step 1, it suffices to show that for any given nonempty subset $S \subset N$ that does not contain i_1, there is another set S' that includes i_1 and dominates S. But which set S' works? Our proof works by policy improvement. We have discovered that if there is some *lower-ranked* college than i_1 in S, then it is optimal to replace it by i_1. Otherwise, if there are no colleges worse than i_1 in S, then it is optimal to *replace the worst better-ranked school* w by i_1. The argument for the induction step is analogous.

An example illustrates the importance of choosing w. Assume three colleges, with $\alpha_1 = 0.1$, $\alpha_2 = 0.9$, $\alpha_3 = 1$, $u_1 = 1$, $u_2 = 0.5$, $u_3 = 0.48$. Notice that $z_3 = 0.48 > z_2 = 0.45 > z_1 = 0.1$. Hence, $3 \in S^*$. In order to prove it, we must show that $\{1, 2\}$ is dominated by a set that contains college 3. Here college 2 plays the role of w. It is easy to show that $\{1, 3\}$ dominates $\{1, 2\}$ but that $\{2, 3\}$ does not.

Note that the true marginal benefit $v(S + i) - v(S)$ of a school i to a portfolio S lies below $v(i) - v(\varnothing) = z_i - c$, whenever $S \neq \varnothing$, by Lemma 2. So the naive rule of applying to all schools N whose expected payoffs exceed their cost c yields too many applications.

Related Algorithms. We believe that the marginal improvement algorithm has not already been explored. This perhaps speaks to the foundation of this literature in combinatorics and graph theory. We have found two main classes of algorithms. The *Greedy Algorithm* successfully solves problems having a linear programming structure: Specifically, assume we wish to maximize bx over all $x \in \mathbb{R}^n$ obeying $\sum_{r \in S} x_r \leq f(S)$ across all subsets $S \subset N$, where $b \geq 0$. This algorithm first ranks the coefficients $b_{i_1} \geq \cdots \geq b_{i_n}$, and then selects $x_{r+1} = f(S^{r+1}) - f(S^r)$, where $S^{r+1} = S^r + i_{r+1}$. Note

[8]So as a function of the number of options, it is not solvable in a polynomial number of steps. Consult the inviting survey article Lovász (1983).

that our submodular optimization problem cannot be written in the linear programming form, due to the multiplicative spill-overs across college choices.[9]

The *Dichotomy Algorithm* applies to more general submodular optimizations over sets, and thus is relevant for our context. This recipe narrows the optimal set from the inside and outside at each stage, shrinking the 'interval' $[S, T]$ of possible schools (sets again ordered by inclusion) that contain the optimal set. We can produce an example in our context where it stops before hitting the optimal set.[10]

4.2 Boundaries of the Marginal Improvement Algorithm

While the MIA may be 'intuitive', any success at all for a submodular maximization problem is remarkable. Does it extend beyond the environment with binary success-or-failure prizes? We first provide a different class of portfolio choice problems with general prize distributions ordered by stochastic dominance where the MIA succeeds. We then temper this success by showing that the MIA fails in a simple enrichment of the success-or-failure binary prize context, or with highly correlated chances.

A. Prize Distributions Ordered by Stochastic Dominance. Our appendicized proof of the optimality of the MIA exploits the property that $\prod_{i \in S - S_{n-1} - w} G_i(u)$ is constant on $[0, w]$. This is not true in general. This property is invoked whenever, at stage n, there is no worse-ranked school than i_n not yet selected. Therefore, a natural starting point is to look for a class of problems in which the algorithm chooses 1 first, and proceeds through $2, 3, \ldots$ until the marginal benefit falls below the application cost.

We now show that the algorithm extends to general prize distributions ordered by stochastic dominance — first order or a quasi-second order one.

Theorem 2 *If* $\int_x^{\bar{u}} G_i(u)du \leq \int_x^{\bar{u}} G_{i+1}(u)du$, $i = 1, \ldots, N-1$, $x \in [0, \bar{u}]$, *with strict inequality at* $x = 0$, *then the MIA is optimal and* $S^* = [1, k]$, *for some* $k \leq N$.

Vishwanath (1992) uses the same condition to show that Weitzman's (1979) solution still holds when, at each stage, more than one option can be tried. It is easy to see that

[9] While this too uses margins, it is not a marginal improvement algorithm. Rather, it chooses controls so that they are 'assortatively mated' to the highest coefficients.

[10] See Goldengorin, Tijssen, and Tso (1998) for a precise statement of this algorithm.

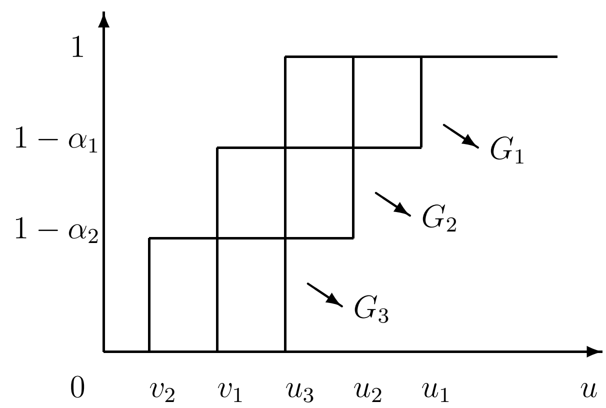

Figure 1: **General Binary Prizes.**

the index I_i of an option solves $c = \int_0^{\bar{u}} \max\{0, u - I_i\} dG_i(u)$; the right side is decreasing in i as $\max\{0, u - I_i\}$ is increasing and convex in u. So $I_i > I_{i+1}$ for all i.

In our success-or-failure binary prize context, the premise of Theorem 2 reduces (using $x = 0$) to $z_1 > z_2 > \cdots > z_n$. This immediately yields an interval rule for those students for whom Harvard is the unconditionally expected best college, etc.[11] Such students include those at the 'top', for whom acceptance rates fall proportionately less than the college payoffs, at better schools. They should just apply to the top schools:

Corollary 1 *Let expected payoffs be ranked $z_1 > \cdots > z_N$. Then the optimal set S^* is an interval around the best school* 1, *for any cost $c > 0$.*

B. General Binary Prizes and Costs. Is the MIA optimal for *any* simultaneous search problem with independent options? The next example shows that the answer is negative. Assume three independent options, $i = 1, 2, 3$, each with two positive prizes, $u_i > v_i > 0$, with an α_i chance of u_i. Thus, $v(i) = \alpha_i u_i + (1-\alpha_i)v_i - c$, assumed positive.

Let $c = 4.9$, $u_1 = 100 > u_2 = 80 > u_3 = 55.4 > v_1 = 50 > v_2 = 49 > v_3 = 0$, and $\alpha_1 = 0.1 < \alpha_2 = 0.2 < \alpha_3 = 1$. (See Figure 1.) Then $v(\{1\}) = 55.0 - 4.9 = 50.1$, $v(\{2\}) = 55.2 - 4.9 = 50.3$, $v(\{3\}) = 55.4 - 4.9 = 50.5$, so that the algorithm chooses

[11] Readers from Princeton, Yale, etc. should assume a typo, and read their school's name here.

college 3 in the first step. Notice, however, that $3 \notin S^* = \{1,2\}$, for

$$\begin{aligned}
v(\{1,2\}) &= \alpha_1 u_1 + (1-\alpha_1)\alpha_2 u_2 + (1-\alpha_1)(1-\alpha_2)v_1 - 2c = 60.40 - 9.8 = 50.60 \\
v(\{1,3\}) &= \alpha_1 u_1 + (1-\alpha_1)\alpha_3 u_3 + (1-\alpha_1)(1-\alpha_3)v_1 - 2c = 59.86 - 9.8 = 50.06 \\
v(\{2,3\}) &= \alpha_2 u_2 + (1-\alpha_2)\alpha_3 u_3 + (1-\alpha_2)(1-\alpha_3)v_2 - 2c = 60.32 - 9.8 = 50.52 \\
v(\{1,2,3\}) &= \alpha_1 u_1 + (1-\alpha_1)\alpha_2 u_2 + (1-\alpha_1)(1-\alpha_2)u_3 - 3c = 64.28 - 14.7 = 49.58
\end{aligned}$$

We omit an example showing that the restriction to identical application costs is also important: The MIA can be suboptimal for sufficiently disparate application costs c_i.

C. Correlated Values. We now very briefly discuss relaxing the assumption that the acceptance decisions by any group of colleges are independent. Specifically, we suppose that students do not fully know their type: Some common unobserved attribute of one's type affects one's admission prospects into all schools. Colleges still observe noisy conditionally iid signals of a student's type; however, from a student's unconditional perspective, the acceptance events are positively correlated; therefore, getting into any school weakly raises the admission chance to any other school.

Section 5.1-A gives an example with imperfect correlation, but where the MIA works. In general, the MIA is defined by inequalities, and so by continuity, a small amount of correlation does not affect the optimality of a strictly optimal set — barring ties. Here, seeking a generic failure, we consider an extreme world with perfect correlation, where a student's admission to college i assures his admission to college $i+1$. In this case,

$$f(S) = \sum_{i=1}^{|S|} \Pr(\text{rejected by } (i-1), \text{ accepted by } (i)) u_{(i)} = \sum_i (\alpha_{(i)} - \alpha_{(i-1)}) u_{(i)}$$

Now, assume college payoffs $u_1 = 11, u_2 = 4, u_3 = 1$, and admission chances $\alpha_1 = 0.1, \alpha_2 = 0.3, \alpha_3 = 1$. Then the MIA chooses $i_1 = 2$ (maximal z_i), which does not belong to the best sample size-two set $\{1,3\}$, with value $f(\{1,3\}) = \alpha_1 u_1 + (\alpha_3 - \alpha_1)u_3 = 2$.

? studies another world with perfect correlation, which is different in many ways. She assumes a continuum of colleges, each perfectly informed of the students' types; the students' are only partially informed, with normally distributed beliefs. Her main results are local characterizations of comparative statics on the optimal policy.

5 Properties of the Optimal Set

5.1 Riskiness of the Optimal Choices

We first address head-on the main stylized street insight into this problem: Insure yourself. Add that (worse) safety school. Is this wise? Or should students 'gamble'?

First, what does it mean to gamble or insure? To flesh this out, we employ *vector* first order stochastic dominance (FSD). The set $S \subseteq N$ is *riskier*[12] than the same-size set $S' \subseteq N$ in the sense of FSD when $s_i \leq s'_i$ for all i, where s_i is the ith best school in S and s'_i in S'. Write this as $S \succeq S'$, and as $S \succ S'$ if also $S \neq S'$. Thus, $\{1, 2\} \succ \{2, 3\}$.

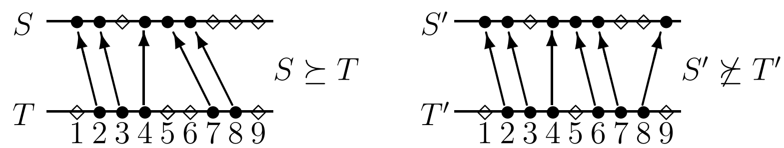

Figure 2: **Riskiness Order on Sets of Colleges.** At the left are two ordered sets of colleges, $S \succeq T$ — as reflected in the non-positive slopes of the vectors joining corresponding first, second, etc. order statistics. At the right, we have added one college to each set, and now $S' \not\succeq T'$. Such pictures underlie our proofs of Theorems 3, 4, and 6.

We now consider two compelling comparison sets. Firstly, there are the colleges $Z_n \subset N$ with the n highest unconditionally expected payoffs $z_i = \alpha_i u_i$. Unlike the optimal portfolio S_n, this set ignores the subtle web of cross college external effects. Secondly, there are the n colleges $W_n(c) \subset N$ with the highest indices I_i; these are *chosen initially by Weitzman's sequentially-optimal rule.* We establish the twin sandwich inequalities $W_n(c) \succeq S_n$ and $S_n \succeq Z_n$. Comparing S_n to these sets is natural because they are (a) easily computed benchmarks, and (b) transparently ordered. Indeed:

Lemma 3 *For any application cost $c > 0$, $W_n(c) \succeq Z_n$ for all n.*

Proof The ranking can only fail if for some $i < j$, the higher ranked school i is unconditionally better, or $z_i \geq z_j$, but yet is sequentially searched last, given $I_j \geq I_i$. Thus,

$$\frac{z_j - c}{\alpha_j} \geq \frac{z_i - c}{\alpha_i}$$

[12]Typically, SSD is associated with riskiness. To clarify, we are using riskiness here in the street sense of being less likely to succeed. For instance, it is riskier to apply to $\{1, 2\}$ than $\{2, 3\}$.

But this inequality is obviously impossible, given $\alpha_j > \alpha_i > 0$ and $z_i \geq z_j > c$. $\square$

This is consistent with the message of Weitzman (1979), that dynamic optimization encourages risk-taking. But we now argue that even static optimization of our portfolio variety encourages risk taking in the exact same sense, but to a slightly muted extent.

A. Portfolio Choices are Riskier than Top Singletons. Consider for a moment a three college world. Assume that the expected payoffs obey $z_2 > z_1 > z_3$, so that $Z_2 = \{1, 2\}$. Then weakly $S_2 = \{1, 2\} \succeq Z_2$. Indeed, for a low enough $c > 0$, the algorithm first calls for an application to college 2. At that point, adding college 3 cannot be optimal, because if it were, it would confer greater marginal benefit than 1:

$$(1-\alpha_2)z_3 \geq z_1 - \alpha_1 z_2 \Leftrightarrow (1-\alpha_2) \geq (z_1/z_3)[1-\alpha_1(z_2/z_1)] > 1-\alpha_1(z_2/z_1) \Leftrightarrow u_2 > u_1$$

which is false, as college 2 is lower ranked than college 1, and yields a lower payoff. A strict ordering $S_2 \succ Z_2$ is impossible if $Z_2 = \{1, 2\}$. But let's add a new top college $0 \succ 1$ with $z_0 < z_2$. Then the proof of $\{1, 2\} \succ \{2, 3\}$ still obtains. But now $S_2 = \{0, 2\} \succ Z_2$ when $\{0, 2\} \succ \{1, 2\}$ — namely, if $z_0 - \alpha_0 z_2 > z_1 - \alpha_1 z_2 \Leftrightarrow z_0 - z_1 > (\alpha_0 - \alpha_1) z_2$. This is true for $z_0 - z_1$ big enough. We now state the general result (proof is in the appendix).

Theorem 3 *The set of the first n colleges S_n chosen by the MIA is riskier than the set Z_n of colleges i with largest expected payoffs z_i. Fix a cost $c > 0$. The optimal set $S^*(c)$ is riskier than the set $Z_{|S^*(c)|}$ of top expected payoff schools.*

Theorem 3 also yields the interval rule in Corollary 1: $S_n \succeq Z_n = [1, n]$ implies $S_n = [1, n]$.

Static portfolio maximization thus precludes 'safety schools'. One never applies to a school for its high admissions rate, when not otherwise justified by its expected payoff. But one might apply to a high-ranked 'stretch school', despite the low expected payoff.

Insight into the proof of Theorem 3 is afforded by the expression (2) for expected payoffs as $\sum_i \rho_{i-1}(S) z_{(i)}$. So inasmuch as the optimal set S^* differs from that with highest expected values z_i, it compensates with a higher rejection rate ρ_{i-1} from better schools. Acceptance chances must be lower, and these schools must be better ranked.

The 'no safety school' substance of Theorem 3 is undermined by assuming correlated acceptances. For an extreme example with perfect correlation, assume three colleges, with payoffs $u_1 = 1$, $u_2 = u < 1$, $\alpha_1 = \alpha_2 = \alpha$, $u_3 = v < u$, $\alpha_3 = \beta > \alpha$, and $\alpha u > \beta v$. Suppose the student is either accepted in both 1 and 2 (chance α), or rejected in both. Then $f(\{1\}) = f(\{1,2\}) = \alpha$, $f(\{2\}) = \alpha u$, $f(\{1,3\}) = f(\{1,2,3\}) = \alpha + (1-\alpha)\beta v$, and $f(\{2,3\}) = \alpha u + (1-\alpha)\beta v$. So $S_1 = Z_1 = \{1\}$. Also, exhaustive checking reveals that $S_2 = \{1,3\}$, while $Z_2 = \{1,2\} \succ S_2$.

B. Portfolio Choices are Less Risky than Sequential Choices. The solution of the static problem that we study substantially differs from the sequential approach. For instance, we have shown that one must above all else be sure to apply to the college yielding the largest expected payoff z_i. Easily, this need not coincide with the one having the highest Gittins index I_i. Indeed, consider two colleges with payoffs $u_1 = 1 > 4/5 = u_2$ having acceptance chances $\alpha_1 = 1/4 < 1/3 = \alpha_2$. Their expected values are thus $z_1 = 1/4 < 4/15 = z_2$, while the indices are ordered from (1) by

$$I_1 = u_1 - c/\alpha_1 = 1 - 4c > 4/5 - 3c = u_2 - c/\alpha_2 = I_2$$

Provided the application cost obeys $c < 1/4$, both colleges are acceptable, as $z_1, z_2 > c$. In other words, the sequentially optimal rule calls for choosing college 1 first, and then college 2, if rejected. The optimal portfolio, by contrast, will only select college 2 provided $c > 11/60$. For then adding college one yields a greater marginal benefit:

$$MB \text{ of } \{1{,}2\} \text{ over } \{2\} = [z_1 + (1-\alpha_1)z_2] - z_2 = z_1 - \alpha_1 z_2 = 1/4 - (1/4)4/15 = 11/60 < 1/5$$

In the above example, the sequential decision-maker plays a more high-risk strategy than does our portfolio one. This turns out to be generally true, as we now assert.

Theorem 4 *Fix a cost $c > 0$. The set S_n of the first n elements chosen by the MIA is less risky than the set $W_n(c)$ of the first n elements with the n highest indices I_i. Also, the optimal portfolio $S^*(c)$ is weakly smaller than the set $W(c)$ of sequentially chosen schools, and is less risky than $W_{|S^*(c)|}(c)$.*

For the size comparison, consider that the sequential rule continues as long as $I_i \geq 0$, or $z_i \geq c$. The static decision-maker, by contrast, stops when the marginal value of the last college i — which is at most $z_i - c$, due to the externalities — turns negative.

The riskiness ordering $S_n \succeq W_n(c)$ is trivial with $c = 0$, for then $I_i = u_i$. We now summarize the spirit of the general proof. Consider the simplest possible case with three colleges such that $z_2 > z_3 > z_1$. For a contradiction, suppose $S_1 = W_1(c) = \{2\}$, but that the ordering fails at the next stage: $S_2 = \{1, 2\} \succ W_2(c) = \{2, 3\}$. Since $(\alpha_2 z_3/\alpha_3) - z_2 = \alpha_2(u_3 - u_2) < 0$,

$$I_1 - I_3 = \frac{z_1 - c}{\alpha_1} - \frac{z_3 - c}{\alpha_3} > \frac{z_1 - \alpha_1 z_2 - c}{\alpha_1} - \frac{z_3 - \alpha_2 z_3 - c}{\alpha_3} \tag{7}$$

But given $S_1 = \{2\}$ and $S_2 = \{1, 2\}$, the marginal benefit of adding college 1 must exceed that of adding college 3, or $z_1 - \alpha_1 z_2 > (1 - \alpha_2)z_3$. Thus, expression (7) is positive, so that $I_1 > I_3$. But $W_1(c) = \{2\}$ and $W_2(c) = \{2, 3\}$ requires $I_3 \geq I_1$, a contradiction.

5.2 Portfolio Choice Sets are Upwardly Diverse

We turn to another key characteristic of the statically optimal set. How similar should be the colleges in the application set? For instance, suppose that a very good but not stellar student is applying to college. Should she then apply just to an *interval* of schools in the 'very good but not stellar' category? Namely, if any two colleges lie in the optimal set $S^*(c)$, are all intermediate schools? Note that Theorem 3 does not preclude interval rules, since the optimal set could be a range of similar schools that are riskier than Z.

We now show that beyond the interval rules for the best students in Corollary 1, the optimal portfolio is not an interval, and diversity is natural. The reason is that a force to gamble upwards emerges for non-top students.

Theorem 5 (Upward Diversity) *For any sufficiently dense and diverse collection of colleges, any non-top student does not apply to an interval of colleges,*[13] *for small enough $c > 0$. There is always at least one higher-ranked school above any choice interval.*

Proof Let the unique maximal expected payoff be z^*. We consider a stylized case with just one college i and $N - 1$ copies of college $j > i$, with $z_j = z^*$. If the result is strictly true in this extreme environment, by continuity it obtains with (a) densely distributed colleges (captured by the college j copies) and (b) sufficient diversity (college i).

[13] Density is measured by the partition size $\max |u_i - u_{i+1}|$, while diversity is captured by $|u_1 - u_N|$.

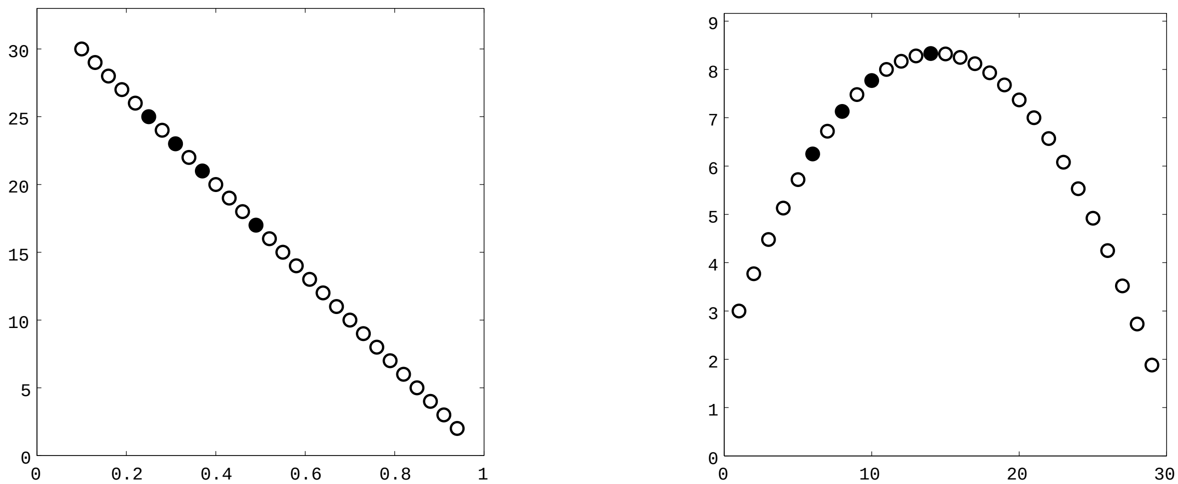

Figure 3: **Portfolio Gambling Illustrated.** The computation represents an application of our optimal algorithm using $c = 2$. At the left is the graph of college payoff u against acceptance chance α. At the right is the graph of expected college payoff $z = \alpha u$ against actual college payoff u. This corresponds to a student with linear payoffs u, and therefore single-peaked expected payoffs z. Chosen schools are represented by filled circles. While she does pick the best expected college, she does not exclusively choose the highest expected schools; rather, she generally gambles upward. Observe how no safety school (high expected value z, but lesser ranked) is chosen. This illustrates Theorem 3 and 5.

The algorithm starts with j, as $z_j = z^*$. We claim that for N large enough and $c > 0$ small enough, the algorithm chooses college i before exhausting college j copies. Indeed, suppose the algorithm has chosen the $n-1$ colleges S_{n-1}, but not yet college i, so that $i \notin S_{n-1}$. The marginal benefit of choosing another college j copy is

$$\Pr(\text{rejected } n-1 \text{ times by } j)\alpha_j u_j - c = (1-\alpha_j)^{n-1}\alpha_j u_j - c, \tag{8}$$

which is geometrically falling in n. But the marginal benefit of choosing college i is

$$\alpha_i u_i - \alpha_i f(S_{n-1}) - c = \alpha_i u_i - \alpha_i u_j(1-(1-\alpha_j)^n) - c, \tag{9}$$

which tends to $\alpha_i(u_i - u_j) - c > 0$ for small $c > 0$. This exceeds (8) for large n.

By continuity, this result obviously holds even when all the college j copies are not literally identical. So for low enough application costs, one always has an incentive to gamble upward, and apply to a discretely higher college than the rest! $\square$

Thus, the marginal value of a 'safety school' is geometrically vanishing in the number of schools given independent acceptances. Conversely, the marginal value of a risky school is boundedly positive, depending only on the (possibly small) admission chance. The event that one is rejected by all better schools need no longer be vanishingly unlikely — so that the marginal gain from insurance needn't be exponentially vanishing.

6 Comparative Statics

We finally consider some natural comparative statics. Obviously, the size of the optimal set $S^*(c)$ decreases in the application cost c, for our algorithm stops sooner.

More interestingly, how will the student's behavior change when acceptance rates change? Assume that a student's prospects change. Will her application set, for instance, grow riskier? Let us parameterize such changes by $\theta = (\alpha_1, \ldots, \alpha_N)$. We first argue that comparative statics here are far from obvious. Topkis (1998) gives the simplest cardinal condition for the optimization of $v(S, \theta)$ in S to be monotone in θ — namely, $v(S, \theta)$ must be supermodular in (S, θ). While Topkis provides Milgrom and Shannon's various (1994) ordinal relaxations of this condition, neither are relevant in our case. For we have the *opposite* cardinal structure here, since Lemma 2 shows that $v(S, \theta)$ is *submodular* in S alone. Summarizing, *we cannot deduce comparative statics by standard monotone methods.* Instead, we shall use induction, and the MIA.

Theorem 6 (Increasing Riskiness) *Assume β proportionately favors better schools more, or $\beta_i/\alpha_i > \beta_{i+1}/\alpha_{i+1}$, for $i = 1, \ldots, N-1$, and that $\beta_i \geq \alpha_i$ for all i.*
(a) Fix $n \geq 1$. The optimal size n sample S_n^β is riskier than S_n^α, or $S_n^\beta \succeq S_n^\alpha$.
(b) Fix $c > 0$. Let $z_1^\alpha > \cdots > z_N^\alpha$ and $z_1^\beta > \cdots > z_N^\beta$, so that $S^\alpha(c) = [1, n^\alpha]$ and $S^\beta(c) = [1, n^\beta]$. If also $(1-\alpha_1)\alpha_2 > (1-\beta_1)\beta_2$, then $n^\beta < n^\alpha$.

One might suppose that the only necessary ingredient for the proof of part (a) is that α favors better schools proportionately more than does β. Indeed, this suffices for sample size $n = 1$. However, we omit a three college example showing necessity of the absolute relation too — that $\beta_i \geq \alpha_i$ for all i.

Notice that in the college application context, part (b) applies to the 'top' students. The top student applies to fewer schools when the acceptance rates rise from α to β. We omit a simple example showing that the condition $(1-\alpha_1)\alpha_2 > (1-\beta_1)\beta_2$ is needed.

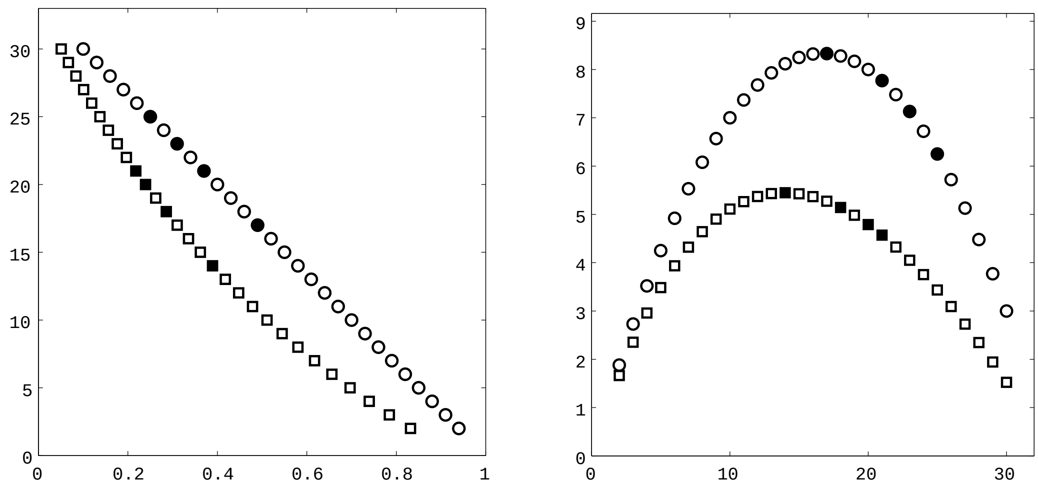

Figure 4: **Comparative Statics Illustrated.** We build on Figure 3. An upward shift in acceptance rates (rightward, in the left graph) that proportionately favors higher colleges improves the acceptance set. Either $n = 4$ for both choice sets, or we posit $c = 2$.

Proof of (a): The proof is instructive — a simultaneous double induction on n and N. Let $S_n^\alpha(N)$ be the optimal n-choice set from N for acceptance chances α. Define $\sigma_{n,k}^\alpha(N) = |S_n^\alpha(N) \cap [1,k]|$. Then $S_n^\beta \succeq S_n^\alpha$ if and only if $\sigma_{n,k}^\beta(N) \geq \sigma_{n,k}^\alpha(N)$ for all k.

The result holds for $n = 1$ and all N, by the MIA. Otherwise, if $j = \arg\max_i \beta_i u_i > \arg\max_i \alpha_i u_i = k$, then $\beta_j u_j \geq \beta_k u_k$ and $\alpha_k u_k \geq \alpha_j u_j$ imply $\beta_j/\beta_k \geq u_k/u_j \geq \alpha_j/\alpha_k$, contrary to our premise.

Assume the result holds for all $n' \leq n$ and $N' \leq N$, with one inequality strict, for all utility payoffs and acceptance chances. We use this induction hypothesis in many ways.

First, if some $j \in S_n^\beta(N) \cup S_n^\alpha(N)$ is not chosen by either optimization, then the problem reduces to choosing from $N-1$ colleges, where $S_n^\beta(N-1) \succeq S_n^\alpha(N-1)$ holds by induction hypothesis. So we may assume there are no such omitted colleges j.

Next, if the result fails at n, N, we now argue that we can decompose

$$\begin{array}{llll} S_n^\alpha(N) = 1 + M + L^\alpha, & \text{where} \quad M = [2, \underline{k}-1], & \text{and} & L^\alpha = S_n^\alpha(N) \cap [\underline{k}, N] \\ S_n^\beta(N) = M + \underline{k} + L^\beta, & \text{where} \quad M = [2, \underline{k}-1], & \text{and} & L^\beta = S_n^\beta(N) \cap [\underline{k}, N] \end{array}$$

where $|L^\alpha| = |L^\beta|$. Define $\Delta(k) = \sigma_{n,k}^\beta(N) - \sigma_{n,k}^\alpha(N)$. Then our decomposition is equivalent to $\Delta(k) = -1$ for $k = 1, \ldots, \bar{k}$, after which $\Delta(k) \geq 0$. Indeed, $\Delta(k)$ changes by at most 1 each time k increments. As stochastic dominance held for $S_n(N-1)$ by assumption, and $S_n(N-1) \subseteq S_n(N)$ by Theorem 1, we have $\Delta(k) \geq -1$ for all k. For a contradiction, assume $0 = \Delta(\underline{k}-1) > \Delta(\underline{k}) = -1$ at some $\underline{k} > 1$. Since by Lemma 1-(a), L^α and L^β are each optimal selections from $[\underline{k}, N]$, and $|L^\alpha| = |L^\beta|$, they are FSD-ranked by induction assumption: $\Delta(\underline{k}) \geq 0$.

The appendix deduces $f^\beta(1 + M + L^\beta) > f^\beta(S_n^\beta(N))$, contradicting S^β optimal. $\square$

7 Conclusion

Static optimization is rapidly becoming yesterday's struggle in economics. In this paper, we have identified a common and yet unsolved class of static portfolio choice problems, where one earns only the best prize from the portfolio. The main lesson of this paper is that such portfolio choices are intriguing, as *the value of a portfolio is subtly less than the sum of its parts.* Such problems are also practically important, being faced by millions of college applicants, thousands of employers competing to hire in student-driven job markets, as well as firms making choosing among uncertain technologies to explore. In other words, there is a wealth of accumulated conventional wisdom by teenagers, academics, and company executives alike.

We have derived a natural algorithm for computing the optimal portfolio, and used it to characterize the optimum. The first concrete lesson here is to apply to one's best singleton school. We have shown that one should proceed more cautiously than sequentially, but still gamble upward relative to a selection from the best singletons. So inasmuch as other included schools are not good singletons, they must be highly ranked. Further, the portfolio generally need not consist of an interval of like choices. Unless one's admission chances to schools are affected by some common component, gamble upward within the portfolio. One's bottom tier of schools must separately optimize among the bottom schools. Finally, when one's success chances rise, and proportionately more at better schools, apply to a higher-ranked set of schools.

Maximizing — as opposed to minimizing — submodular functions is a currently unexplored topic in economics worthy of future research. The problem is theoretically quite rich, as we have underscored. We have completely solved the success-or-failure world, and shown that it may fail with richer prize spaces. This paper therefore opens this new door to research on this exciting open problem, where a richer algorithm is needed. Furthermore, high correlation undoes our algorithm — although this failure is less surprising, because correlation is also an unresolved problem in the closely-related multi-armed bandit problem in dynamic choice theory.

We hope that our results not only have practical value, but can also prove to be a building block for directed search applications, or other foundational work. One very natural extension of this work would allow the decision maker to accept more than one option. This naturally arises in hiring contexts where multiple positions must be filled. Another more intriguing extension would ask about a combination of our work with Weitzman, combining static and dynamic optimality. For instance, Weitzman *restricted* the decision-maker to a single choice per period, but that is not necessarily optimal. For instance, the growing phenomenon of early admission has such a static-dynamic feature: future research beckons. Just as well, one might imagine a generalization of the multi-armed bandit, where one is permitted to choose a portfolio of arms.

A Appendix: Omitted Proofs

A.1 Optimality of the MIA: Proof of Theorem 1

This proof ignores the non-generic possibility of tied values of two choice subsets. This is without loss of generality, because any choice S made is optimal in the event of a sequence of ties. To see this, assume the MIA is valid absent ties. Then there clearly exists a vanishing sequence of ε-perturbations of the payoffs that render the choice S made strictly optimal at each stage. By the Theorem of the Maximum, the maximized values of the ε-perturbed problems converge to the maximized value of the unperturbed problem. But by continuity, the values of the perturbed problems converge to the value of the unperturbed problem with choice S. I.e., the choice S is optimal.

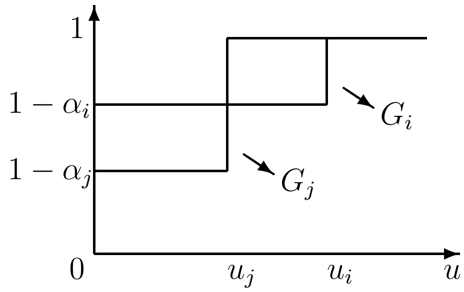

Figure 5: **Prize Distributions.** Here, we depict two cdf's for the binary prize case, and show that they enjoy a simple, but useful, single crossing property.

A.1.1 Optimality of the Selection Made in Step 1

Let $S \subset N$, and suppose $i_1 \notin S$. Suppose first that $(i_1, N] \cap S \neq \varnothing$, and let b be the best college in this set. Define $S' = S - b + i_1$. Easy algebra reveals that

$$v(S') - v(S) = \int_0^{\bar{u}} (G_b(u) - G_{i_1}(u)) \prod_{i \in S-b} G_i(u) du \tag{10}$$

We will use (10) to show that $v(S') > v(S)$. Notice that (i) $\int_0^{\bar{u}} G_b(u)du > \int_0^{\bar{u}} G_{i_1}(u)du$ since $i_1 = \text{argmax}_{i \in S} \int_0^{\bar{u}} (1 - G_i(u))\, du$; (ii) $G_b(u) - G_{i_1}(u)$ (weakly single) crosses zero from below at u_b (see Figure 5); (iii) $\prod_{i \in S-b} G_i(u)$ is increasing in u. It follows from an inequality in **?** that the right side of (10) is positive.

If $(i_1, N] \cap S = \varnothing$, then let $S' = S - w + i_1$, where w is the worst element of $[1, i_1) \cap S$. Then

$$v(S') - v(S) = \int_0^{\bar{u}} (G_w(u) - G_{i_1}(u)) \prod_{i \in S-w} G_i(u) du. \tag{11}$$

Notice that (i) $\int_0^{\bar{u}} G_w(u)du > \int_0^{\bar{u}} G_{i_1}(u)du$ since $i_1 = \text{argmax}_{i \in S} \int_0^{\bar{u}} (1 - G_i(u))\, du$; (ii) $G_w(u) - G_{i_1}(u) \neq 0$ if $u \in [0, w]$ and vanishes outside this interval; (iii) $S - w$ consists of colleges that are better than w and i_1, and therefore $\prod_{i \in S-w} G_i(u)$ remains constant on $[0, w]$. Thus, for any $u^* \in [0, w]$,

$$v(S') - v(S) = \prod_{i \in S-w} G_i(u^*) \int_0^{\bar{u}} (G_w(u) - G_{i_1}(u))\, du > 0, \tag{12}$$

A.1.2 Optimality of the Selection Made in the Induction Step

Let $S \subset N$, and suppose $S_{n-1} \subset S$ but $i_n \notin S$. There are two cases:

Case 1. $(i_n, N] \cap (S - S_{n-1}) \neq \varnothing$

Let $S' = S - b + i_n$, where b is the best element of $(i_n, N] \cap (S - S_{n-1})$. Then

$$\begin{aligned} v(S') - v(S) &= \int_0^{\bar{u}} (G_b(u) - G_{i_n}(u)) \prod_{i \in S-b} G_i(u) du \\ &= \int_0^{\bar{u}} (G_b(u) - G_{i_n}(u)) \prod_{i \in S - S_{n-1} - b} G_i(u) \prod_{j \in S_{n-1}} G_j(u) du. \qquad (13) \end{aligned}$$

Notice that:

- $\int_0^{\bar{u}} G_b(u) \prod_{j \in S_{n-1}} G_j(u) du > \int_0^{\bar{u}} G_{i_1}(u) \prod_{j \in S_{n-1}} G_j(u) du$ since

$i_n = \text{argmax}_{i \in S - S_{n-1}} \int_0^{\bar{u}} (1 - G_i(u)) \prod_{j \in S_{n-1}} G_j(u) du$;

- $G_b(u) - G_{i_n}(u)$ (weakly) single crosses zero from below at u_b;
- $\prod_{i \in S - S_n - b} G_i(u)$ is increasing in u.

By another application of Beesack's inequality, (13) is positive and so S' improves S.

Case 2. $(i_n, N] \cap (S - S_{n-1}) = \varnothing$

Let $S' = S - w + i_n$, where w is the worst element of $[1, i_n) \cap (S - S_{n-1})$. We will show that $v(S') > v(S)$. Then

$$\begin{aligned} v(S') - v(S) &= \int_0^{\bar{u}} (G_w(u) - G_{i_n}(u)) \prod_{i \in S-w} G_i(u) du \\ &= \int_0^{\bar{u}} (G_w(u) - G_{i_n}(u)) \prod_{i \in S - S_{n-1} - w} G_i(u) \prod_{j \in S_{n-1}} G_j(u) du \qquad (14) \end{aligned}$$

Notice that (i) $\int_0^{\bar{u}} G_w(u) \prod_{j \in S_{n-1}} G_j(u) du > \int_0^{\bar{u}} G_{i_n}(u) \prod_{j \in S_{n-1}} G_j(u) du$ since

$$i_n = \text{argmax}_{i \in S - S_{n-1}} \int_0^{\bar{u}} (1 - G_i(u)) \prod_{j \in S_{n-1}} G_j(u) du;$$

(ii) $G_w(u) - G_{i_n}(u) \neq 0$ if $u \in [0, w]$ and vanishes outside this interval; (iii) $S - S_{n-1} - w$

consists of colleges that are better than w and i_n, and therefore $\prod_{i\in S-S_{n-1}-w} G_i(u)$ remains constant on $[0, w]$. Thus, for any $u^* \in [0, w]$,

$$v(S')-v(S) = \prod_{i\in S-S_{n-1}-w} G_i(u^*) \int_0^{\bar{u}} (G_w(u) - G_{i_1}(u)) \prod_{j\in S_{n-1}} G_j(u)du > 0. \qquad \square$$

Observe that the choice of b is not crucial for constructing a set that improves upon S; any school in $S - S_{n-1}$ worse than i_n will do. The choice of w, however, is critical; otherwise, we could not use the property that $\prod_{i\in S-S_{n-1}-w} G_i(u)$ is constant on $[0, w]$.

Further note that the counterexample of §4.2 violates two key properties exploited in this proof: The cdf's should not cross more than once and should be constant over the set of prizes of lower-ranked options.

A.2 MIA and General Ranked Options: Proof of Theorem 2

Modifying the proof of the Second Order Stochastic Dominance Ranking Theorem, the condition imposed on the G_i is equivalent to $\int_0^{\bar{u}} h(u)dG_i(u) > \int_0^{\bar{u}} h(u)dG_{i+1}(u)$ for all increasing and convex functions h.

Notice that $i_1 = 1$, and therefore $1 \in S^*$ if $v(S') - v(S) > 0$, where $S' = S - b + i_1$, and $b > 1$. Integrating (10) by parts yields

$$v(S') - v(S) = \textstyle\int_0^{\bar{u}} \left(\int_0^u \prod_{i\in S-b} G_i(\xi)d\xi\right) dG_1(u) - \int_0^{\bar{u}} \left(\int_0^u \prod_{i\in S-b} G_i(\xi)d\xi\right) dG_b(u), \quad (15)$$

which is positive since $\int_0^u \prod_{i\in S-b} G_i(\xi)d\xi$ is increasing and convex in u.

Suppose $S_{n-1} = [1, n-1]$. If $j \in [n, N]$, then

$$v(S_{n-1} + j) - v(S_n) = \textstyle\int_0^{\bar{u}} \left(\int_0^u \prod_{i\in S_{n-1}} G_i(\xi)d\xi\right) dG_j(u), \quad (16)$$

which is clearly maximized at $j = n$.

To show that $n \in S^*$, it suffices (by the induction hypothesis) to show that

$$v(S') - v(S) = \textstyle\int_0^{\bar{u}} \left(\int_0^u \prod_{i\in S-b} G_i(\xi)d\xi\right) dG_n(u) - \int_0^{\bar{u}} \left(\int_0^u \prod_{i\in S-b} G_i(\xi)d\xi\right) dG_b(u) \quad (17)$$

is positive, which holds since $\int_0^u \prod_{i\in S-b} G_i(\xi)d\xi$ is increasing and convex in u. $\square$

A.3 Portfolio vs. Top Singletons: Proof of Theorem 3

Fixed Sample Size. Consider first the case in which the DM is restricted to choice sets of size n. We will show that $S_n^* \succeq Z_n$. Proceeding by contradiction, suppose not. Since $S_n^* \not\succeq Z_n$, and $|S_n^*| = |Z_n| = n$, there exists $i \in Z_n - S_n^*$, and $j \in S_n^* - Z_n$ such that $j > i$.[14] Since $j \notin Z_n$ and $i \in Z_n$, it follows that $z_i > z_j$. Apply Lemma 1-(b).

Endogenous Sample Size. If the DM can choose sets of any size, and each selection costs $c > 0$, then the optimal algorithm determines the cardinality of the optimal set S^*, denoted by $|S^*|$. The result for the fixed sample size case implies that $S^* \succeq Z_{|S^*|}$. $\square$

A.4 Static vs. Dynamic Choices: Proof of Theorem 4

We'll show that $W_{|S^*(c)|}(c) \succeq S^*(c)$. Proceeding as in the proof of Theorem 3, there exists a college $i \in S^*(c) - W_{|S^*(c)|}(c)$ and a college $j \in W_{|S^*(c)|}(c) - S^*(c)$ such that $j > i$.

Consider the ranked sets $U = S^*(c) \cap [1, i)$, $M = S^*(c) \cap (i, j)$, $L = S^*(c) \cap (j, N]$. Since $S^*(c)$ is optimal, it must be true that $f(U + i + M + L) \geq f(U + M + j + L)$, which, by a repeated application of (3), can be written as (noting $f(i) = z_i$):

$$\begin{array}{rrcl}
 & f(i) + (1 - \alpha_i) f(M + L) & \geq & f(M) + \rho(M)[f(j) + (1 - \alpha_j) f(L)] \\
 & & = & f(M + L) + \rho(M)[f(j) - \alpha_j f(L)] \\
\Rightarrow & z_i - \alpha_i f(M + L) & \geq & \rho(M)[z_j - \alpha_j f(L)] \\
\Rightarrow & \dfrac{z_i - \alpha_i f(M + L) - c}{\alpha_i} & > & \dfrac{\rho(M)[z_j - \alpha_j f(L)] - c}{\alpha_j} \\
\Rightarrow & I_i - f(M + L) & > & I_j - [1 - \rho(M)] u_j - \rho(M) f(L) \\
\Rightarrow & I_i - I_j & > & f(M) - [1 - \rho(M)] u_j \\
 & & = & \displaystyle\sum_M u_{(k)} \alpha_{(k)} \rho_{(k-1)} - u_j \sum_M \alpha_{(k)} \rho_{(k-1)} > 0
\end{array}$$

as $u_k > u_j$ for all $k \in M$. So $I_i > I_j$ — impossible if $j \in W_{|S^*(c)|}(c)$ but $i \notin W_{|S^*(c)|}(c)$. $\square$

[14] This intuitive result follows from an equivalent definition of $S \succeq S'$ that there exists a mapping $\mu : N \mapsto N$ with $\mu(i) \leq i$ such that for each element $i \in S'$, there is an element $\mu(i) \in S$.

A.5 Comparative Statics: Proof of Theorem 6

Part (a). We need $f^\beta(\hat{S}_n^\beta(N)) > f^\beta(S_n^\beta(N))$, where $\hat{S}_n^\beta(N) = 1 + M + L^\beta$. Now,

$$\begin{aligned} f^\alpha(S_n^\alpha(N)) &= \alpha_1 u_1 + (1-\alpha_1)[f^\alpha(M) + \rho^\alpha(M) f^\alpha(L^\alpha)] \\ &\geq f^\alpha(M) + \rho^\alpha(M)[\alpha_{\underline{k}} u_{\underline{k}} + (1-\alpha_{\underline{k}}) f^\alpha(L^\alpha)] = f^\alpha(S_n^\beta(N)), \end{aligned}$$

since $S_n^\alpha(N)$ is optimal. This holds if and only if

$$\frac{\alpha_1}{\alpha_{\underline{k}}}\left(\frac{u_1 - f^\alpha(M)}{\rho^\alpha(M)}\right) + \left(1 - \frac{\alpha_1}{\alpha_{\underline{k}}}\right) f^\alpha(L^\alpha) \geq u_{\underline{k}}. \tag{18}$$

Replacing α by β in (18) is equivalent to $f^\beta(\hat{S}_n^\beta(N)) > f^\beta(S_n^\beta(N))$. We now justify this:

- Since 1 dominates every college in $M + L^\alpha$ and $M + L^\beta$, we have $u_1 > f^\alpha(M + L^\alpha)$ and $u_1 > f^\beta(M + L^\beta)$. These are equivalent to

$$\frac{u_1 - f^\alpha(M)}{\rho^\alpha(M)} > f^\alpha(L^\alpha) \quad \text{and} \quad \frac{u_1 - f^\beta(M)}{\rho^\beta(M)} > f^\beta(L^\beta)$$

- Since $\beta_1/\beta_{\underline{k}} > \alpha_1/\alpha_{\underline{k}}$, the weight on the first term of (18) increases.
- $f^\beta(L^\beta) \geq f^\beta(L^\alpha) > f^\alpha(L^\alpha)$ by Lemma 1-(a) and $\beta_i \geq \alpha_i$, for all i, respectively
- Finally,

$$\frac{\partial}{\partial \alpha_\ell}\left(\frac{u_1 - f^\alpha(M)}{\rho^\alpha(M)}\right) > 0 \quad \forall \ell \in M \quad \Rightarrow \quad \frac{u_1 - f^\beta(M)}{\rho^\beta(M)} \geq \frac{u_1 - f^\alpha(M)}{\rho^\alpha(M)} \tag{19}$$

where the inequality on the RHS of (19) is desired. Applying (3), we have $f^\alpha(M) = f^\alpha(U) + \rho^\alpha(U)[\alpha_\ell u_\ell + (1-\alpha_\ell) f^\alpha(L)]$, where $L = (\ell, N] \cap M$ and $U = [1, \ell) \cap M$. Thus,

$$\begin{aligned} u_1 - f^\alpha(M) &= u_1 - (f^\alpha(U) + \rho^\alpha(U)[\alpha_\ell u_\ell + (1-\alpha_\ell) f^\alpha(L)]) \\ &= [u_1 - f^\alpha(U) - \rho^\alpha(U) f^\alpha(L)] - \rho^\alpha(U)[u_\ell - f^\alpha(L)]\alpha_\ell \\ &= A - B\alpha_\ell \end{aligned}$$

hereby implicitly defining A and B. The derivative on the LHS of (19) has the sign of

$$\frac{\partial}{\partial \alpha_\ell} \frac{A - B\alpha_\ell}{\rho^\alpha(M-\ell)(1-\alpha_\ell)} = \frac{A - B}{\rho^\alpha(M-\ell)(1-\alpha_\ell)^2}$$

since $\rho^\alpha(M) = \rho^\alpha(M-\ell)(1-\alpha_\ell)$. But this is positive given

$$A - B = [u_1 - f^\alpha(U) - \rho^\alpha(U)f^\alpha(L)] - \rho^\alpha(U)[u_\ell - f^\alpha(L)] = u_1 - f^\alpha(U) - \rho^\alpha(U)u_\ell > 0$$

because college 1 dominates colleges in $[1,\ell] \cap M$, and using decomposition (3).

Part (b). We reinterpret this result using standard monotone methods. Consider the optimization problem: $\max_{S \subseteq C} v(S,\theta)$, where $C = \{S \subseteq N | S = [1,n], n \leq N\}$. Note that restricting attention to C has no loss of generality for 'top' students, by Corollary 1.

Since C is a chain (i.e. a totally-ordered set) $v(S,\theta)$ is quasi-supermodular in S. Thus, to show that the maximizer is increasing in θ, we need to show that the single crossing property holds (Milgrom and Shannon (1994)), namely

$$v(S_H,\theta_L) - v(S_L,\theta_L) \geq 0 (>0) \;\; \Rightarrow \;\; v(S_H,\theta_H) - v(S_L,\theta_H) \geq 0 (>0), \tag{20}$$

where θ_H is above θ_L, $S_H = [1,n_H]$, $S_L = [1,n_L]$, and $n_L > n_H$. Rewrite (20) as

$$\rho^{\theta_L}([1,n_H])f^{\theta_L}((n_H,n_L]) \leq c(n_L - n_H) \;\; \Rightarrow \;\; \rho^{\theta_H}([1,n_H])f^{\theta_H}((n_H,n_L]) \leq c(n_L - n_H)$$

for which a sufficient condition is

$$\rho^{\theta_H}([1,n_H])f^{\theta_H}((n_H,n_L]) \leq \rho^{\theta_L}([1,n_H])f^{\theta_L}((n_H,n_L]). \tag{21}$$

Claim 1 *Inequality* (21) *holds if* $\alpha_2^{\theta_H}(1-\alpha_1^{\theta_H}) \leq \alpha_2^{\theta_L}(1-\alpha_1^{\theta_L})$ *and* $\alpha_i^{\theta_H}/\alpha_i^{\theta_L} > \alpha_{i+1}^{\theta_H}/\alpha_{i+1}^{\theta_L}$.

Proof First, (21) holds if and only if

$$\rho^{\theta_H}([1,n])z_{n+1}^{\theta_H} \leq \rho^{\theta_L}([1,n])z_{n+1}^{\theta_L} \tag{22}$$

for all n. For (21) implies (22). To prove the converse, notice that

$$\begin{aligned} \rho^{\theta_H}([1,n_H])f^{\theta_H}((n_H,n_L]) &= \rho^{\theta_H}([1,n_H])z_{n_H+1}^{\theta_H} + \cdots + \rho^{\theta_H}([1,n_L-1])z_{n_L}^{\theta_H} \\ &\leq \rho^{\theta_L}([1,n_H])z_{n_H+1}^{\theta_L} + \cdots + \rho^{\theta_L}([1,n_L-1])z_{n_L}^{\theta_L} \\ &= \rho^{\theta_L}([1,n_H])f^{\theta_L}((n_H,n_L]), \end{aligned}$$

where the inequality follows from repeated application of (22).

Next, we claim that (22) is equivalent to

$$\rho^{\theta_H}([1,n])\alpha_{n+1}^{\theta_H} \leq \rho^{\theta_L}([1,n])\alpha_{n+1}^{\theta_L} \tag{23}$$

If $n = 1$, then (23) reduces to $\alpha_2^{\theta_H}(1-\alpha_1^{\theta_H}) \leq \alpha_2^{\theta_L}(1-\alpha_1^{\theta_L})$, which holds by assumption. Suppose the result holds for $n-1$, i.e., $\rho^{\theta_H}([1,n-1])\alpha_n^{\theta_H} \leq \rho^{\theta_L}([1,n-1])\alpha_n^{\theta_L}$. Then

$$\begin{aligned}
\rho^{\theta_H}([1,n])\alpha_{n+1}^{\theta_H} &< \rho^{\theta_H}([1,n])\alpha_{n+1}^{\theta_H}\frac{\alpha_n^{\theta_H}}{\alpha_n^{\theta_L}}\frac{\alpha_{n+1}^{\theta_L}}{\alpha_{n+1}^{\theta_H}}\frac{(1-\alpha_n^{\theta_L})}{(1-\alpha_n^{\theta_H})} \\
&= \rho^{\theta_H}([1,n-1])\frac{\alpha_n^{\theta_H}}{\alpha_n^{\theta_L}}(1-\alpha_n^{\theta_L})\alpha_{n+1}^{\theta_L} \\
&< \rho^{\theta_L}([1,n-1])(1-\alpha_n^{\theta_L})\alpha_{n+1}^{\theta_L} \\
&= \rho^{\theta_L}([1,n])\alpha_{n+1}^{\theta_L}
\end{aligned}$$

where the first inequality owes to $\alpha_i^{\theta_H}/\alpha_i^{\theta_L} > \alpha_{i+1}^{\theta_H}/\alpha_{i+1}^{\theta_L}$ and $\alpha_i^{\theta_H} > \alpha_i^{\theta_L}$, the second to the induction hypothesis, and the last equality to $\rho^{\theta_L}([1,n]) = \rho^{\theta_L}([1,n-1])(1-\alpha_n^{\theta_L})$. Finally, set $\beta_i = \alpha_i^{\theta_H}$ and $\alpha_i = \alpha_i^{\theta_L}$.

References

ALBRECHT, J., P. GAUTIER, AND S. VROMAN (2002): "Equilibrium Directed Search with Multiple Applications," Georgetown mimeo.

BURDETT, K., AND K. JUDD (1983): "Equilibrium Price Dispersion," *Econometrica*, 51(4), 955–69.

BURDETT, K., S. SHI, AND R. WRIGHT (2001): "Pricing and Matching with Frictions," *Journal of Political Economy*, 109(5), 1060–1085.

CHADE, H., G. LEWIS, AND L. SMITH (2003): "The College Admissions Problem with Uncertainty," Michigan mimeo.

FUJISHIGE, S. (1991): *Submodular Functions and Optimization*, vol. 47 of *Annals of Discrete Mathematics*. North Holland, Amsterdam.

Gittins, J., and D. Jones (1974): "A Dynamic Allocation Index for the Sequential Design of Experiments," in *Progress in Statistics*, ed. by J. Gani, pp. 241–266. North-Holland, Amsterdam.

Goldengorin, B., G. Tijssen, and M. Tso (1998): "The Maximization of Submodular Functions: Old and New Proofs for the Correctness of the Dichotomy Algorithm," SOM Report No. 98A08, University of Groningen.

Gul, F., and E. Stacchetti (1999): "Walrasian Equilibrium with Gross Substitutes," *Journal of Economic Theory*, 87(1), 95–124.

Kelso, A. S. J., and V. P. Crawford (1982): "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 50(6), 1483–1504.

Lovász, L. (1983): "Submodular Functions and Convexity," in *Mathematical Programming: The State of the Art*, ed. by A. Bachem, M. Grotschel, and B. Korte, pp. 235–257. Springer, Berlin.

Milgrom, P., and C. Shannon (1994): "Monotone Comparative Statics," *Econometrica*, 62(1), 157–180.

Stigler, G. (1961): "The Economics of Information," *Journal of Political Economy*, 69(3), 213–25.

Topkis, D. (1998): *Supermodularity and Complementarity.* Princeton University Press, Princeton, NJ.

Vishwanath, T. (1992): "Optimal Orderings for Parallel Selection," *International Economic Review*, 33(1), 79–89.

Weitzman, M. (1979): "Optimal Search for the Best Alternative," *Econometrica*, 47(3), 641–654.