*Simultaneous Search**

Hector Chade[†] Lones Smith[‡]

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

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

this version: August 2005

Abstract

We introduce and solve a new class of "downward-recursive" static portfolio choice problems. An individual simultaneously chooses among ranked stochastic options, and each choice is costly. In the motivational application, just one may be exercised from those that succeed. This often emerges in practice, such as when a student applies to many colleges.

We show that a greedy algorithm finds the optimal set. The optimal choices are "less aggressive" than the sequentially optimal ones, but "more aggressive" than the best singletons. The optimal set in general contains gaps. We provide a comparative static 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 2003 Midwest Economic Theory Meetings, ITAM, LBS, Penn, Duke, Michigan, Toronto Matching Conference, 2004 Society for Economic Dynamics, 2004 North American Econometric Society Meetings, 2004 Latin American Econometric Society Meetings, Yale, Texas, Stanford, and NYU. We are very grateful for research assistance of Kan Takeuchi, and the feedback of Miles Kimball, Steve Salant, and Ennio Stacchetti. Lones is grateful for the financial support of the National Science Foundation.

[†]Email: hector.chade@asu.edu

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

1 Introduction

We introduce and solve a new class of "downward recursive" portfolio choice problems. For instance, a decision maker (DM) simultaneously chooses among ranked stochastic options, each choice is costly, and only the best realized alternative is exercised.

Our paper generalizes Stigler (1961), who analyzes optimal static wage search. Unlike Stigler, we do not assume *a priori* identical prizes, and characterize both the optimal sample size and *choice composition.* Weitzman (1979) also explores a problem with *a priori* distinct prizes — but in the sequential world. His is a nice application of Gittins' solution of the bandit problem. Each option can be assigned an index in isolation of all others; sequentially, one simply chooses the unexplored option with the highest index.

In our problem, no such simple index rule presents itself. Instead, we find ourselves faced with the maximization of a submodular function of sets of alternatives — to be sure, a complex combinatorial optimization problem. Nevertheless, we show that an economically natural algorithm produces the optimal set in a quadratic number of steps.

We then explore properties of the optimal set. We ask, for instance, how much risk should one take. We show that the optimal portfolio is 'more aggressive' than the set of best options taken individually, but 'less aggressive' than Weitzman's sequential choices.

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 options and low enough costs, a connected 'interval' of similarly-risky prospects is not optimal. We next provide a key comparative static, 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. 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 to explore.

Example 3. An economics department must fly out new PhD job candidates; the fly-outs are costly. Each school ranks the candidates, and better ones are harder to hire.

Our paper may also be more topically viewed as a foundation for the recent literature on directed search (eg., allowing employees to choose which jobs to apply to). We solve this decision problem for multiple applications and heterogeneous jobs.[1]

We first describe the problem. We introduce the algorithm and prove its optimality. We then explore the properties of the optimal set: Does the DM insure herself or gamble? Are the optimal choices similar, or disparate? What if success rates increase? The appendix contains the proofs omitted in the text.

2 The Portfolio Problem

A decision maker (DM) can consume prizes from a finite set $N = \{1, 2, \dots, N\}$. Here, N is a natural number, but abusing notation, we denote this set by N too, and its subsets by 2^N (with the subset inclusion order). Let $f : 2^N \mapsto \mathbb{R}_+$ be a strictly increasing function, with $f(\varnothing) = 0$. Interpret $f(S)$ as the expected value of subset S, and put $z_i \equiv f(i) > 0$.[2] Prizes are random, and the prize set S has failure chance $\rho(S) \in [0, 1)$ for all $S \neq \varnothing$ (and $\rho(\varnothing) = 1$). Since $\alpha_i \equiv 1 - \rho(i)$ is the success chance of prize i, the ex post payoff is $u_i \equiv f(i)/\alpha_i$. We assume that prize 1 is *ex post* the best, 2 the next, etc. so that $u_1 > u_2 > \cdots > u_N$.

Say that U is *above* L, written $U \sqsupseteq L$, if the worst prize in U beats the best in L. We assume that the portfolio S is worth less than the sum of its parts. Specifically, this payoff function is *downward recursive* (*DR*), so that for all sets $U \sqsupseteq L$ in N:

$$f(U + L) = f(U) + \rho(U)f(L) \tag{1}$$

We observe that ρ is multiplicative in a DR payoff structure, since for all $U \sqsupseteq M \sqsupseteq L$:

$$f(U + M + L) = f(U + M) + \rho(U + M)f(L) = [f(U) + \rho(U)f(M)] + \rho(U)\rho(M)f(L)$$

so that $\rho(U+M) = \rho(U)\rho(M)$. Since $\rho < 1$, and is multiplicative, *ρ is strictly decreasing.*

The cost of a portfolio S is given by a function $c(|S|)$, where S has cardinality $|S|$, $c(0) = 0$, and c is increasing and convex on $\mathbb{R}_+$. We assume $z_i > c(1)$ for all i, thereby

[1] See Burdett, Shi, and Wright (2001) and Albrecht, Gautier, and Vroman (2002). Perhaps the first equilibrium paper with multiple simultaneous searches is Burdett and Judd (1983).

[2] We avoid set notation, writing $i = \{i\}$, $A+B = A \cup B$, $A-B = A \backslash B$, $(i, j) = \{k \in N | i < k < j\}$, etc.

pruning weakly dominated prizes. This paper studies a one-shot maximization of $v(S) = f(S) - c(|S|)$. Of course, $v(\varnothing) = 0$. Denote by $\Sigma^* \equiv \Sigma^*(N)$ the solution of:

$$\max_{S \subseteq N} v(S) \tag{2}$$

We also explore two prominent special cases of (2). By a *fixed cost per application*, we mean that $c(|S|) = \bar{c}\,|S|$, for some $\bar{c} > 0$. In the *fixed sample size* n case, $c(|S|) = 0$ if $|S| \leq n$, and $c(|S|) = \infty$ if $|S| > n$. Problem (2) becomes $\max_{S \subseteq N \& |S|=n} f(S)$ with solution $\Sigma_n(N)$. We then define $\Sigma_n \equiv \Sigma_n(N)$. Notice that $\Sigma^* = \Sigma_n$ for some n.

3 Applications

A. The College Problem. A student must choose once and for all a portfolio $S \subseteq N$ of colleges to apply for admission, at cost $c(|S|)$. The best is 1, the second best 2, and so on. The student's cardinal utility (*ex post payoff*) from attending college i is u_i, where $u_1 > u_2 > \cdots > u_N$. Her *admission chance* at college i is $\alpha_i \in (0, 1]$. Intuitively one might imagine the *inverse ordering* $\alpha_1 < \alpha_2 < \cdots < \alpha_N$, but this is inessential, as we shall see. The acceptance decisions by any set of colleges are independent. For instance, this arises when colleges perceive noisy conditionally iid signals of a student's type, and she fully knows her true type. The expected payoff of college i *alone* is $z_i = \alpha_i u_i$.

Working recursively, one either gets into the best college in S, or one does not; if rejected, one either gets into the next best, or not, etc. Since $\rho(S) \equiv \Pi_{i \in S}(1 - \alpha_i)$ is the chance of rejection by all colleges in the set S, the *gross payoff* may be decomposed as:

$$f(S) = \sum_{i=1}^{|S|} z_{(i)} \prod_{\ell=1}^{i-1}(1 - \alpha_{(\ell)}) = \sum_{i=1}^{|S|} z_{(i)} \rho_{(i-1)}(S) \tag{3}$$

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

This college structure contains the generality of the DR payoff structure of §2, and we sometimes cast our results in the language of this application, for definiteness.

B. Other Singleton Prize Models. The technology choice clearly has this structure. Hiring at the economics department assumes this form after some reworking. Indeed, assume that (i) fly-outs do not inform the hiring decision; (ii) each department needs at most one job candidate; (iii) after the fly-out stage, the market clears top to bottom, so that the better recruits are available with smaller chance to any school below the top.

C. Correlated Rejection Chances. Modify the college problem so that rejection from school i scales down the acceptance chance at colleges $j > i$ by a factor $\beta_i \in [0,1]$. Portfolio S then has value $f(S) = \sum_{i=1}^{|S|} z_{(i)} \prod_{l=1}^{i-1}(1 - \alpha_{(l)})\beta_{(l)}$. This derives from a consistent probability distribution over N for large enough β_i (all i), and reduces to (2) when $\beta_i \equiv 1$. Because of the DR structure, an equivalent college problem exists with independent admission events: Assume acceptance rates $\bar{\alpha}_i = 1 - (1 - \alpha_i)\beta_i$ and college payoffs $\bar{u}_i = \alpha_i u_i / \bar{\alpha}_i$. The ex post payoffs $\bar{u}_i$ fall in i also when the β_i's are large enough.

D. One Shot Multi-Decisions for Dynamic Choice with Payoff Discounting. The DM enjoys payoffs from all successful options, but can only consume one per period. He thus eats the best first, etc. Future payoffs are discounted by the factor $\delta \in [0,1)$. One can show that the expected payoff of portfolio S is $f(S) = \sum_{i=1}^{|S|} z_{(i)} \prod_{l=1}^{i-1}(1-\alpha_{(l)}+\alpha_{(l)}\delta)$. Here, an equivalent college problem requires $\bar{\alpha}_i = (1-\delta)\alpha_i$ and $\bar{u}_i = u_i/(1-\delta)$.

4 The Solution

4.1 Consistency Checks on the Optimal Set

Computing the optimal set is a complex task, but we are able now to provide two useful tests that it must obey. The DR equality (1) implies a key *ordinal* property, *downward maximization* — optimizations on sets imply optimizations on lower ends of those sets:

Lemma 1 *Let $\Sigma_n = U + L$, where $U \sqsupseteq L$ and L has k elements. Then $\Sigma_k(D) = L$ where D are those options in N not better ranked than the best in L.*

If ex ante and ex post ranks of options agree, their marginal values are likewise ranked.

Lemma 2 *Assume $z_i > z_j$ and $i < j$. Then the marginal benefits of i, j are ordered $MB_i(S) \equiv f(S+i) - f(S) > f(S+j) - f(S) = MB_j(S)$ for any set $S \subset N \setminus \{i,j\}$.*

Proof As $i < j$, we may write $S = U + M + L$, for sets (upper) $U = [1, i) \cap S$, (middle) $M = (i, j) \cap S$, and (lower) $L = (j, N] \cap S$. So $U \sqsupseteq M \sqsupseteq L$. Consider the *suboptimal* implementation policy for $S + i$: Accept the best available option, unless it is i, in which case accept the best option in M (if available) over i. So by (1),

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

since $z_i - \alpha_i f(L) > z_j - \alpha_j f(L)$, given $z_i = \alpha_i u_i > \alpha_j u_j = z_j$ and $u_i > u_j > f(L)$. □

If $j \in \Sigma_n(N)$, then setting $S = \Sigma_n(N) - j$ yields at once a simple insight into Σ^*. For any chosen option, any better-ranked one with greater expected payoff is also chosen:

Lemma 3 *Assume $z_i > z_j$ and $i < j$. If $j \in \Sigma_n(N)$, then $i \in \Sigma_n(N)$.*

4.2 An Optimal Marginal Improvement Algorithm

A "greedy algorithm" at each step makes the locally optimal choice, with the hope of finding the global optimum. The next greedy algorithm, which we call the *Marginal Improvement Algorithm* (MIA), identifies Σ^* via an inductive procedure. Let $\Upsilon_0 = \varnothing$.

Step 1 *Choose any $i_n \in \arg\max_{i \in N \setminus \Upsilon_{n-1}} \; f(\Upsilon_{n-1} + i)$.*

Step 2 *If $f(\Upsilon_{n-1} + i_n) - f(\Upsilon_{n-1}) < c(n) - c(n-1)$, then stop*

Step 3 *Set $\Upsilon_n = \Upsilon_{n-1} + i_n$ and go to* Step 1.

So one first identifies the option i_1 whose expected payoff z_i is largest.[3] At any stage n, one finds the option i_n affording the largest marginal benefit over the college set constructed so far. The algorithm stops if the *net* marginal benefit turns negative.

Theorem 1 *The MIA identifies the optimal set Σ^* for problem* (2).

[3]The proof actually ignores the non-generic possibility of tied values of multiple argmax. With tied values, there exists a vanishing sequence of ε payoff-perturbations that renders Σ_n uniquely optimal along the ε-sequence. By the Theorem of the Maximum, this constant solution correspondence of the ε-perturbed problems gives the solution of the unperturbed limit problem. So the choice Σ_n is optimal.

Proof Let $P(n)$ be the statement "*$\Upsilon_n \subseteq \Sigma_k$, for all $k \geq n$*". We first show by induction that $P(n)$ holds for every n, thereby proving that $\Sigma_n = \Upsilon_n$ for $n = 1, ..., N$.

Assume $i_1 \notin \Sigma_k$ for some $k \geq 1$. If $(i_1, N] \cap \Sigma_k \neq \varnothing$, let b be the best option in this set and let $S = \Sigma_k - b + i_1$. Since $i_1 < b$ and $z_{i_1} > z_b$, $f(S) > f(\Sigma_k)$ by Lemma 2. Contradiction. If $(i_1, N] \cap \Sigma_k = \varnothing$, let w be the worst option in $[1, i_1) \cap \Sigma_k$ and let $S = \Sigma_k - w + i_1$. Then $z_{i_1} > z_w$ implies that $f(S) > f(\Sigma_k)$. Hence, $P(1)$ is true.

Suppose $P(n-1)$ holds; we must show that $P(n)$ holds as well. Assume $i_n \notin \Sigma_k$ for some $k \geq n$. By the induction hypothesis, $\Upsilon_{n-1} \subseteq \Sigma_k$.

If $(i_n, N] \cap (\Sigma_k - \Upsilon_{n-1}) \neq \varnothing$, let b be the best college in this set and $S = \Sigma_k - b + i_1$. If $\alpha_{i_n} \geq \alpha_b$, then $f(S) > f(\Sigma_k)$ by Lemma 2. Contradcition. Assume $\alpha_{i_n} < \alpha_b$. Let $M = (i_n, b) \cap \Sigma_k$, $L = (b, N] \cap \Sigma_k$, $M_{n-1} = (i_n, b) \cap \Upsilon_{n-1}$, and $L_{n-1} = (b, N] \cap \Upsilon_{n-1}$. Notice that $M = M_{n-1}$ and $L_{n-1} \subset L$. Using the DR decomposition (1), $f(L) > f(L_{n-1})$, and $\rho(i_n + M) > \rho(M + b)$, we obtain the following contradiction

$$\begin{aligned} f(\Upsilon_n) > f(\Upsilon_{n-1} + b) \quad &\Rightarrow \quad f(i_n + M) + \rho(i_n + M) f(L) > f(M + b) + \rho(M + b) f(L) \\ &\Leftrightarrow \quad f(S) > f(\Sigma_k). \end{aligned}$$

If $(i_n, N] \cap (\Sigma_k - \Upsilon_{n-1}) = \varnothing$, let w be the worst college in $[1, i_n) \cap (\Sigma_k - \Upsilon_{n-1})$ and let $S = \Sigma_k - w + i_1$. Define $U = [1, w) \cap \Sigma_k$, $M = (w, i_n) \cap \Sigma_k$, $L = (i_n, N] \cap \Sigma_k$, $U_{n-1} = [1, w) \cap \Upsilon_{n-1}$, $M_{n-1} = (w, i_n) \cap \Upsilon_{n-1}$, and $L_{n-1} = (i_n, N] \cap \Upsilon_{n-1}$. Notice that $M = M_{n-1}$ and $L = L_{n-1}$. Repeated application of the DR decomposition (1) yields

$$f(S) - f(\Sigma_k) \quad = \quad \frac{\rho(U)}{\rho(U_{n-1})} (f(\Upsilon_{n-1} + i_n) - f(\Upsilon_{n-1} + w)).$$

Since $f(\Upsilon_{n-1} + i_n) > f(\Upsilon_{n-1} + w)$, $f(S) > f(\Sigma_k)$, contradiction. Hence, $P(n)$ holds.

As the cost of a portfolio depends *only* on its size, $\Sigma^* = \Sigma_n$ for some n. The stopping rule is optimal since the cost $c(n)$ is convex in n and because f has *diminishing returns* — $f(S + k) - f(S)$ is decreasing in S for any $k \notin S \subseteq N$ — as we see below. □

To see diminishing returns, let us introduce the *marginal benefit* of adding college k to a set S. Partition $S = U_k + L_k$, where $U_k = [1, k) \cap S$ and $L_k = (k, N] \cap S$. Then

$$MB_k(S) = f(S + k) - f(S) = \rho(U_k)[z_k - \alpha_k f(L_k)] \tag{4}$$

Because $u_k > f(L_k)$, and ρ is decreasing and f increasing, we have:

Lemma 4 *Any DR function $f : 2^N \mapsto \mathbb{R}$ has diminishing returns.*

Intuitively, additions to the current portfolio grow less valuable as more options are added. Note that $v(S+i) - v(S) < v(i) - v(\varnothing) = z_i - c(1)$, whenever $S \neq \varnothing$, by Lemma 4. So choosing all options with $z_i > c(1)$ yields a suboptimally large portfolio.

4.3 Submodular Optimization

As noted, the value of a portfolio is less than the sum of its parts, since each option exerts a negative externality on the others. To cleanly capture this notion, call a function f on 2^N *submodular* if $f(S \cap T) + f(S \cup T) \leq f(S) + f(T)$ for any two subsets S and T of N.

Lemma 5 *Any DR function f is submodular, and thus so is $v : 2^N \mapsto \mathbb{R}$ in* (2)*.*

Proof First, f is submodular as it has diminishing returns.[4] Next, $-c(|S|)$ is a concave function, and so a submodular function (Proposition 5.1 in Lovász (1982) p.251).[5] □

It is well-known that the maximization of a *general* submodular set function is NP-hard and thus computationally intractable. Indeed, no polynomial algorithm exists for it (this is independent of the P≠ NP problem; see Lovász (1982), p. 252). By exploiting the special functional form of our objective function v, the MIA quickly finds the optimal set Σ^* for all DR submodular functions. One must in principle calculate the values of all 2^N college application patterns. Yet our algorithm succeeds in polynomial time: Initially, one examines N options and finds the best one. One then examines the next $N-1$ and finds the second best, etc. This amounts to $\sum_{i=0}^{N-1}(N-i) = N(N+1)/2 = O(N^2)$ steps.

Let us step back and ask whether the MIA's success pre-destined, in light of the recent theory of combinatorial optimization. One can show that $f(S)$ is (what is known as) *semi-strictly quasi-concave.*[6] As with standard quasiconcavity, local then implies global optimization. It does not, however, imply that a 'steepest ascent' algorithm like the MIA will succeed, as we prove it does for the class of DR payoff functions.

[4]See Proposition 1.1 in Lovász (1982). Gul and Stacchetti (1999) recently used this property in the economics literature. See related work by Kelso and Crawford (1982) on the gross substitutes condition.

[5]Observe that a sum of submodular functions, like $f + (-c)$, is submodular.

[6]See Murota and Shioura (2003). More precisely, $f(S)$ only satisfies a weak notion of semi-strictly quasi $M^\natural$-concavity, given by property $(SSQM_w)$ on p. 472. See the Appendix for a proof.

5 Properties of the Optimal Set

5.1 Aggressiveness of the Optimal Choices

How "risk-taking" should the portfolio be? To flesh this out, we employ *vector* first order stochastic dominance (FSD). The set $S \subseteq N$ is *more aggressive* 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, $\{1, 2\} \succ \{2, 3\}$.

We now compare the best set Σ^* against two easily computed benchmarks.

A. Portfolio Choices are more Aggressive than Top Singletons. Consider the set $Z_{|\Sigma^*|} \subseteq N$ of options with the $|\Sigma^*|$ highest expected payoffs $z_i = \alpha_i u_i$. Unlike the portfolio Σ^*, this set ignores the web of cross college external effects, as captured by (3).

Theorem 2 *The best portfolio Σ^* is more aggressive than the best singletons $Z_{|\Sigma^*|}$.*

Proof It suffices to show that if $i < j$ and $z_i > z_j$, then MIA picks i before j. By Lemma 2, for *any* portfolio S excluding i, j, we have $MB_i(S) > MB_j(S)$. $\square$

For an intuition, consider expression (3) for expected payoffs, i.e. $\sum_i z_{(i)} \rho_{i-1}(S)$. If options in Σ^* do not have the highest z_i's, then they must compensate with a higher $\rho_{i-1}(S)$. So acceptance chances are lower, and these options must be better ranked.

To see that the order can be strict, 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 $z_3 = 0.48 > z_2 = 0.45 > z_1 = 0.1$. One can show that $\Sigma_2(3) = \{1, 3\}$ which is strictly more aggressive than $Z_2 = \{2, 3\}$.

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.

The 'no safety school' substance of Theorem 2 is undermined with arbitrary correlation (not of the form in §3-C). For an extreme example with perfect correlation, assume three colleges, with payoffs $u_1 = 1$, $u_2 = u < 1$, $\alpha_1 = \alpha_2 = \alpha$, $u_3 = v < u$, $\alpha_3 = \alpha' > \alpha$, and $\alpha u > \alpha' v$. Suppose the student is either accepted in both 1 and 2 (chance α), or rejected in both. Exhaustive checking reveals that $\Sigma_2(2) = \{1, 3\}$, while $Z_2 = \{1, 2\}$.

B. Portfolio Choices are Less Aggressive than Sequential Choices. Consider the case where a student can apply to the colleges sequentially, observing whether one accepts her before she applies to the next. For a fair comparison, let us restrict to constant marginal costs $c(|S|)=\bar{c}|S|$, $\bar{c}>0$. The optimal policy in Weitzman (1979) is derived as follows. To each college i, associate an intrinsic *index* or reservation value I_i; this leaves the student indifferent between a final payoff I_i, and first applying to college i and then earning payoff I_i if rejected. Then $I_i = z_i - \bar{c} + (1-\alpha_i)I_i$, and thus $I_i = (z_i - \bar{c})/\alpha_i = u_i - \bar{c}/\alpha_i$. The optimal policy orders the colleges by their indices I_i; the student proceeds down the list, stopping when one accepts him (since $u_i > I_i$).

The solution of our static problem substantially differs from the sequential approach. For instance, we have shown that one must apply to the college with the largest expected payoff z_i. Easily, this needn't coincide with the one having the highest Gittins index I_i.

In general, the sequential decision-maker employs a more aggressive strategy than does our portfolio one. Let W be the list of colleges for which it is sequentially optimal to search, given continued failure, and $W_{|\Sigma^*|}$ the set with the $|\Sigma^*|$ highest indices I_i.

Theorem 3 *The best portfolio Σ^* is not larger than W, and less aggressive than $W_{|\Sigma^*|}$.*

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

We now show that $W_{|\Sigma^*|} \succeq \Sigma^*$. It suffices to show that if $i < j$, and S is *any* portfolio for which MIA picks i over j, then the Gittins indices are likewise ranked $I_i > I_j$. This is obvious if $\alpha_i > \alpha_j$, for then $I_i = u_i - \bar{c}/\alpha_i > u_j - \bar{c}/\alpha_j = I_j$. Otherwise, using the marginal benefit expression $MB_k(S) = \rho(U_k)[z_k - \alpha_k f(L_k)] > 0$ from (4), we find that:

$$I_i - I_j = \frac{z_i - \bar{c}}{\alpha_i} - \frac{z_j - \bar{c}}{\alpha_j} = \frac{1}{\alpha_i}\left(\frac{MB_i(S)}{\rho(U_i)} - \bar{c}\right) - \frac{1}{\alpha_j}\left(\frac{MB_j(S)}{\rho(U_j)} - \bar{c}\right) + [f(L_i) - f(L_j)]$$

If $MB_i(S) \geq MB_j(S)$, then $I_i > I_j$ as: $\rho(U_i) < \rho(U_j)$, $\alpha_i < \alpha_j$, and $f(L_i) > f(L_j)$. $\square$

To see that the order can be strict, assume three colleges, again with $\alpha_1 = 0.1$, $\alpha_2 = 0.9$, $\alpha_3 = 1$, $u_1 = 1$, $u_2 = 0.5$, $u_3 = 0.48$, but now $\bar{c} = 0.05$. One can show that $W_{|\Sigma^*|} = \{1, 2\}$, which is strictly more aggressive than $\Sigma^* = \{1, 3\}$.

5.2 Portfolio Choice Sets are Upwardly Diverse

We turn to another key characteristic of the statically optimal set. How similar should be the chosen options? Is the optimal set an "interval", say $[i, j]$?

Assume first $z_1 > z_2 > \cdots > z_n$. It follows from Theorem 2 that they should just apply to an interval of top schools. Indeed, $\Sigma^* \succeq Z_{|\Sigma^*|} = [1, |\Sigma^*|\,]$ implies $\Sigma^* = [1, |\Sigma^*|\,]$.

Apart from this case, a force to gamble upwards emerges and the optimal portfolio is not in general an interval. To see this, consider a stylized world with constant marginal cost $\bar{c} > 0$, one college i, and $N - 1$ copies of college $j > i$, with $z_j > z_i$. The algorithm starts with j. We claim that for N large enough and $\bar{c}$ small enough, the algorithm chooses college i before exhausting college j copies. Indeed, suppose the algorithm has chosen $n - 1$ copies of college j, but not yet college i. The marginal benefit of choosing another college j copy is $(1-\alpha_j)^{n-1}\alpha_j u_j - \bar{c}$. While this vanishes geometrically fast in n, the marginal benefit of choosing college i, namely $\alpha_i u_i - \alpha_i u_j(1 - (1 - \alpha_j)^{n-1}) - \bar{c}$, tends to $\alpha_i(u_i - u_j) - \bar{c}$. For small $\bar{c}$, this is positive. Thus, for large n, it is optimal to choose i over another copy of j. By continuity, this result obviously holds even when all copies of j are not literally identical and there is a sufficiently dense and diverse collection of colleges. So for low enough application costs, one always has an incentive to gamble upward, and apply to a discretely higher college than the rest.

5.3 Comparative Statics

We now consider some natural comparative statics. Obviously, with greater costs $c(\cdot)$, the size of Σ^* decreases, for the algorithm stops sooner.

More interestingly, how will choices change when acceptance chances $(\alpha_1, \ldots, \alpha_N)$ change? The answer is far from obvious, for the submodular character of f precludes any direct application of the monotone comparative statics results (see Topkis (1998)).

Theorem 4 *Assume $\beta = (\beta_1, ..., \beta_N)$ is higher than α and proportionately favors better options more than α. Namely, $\beta_i \geq \alpha_i$ for all i and $\beta_i/\alpha_i > \beta_{i+1}/\alpha_{i+1}$, for all $i < N$.*
(a) The best n-portfolio Σ_n^β is more aggressive than Σ_n^α, or $\Sigma_n^\beta \succeq \Sigma_n^\alpha$.
(b) Let $z_i^\alpha = \alpha_i u_i$ and $z_i^\beta = \beta_i u_i$, for all i, and let $z_1^\alpha > \cdots > z_N^\alpha$ and $z_1^\beta > \cdots > z_N^\beta$. The optimal sets are thus $[1, n^\alpha]$ and $[1, n^\beta]$. If also $(1 - \alpha_1)\alpha_2 > (1 - \beta_1)\beta_2$, then $n^\beta \leq n^\alpha$.

Proof of (a): The proof is a double induction on n and N. Let $\Sigma_n^\alpha(N)$ be the optimal n-choice set from N for acceptance chances α. 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 $\hat{n} \leq n$ and $\hat{N} \leq N$, with one inequality strict. If some $j \notin \Sigma_n^\alpha(N) \cup \Sigma_n^\beta(N)$, then the result holds by induction on the domain $N - j$. Assume there are no omitted options j. Thus, $1 \in \Sigma_n^\alpha(N) \cup \Sigma_n^\beta(N)$. If $1 \in \Sigma_n^\alpha(N) \cap \Sigma_n^\beta(N)$, then

$$\Sigma_n^\beta(N) = 1 + \Sigma_n^\beta(N-1) \succeq 1 + \Sigma_n^\alpha(N-1) = \Sigma_n^\alpha(N)$$

by Lemma 1. If $1 \notin \Sigma_n^\alpha(N)$, pick the least $k \notin \Sigma_n^\beta(N)$. Putting $M = [2, k-1]$, we have

$$\Sigma_n^\beta(N) = 1 + M + L^\beta \succeq M + k + L^\alpha = \Sigma_n^\alpha(N)$$

where $L^\alpha = \Sigma_n^\alpha(N) \cap [k, N]$ and $L^\beta \equiv \Sigma_n^\beta(N) \cap [k, N]$, by Lemma 1. Since $|L^\alpha| = |L^\beta|$, we have $L^\beta \succeq L^\alpha$ by the induction assumption. If $1 \notin \Sigma_n^\beta(N)$, then we can likewise decompose

$$\Sigma_n^\alpha(N) = 1 + M + L^\alpha \qquad \text{and} \qquad \Sigma_n^\beta(N) = M + k + L^\beta$$

where $L^\beta \succeq L^\alpha$. The appendix proves $f^\beta(1+M+L^\beta) > f^\beta(\Sigma_n^\beta(N))$, contradicting $\Sigma^\beta(N)$ optimal. This case cannot therefore arise. □

6 Boundaries of the MIA

Does the MIA extend beyond the class of DR problems considered in this paper? We now provide a different class of portfolio choice problems with general prize distributions ordered by stochastic dominance where the MIA succeeds. But we also temper this success by showing that the MIA can fail in problems that are not DR, or when costs are disparate and option-dependent, or with arbitrary correlation.

It will be useful to consider a generalization of problem (2). 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 $S \subseteq N$; after the choice is made, he draws a prize from each G_i, $i \in S$; and chooses the maximum of the realized prizes: That is, $f(S)$ is

$$f(S) = \int_0^{\bar{u}} \left(1 - \prod_{i \in S} G_i(u)\right) du. \tag{5}$$

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

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

Theorem 5 *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 $\Sigma^* = [1, k]$, for some $k \leq N$.*[7]

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 the index I_i of an option solves $\bar{c} = \int_0^{\bar{u}} \max\{0, u - I_i\} dG_i(u)$; 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.

Albeit not DR, the MIA works in this class of problems. The reason is that, under the stochastic dominance condition of Theorem 5, it is easy to show that $f(S)$ satisfies a strong notion of quasiconcavity that automatically guarantees the success of the MIA.

B. General 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, $u_i > w_i > 0$, with an α_i chance of u_i. Thus, $v(i) = \alpha_i u_i + (1 - \alpha_i) w_i - \bar{c}$, assumed positive.

Let $\bar{c} = 4.9$, $u_1 = 100 > u_2 = 80 > u_3 = 55.4 > w_1 = 50 > w_2 = 49 > w_3 = 0$, and $\alpha_1 = 0.1 < \alpha_2 = 0.2 < \alpha_3 = 1$. 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 option 3 in the first step.

[7]In the college application problem, the premise is equivalent to $z_1 > z_2 > \ldots > z_N$, which clearly yields $\Sigma^* = [1, k]$ for some k.

However, notice that $3 \notin \Sigma^* = \{1, 2\}$, for

$$\begin{aligned} v(\{1,2\}) &= \alpha_1 u_1 + (1-\alpha_1)\alpha_2 u_2 + (1-\alpha_1)(1-\alpha_2) w_1 - 2c = 60.40 - 9.8 = 50.60 \\ v(\{1,3\}) &= \alpha_1 u_1 + (1-\alpha_1)\alpha_3 u_3 + (1-\alpha_1)(1-\alpha_3) w_1 - 2c = 59.86 - 9.8 = 50.06 \\ v(\{2,3\}) &= \alpha_2 u_2 + (1-\alpha_2)\alpha_3 u_3 + (1-\alpha_2)(1-\alpha_3) w_2 - 2c = 60.32 - 9.8 = 50.52 \\ v(\{1,2,3\}) &= \alpha_1 u_1 + (1-\alpha_1)\alpha_2 u_2 + (1-\alpha_1)(1-\alpha_2) u_3 - 3c = 64.28 - 14.7 = 49.58 \end{aligned}$$

Note that this problem is *not* DR and does *not* satisfy the stochastic dominance order.

We omit an example showing that the MIA can also be suboptimal for sufficiently disparate option-dependent costs $\bar{c}_i > 0$.

C. Correlated Values. MIA can also fail with arbitrary correlation. Here, seeking a generic failure, we consider an extreme world with perfect correlation, where if option i pans out, option $i+1$ pans out as well. 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 $u_1 = 11, u_2 = 4, u_3 = 1$, and success 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 $\Sigma_2 = \{1, 3\}$, with value $\alpha_1 u_1 + (\alpha_3 - \alpha_1) u_3 = 2$.

Nagypál (2004) studies another world with perfect correlation in a college application setting, 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.

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 downward recursive static portfolio choice problems, such as where one earns only the best prize from a portfolio. Such portfolio choices are intriguing, insofar 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 choosing among uncertain technologies to explore.

We have shown that a greedy algorithm finds the optimal portfolio, and have identified the key properties that account for its success. This defines a useful class of submodular functions that can be efficiently maximized. We have also provided some interesting properties that the optimal set possesses.

It is an exciting open problem to find an algorithm that works efficiently with option-dependent costs and a richer set of prizes: future research beckons.

A Appendix

A.1 Proof of Theorem 4 Finished

Part (a). We need $f^\beta(1+M+L^\beta) > f^\beta(\Sigma_n^\beta(N))$. If $\alpha_1 > \alpha_k$ then $z_1 > z_k$, and the claim follows from Lemma 2. Assume hereafter $\alpha_1 < \alpha_k$. Then $f^\alpha(\Sigma_n^\alpha(N)) \geq f^\alpha(M+k+L^\alpha)$, since $\Sigma_n^\alpha(N)$ is optimal for f^α. Hence,

$$\alpha_1 u_1 + (1-\alpha_1)[f^\alpha(M) + \rho^\alpha(M) f^\alpha(L^\alpha)] \geq f^\alpha(M) + \rho^\alpha(M)[\alpha_k u_k + (1-\alpha_k) f^\alpha(L^\alpha)].$$

This holds if and only if

$$\frac{\alpha_1}{\alpha_k}\left(\frac{u_1 - f^\alpha(M)}{\rho^\alpha(M)}\right) + \left(1 - \frac{\alpha_1}{\alpha_k}\right) f^\alpha(L^\alpha) \geq u_k. \tag{6}$$

We now argue that replacing α by β yields a strict inequality in (6), which is likewise equivalent to $f^\beta(1+M+L^\beta) > f^\beta(\Sigma_n^\beta(N))$. We now justify this assertion:

- Since 1 dominates every option in $M+L^\alpha$ and $M+L^\beta$, we have $u_1 > f^\alpha(M+L^\alpha)$ and $u_1 > f^\beta(M+L^\beta)$. Using (1), 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 $\beta_1/\beta_k > \alpha_1/\alpha_k$, the weight on the first term of (6) strictly increases.
- $f^\beta(L^\beta) \geq f^\beta(L^\alpha) > f^\alpha(L^\alpha)$, respectively by Lemma 1, and because $\beta_i \geq \alpha_i$, for

all i, and $\beta_i > \alpha_i$ for some i (since the ratio ordering $\beta_i/\alpha_i > \beta_{i+1}/\alpha_{i+1}$ is strict)

- Finally, the first term in (6) increases as well, since

$$\frac{\partial}{\partial \alpha_\ell}\left(\frac{u_1 - f^\alpha(M)}{\rho^\alpha(M)}\right) > 0 \quad \forall \ell \in M \quad \Rightarrow \quad \frac{u_1 - f^\beta(M)}{\rho^\beta(M)} > \frac{u_1 - f^\alpha(M)}{\rho^\alpha(M)} \tag{7}$$

To see this, write $f^\alpha(M) = f^\alpha(U) + \rho^\alpha(U)[\alpha_\ell u_\ell + (1-\alpha_\ell)f^\alpha(L)]$ using (1), 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 (7) 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 option 1 dominates options in $[1, \ell] \cap M$, so that $u_1 > f(U + \ell)$.

Part (b). Parameterize $\theta = \alpha, \beta$, where $\alpha = \theta_L, \beta = \theta_H$. When $z_1^\theta > z_2^\theta > \cdots > z_N^\theta$, the restriction to $C = \{S \subseteq N | S = [1, n], n \leq N\}$ is wlog. So consider $\max_{S \subseteq C} v(S, \theta)$.

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

$$v(S_H, \theta_L) - v(S_L, \theta_L) \begin{cases} \geq 0 \\ > 0 \end{cases} \Rightarrow \; v(S_H, \theta_H) - v(S_L, \theta_H) \begin{cases} \geq 0 \\ > 0 \end{cases}$$

where $S_H = [1, n_H]$, $S_L = [1, n_L]$, and $n_L > n_H$. Rewrite the above as

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

for which a sufficient condition is

$$\rho^{\theta_H}([1,n_H])f^{\theta_H}((n_H,n_L]) \leq \rho^{\theta_L}([1,n_H])f^{\theta_L}((n_H,n_L]). \tag{8}$$

We now show that inequality (8) holds if $\beta_2(1-\beta_1) \leq \alpha_2(1-\alpha_1)$ and $\beta_i/\alpha_i > \beta_{i+1}/\alpha_{i+1}$. To prove this claim, we first show that (8) holds if and only if for all n

$$\rho^{\theta_H}([1,n])z_{n+1}^{\theta_H} \leq \rho^{\theta_L}([1,n])z_{n+1}^{\theta_L}. \tag{9}$$

Obviously, (8) implies (9). 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 (9).

To complete the proof, notice that (9) is equivalent to

$$\rho^{\theta_H}([1,n])\beta_{n+1} \leq \rho^{\theta_L}([1,n])\alpha_{n+1}, \tag{10}$$

so it suffices to show that (10) holds for all n under our assumptions. If $n = 1$, then (10) reduces to $\beta_2(1-\beta_1) \leq \alpha_2(1-\alpha_1)$, which holds by assumption. Suppose the result holds for $n-1$, i.e., $\rho^{\theta_H}([1,n-1])\beta_n \leq \rho^{\theta_L}([1,n-1])\alpha_n$. Then

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

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

A.2 Proof of Theorem 5

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 $\Upsilon_1 = \{i_1\} = \{1\}$, and therefore $1 \in \Sigma_k$ for all $k \geq 1$ if for any S such that $1 \notin S$, $v(S-b+1) - v(S) > 0$, where $b > 1$, $b \in S$. Integration by parts yields

$$v(S-b+1) - v(S) = \int_0^{\bar{u}} \left(\int_0^u \prod_{i \in S-b} G_i(\xi)d\xi\right) dG_1(u) - \int_0^{\bar{u}} \left(\int_0^u \prod_{i \in S-b} G_i(\xi)d\xi\right) dG_b(u),$$

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

Suppose $\Upsilon_{n-1} = [1, n-1] \subseteq \Sigma_k$ for all $k \geq n-1$. If $j \in [n, N]$, then

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

which is clearly maximized at $j = n$.

To show that $n \in \Sigma_k$ for all $k \geq n$, it suffices to show that if $\Upsilon_n - 1 \subseteq S$, $n \notin S$, $b \in S$, and $b > n$, then $v(S-b+n) - v(S) > 0$. But

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

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

We have thus proved that $\Upsilon_n = \Sigma_n$ for all n. Since $c(|S|)$ depends on the cardinality of S, it follows that $\Sigma^* = \Sigma_n = [1,n]$ for some n, thereby completing the proof. $\square$

A.3 Remarks on Concavity, Quasiconcavity, and the MIA

In these section we show that our objective function is not $M^\natural$-concave or semistrictly quasi $M^\natural$-concave, thus precluding a justification of the MIA based on these notions. We also show that our objective function satisfies a weaker notion of quasiconcavity, which only guarantees that local and global optimization coincide (but does not justify the MIA). Thus, we still need to prove that the MIA is optimal, as we do in Theorem 1.

All the concepts introduced below (as well as the notation) are borrowed from the aforementioned references Murota (2003), and Murota and Shioura (2003).[8]

A.3.1 M and $M^\natural$ Concave and Quasiconcave Functions

Let N be a finite set, Z be the set of integers, and $f : Z^N \to \mathbb{R} \cup \{-\infty\}$ be a function. We shall denote by *dom* f the set $\{x \in Z^N \mid f(x) > -\infty\}$, and by $supp^+(x)$ and $supp^-(x)$ the sets $\{i \in N \mid x(i) > 0\}$ and $\{i \in N \mid x(i) < 0\}$, respectively. Finally, for any vector $x \in Z^N$, let $x(N) = \sum_{i \in N} x(i)$ and $||x|| = \sum_{i \in N} |x(i)|$; and for any $i \in N$, let $\chi_i \in \{0,1\}^N$ be its characteristic vector (i.e., it is an N-dimensional vector with a 1 in the i-th coordinate and 0 everywhere else).

Remark 1 Notice that our function $f(S)$ can be written as $f : \{0,1\}^N \to \mathbb{R}$ by identifying each set $S \in N$ with its characteristic vector.

Definition 1 *$f : Z^N \to \mathbb{R} \cup \{-\infty\}$ is M-concave if $domf \neq \varnothing$ and $\forall\, x, y \in domf$, $\forall\, u \in supp^+(x-y)$, $\exists\, v \in supp^-(x-y)$ such that*

$$f(x) + f(y) \leq f(x - \chi_u + \chi_v) + f(y - \chi_v + \chi_u).$$

It is easy to show (Murota (2003), p.134) that if f is M-concave, then $domf \subset \{x \in Z^N \mid x(N) = r\}$ for some integer r. In other words, the sum of the elements is the same for all the vectors in the domain of f.

Remark 2 In the fixed sample size k of our problem the characteristic vectors of all the subsets in the domain of f have the same number of 'ones'; but if we allow for any size less than or equal to k (or if the sample size is endogenous), then the characteristic vectors of the feasible subsets do not have a constant number of ones.

The notion of $M^\natural$-concavity allows us to handle cases like the ones in the remark.

Definition 2 *Let 0 be an element not in N. A function $f : Z^N \to \mathbb{R} \cup \{-\infty\}$ is $M^\natural$-concave if the function $\tilde{f} : Z^{\{0\} \cup N} \to \mathbb{R} \cup \{-\infty\}$ defined by $\tilde{f}(x_0, x) = f(x)$ if $x_0 = -x(N)$ and $\tilde{f}(x_0, x) = -\infty$ otherwise, is M-concave.*

[8] We are grateful to Ennio Stacchetti for pointing out this literature to us.

For set functions, there is a nice characterization of $M^\natural$-concavity, which uses a property introduced by Gul and Stacchetti (1999).

Definition 3 *A set function $f(S)$ satisfies the single improvement property (SIP) if for any $p \in \mathbb{R}^N$, if $S \notin argmax\ f(S) - \sum_{i \in S} p(i)$, there exists $S' \subseteq N$ such that $|S - S'| \leq 1$, $|S' - S| \leq 1$, and $f(S) - \sum_{i \in S} p(i) < f(S') - \sum_{i \in S'} p(i)$.*

In words, f satisfies SIP if, for any perturbation of f by p, if a set is not optimal, then it can be improved by another set that either replaces an element for another, or adds an element, or subtracts an element from the original set. Murota (2003) p. 333 shows:

Proposition *$f(\chi_S)$ is $M^\natural$-concave* if and only if *$f(S)$ satisfies SIP.*

The notion of quasi M-concave function for discrete optimization was introduced by Murota and Shioura (2003). They propose several definitions, but the following are relevant for our purposes:

Definition 4 *$f : Z^N \to \mathbb{R} \cup \{-\infty\}$ is semistrictly quasi M-concave if $dom f \neq \varnothing$ and $\forall\, x, y \in dom f$, $\forall\, u \in supp^+(x - y)$, $\exists\, v \in supp^-(x - y)$ such that one of the following holds:*

(a) $f(x - \chi_u + \chi_v) - f(x) > 0$;
(b) $f(y - \chi_v + \chi_u) - f(y) > 0$;
(c) $f(x - \chi_u + \chi_v) - f(x) = f(y - \chi_v + \chi_u) - f(y) = 0$.

Murota and Shioura (2003) *do not* define the corresponding notion of quasi $M^\natural$-concavity, but it is straightforward to do so:

Definition 5 *Let 0 be an element not in N. A function $f : Z^N \to \mathbb{R} \cup \{-\infty\}$ is semistrictly quasi $M^\natural$-concave if the function $\tilde{f} : Z^{\{0\} \cup N} \to \mathbb{R} \cup \{-\infty\}$ defined by $\tilde{f}(x_0, x) = f(x)$ if $x_0 = -x(N)$ and $\tilde{f}(x_0, x) = -\infty$ otherwise, is semistrictly quasi M-concave.*

There are two important *implications* of these notions of concavity and quasiconcavity that are relevant for our purposes:

(i) Local optimality implies global optimality. That is, for an M-concave function,

$$f(x) \geq f(y) \, \forall \, y \in domf \iff f(x) \geq f(x - \chi_u + \chi_v) \, \forall \, u, v \in N,$$

and for an $M^\natural$-concave function,

$$f(x) \geq f(y) \, \forall \, y \in domf \iff f(x) \geq f(x - \chi_u + \chi_v) \, \forall \, u, v \in N, f(x) \geq f(x \pm \chi_v) \, \forall \, v \in N.$$

(ii) Success of Steepest Ascent Algorithm (SAA). The following algorithm, which is a greedy algorithm, reaches the global maximum.

Step 0: Let $x \in domf$, set $B = domf$.

Step 1: If $f(x)$ is a local maximum stop.

Step 2: Find u, v such that $x - \chi_u + \chi_v \in B$ and $f(x - \chi_u + \chi_v) = \max_{s,t} f(x - \chi_s + \chi_t)$.

Step 3: Set $x := x - \chi_u + \chi_v$ and $B := B \cap \{y \,|\, y(u) \leq x(u) - 1, \; y(v) \geq x(v) + 1\}$. Go to Step 1.

Notice that SAA starts from an arbitrary element in the domain of the function and selects, at each stage, the 'best local improvement.' Moreover, Step 3 shows that the domain is 'reduced' in each iteration (roughly, whatever is 'put in' remains and whatever is 'taken out' stays out from that point onward). We shall see shortly that MIA is a special case of SAA.

Murota and Shioura (2003) also define the following weaker notion of quasiconcavity. To distinguish it from the previous definition, let us add the qualifier 'weak.'

Definition 6 *$f : Z^N \to \mathbb{R} \cup \{-\infty\}$ is weak semistrictly quasi M-concave if $domf \neq \varnothing$ and $\forall \, x, y \in domf$, $\exists \, u \in supp^+(x - y)$, $\exists \, v \in supp^-(x - y)$ such that one of the following holds:*

(a) $f(x - \chi_u + \chi_v) - f(x) > 0$;
(b) $f(y - \chi_v + \chi_u) - f(y) > 0$;
(c) $f(x - \chi_u + \chi_v) - f(x) = f(y - \chi_v + \chi_u) - f(y) = 0$.

Notice that we only need to find a pair of u and v such that their exchange leads to an improvement. The extension of this definition to the $M^\natural$ case is the same as before.

Murota and Shioura (2003) show that if f is weak semistrictly quasi M (or $M^\natural$) concave, implication (i) holds, but it need not imply that SAA works (they use the stronger version to show implication (ii)). Inspection of their proof reveals that the problem is with Step 3: this notion of quasiconcavity need not guarantee that we can reduce the domain by keeping what we put in and throwing away what we take out.

A.3.2 Application of these Notions to our Problem

Consider the following optimization problem, which is the one that we analyze in the fixed sample size case:

$$\max_{S \subseteq N \,\&\, |S| \leq k} f(S). \tag{11}$$

By solving this problem using MIA for each $k = 1, ..., N$, we obtain the chain $\{i_1\}, \{i_1, i_2\}$, ..., $\{i_1, ..., i_N\}$. Then the 'stopping rule' (i.e., taking into account the cost function) determines what member of the chain is the optimal set.

Let us write this problem in terms of vectors of integers. With some abuse of notation, set $f(S) = f(\chi_S)$, where χ_S is the characteristic N-dimensional vector of the set S, with a coordinate 'one' on i if $i \in S$ and 'zero' otherwise.

Similarly, the set of subsets $S \subseteq N \,\&\, |S| \leq k$ can be written as:

$$C_k = \{x \in \{0,1\}^N \mid x = \chi_S \text{ } for \text{ } some \text{ } S \subseteq N \,\&\, |S| \leq k\}.$$

Remark 3 For example, if $N = 3$ and $k = 2$, then $C_k = \{000, 100, 010, 001, 110, 101, 011\}$.

Notice that $x(N)$ is not constant on C_k. In order to apply the concepts of the previous section, let us 'add' a coordinate 0 as follows:

$$\tilde{C}_k = \{(x_0, x) \mid x \in C_k, x_0 = -x(N)\},$$

and let $\tilde{f}(x_0, x) = f(x)$ if $x_0 = -x(N)$ and $\tilde{f}(x_0, x) = -\infty$.

Remark 4 If $N = 3$ and $k = 2$, then $\tilde{C}_k = \{0000, -1100, -1010, -1001, -2110, -2101, -2011\}$.

Thus, (11) can be written as

$$\max_{\tilde{x} \in \tilde{C}_k} \tilde{f}(\tilde{x}).$$

If we could show that $\tilde{f}(\tilde{x})$ is $M^\natural$-concave or semistrictly quasi $M^\natural$-concave, then this would explain why MIA works (see (i) and (ii) above). To see that MIA is a special case of SAA, simply set the initial vector equal to 00...0. In Step 2, it will choose the element with the largest z (and add a minus one to the first coordinate), and in Step 3 it will keep that element and only allow for inspection of the remaining elements in subsequent iterations.

Proposition *$f(S)$ is not $M^\natural$-concave or semistrictly quasi $M^\natural$-concave.*

Proof To show that $f(S)$ is not $M^\natural$-concave, we show that it does not satisfy SIP (see Proposition 1). It suffices to provide an example where SIP fails.

Let $N = 3$, $u_1 = 8.5$, $u_2 = 8$, $u_3 = 5$, $\alpha_1 = 0.4$, $\alpha_2 = 0.8$, $\alpha_3 = 0.9$, $p_1 = 1$, $p_2 = 2.5$, $p_3 = 1$. Notice that $\{1, 3\}$ maximizes $f(S) - p \cdot |S|$, with value equal to 4.1.

Take $S = 2$, so $f(2) - p_2 = 3.9$. The only candidates for a single improvement are $S' = \varnothing$, $S' = \{1, 2\}$, or $S' = \{2, 3\}$, with $f(\varnothing) = 0$, $f(1, 2) - p_1 - p_2 = 3.74$, and $f(2, 3) - p_2 - p_3 = 3.8$. Thus, no single improvement exists in this case, thereby showing that f is not $M^\natural$-concave.

To show that $f(S)$ is not semistrictly quasi $M^\natural$-concave, consider the following example. Let $N = 5$, and $u_1 = 10$, $u_2 = 5.6$, $u_3 = 4.91666$, $u_4 = 2$, $u_5 = 1$, $\alpha_1 = 0.2$, $\alpha_2 = 0.5$, $\alpha_3 = 0.6$, $\alpha_4 = 0.8$, $\alpha_5 = 1$. Thus, $z_1 = 2$, $z_2 = 2.8$, $z_3 = 2.95$, $z_4 = 1.6$, $z_5 = 1$.

Consider the sets $\{2, 4, 5\}$ and $\{1, 3, 5\}$, or in vector notation (recall the coordinate x_0 that we add) $x = -301011$ and $y = -310101$. Hence, $supp^+(x - y) = \{2, 4\}$ and $supp^-(x - y) = \{1, 3\}$. Notice that $f(x) = f(2, 4, 5) = 3.7$, and $f(y) = f(1, 3, 5) = 4.68$.

Take $u = 2$. We will show that for $v = 1$ and $v = 3$, it is the case that both $f(x - \chi_u + \chi_v) - f(x) < 0$ and $f(y - \chi_v + \chi_u) - f(y) < 0$, thereby showing the semistrictly quasi $M^\natural$-concavity fails. If $v = 1$, then $f(x - \chi_u + \chi_v) = f(1, 4, 5) = 3.44 < 3.7$ and $f(y - \chi_v + \chi_u) = f(2, 3, 5) = 4.475 < 4.68$. If $v = 3$, then $f(x - \chi_u + \chi_v) = f(3, 4, 5) = 3.67 < 3.7$ and $f(y - \chi_v + \chi_u) = f(1, 2, 5) = 4.64 < 4.68$. This completes the proof. $\square$

Although $f(S)$ is not $M^\natural$-concave or semistrictly quasi $M^\natural$-concave, it does satisfy the weak notion of semistrict quasiconcavity.

Proposition *f is weak semistrictly quasi $M^\natural$-concave for all $k = 1, 2, ..., N$.*

Proof Fix k. Since the result is trivial for $k = 1$ (it is just a matter of comparing singletons), we may assume that $k > 1$.

Take $x, y \in \tilde{C}_k$. We must show that $\exists\, u \in supp^+(x - y)$, $\exists\, v \in supp^-(x - y)$ such that (a), (b) or (c) holds. Let S_x be the set represented by x and S_y the one by y.

If $x_0 < y_0$, which means that S_x has more elements than S_y, the result is true by the monotonicity of f. In this case, take any $u \in supp^+(x - y)$ and $v = 0 \in supp^-(x - y)$. Then $f(y - \chi_v + \chi_u) - f(y) > 0$ (we are adding an element to S_y) and (b) holds.

If $x_0 > y_0$, which means that S_y has more elements than S_x, the result is true by the monotonicity of f. In this case, take any $u = 0 \in supp^+(x-y)$ and any $v \in supp^-(x-y)$. Then $f(x - \chi_u + \chi_v) - f(x) > 0$ (we are adding an element to S_x) and (a) holds.

The only remaining case to consider is when $x_0 = y_0$, i.e., $|S_x| = |S_y|$.

By Theorem 3.19 in Murota and Shioura (2003), it is enough to restrict attention to $x, y \in \tilde{C}_k$ such that $||x - y|| = \sum_{i=1}^{N} |x(i) - y(i)| = 4$. Since we are assuming $x_0 = y_0$, S_x and S_y differ only by two elements; i.e., $supp^+(x - y)$ contains two elements, say u_1 and u_2 with $u_1 < u_2$, and $supp^-(x - y)$ contains two elements, say v_1 and v_2 with $v_1 < v_2$.

Let $\hat{k}$ be the largest k such that $\sum_{i=k}^{N} |x(i) - y(i)| = 3$. Wlog, suppose that $\hat{k} \in supp^+(x - y)$ (the arguments can be trivially adapted if instead we assume that $\hat{k} \in supp^-(x - y)$). Then, $\hat{k}$ is either equal to u_1 or to u_2.

Suppose $\hat{k} = u_2$. Since $\sum_{i=u_2}^{N} |x(i) - y(i)| = 3$, it follows that $u_2 < v_1 < v_2$. Moreover, $\sum_{i=u_2}^{N} y(i) - \sum_{i=u_2}^{N} x(i) = 1$ (i.e., the 'tail' of vector y contains one more element than the tail of vector x). We shall show that if we choose $u = u_2$ and $v = v_1$, then either (a) or (b) of Definition 6 hold. Let $L_x = S_x \cap [u_2, N]$ and $L_y = S_y \cap [u_2, N]$. Then $|L_y| = |L_x| + 1$. Consider L_x and $L_x - u_2 + v_1$.

If $f(L_x - u_2 + v_1) > f(L_x)$, then by downward recursiveness $f(S_x - u_2 + v_1) > f(S_x)$. But, by definition, $f(x - \chi_{u_2} + \chi_{v_1}) = f(S_x - u_2 + v_1)$ and $f(x) = f(S_x)$. Hence, $f(x - \chi_{u_2} + \chi_{v_1}) > f(x)$ and (a) holds.

If $f(L_x) > f(L_x - u_2 + v_1)$, then $f(L_x + v_2) > f(L_x - u_2 + v_1 + v_2) = f(L_y)$. The proof is as follows. We can write $f(L_x + v_2)$ and $f(L_y)$ as

$$f(L_x) + \rho([u_2, v_2] \cap L_x)(z_{v_2} - \alpha_{v_2} f([u_2, N] \cap L_x)),$$

$$f(L_x - u_2 + v_1) + \rho([u_2, v_2] \cap L_x - u_2 + v_1)(z_{v_2} - \alpha_{v_2} f([u_2, N] \cap L_x - u_2 + v_1)).$$

But $f([u_2, N] \cap L_x) = f([u_2, N] \cap L_x - u_2 + v_1)$; $z_{v_2} - \alpha_{v_2} f([u_2, N] \cap L_x) > 0$; $f(L_x) > f(L_x - u_2 + v_1)$; and $\rho([u_2, v_2] \cap L_x) > \rho([u_2, v_2] \cap L_x - u_2 + v_1)$.

Hence, $f(L_x + v_2) > f(L_y)$. By downward recursiveness, it follows that $f(S_y - v_1 + u_2) > f(S_y)$. Since, by definition, $f(y - \chi_{v_1} + \chi_{u_2}) = f(S_y - v_1 + u_2)$ and $f(y) = f(S_y)$, it follows that $f(y - \chi_{v_1} + \chi_{u_2}) > f(y)$ and hence (b) holds.

We have shown that if $\hat{k} = u_2$, then either (a) or (b) holds. If instead $\hat{k} = u_1$, then $v_1 < u_1 < v_2$. But then there is an alternative k, equal to u_2 or to v_2, such that the tails of S_x and S_y below k, say L'_x and L'_y, contain the same number of elements. Then either $f(L'_x - u_2 + v_2) > f(L'_x)$, or $f(L'_y - v_2 + u_2) > f(L'_y)$. By downward recursiveness, either (a) or (b) hold. This completes the proof. □

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.

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. (1982): "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.

Murota, K., and A. Shioura (2003): "Quasi M-convex and L-convex Functions—Quasiconvexity in Discrete Optimization," *Discrete Applied Mathematics*, 131, 467–494.

NAGYPÁL, E. (2004): "Applications with Incomplete Information," Northwestern University mimeo.

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): "Parallel Search for the Best Alternative," *Economic Theory*, 2(1), 495–507.

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