Supplemental Appendix to
Hedging When Applying: Simultaneous Search with Correlation

S. Nageeb Ali   Ran Shorrer

The supplemental appendices are organized as follows:

- Appendix B proves Lemma 1.
- Appendix C collects results pertaining to the the algorithm.
- Appendix D collects additional results for the common-score model.
- Appendix E collects results for the simultaneous search framework with independent success.
- Appendix F collects all examples and proofs for Section 4.1.
- Appendix G collects all examples and proofs for Section 4.2.

B Proof of Lemma 1

We proceed by induction on k. To simplify notation, we write G for G' and H for H'. Observe that if either $P^*(k, G)$ or $P^*(k, H)$ has fewer than k elements, the conclusion follows from Cases 1 and 2 of the proof of Proposition 1, each of which was proven independently. Hence, we assume below that $|P^*(k, G)| = |P^*(k, H)| = k$.

Base Step ($k = 1$): Suppose $P^*(1, H) = \{i\}$. Suppose towards a contradiction that $P^*(1, G) = \{j\}$ for some $j > i$. Then both Colleges i and j are strictly preferred to the outside option and

$$(1 - G(\tau_i))(u_i - u_o) < (1 - G(\tau_j))(u_j - u_o)$$
$$(1 - H(\tau_j))(u_j - u_o) < (1 - H(\tau_i))(u_i - u_o).$$

Multiplying the inequalities and dividing by $(u_i - u_o)(u_j - u_o)$ yields

$$(1 - G(\tau_j))(1 - H(\tau_i)) < (1 - G(\tau_i))(1 - H(\tau_j)).$$

However, since $G \geqslant_{LR} H$ and $j > i$, we have that

$$\begin{aligned}(1 - G(\tau_i))\left(1 - H\left(\tau_j\right)\right) &= \sum_{l,p \leqslant i} \mu(l, G)\mu(p, H) + \sum_{l \leqslant i < p \leqslant j} \mu(l, G)\mu(p, H) \\ &\geqslant \sum_{l,p \leqslant i} \mu(l, G)\mu(p, H) + \sum_{l \leqslant i < p \leqslant j} \mu(p, G)\mu(l, H) \\ &= (1 - H(\tau_i))\left(1 - G\left(\tau_j\right)\right),\end{aligned}$$

resulting in a contradiction.

Inductive Step: We consider the following inductive hypothesis:

> *For an integer $k > 1$, for all $k' \in \{1, \ldots, k-1\}$, the top choice in the optimal k'-portfolio under distribution $\tilde{G}$ is higher ranked than that under distribution $\tilde{H}$ whenever $\tilde{G} \geqslant_{LR} \tilde{H}$.*

We prove that if the inductive hypothesis is satisfied, then it must also be true that the top choice in the optimal k-portfolio under distribution G is higher ranked than that under H.

Suppose towards a contradiction that there is an integer $k > 1$, and distributions G and H such that $G \succcurlyeq_{LR} H$, $P^*_{(1)}(k, G)$ is lower ranked than $P^*_{(1)}(k, H)$, and for which the italicized statement above is true. We denote $P^*(k, G)$ by P and $P^*(k, H)$ by R.

Observe that by (4), (5), and the inductive hypothesis, there exists $m \in \{1, \ldots, k\}$ such that $P_{(j)} > R_{(j)}$ for all $j \leqslant m$ and $P_{(j)} \leqslant R_{(j)}$ for all $j > m$. Consider a chain of portfolios $Q^0, Q^1, \ldots, Q^m$ in which $Q^0 := R$, and for $i \in \{1, \ldots, m\}$, Q^i consists of $\lfloor R \rfloor_{k-i}$ along with $\lceil P \backslash (\lfloor R \rfloor_{k-i}) \rceil^i$. When P and R are disjoint, Q^i is the portfolio that comprises the top i colleges from portfolio P and the $(i+1)$ to k top colleges from portfolio R. We erase repetitions on the chain of portfolios and re-index the sequence $\{Q^0, \ldots, Q^{m'}\}$. Our argument uses this chain of portfolios to find a portfolio that outperforms portfolio P under score distribution G, contradicting the optimality of $P = P^*(k, G)$.

Claim 1. *For each $j \in \{1, \ldots, m'\}$, $V(Q^j, G) < V(Q^0, G)$.*

Proof of Claim 1. We proceed by induction.

Base Step ($j = 1$): As $R = Q^0$ is uniquely optimal under distribution H, $V(Q^1, H) < V(Q^0, H)$. Let w be the index such that $Q^1_{(w)} = Q^1 \backslash Q^0$ (and recall that $Q^0_{(0)} = Q^0 \backslash Q^1$).[27] Writing out each value explicitly, canceling common terms, and re-arranging, one obtains

$$\left(1 - H(\tau_{Q^0_{(1)}})\right)(u_{Q^0_{(1)}} - u_{Q^1_{(1)}}) > \left(H\left(\min\{\tau_{Q^0_{(w)}}, \tau_{Q^1_{(w-1)}}\}\right) - H(\tau_{Q^1_{(w)}})\right)(u_{Q^1_{(w)}} - u_{Q^1_{(w+1)}}).$$

Because $G \succcurlyeq_{LR} H$, and $Q^0_{(1)} = R_{(1)}$ is higher ranked than $P_{(1)}$ which, in turn is weakly higher ranked than $Q^1_{(w)}$, we have that

$$\left(1 - G(\tau_{Q^0_{(1)}})\right)\left(H\left(\min\{\tau_{Q^0_{(w)}}, \tau_{Q^1_{(w-1)}}\}\right) - H(\tau_{Q^1_{(w)}})\right) \geqslant \left(G\left(\min\{\tau_{Q^0_{(w)}}, \tau_{Q^1_{(w-1)}}\}\right) - G(\tau_{Q^1_{(w)}})\right)\left(1 - H(\tau_{Q^0_{(1)}})\right).$$

Note that the RHS of the above inequality is strictly positive.[28] Multiplying the two inequalities and dividing each side by common terms (all of which are strictly positive), we obtain

$$\left(1 - G(\tau_{Q^0_{(1)}})\right)(u_{Q^0_{(1)}} - u_{Q^0_{(1)}}) > \left(G\left(\min\{\tau_{Q^0_{(w)}}, \tau_{Q^1_{(w-1)}}\}\right) - G(\tau_{Q^1_{(w)}})\right)(u_{Q^1_{(w)}} - u_{Q^1_{(w+1)}})$$

which implies that $V(Q^1, G) < V(Q^0, G)$.

Inductive Step ($j > 1$): We assume that $V(Q^\ell, G) < V(Q^0, G)$ for all $\ell < j \leqslant m'$, and show that this inequality also holds for $\ell = j$.

For $\ell \in \{1, \ldots, m'\}$ and distribution $D \in \{G, H\}$, we denote $Q^\ell_{(w^\ell)} := Q^\ell \backslash Q^{\ell-1}$, $Q^{\ell-1}_{(z^\ell)} := Q^{\ell-1} \backslash Q^\ell$,

$$W^\ell_D := \left(D\left(\min\left\{\tau_{Q^\ell_{(w^\ell-1)}}, \tau_{Q^{\ell-1}_{(z^\ell)}}\right\}\right) - D\left(\tau_{Q^\ell_{(w^\ell)}}\right)\right)\left(u_{Q^\ell_{(w^\ell)}} - u_{Q^\ell_{(w^\ell+1)}}\right),$$

[27] To simplify notation, we abuse notation by omitting some braces.

[28] Were the first term equal to 0, replacing $Q^1_{(w)}$ with $R_{(1)}$ would achieve a higher value under distribution G; the second term is strictly positive as $R_{(1)} = Q^0_{(1)}$ is rationalizable under distribution H.

and

$$L_D^\ell := \left(D\left(\tau_{Q^{\ell-1}_{(z^\ell-1)}}\right) - D\left(\tau_{Q^{\ell-1}_{(z^\ell)}}\right)\right)\left(u_{Q^{\ell-1}_{(z^\ell)}} - u_{Q^{\ell}_{(z^\ell)}}\right).$$

Then, for each $\ell \in \{1, \dots, m'\}$ and $D \in \{H, G\}$, $V(Q^\ell, D) - V(Q^{\ell-1}, D) = W_D^\ell - L_D^\ell$.

We write $q_\ell := \frac{\mu(Q^{\ell-1}_{(z^\ell)}, G)}{\mu(Q^{\ell-1}_{(z^\ell)}, H)}$.[29] Then,

$$\frac{L_G^\ell}{L_H^\ell} = \frac{G\left(\tau_{Q^{\ell-1}_{(z^\ell-1)}}\right) - G\left(\tau_{Q^{\ell-1}_{(z^\ell)}}\right)}{H\left(\tau_{Q^{\ell-1}_{(z^\ell-1)}}\right) - H\left(\tau_{Q^{\ell-1}_{(z^\ell)}}\right)} \geqslant q_\ell \geqslant \frac{G\left(\min\left\{\tau_{Q^{\ell}_{(w^\ell-1)}}, \tau_{Q^{\ell-1}_{(z^\ell)}}\right\}\right) - G\left(\tau_{Q^{\ell}_{(w^\ell)}}\right)}{H\left(\min\left\{\tau_{Q^{\ell}_{(w^\ell-1)}}, \tau_{Q^{\ell-1}_{(z^\ell)}}\right\}\right) - H\left(\tau_{Q^{\ell}_{(w^\ell)}}\right)} = \frac{W_G^\ell}{W_H^\ell}. \tag{15}$$

As $Q^0 = R$ is the uniquely optimal k-portfolio under distribution H, $V(Q^j, H) < V(Q^0, H)$ and therefore, $\sum_{\ell=1}^j L_H^\ell > \sum_{\ell=1}^j W_H^\ell$.

Observe that

$$\begin{aligned}
\sum_{\ell=1}^{j} L_G^\ell &= L_G^j + \sum_{\ell=1}^{j-1}\left(L_G^\ell - W_G^\ell + W_G^\ell\right) = L_G^j + \sum_{\ell=1}^{j-1} W_G^\ell + \sum_{\ell=1}^{j-1}\left(L_G^\ell - W_G^\ell\right) \\
&\geqslant q_j L_H^j + \sum_{\ell=1}^{j-1} W_G^\ell + \sum_{\ell=1}^{j-1} q_\ell\left(L_H^\ell - W_H^\ell\right) \\
&= q_j L_H^j + \sum_{\ell=1}^{j-1} W_G^\ell + \sum_{\ell=1}^{j-1} q_j\left(L_H^\ell - W_H^\ell\right) + \sum_{b=1}^{j-1}(q_b - q_{b+1})\left(V(Q^0, H) - V(Q^b, H)\right) \\
&> q_j L_H^j + \sum_{\ell=1}^{j-1} W_G^\ell + \sum_{\ell=1}^{j-1} q_j\left(L_H^\ell - W_H^\ell\right) = \sum_{\ell=1}^{j-1} W_G^\ell + \sum_{\ell=1}^{j} q_j\left(L_H^\ell - W_H^\ell\right) + q_j W_H^j \\
&= q_j\left(V(Q^0, H) - V(Q^j, H)\right) + \sum_{\ell=1}^{j-1} W_G^\ell + q_j W_H^j > \sum_{\ell=1}^{j-1} W_G^\ell + q_j W_H^j \geqslant \sum_{\ell=1}^{j} W_G^\ell.
\end{aligned}$$

where the first inequality uses (15), the second inequality follows from q_b being decreasing in b and $V(Q^0, H) > V(Q^b, H)$, the third inequality uses $V(Q^0, H) > V(Q^j, H)$ and that q_j is nonnegative, and the final inequality uses (15). Hence, we obtain that $V(Q^j, G) - V(Q^0, G) = \sum_{\ell=1}^j W_G^\ell - \sum_{\ell=1}^j L_G^\ell < 0$, establishing the inductive step. □

We complete the proof of Lemma 1 by noting that

$$\begin{aligned}
&V\left(\{R_{(1)}, R_{(2)}, \dots, R_{(m)}, P_{(m+1)}, \dots, P_{(k)}\}, G\right) - V(P, G) \\
&= \sum_{\ell=1}^{m}\left(L_G^\ell - W_G^\ell\right) + \left(G(\tau_{R_{(m)}}) - G(\tau_{P_{(m)}})\right)\left(u_{P_{(m+1)}} - u_{R_{(m+1)}}\right)\mathbb{1}_{m<k} \\
&= \left[V(Q^0, G) - V(Q^m, G)\right] + \left(G(\tau_{R_{(m)}}) - G(\tau_{P_{(m)}})\right)\left(u_{P_{(m+1)}} - u_{R_{(m+1)}}\right)\mathbb{1}_{m<k}.
\end{aligned}$$

Claim 1 establishes that the first term is strictly positive. By inspection, the second term is non-

[29] This expression is well defined since $Q^{\ell-1}_{(z_\ell)} \in P^*(k, H)$, and so the denominator is greater than zero.

negative.[30] Therefore, the portfolio $\{R_{(1)}, R_{(2)}, \ldots, R_{(m)}, P_{(m+1)}, \ldots, P_{(k)}\}$ performs strictly better than portfolio P under distribution G, yielding a contradiction.

C Algorithms for the Optimal Portfolio

C.1 Speeding up the Algorithm Using Proposition 1

The main text presents an algorithm for calculating the optimal portfolio in $O(n^3)$ computation steps, where each step of the algorithm requires $O(n^2)$ computation steps and the algorithm comprises n steps. We use Proposition 1 to reduce the number of computation steps required for each step of the algorithm to $O(n \log n)$. Therefore only $O(n^2 \log n)$ computation steps are required in total.

We proceed with a faster routine for executing Step k of our algorithm. Recall that $\mathcal{C}^\dagger := \mathcal{C} \cup \{0\}$, where 0 is a fictitious college that rejects all applications (i.e., $\tau_0 = 1$).

Stage 1. For $i_1 = \mathtt{median}(\mathcal{C}^\dagger)$ (throughout, when the median is not an integer, we round it down to the nearest integer) find the optimal k-college portfolio following rejection from i:

$$j_1^* := \underset{j \in \mathcal{C}^\dagger \text{ s.t. } \{j\} \leqslant_A \{i_1\}}{\operatorname{argmax}} \left\{(1 - F_{i_1}(\tau_j))u_j + F_{i_1}(\tau_j)V(C(j, k-1), F_{\{j\}})\right\}.$$

where the continuation $C(j_1^*, k-1)$ is solved in step $(k-1)$. The *optimal k-college continuation* following i_1 is then $C(i_1, k) := \{j_1^*\} \cup C(j^*, k-1)$.

Stage 2.1. For $i_{2.1} = \mathtt{median}(\{0, \ldots, i_1 - 1\})$, find the optimal k-college portfolio following rejection from $i_{2.1}$:

$$j_{2.1}^* := \underset{j \in \mathcal{C}^\dagger \text{ s.t. } \{j_1^*\} \leqslant_A \{j\} \leqslant_A \{i\}}{\operatorname{argmax}} \left\{(1 - F_{i_{2.1}}(\tau_j))u_j + F_{i_{2.1}}(\tau_j)V(C(j, k-1), F_{\{j\}})\right\}.$$

where the continuation $C(j, k-1)$ is solved in step $(k-1)$. The restriction to colleges at least as aggressive as j_1^* is justified by Proposition 1. The *optimal k-college continuation* following $i_{2.1}$ is then $C(i_{2.1}, k) := \{j_{2.1}^*\} \cup C(j_{2.1}^*, k-1)$.

Stage 2.2. For $i_{2.2} = \mathtt{median}(\{i_1 + 1, \ldots, n\})$, find the optimal k-college portfolio following rejection from $i_{2.2}$:

$$j_{2.2}^* := \underset{j \in \mathcal{C}^\dagger \text{ s.t. } \{j\} \leqslant_A \{j_1^*\}}{\operatorname{argmax}} \left\{(1 - F_{i_{2.2}}(\tau_j))u_j + F_{i_{2.2}}(\tau_j)V(C(j, k-1), F_{\{j\}})\right\}.$$

where the continuation $C(j, k-1)$ is solved in step $(k-1)$. The restriction of colleges no more aggressive than j_1^* is justified by Proposition 1. The *optimal k-college continuation* following $i_{2.2}$ is then $C(i_{2.2}, k) := \{j_{2.2}^*\} \cup C(j_{2.2}^*, k-1)$.

⋮

[30] If $m < k$, then $P_{(m+1)}$ is a higher-ranked college than $R_{(m+1)}$. Thus, $u_{P_{(m+1)}} - u_{R_{(m+1)}} > 0$. Moreover, college $R_{(m)}$ is higher ranked than $P_{(m)}$, and therefore, $H(\tau_{R_{(m)}}) - H(\tau_{P_{(m)}}) \geqslant 0$.

Stage $m.1$. For $i_{m.1} = \texttt{median}\left(\{0, \ldots, i_{(m-1).1} - 1\}\right)$, find the optimal k-college portfolio following rejection from $i_{m.1}$:

$$j^*_{m.1} := \underset{j \in \mathcal{C}^\dagger \text{ s.t. } \{j^*_{(m-1).1}\} \preccurlyeq_A \{j\} \preccurlyeq_A \{i\}}{\text{argmax}} \left\{(1 - F_{i_{m.1}}(\tau_j))u_j + F_{i_{m.1}}(\tau_j)V(C(j, k-1), F_{\{j\}})\right\}.$$

where the continuation $C(j, k-1)$ is solved in step $(k-1)$. The restriction of colleges no more aggressive than $j^*_{(m-1).1}$ is justified by Proposition 1. The *optimal k-college continuation* following $i_{m.1}$ is then $C(i_{m.1}, k) := \{j^*_{m.1}\} \cup C(j^*_{m.1}, k-1)$.

Stage $m.2$. For $i_{m.2} = \texttt{median}\left(\{i_{(m-1).1} + 1, \ldots, i_{(m-1).2} - 1\}\right)$, find the optimal k-college portfolio following rejection from $i_{m.2}$:

$$j^*_{m.2} := \underset{j \in \mathcal{C}^\dagger \text{ s.t. } \{j^*_{(m-1).2}\} \preccurlyeq_A \{j\} \preccurlyeq_A \{j^*_{(m-1).1}\}}{\text{argmax}} \left\{(1 - F_{i_{m.2}}(\tau_j))u_j + F_{i_{m.2}}(\tau_j)V(C(j, k-1), F_{\{j\}})\right\}.$$

where the continuation $C(j, k-1)$ is solved in step $(k-1)$. The restriction of colleges no more aggressive than $j^*_{(m-1).2}$ and no less aggressive than $j^*_{(m-1).1}$ is justified by Proposition 1. The *optimal k-college continuation* following $i_{m.2}$ is then $C(i_{m.2}, k) := \{j^*_{m.2}\} \cup C(j^*_{m.2}, k-1)$.

⋮

Stage $m.2^{m-1}$. For $i_{m.2^{m-1}} = \texttt{median}\left(\left\{i_{(m-1).2^{m-2}} + 1, \ldots, n\right\}\right)$, find the optimal k-college portfolio following rejection from $i_{m.2^{m-1}}$:

$$j^*_{m.2^{m-1}} := \underset{j \in \mathcal{C}^\dagger \text{ s.t. } \{j\} \preccurlyeq_A \{j^*_{(m-1).2^{m-2}}\}}{\text{argmax}} \left\{(1 - F_{i_{m.2^{m-1}}}(\tau_j))u_j + F_{i_{m.2^{m-1}}}(\tau_j)V(C(j, k-1), F_{\{j\}})\right\}.$$

where the continuation $C(j, k-1)$ is solved in step $(k-1)$. The restriction of colleges no less aggressive than $j^*_{(m-1).2^{m-2}}$ is justified by Proposition 1. The *optimal k-college continuation* following $i_{m.2^{m-1}}$ is then $C(i_{m.2^{m-1}}, k) := \{j^*_{m.2^{m-1}}\} \cup C(j^*_{m.2^{m-1}}, k-1)$.

Note that by the m-th stage, the routine solves for $1 + 2 + \cdots + 2^{m-1} = 2^m - 1$ optimal k-college continuations. Hence, the routine requires at most $\lceil \log_2(n+1) \rceil$ stages to complete Step k of the algorithm. Furthermore, using Proposition 1 we restricted the arguments under the argmax to be such that in each stage of the routine no more than $3n/2$ options must be considered. For example, in Stage 2.1 the routine only searches colleges that are at least as aggressive as j^*_1 while in Stage 2.2 it only searches colleges that are no more aggressive than j^*_1, and so j^*_1 is considered twice and each of the other $n-1$ colleges is considered at most once.

In sum, the routine requires only $O(n \log n)$ calculation steps for each step of the algorithm, bringing the number of calculation steps required by all n steps of the algorithm to $O(n^2 \log n)$.

C.2 Accommodating Tier Constraints

Some education systems impose "tier constraints" that limit applicants to a certain number of schools in each tier. For example, in Kenya, secondary schools applicants are restricted to rank two national schools, two provincial schools, and two district schools (Lucas and Mbiti, 2012). Similarly, applicants in Ghana can rank four schools including at most one Option 3 school, up to two Option 2 schools, up to four Option 1 schools, and up to four Option 4 and 5 schools (Ajayi, Friedman and Lucas, 2020).

The algorithm of Figure 8 accommodates such constraints. For example, consider a constraint that no more than m colleges from $C' \subset C$ can be ranked. Then, Step k of the algorithm should consider not only each $i \in C^{\dagger}$, but rather each $(i, j) \in C^{\dagger} \times \{0, \dots m\}$ (representing Ann's score being lower than τ_i and that she has previously ranked j schools from C'). Using this approach, the algorithm will terminate within $O(mn^3)$ computation steps.

D Additional Results in the Common Score Framework

D.1 Uncertainty About Thresholds

The baseline analysis presumes that Ann is uncertain about her standing but knows perfectly the thresholds used by colleges. Herein, we show that so long as the relative selectivity of colleges remains unchanged, the optimal portfolio with uncertain thresholds coincides with that of known thresholds where each threshold is the "certainty equivalent," suitably defined.

Let $\tilde{\boldsymbol{\tau}}_{\boldsymbol{i}}$ be the random variable that denotes College i's threshold, with a support that is a subset of $[\underline{\tau}_i, \overline{\tau}_i]$. We denote the joint distribution on $(\tilde{\boldsymbol{\tau}}_{\boldsymbol{1}}, \dots, \tilde{\boldsymbol{\tau}}_{\boldsymbol{n}})$ by Z; thresholds may be drawn with arbitrary correlation. We define the *certainty-equivalent threshold* for College i to be the τ_i^{CE} that solves $F(\tau_i^{CE}) = \mathbb{E}_Z[F(\tilde{\boldsymbol{\tau}}_{\boldsymbol{i}})]$. This is the known threshold under which the probability of acceptance (under score distribution F) coincides with the expected probability of acceptance by College i. A certainty equivalent exists as F is continuous and increasing on its interval support.

We say that **relative selectivity is known** if for every pair of colleges i and $j > i$, $\underline{\tau}_i > \overline{\tau}_j$: in other words, the applicant always anticipates i to be more selective than lower-ranked College j. We view this to be a reasonable assumption as it is consistent with the idea that schools have a known "pecking order." Lucas and Mbiti (2012) and Ajayi (2024) document that the relative selectivity of schools is extremely stable in the context of school admissions in Kenya and Ghana, respectively.

Proposition 7. *Suppose relative selectivity is known. Then the optimal k-portfolio with uncertain thresholds coincides with the optimal k-portfolio in which the threshold of each College i is known to be its certainty-equivalent threshold* τ_i^{CE}.

As the argument is straightforward, we offer it here. Observe that the expected value of a portfolio

P with uncertain thresholds is

$$
\begin{aligned}
&\mathbb{E}_Z\left[\sum_{i=1}^{|P|}\left(F\left(\tilde{\boldsymbol{\tau}}_{P_{(i-1)}}\right)-F\left(\tilde{\boldsymbol{\tau}}_{P_{(i)}}\right)\right)\max\{u_{P_{(i)}},u_o\}\right]\\
=&\sum_{i=1}^{|P|}\left(\mathbb{E}_Z\left[F\left(\tilde{\boldsymbol{\tau}}_{P_{(i-1)}}\right)\right]-\mathbb{E}_Z\left[F\left(\tilde{\boldsymbol{\tau}}_{P_{(i)}}\right)\right]\right)\max\{u_{P_{(i)}},u_o\}]\\
=&\sum_{i=1}^{|P|}\left(F\left(\tau^{CE}_{P_{(i-1)}}\right)-F\left(\tau^{CE}_{P_{(i)}}\right)\right)\max\{u_{P_{(i)}},u_o\}],
\end{aligned}
$$

which coincides with the expected value of a portfolio P under belief F and using the (deterministic) certainty-equivalent threshold for each school. Although we emphasize uncertainty in score thresholds, the result also applies when the applicant is uncertain about the utility of attending each college so long as the relative attractiveness of colleges is known in advance.

D.2 The Marginal Benefit of An Additional Application

We show that the marginal benefit of an additional application falls in the number of applications.[31]

Proposition 8. *For every $k' \geqslant k \geqslant 0$, $V(P^*(k'+1)) - V(P^*(k')) \leqslant V(P^*(k+1)) - V(P^*(k))$.*

Proof. For every $k \geqslant 1$, we establish that

$$V(P^*(k+1)) - V(P^*(k)) \leqslant V(P^*(k)) - V(P^*(k-1)). \tag{16}$$

If $|P^*(k+1)| < k+1$, then $V(P^*(k+1)) - V(P^*(k)) = 0$; as the RHS of (16) is always non-negative, (16) would then hold trivially. So below, we assume that $|P^*(k+1)| = k+1$.

Our approach below identifies two portfolios of k colleges, Q and Q', such that

$$V(P^*(k+1)) + V(P^*(k-1)) \leqslant V(Q) + V(Q'). \tag{17}$$

Let us first argue that once these portfolios are identified, then (16) holds. To see why, observe that as both Q and Q' have k colleges, each has a value no more than $P^*(k)$, the optimal k-college portfolio. Therefore, $V(Q) + V(Q') \leqslant 2V(P^*(k))$. Hence (17) implies that

$$V(P^*(k+1)) + V(P^*(k-1)) \leqslant 2V(P^*(k)),$$

which is equivalent to (16).

We now identify portfolios Q and Q'. Let $\ell > 1$ be the smallest index i such that $P^*_{(i)}(k+1) \leqslant$

[31] As our proof holds the utility profile U and the score distribution F fixed, we simplify our notation by omitting those arguments from functions below. Moreover, if $k = 0$, $P^*(k) = \{\}$ and therefore, $V(P^*(0)) = 0$.

$P^*_{(i-1)}(k-1)$; if this inequality never holds, then we set $\ell = k+1$. We define

$$Q := \lceil P^*(k+1)\rceil^{\ell-1} \bigcup \lfloor P^*(k-1)\rfloor_{k-(\ell-1)},$$
$$Q' := \lceil P^*(k-1)\rceil^{\ell-2} \bigcup \lfloor P^*(k+1)\rfloor_{k+1-(\ell-1)}.$$

Observe that Q is a portfolio of k colleges that includes the $\ell-1$ best colleges in $P^*(k+1)$ and the $(k-\ell+1)$ worst colleges in $P^*(k-1)$. Q' is also a portfolio of k colleges but it includes the $\ell-2$ best colleges in $P^*(k-1)$ and the $(k-\ell+2)$ worst colleges in $P^*(k+1)$.[32]

To compress notation, we write $P := P^*(k+1)$ and $P' := P^*(k-1)$. Observe that

$$\begin{aligned}
&V(Q)+V(Q')-V(P)-V(P') \\
=\;& \left[\sum_{i=1}^{\ell-1}\Big(F(\tau_{P_{(i-1)}})-F(\tau_{P_{(i)}})\Big)u_{P_{(i)}} + \Big(F(\tau_{P_{(\ell-1)}})-F(\tau_{P'_{(\ell-1)}})\Big)u_{P'_{(\ell-1)}} + \sum_{\ell}^{k-1}\Big(F(\tau_{P'_{(i-1)}})-F(\tau_{P'_{(i)}})\Big)u_{P'_{(i)}}\right] \\
&+\left[\sum_{i=1}^{\ell-2}\Big(F(\tau_{P'_{(i-1)}})-F(\tau_{P'_{(i)}})\Big)u_{P'_{(i)}} + \Big(F(\tau_{P'_{(\ell-2)}})-F(\tau_{P_{(\ell)}})\Big)u_{P_{(\ell)}} + \sum_{\ell+1}^{k+1}\Big(F(\tau_{P_{(i-1)}})-F(\tau_{P_{(i)}})\Big)u_{P_{(i)}}\right] \\
&-\sum_{i=1}^{k+1}\Big(F(\tau_{P_{(i-1)}})-F(\tau_{P_{(i)}})\Big)u_{P_{(i)}} - \sum_{i=1}^{k-1}\Big(F(\tau_{P'_{(i-1)}})-F(\tau_{P'_{(i)}})\Big)u_{P'_{(i)}} \\
=\;& \Big(F(\tau_{P_{(\ell-1)}})-F(\tau_{P'_{(\ell-1)}})\Big)u_{P'_{(\ell-1)}} + \Big(F(\tau_{P'_{(\ell-2)}})-F(\tau_{P_{(\ell)}})\Big)u_{P_{(\ell)}} \\
&\quad -\Big(F(\tau_{P_{(\ell-1)}})-F(\tau_{P_{(\ell)}})\Big)u_{P_{(\ell)}} + \Big(F(\tau_{P'_{(\ell-2)}})-F(\tau_{P'_{(\ell-1)}})\Big)u_{P'_{(\ell-1)}} \\
=\;& \Big(F(\tau_{P'_{(\ell-2)}})-F(\tau_{P_{(\ell-1)}})\Big)\Big(u_{P_{(\ell)}}-u_{P'_{(\ell-1)}}\Big),
\end{aligned}$$

where in the first equality, the first line writes out $V(Q)$, the second writes out $V(Q')$, and the third writes out $V(P)$ and $V(P')$; the second equality follows from canceling common terms; and the third equality cancels common terms and re-arranges the remaining terms. We argue that this final expression is positive. Observe that the definition of ℓ implies two facts: (i) College $P_{(\ell)}$ is higher ranked than $P'_{(\ell-1)}$ and hence the second term in the product is non-negative, and that (ii) College $P'_{(\ell-2)}$ is higher ranked than $P_{(\ell-1)}$, which implies that the first term is also non-negative. □

E Additional Results in the Independent-Success Framework

Here, we establish some new results for the independent-success framework of Chade and Smith (2006). These results are either referenced in the main text or used in our subsequent analysis.

So as to be self-contained, the set of college types $C := \{1, \ldots, n\}$ comprises n colleges. Being accepted by a college of type i generates utility u_i, and obtaining her outside option generates utility u_o. If Ann applies to college of type i, then she is admitted by that college with probability α_i independently of her admissions at any other college. As before, we assume that higher indices yield lower utility. However, with independent success, and unlike our framework, "replicas" are valuable for an

[32]The definition of ℓ guarantees that the unions in the definitions of Q and Q' are of disjoint sets.

applicant: if Colleges a and b are replicas, being rejected by College a is no longer informative about the probability with which one is accepted by College b. As in Chade and Smith (2006), we allow colleges to have replicas, and denote the replicas of type i college by $i_1, i_2, \ldots$ (we extend $<$ so that $i_j < i_{j+1}$ and assume that Ann breaks the indifference in favor of lower-indexed copies). Similarly, colleges that are less desirable and more selective than another college are not ruled out.[33] Finally, to accommodate replicas, we require uniqueness only up to replacing replicas.

E.1 Upward Diversity

Section 5.2 of Chade and Smith (2006) alludes to how the optimal portfolio is upwardly diverse when each college has replicas, but they establish it only for the case of two college types. Proposition 9 shows that this conclusion holds generally whenever there are replicas. We prove this result by first obtaining a preliminary result about risk aversion in a setting in which outside options are stochastic. In this setting, we show that having access to a higher number of stochastic outside options makes the applicant more risk-loving.

To be clear about our stochastic outside option setting, let $\{\tilde{u}_j\}_{j=1}^{\overline{r}}$ be independent random variables taking the value L_j with probability β_j and 0 otherwise. Each random variable $\tilde{u}_j$ specifies a stochastic outside option. We do not assume that all the outside options are available to the applicant: instead, we suppose that the set of *available* outside options is $\{\tilde{u}_j\}_{j=1}^{r}$ (where $r \in \{1, \ldots, \overline{r}\}$). We take r as a primitive, and refer to it as the portfolio problem with r stochastic outside options.[34] We contrast this with the baseline framework in which the outside option u_o is deterministic.

Lemma 2 documents two facts. First, for each value of r, there exists a payoff-equivalent problem in which the outside option is deterministic. Second, higher values of r lead to a more risk-loving profile in the sense of Definition 3.

Lemma 2. *For each portfolio problem with r stochastic outside options, there exists a payoff equivalent portfolio problem in which the outside option is deterministic; i.e., there exists a utility profile* $V_r := (v_0^r; v_1^r, \ldots, v_n^r)$ *(with deterministic outside option* v_0^r*) that generates the same expected payoff for each portfolio. Moreover, if* $r' \geqslant r$*, then* $V_{r'} \succcurlyeq_{RL} V_r$*.*

Proof. We prove the first part by construction. We set the outside option v_o^r to be $\mathbb{E}[\max_{j \leqslant r} \tilde{u}_j]$. Denote by G^r the CDF of $\max_{j \leqslant r} \tilde{u}_j$. We also set the utility of attending college i to be

$$v_i^r = \beta^r(u_i) := u_i + \int_0^\infty \max\{z - u_i, 0\} \mathrm{d}G^r(z).$$

The term $\beta^r(u_i)$ embodies the idea that if accepted by College i, the student has the option either to attend that school or choose the best realized outside option (denoted by the variable z). She chooses an outside option only if its realized payoff exceeds u_i, and in that case, she accrues the marginal improvement from the outside option. It follows from integration by parts and some algebra that $\beta^r(u_i) = u_i + \int_{u_i}^\infty (1 - G^r(z)) \mathrm{d}z$. This setup establishes the first part of Lemma 2: a direct calculation

[33] In cases of a tie in utility, we label the dominated college with a higher index.

[34] We emphasize that r does *not* denote the number of outside options that mature.

shows that this utility profile generates the same expected utility for each portfolio as the portfolio problem with r stochastic options.

We prove the second step by induction, relying on the transitivity of $\succcurlyeq_{RL}$. Let $r' = r + 1$. Denote

$$\psi(x) := \begin{cases} \int_0^\infty \left(1 - G^{r'}(z)\right) \mathrm{d}z & \text{if } x \leqslant \int_0^\infty 1 - G^r(z)\mathrm{d}z, \\ x + \int_{\mathrm{inv}\beta^r(x)}^\infty G^r(z) - G^{r'}(z)\mathrm{d}z & \text{otherwise;} \end{cases},$$

where we use the fact that the inverse of $\beta^r(\cdot)$ exists for values greater than $\int_0^\infty 1 - G^r(z)\mathrm{d}z$ since $\beta^r(\cdot)$ is increasing for values greater than $\int_0^\infty 1 - G^r(z)\mathrm{d}z$. We note that $v_i^{r'} = \psi\left(v_i^r\right)$. Leibniz's rule and the Implicit Function Theorem imply that for values of x greater than $\int_0^\infty 1 - G^r(z)dz$, we have

$$\frac{\mathrm{d}\psi}{\mathrm{d}x} = 1 - \frac{G^r(\mathrm{inv}\beta^r(x)) - G^{r'}(\mathrm{inv}\beta^r(x))}{G^r(\mathrm{inv}\beta^r(x))} = \frac{G^{r'}(\mathrm{inv}\beta^r(x))}{G^r(\mathrm{inv}\beta^r(x))}.$$

Observe that $G^{r'}/G^r$ is a non-decreasing step function with range in 0 to 1. Since ψ is constant for values of x lower than $\int_0^\infty 1 - G^r(z)dz$, this implies that ψ is convex. □

Proposition 9. *If each school has m replicas, then for each $k < m$, the optimal $(k+1)$-portfolio is more aggressive than the optimal k-portfolio.*

Proof. Since the parameters of the problem are fixed throughout the proof, for each k, we denote the optimal k-portfolio by $P(k)$. We show that the conclusion obtains so long as $P_{(1)}(k)$ has a replica that is not included in $P(k)$, which is implied by $m > k$.

Chade and Smith (2006) show that there is an optimal portfolio of size $(k+1)$, $P(k+1)$, such that $P(k+1) = P(k) \cup \{x\}$, unless we are in the trivial case that $P(k+1) = P(k)$. Let x denote a college such that $P(k) \cup \{x\}$ is an optimal $(k+1)$-portfolio. Let y denote a replica of $P_{(1)}(k)$ that is not included in $P(k)$. Observe that if Ann must choose a portfolio of k colleges that includes the colleges in $P(k)\backslash\left\{P_{(1)}(k)\right\}$, and can choose an additional college in the set $\left\{x, y, P_{(1)}(k)\right\}$, $P(k)$ remains optimal.[35] This constrained problem is equivalent to the problem of choosing a single-college portfolio from $\left\{x, y, P_{(1)}(k)\right\}$ with a stochastic outside option distributed as the utility from the portfolio $P(k)\backslash\left\{P_{(1)}(k)\right\}$. Both a choice of $P_{(1)}(k)$ and y must be optimal single-college portfolios, because $P_{(1)}(k) \in P(k)$ and y is a replica of $P_{(1)}(k)$. Therefore, y must also be the optimal single-college portfolio when the set of available schools is only $\{x, y\}$ and the stochastic outside option is distributed as the utility from the portfolio $P(k)\backslash\left\{P_{(1)}(k)\right\}$.

Using the same logic, it follows from $\{x\} \cup P(k)$ being an optimal $(k+1)$-portfolio and $y \notin P(k)$ that $\{x\}$ is an optimal single-college portfolio from the menu $\{x, y\}$ with an outside option that is distributed as the utility from the portfolio $P(k)$. By Lemma 2, Ann is more risk loving with the (stochastic) outside option from the portfolio $P(k)$ than with the (stochastic) outside option from the portfolio $P(k)\backslash\left\{P_{(1)}(k)\right\}$. It then follows from Proposition 2 and the definition of x that $\{x\} \succcurlyeq_A \{y\}$, and so $P(k+1) \succcurlyeq_A P(k)$.[36] □

[35] In other words, an unconstrained optimal portfolio of k colleges must also be an optimal k-portfolio when chosen from a smaller menu of portfolios that includes it; this property is the Weak Axiom of Revealed Preference or Sen's α.

[36] We can invoke Proposition 2 because Ann is choosing a single-college portfolio in both cases and hence the correlation

E.2 The Risk-Loving Effect with Independent Success

In Section 5 we claim that a result parallel to Proposition 2 holds in the independent-success framework. Thus, even if admissions probabilities are stochastically independent, unequal outside options lead to more aggressive applications and therefore segregation in the composition of schools.

Proposition 10. *More risk-loving payoffs lead to a more aggressive portfolio:* $U' \succcurlyeq_{RL} U \Rightarrow P^*(k, U', \alpha) \succcurlyeq_A P^*(k, U, \alpha)$.

Lemma 3. *Assume that* $\overline{U} = \overline{U'} = 0$ *and that* $U' \succcurlyeq_{RL} U$ *(i.e., there exists a convex nondecreasing* ψ *such that* $\psi(0) = 0$ *and* $u'_i = \psi(u_i)$ *for each* $i \in C$*). If the agents* U *and* U' *get access to a stochastic outside option that gives them utility* $L > 0$ *(respectively* $\psi(L)$*) with probability* α *and zero otherwise, then their utility from each portfolio can be described by the deterministic profiles* W *and* W' *such that* $\overline{W} = \overline{W'} = 0$ *and* $W' \succcurlyeq_{RL} W$.

Proof. Direct calculation shows that

$$w_i = \begin{cases} u_i - \alpha L & \text{if } u_i > L \\ (1-\alpha)\, u_i(x) & \text{else,} \end{cases}$$

and

$$w'_i = \begin{cases} u'_i - \alpha \psi(L) & \text{if } u'_i > \psi(L) \\ (1-\alpha)\, u'_i(x) & \text{else} \end{cases}$$

are profiles as required by the statement. To see that $W' \succcurlyeq_{RL} W$ observe that the function

$$\beta(x) = \begin{cases} (1-\alpha)\psi(\frac{x}{1-\alpha}) & \text{if } z \leqslant (1-\alpha)L \\ \psi(x + \alpha L) - \alpha\psi(L) & \text{else} \end{cases}$$

maps W to W' (in particular, $\beta(0) = 0$). Since C is finite, we can assume without loss of generality that ψ is smooth. With this assumption, it is straightforward to verify that β is nondecreasing and convex (it is differentialble—including at $(1-\alpha)L$—with a nonegative increasing derivative). □

Proof of Proposition 10. The proof proceeds by induction on k. The case of $k = 1$ follows from Proposition 2 (correlation between admissions decisions does not matter in choosing a single-college portfolio). For $k > 1$, if for some $i, j \leqslant k$ we have $P^*_{(i)}(k, U', \alpha) = P^*_{(j)}(k, U, \alpha)$, we are done by the inductive hypothesis and Lemma 3 (the rest of each portfolio is the optimal size $k-1$ portfolio from $C \setminus \left\{P^*_{(i)}(k, U', \alpha)\right\}$ with the stochastic outside option of $P^*_{(i)}(k, U', \alpha) = P^*_{(j)}(k, U, \alpha)$). Otherwise, $P^*_{(k)}(k, U', \alpha) \neq P^*_{(k)}(k, U, \alpha)$.

Assume that $P^*_{(k)}(k, U', \alpha) < P^*_{(k)}(k, U, \alpha)$ (i.e., the lowest ranked choice of the more risk loving agent is more aggressive). In that case, $P^*_{(k)}(k, U, \alpha)$ is available to U as a last (k-th) choice, which implies that

$$\alpha_{P^*_{(k)}(k,U,\alpha)} u'_{P^*_{(k)}(k,U,\alpha)} \leqslant \alpha_{P^*_{(k)}(k,U',\alpha)} u'_{P^*_{(k)}(k,U',\alpha)}.$$

structure between colleges' admissions decisions is irrelevant.

Imagine constraining U' to include $P^*_{(k)}(k, U, \alpha)$ as the last choice in her portfolio. In that case, by the inductive hypothesis and Lemma 3, she would choose a portfolio of $k-1$ colleges that is more aggressive than $[P^*(k, U, \alpha)]^{k-1}$ (i.e., the first $k-1$ choices on $P^*(k, U, \alpha)$). Next, observe that since U' prefers the "outside option" offered by her last choice $P^*_{(k)}(k, U', \alpha)$ to $P^*_{(k)}(k, U, \alpha)$ she only becomes more aggressive in her choosing the optimal $k-1$ colleges to add to this college.[37]

Finally, assume toward contradiction that $P^*_{(k)}(k, U', \alpha) > P^*_{(k)}(k, U, \alpha)$. Since $P^*(k, U', \alpha)$ and $P^*(k, U, \alpha)$ are disjoint, $P^*_{(k)}(k, U', \alpha)$ is available to U as a last choice and $P^*_{(k)}(k, U, \alpha)$ is available to U' as a last choice. Since rejections convey no information this means that the optimal portoflio of size 1 from the menu $\left\{P^*_{(k)}(k, U', \alpha), P^*_{(k)}(k, U, \alpha)\right\}$ is $P^*_{(k)}(k, U', \alpha)$ for U' and $P^*_{(k)}(k, U, \alpha)$ for U, contradicting the base case (and Proposition 2). □

E.3 Algorithm for Solving for the Optimal Portfolio

Here, we adapt the algorithm of Section 3.4 to find the optimal portfolio for the Chade and Smith setting. The key idea is that we build the optimal portfolio top-down, starting with Ann's first choice.

Let $P^*(k, U, \alpha, c)$ denote the optimal portfolio with utility profile U and admission probabilities α, where Ann is restricted to apply to colleges in $\{x \in C \mid x > c\}$. Since rejections from colleges in $\{x \in C \mid x \leqslant c\}$ convey no information about admissions at colleges $\{x \in C \mid x > c\}$, $P^*(k, U, \alpha, c)$ is the optimal continuation of size k for any "history" where College c is the least aggressive choice that has rejected Ann. We therefore have

$$V\left(P^*(k, U, \alpha, i)\right) := \max_{j \in \mathcal{C} \text{ s.t. } \{i\} >_A \{j\}} \left\{\alpha_j u_j + (1-\alpha_j) V\left(P^*(k-1, U, \alpha, j)\right)\right\}. \tag{18}$$

Using this dynamic program, one can run a routine analogous to Figure 8: we find the optimal continuation where one has to find colleges less aggressive than the least aggressive college that has rejected one thus far. The algorithm continues to be computationally efficient, since only $n+1$ histories must be considered at any step, just as in our baseline framework. The routine from Appendix C also remains valid, and so the algorithm can be sped up to $n^2 \log n$ steps.

If application costs, $\phi(\cdot)$, depend only on the number of colleges, our algorithm requires more computation steps than the Marginal Improvement Algorithm of Chade and Smith (2006). However, as we discuss in Appendix C.2, our approach can expand the scope of their analysis by accommodating tier constraints, unlike that algorithm.

F Results for Section 4.1

F.1 Preliminaries

We denote by Φ (resp., ϕ) the CDF (resp., PDF) of the standard univariate normal distribution and by $\Phi_k(\cdot, \cdot, \rho)$ (resp., $\phi_k(\cdot, \cdot, \rho)$) the CDF (resp., PDF) of the standard multivariate normal distribution with correlation ρ.

[37]This follows since the corresponding profiles V from Lemma 3 are $\succcurlyeq_{RL}$-ranked.

Lemma 4. *For any* $\Delta > 0$, $R(x) := \frac{1-\Phi(x)}{1-\Phi(x+\Delta)}$ *is increasing.*

Proof. Observe that

$$\frac{\mathrm{d}R}{\mathrm{d}x} = \frac{-\phi(x)\,(1-\Phi(x+\Delta)) + \phi(x+\Delta)\,(1-\Phi(x))}{(1-\Phi(x+\Delta))^2} = \frac{1-\Phi(x)}{1-\Phi(x+\Delta)}\left(\frac{\phi(x+\Delta)}{1-\Phi(x+\Delta)} - \frac{\phi(x)}{1-\Phi(x)}\right)$$

The term of the product outside parentheses is the ratio of nonzero probabilities and hence positive. The term in parentheses is also non-negative because it is the difference between two inverse Mills ratios (which is an increasing function).[38] □

Lemma 5. *For any* $i < j < k$, *the ratio* $\frac{\Pr\{\textit{accepted at } k, \textit{ rejected at } i\}}{\Pr\{\textit{accepted at } j, \textit{ rejected at } i\}}$ *increases with* ρ.

Proof. Let $B := \Phi_2(\tau_i, \infty, \rho) = \Phi(\tau_i)$ be the probability of rejection from school i; observe that B is independent of ρ. Let $g(\rho) := \Phi_2(\tau_i, \tau_j, \rho)$, and $f(\rho) := \Phi_2(\tau_i, \tau_k, \rho)$, the probabilities of being rejected from both i and j (resp. i and k). With this notation, our goal is to show that $H(\rho) := (B - f(\rho))/(B - g(\rho))$ is increasing. We will show that $\dot{H} > 0$.[39]

To begin with, note that $\dot{H}$ has the same sign as $\dot{g}(B-f) - \dot{f}(B-g)$, and so it suffices to show that $\frac{(B-f)}{(B-g)} > \frac{\dot{f}}{\dot{g}}$. Conditional on $S_i = x$, the marginal distributions of S_j and of S_k are governed by the CDF $\Phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})$. Hence, by Fubini's theorem,

$$\frac{(B-f)}{(B-g)} = \frac{\int_{-\infty}^{\tau_i}\left(1-\Phi\left(\frac{\tau_k-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)\mathrm{d}x}{\int_{-\infty}^{\tau_i}\left(1-\Phi\left(\frac{\tau_j-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)\mathrm{d}x}$$

Lemma 4 implies that, on the domain$(-\infty, \tau_i]$, the ratio $1-\Phi\left(\frac{\tau_k-\rho x}{\sqrt{1-\rho^2}}\right)/1-\Phi\left(\frac{\tau_j-\rho x}{\sqrt{1-\rho^2}}\right)$ is minimized at $x = \tau_i$ (since $\tau_j > \tau_k$ and $\rho \geqslant 0$). Denote the minimal value by λ. We have

$$\frac{\int_{-\infty}^{\tau_i}\left(1-\Phi\left(\frac{\tau_k-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)\mathrm{d}x}{\int_{-\infty}^{\tau_i}\left(1-\Phi\left(\frac{\tau_j-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)\mathrm{d}x} > \frac{\int_{-\infty}^{\tau_i}\lambda\left(1-\Phi\left(\frac{\tau_j-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)\mathrm{d}x}{\int_{-\infty}^{\tau_i}\left(1-\Phi\left(\frac{\tau_j-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)\mathrm{d}x} = \lambda. \tag{19}$$

Next, note that

$$\frac{1-\Phi\left(\frac{\tau_k-\rho\tau_i}{\sqrt{1-\rho^2}}\right)}{1-\Phi\left(\frac{\tau_j-\rho\tau_i}{\sqrt{1-\rho^2}}\right)} \geqslant \frac{\phi\left(\frac{\tau_k-\rho\tau_i}{\sqrt{1-\rho^2}}\right)}{\phi\left(\frac{\tau_j-\rho\tau_i}{\sqrt{1-\rho^2}}\right)} = \frac{\phi_2(\tau_i,\tau_k,\rho)}{\phi_2(\tau_i,\tau_j,\rho)}$$

where the inequality uses again the monotonicity of the inverse Mill's ratio.

[38]A straightforward way to see this is that $\frac{\phi(x+\Delta)}{1-\Phi(x+\Delta)} - \frac{\phi(x)}{1-\Phi(x)} = E[X|X > x+\Delta] - E[X|X > x]$.

[39]In line with the literature deriving these heat equations, we represent the derivative with respect to ρ as a dot.

Finally, Plackett (1954) shows that

$$\dot{\phi}_2(x,y,\rho) = \frac{\partial^2 \phi_2(x,y,\rho)}{\partial x \partial y}$$

which implies that $\dot{g} = \phi_2(\tau_i, \tau_j, \rho)$ and $\dot{f} = \phi_2(\tau_i, \tau_k, \rho)$. Altogether, we obtain:

$$\frac{B-f}{B-g} = \frac{\int_{-\infty}^{\tau_i} \left(1 - \Phi\left(\frac{\tau_k - \rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)\mathrm{d}x}{\int_{-\infty}^{\tau_i} \left(1 - \Phi\left(\frac{\tau_j - \rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)\mathrm{d}x} > \lambda = \frac{1 - \Phi\left(\frac{\tau_k - \rho\tau_i}{\sqrt{1-\rho^2}}\right)}{1 - \Phi\left(\frac{\tau_j - \rho\tau_i}{\sqrt{1-\rho^2}}\right)} \geqslant \frac{\phi_2(\tau_i, \tau_k, \rho)}{\phi_2(\tau_i, \tau_j, \rho)} = \frac{\dot{f}}{\dot{g}}$$

. □

F.2 Proof of Proposition 3

We prove Proposition 3 using several intermediate results, Lemmas 6 to 8, established below.

First, we fix an ambient problem with a continuum of schools, and denote the optimal k-college portfolio in this problem by $P(k)$. Throughout our analysis, we assume that all schools place some weight on the common component, $\rho > 0$.

Lemma 6. *If $k > 1$, the lowest-ranked college in the optimal k-portfolio is a safety: $P_{(k)}(k) < m$.*

Proof. We start with the case of $k = 2$ and then consider $k > 2$.

Step 1 ($k = 2$). We establish a stronger claim than that asserted: *following rejection from a college whose threshold is $w \in \mathbb{R}$, the optimal backup college is always a safety.*

Denote by $U(z, x)$ the expected utility from applying only to a college of quality z conditional on receiving a score of x in another college. Since the correlation between scores is ρ we have that

$$U(z,x) = \left(1 - \Phi\left(\frac{\tau(z) - \rho x}{\sqrt{1-\rho^2}}\right)\right) u(z).$$

Define $\mathcal{V}(z,w) := \int_{-\infty}^{w} U(z,x)\phi(x)\mathrm{d}x$ to be the (ex ante) expected marginal value from applying to college z as a backup for a college whose threshold is w. Denote $z^*(w) := \operatorname{argmax}_z \mathcal{V}(z,w)$ to be the best backup when rejected by a college with threshold w;[40] by definition, $z^*(w)$ then solves the first order condition $\frac{\partial \mathcal{V}}{\partial z} = 0$.

Observe that $m = z^*(\infty)$: the match is the best backup for a school that rejects *all* scores (as that rejection conveys no information). Furthermore, since u, τ, and Φ are all smooth,

$$\frac{\partial \mathcal{V}(z,w)}{\partial z} = \int_{-\infty}^{w} \frac{\partial U}{\partial z}\phi(x)\mathrm{d}x,$$

[40] We assume for expositional simplicity that the solution is unique; if otherwise, the arguments apply with a tie-breaking rule that selects the highest solution.

and $z^*(\infty)$ solves

$$\frac{\partial \mathcal{V}(z,\infty)}{\partial z} = \lim_{w\to\infty} \frac{\partial \mathcal{V}(z,w)}{\partial z} = 0.$$

Below, we argue that for any finite w, $z^*(w) < z^*(\infty) = m$, which implies that the optimal backup following any first choice is a safety school.

We first note that if $\rho = 0$, then success is independent and then $z^*(w) = m$: the optimal backup for any w is the match m. By Lemma 5, increasing the correlation to $\rho > 0$ implies that $z^*(w) \leqslant m$.

First, we consider high values of w. Observe that the derivative of $U(z,x)$ with respect to z is

$$\frac{\partial U}{\partial z} = \left(1 - \Phi\left(\frac{\tau(z) - \rho x}{\sqrt{1-\rho^2}}\right)\right)\frac{\mathrm{d}u}{\mathrm{d}z} - \left(\phi\left(\frac{\tau(z) - \rho x}{\sqrt{1-\rho^2}}\right)\right)u(z)\frac{\mathrm{d}\tau}{\mathrm{d}z}.$$

Fixing z, this expression is strictly positive for sufficiently large x given that $\rho > 0$: as $x \to \infty$, the coefficient on the positive term, namely $\frac{\mathrm{d}u}{\mathrm{d}z}$, increases to 1 and that on the negative term, namely $-u(z)\frac{\mathrm{d}\tau}{\mathrm{d}z}$ decreases to 0. Therefore, for sufficiently large $\hat{w}$, for any $\tilde{w} \geqslant \hat{w}$, $\frac{\partial U}{\partial z}|_{z=m,x=\tilde{w}} > 0$.

It then follows that

$$\begin{aligned} 0 = \frac{\partial \mathcal{V}}{\partial z}\Big|_{z=m,w=\infty} = \int_{-\infty}^{\infty} \frac{\partial U(z,x)}{\partial z}\Big|_{z=m}\phi(x)\mathrm{d}x &= \int_{-\infty}^{\tilde{w}} \frac{\partial U(z,x)}{\partial z}\Big|_{z=m}\phi(x)\mathrm{d}x + \int_{\tilde{w}}^{\infty} \frac{\partial U(z,x)}{\partial z}\Big|_{z=m}\phi(x)\mathrm{d}x \\ &> \int_{-\infty}^{\tilde{w}} \frac{\partial U(z,x)}{\partial z}\Big|_{z=m}\phi(x)\mathrm{d}x = \frac{\partial \mathcal{V}}{\partial z}\Big|_{z=m,w=\tilde{w}}, \end{aligned}$$

where the equalities follow from $z^*(\infty) = m$ and the strict inequality follows from $\tilde{w} \geqslant \hat{w}$. The expression above shows that for any $\tilde{w} \geqslant \hat{w}$, $z^*(\tilde{w}) \neq m$. As we have from before that $z^*(\tilde{w}) \leqslant m$, it follows that $z^*(\tilde{w}) < m$.

To complete the proof, consider a two-college portfolio in which the top choice has a threshold w lower than $\hat{w}$. We compare $z^*(w)$ with $z^*(\hat{w})$. Observe that the former corresponds to choosing a single-college portfolio after being rejected by school with threshold w and the latter to a single-college portfolio after being rejected by a school with higher threshold $\hat{w}$. Since the former has lower beliefs in an LR-sense (Karlin and Rinott, 1980), Proposition 1 implies that for any $w < \hat{w}$, $z^*(w) \leqslant z^*(\hat{w})$. This completes the proof as we have shown above that $z^*(\hat{w})$ is strictly less than m.

Step 2 ($k > 2$). The proof for the case $k > 2$, is nearly identical. We define

$$\mathcal{V}(z; w_1, \ldots, w_{k-1}) = \int_{-\infty}^{w_1} \cdots \int_{-\infty}^{w_{k-1}} U(z; x_1, \ldots, x_{k-1})\phi_{(k-1)}(x)\mathrm{d}x_1 \ldots, \mathrm{d}x_{k-1}$$

where

$$U(z; x_1, \ldots, x_{k-1}) = \left(1 - \Phi\left(\frac{\tau(z) - \alpha \sum x_i}{\beta}\right)\right)u(z),$$

for the appropriate positive α, β.[41] $\mathcal{V}(z; w_1, \ldots, w_{k-1})$ is the expected added value from applying to college z as a backup after being rejected by colleges whose thresholds are $w_1, \ldots, w_{k-1}$. We denote

[41] Positivity follows from the covariance matrix being positive semi-definite.

$z^*(w_1, \ldots, w_{k-1}) := \operatorname{argmax} \mathcal{V}(z; w_1, \ldots, w_{k-1})$.

We assume, by induction, that for any $\overline{w} := (w_1, \ldots, w_{k-2}) \in \mathbb{R}^{k-2}$ we have $z^*(\overline{w}) < m$. The base case of this inductive argument was established in Step 1. Observe that

$$z^*(\infty, w_2, \ldots, w_{k-1}) = z^*(w_2, \ldots, w_{k-1}) < m$$

in which the equality follows from the property that being rejected by a school whose threshold is ∞ is completely uninformative and does not change the optimal backup; the inequality follows from the inductive hypothesis. Finally, observe that for $w_1 < \infty$, $z^*(w_1, w_2, \ldots, w_{k-1}) \leqslant z^*(\infty, w_2, \ldots, w_{k-1})$: being rejected by a school with threshold $w_1 < \infty$ leads to LR-lower beliefs than a school with threshold ∞ (Karlin and Rinott, 1980), and hence, Proposition 1 implies that the optimal single-college with the former beliefs is a lower-ranked school than those with the latter. □

Lemma 7. *If $k > 1$, the highest-ranked college in the optimal k-portfolio is a reach: $P_{(1)}(k) > m$.*

Proof. Using $P(k)$ to denote the optimal portfolio, we write $\tilde{P} := (P_{(2)}(k), \ldots, P_{(k)}(k))$ to denote the backups if Ann's top choice $P_{(1)}(k)$ rejects her. For a school i, we denote q_i to be its quality.

Define $W(q_i) := \mathbb{E}\left[\max_{\{j \in \tilde{P}: S_j \geqslant \tau(q_j)\}} u(q_j) \mid S_i < \tau(q_i)\right]$ to be the expected payoff from the portfolio $\tilde{P}$ conditional on being rejected by College i. The interpretation of $W(q_i)$ is that if Ann selects College i to be her top choice and the portfolio $\tilde{P}$ as her $(k-1)$ backup options, her conditional expected payoff following rejection from College i is $W(q_i)$. Observe that W is differentiable and strictly increasing in its argument, and hence $W'(q) > 0$ for every quality q.[42]

The optimal top choice $P_{(1)}(k)$ chooses a quality q to maximize

$$\underbrace{(1 - \Phi(\tau(q)))\, u(q)}_{\text{Accepted by top choice}} + \underbrace{\Phi(\tau(q))\, W(q)}_{\text{Continuation } \tilde{P} \text{ after rejection}}. \tag{20}$$

Taking the first-order condition yields that for a solution q^*

$$(1 - \Phi(\tau(q^*)))\, u'(q^*) - \phi(\tau(q^*))\, \tau'(q^*) u(q^*) + \phi(\tau(q^*))\, \tau'(q^*) W(q^*) + \Phi(\tau(z))\, W'(q^*) = 0 \tag{21}$$

Observe that the third and fourth terms of the LHS are strictly positive for every value of q.

We argue that q^* must be strictly higher than the quality of the match, m. Suppose towards a contradiction that $q^* \leqslant m$. By definition, College m maximizes $(1 - \Phi(\tau(q)))\, u(q)$. Thus, setting $q^* = m$ in the LHS of (21) would result in the first two terms equaling zero but the third and fourth terms remaining strictly positive, invalidating m as an optimal choice. Now suppose that $q^* < m$. Relative to q^*, College m achieves a strictly higher payoff—by definition, the first term of (20) is strictly higher, and the second term is as well since Φ and $W(\cdot)$ are strictly increasing—contradicting that any $q^* < m$ is optimal. □

[42] To see why it is strictly increasing, suppose that College i is of strictly lower quality than College ℓ ($q_i < q_\ell$). Because College i is less selective than College ℓ, being rejected by it is worse news for Ann's scores S_j for $j \in \tilde{P}$ in a likelihood-ratio and thus in the first-order stochastically dominance sense (Karlin and Rinott, 1980). Therefore, Ann's expected payoff from $\tilde{P}$ is strictly lower if she is rejected by College i than if she is rejected by College ℓ.

Our results establish that given a continuum of qualities, every optimal portfolio of two or more colleges includes a safety and reach. We now show that optimal portfolios chosen from any sufficiently fine grid share these properties.

Lemma 8. *For any* k *and* $\delta > 0$, *there exists* $\bar{\varepsilon} > 0$ *such that* $\max_{1 \leqslant i \leqslant k} \left\{ \left| P^{\varepsilon}_{(i)}(k) - P_{(i)}(k) \right| \right\} < \delta$ *whenever* $\varepsilon < \bar{\varepsilon}$.

Proof. Assume towards a contradiction that this is not the case. Then, since $[0,1]^k$ is compact, there is a sequence of positive values $\varepsilon_n \to 0$ such that $P^{\varepsilon_n}(k) \subset [0,1]^k$ that converges to $Q \in [0,1]^k$, with $Q \neq P(k)$. Since $V(P)$ is continuous in P, the expected utility from portfolios in this subsequence converges to $V(Q)$. Since we assumed that Q is not optimal from the continuum menu, $V(Q) < V(P(k))$. Since C_{ε_n} is increasingly fine, one can construct a sequence of k-college portfolios $T_n \subset C_{\varepsilon_n}$, such that $T_n \to P(k)$. But, by continuity of V, this implies that for sufficiently large n, T_n achieves strictly higher utility than $P^{\varepsilon_n}(k)$, contradicting the optimality of $P^{\varepsilon_n}(k)$. □

F.3 Proof of Proposition 4

The proof comprises two parts. Lemma 9 shows that the number of College 1 replicas in the optimal k-college portfolio approaches infinity as $k \to \infty$; the main idea is that if this were not to happen, then an additional application to College 1 would accrue a higher marginal benefit than an additional application to colleges that feature often in the k-college portfolio, suggesting a profitable deviation. Lemma 10 establishes that the number of College $\underline{m}$ replicas in the optimal k-college portfolio also approaches infinity as $k \to \infty$; this argument is considerably more involved, where we compare the rates of bad-news generated by rejections to show that if it were otherwise, there is a greater marginal benefit from applying to more replicas of College $\underline{m}$ then the least selective school that features often in the optimal k-portfolio. Throughout our analysis, we denote the optimal k-portfolio by $P(k)$. We also assume that there are multiple rationalizable college types ($\underline{m} > 1$) because otherwise, the result holds trivially.

Lemma 9. *The number of College* 1 *replicas in* $P(k)$ *approaches infinity as* k *approaches infinity.*[43]

Proof. Towards a contradiction, suppose that $\liminf |\{x \in P(k) \mid x \text{ is a replica of College } 1\}| = \tilde{k} < \infty$. Then, there is an increasing sequence of portfolio sizes $\{k_i\}_{i=1}^{\infty}$ such that in each portfolio $P(k_i)$, the number of replicas of College 1 is at most $\tilde{k}$. By the Pigeonhole Principle, for this sequence, there exists $i > 1$ such that $\limsup |\{x \in P(k) \mid x \text{ is a replica of College } i\}| = \infty$. Let $\ell > 1$ be the lowest such index.

Take a subsequence of portfolio sizes such that the number of replicas of College ℓ increases to infinity and the number of replicas of each higher-ranked college remains constant along the subsequence.[44] For a sufficiently large portfolio in this sequence, admission to a replica of College ℓ is

[43] Our argument allows for the case of independent success ($\rho = 0$); for that model, Lemma 9 combined with Proposition 9 implies that there exists finite integers M and k such that for all $k' > k$, the optimal k'-portfolio has at least $(k' - M)$ replicas of College 1.

[44] This is possible as, by construction, the number of replicas of such colleges is uniformly bounded across all portfolio sizes in the sequence.

nearly guaranteed, even when conditioning on all better schools rejecting one's application.[45] Therefore, the marginal benefit of an additional application to a replica of College ℓ is vanishing.

Next, observe that since the number of applications to replicas of Colleges $1, \ldots, (\ell-1)$ is constant on the subsequence, there exists $K \in (0,1)$ such that, for any portfolio in the subsequence, Ann is rejected from all replicas of Colleges $1, \ldots, (\ell-1)$ with probability K. Furthermore, there exists $\Delta > 0$ such that, conditional on being rejected from all replicas of Colleges $1, \ldots, (\ell-1)$, Ann believes that she will be admitted to College 1 with probability Δ.[46] Hence, for sufficiently large portfolios in the subssequence, Ann strictly prefers to replace a replica of College ℓ with a replica of College 1, contradicting the optimality of $P(k)$. $\square$

Lemma 10. *The number of College $\underline{m}$ replicas in $P(k)$ approaches infinity as k approaches infinity.*

As the proof for this result is intricate, we first sketch the intuition. The key idea is that if Lemma 10 were false, then most of the applicant's backup options are schools more selective than $\underline{m}$. These backup options are useful only if an applicant is rejected by all other schools, which conveys bad news about her common score. We show under this bad news, the applicant obtains a much higher marginal benefit from applying to a less selective rationalizable school than the least selective one that features often in her portfolio, resulting in a contradiction. Our argument is notationally intensive because making these comparisons requires comparing and bounding the relative likelihood of various tail events.

Proof. Suppose towards a contradiction that there exists an increasing sequence of portfolio sizes, $\{k^j\}_{j=1}^{\infty} \to \infty$ and finite integer $\tilde{k} < \infty$ such that the number of College $\underline{m}$ replicas in $P(k^j)$ is lower than $\tilde{k}$ for each k^j. Let $\ell < \underline{m}$ denote the largest index such that

$$\limsup \left| \left\{ x \in P(k^j) \mid x \text{ is a replica of College } \ell \right\} \right| = \infty,$$

where the existence of such an ℓ is guaranteed by the Pigeonhole Principle. The sequence $\{k^j\}_{j=1}^{\infty}$ has a subsequence such that i) the number of College ℓ replicas is positive and increasing along the sequence, and ii) the number of applications to replicas of each college in $\{\ell+1, \ldots, \underline{m}\}$ is constant along the sequence. Without loss of generality, we assume that $\{k^j\}_{j=1}^{\infty}$ has these properties.

For a portfolio of k schools, let $k_1, k_2, \ldots k_\ell$ denote the number of applications to replicas of Colleges $1, 2, \ldots, \ell$, respectively. In what follows, we will show that, for sufficiently large portfolios,

[45] Recall that, on the subsequence, the number of copies of each of Colleges $\{1, \ldots, \ell-1\}$ is fixed. Conditional on being rejected from all these colleges, Ann's beliefs about her score are distributed on $\mathbb{R}$, with full support. For every $\varepsilon > 0$ there exists a sufficiently low s^* such that she believes that her score is above s^* with probability higher than $\sqrt{1-\varepsilon}$. Observe that there exists $\tilde{k}$ such that if Ann conditioned on the common component of her score being some $s \geqslant s^*$, she would be accepted by a replica of College ℓ with probability higher than $\sqrt{1-\varepsilon}$ if she were to submit $\tilde{k}$ applications to schools of that type. Hence, conditional on $S > s^*$ admission occurs at least with probability $\sqrt{1-\varepsilon}$, and since the conditional probability that $S > s^*$ is greater than $\sqrt{1-\varepsilon}$, Ann's believes, conditional on rejections from all higher ranked colleges, that she will be admitted to College ℓ with probability greater than $1-\varepsilon$.

[46] To see this, note that for sufficiently low $\varepsilon > 0$, there is a probability higher than ε that the common component of Ann's score $\rho S \in [\tau_\ell, \tau_\ell + 1]$ but, for all of the independent draws for replicas of Colleges $1, \ldots, (\ell-1)$ are such that $\left(\sqrt{1-\rho^2}\right) \varepsilon_c < -1$. Conditional on this positive probability event, Ann's probability of admission to College ℓ is greater than $1/2$. This allows us to set $\Delta = \varepsilon/2$.

replacing an application to a replica of College ℓ in $P(k)$ with an application to a replica of College $(\ell+1)$ will be strictly beneficial to Ann (contradicting the optimality of this portfolio).

To simplify notation, we denote the CDF of the common component, ρS, by G, and its PDF by g, and similarly, the CDF of schools specific component, $\sqrt{1-\rho^2}\varepsilon_c$ by F and the PDF by f. Without loss of generality, we also normalize the outside option to zero. Using this notation, for any $k \in \{k^j\}_{j=1}^{\infty}$, the benefit from the k_ℓ-th application to a replica of College ℓ is bounded above by

$$\int_{-\infty}^{\infty} g(z) \left(\prod_{i=1}^{\ell-1} (F(\tau_i - z))^{k_i} \right) (F(\tau_\ell - z))^{k_\ell - 1} \left(1 - F(\tau_\ell - z)\right) u_\ell dz. \tag{22}$$

The expression (22) is an upper bound on the value of the marginal application to a replica of College ℓ: this application is beneficial if Ann is rejected by all (weakly) preferred colleges, and it is an upper bound as it ignores the fact that if Ann is rejected by all these colleges, she may still be accepted by a college in her portfolio that is lower ranked than ℓ.

Let $\bar{k}$ denote the total number of applications to colleges of types in $\{l+1, \ldots, \underline{m}\}$ in the sequence of optimal portfolios. The marginal benefit from an application to College $(\ell+1)$ is no less than

$$\int_{-\infty}^{\infty} g(z) \left(\prod_{i=1}^{\ell-1} (F(\tau_i - z))^{k_i} \right) (F(\tau_\ell - z))^{k_\ell - 1} \left(1 - F(\tau_{\ell+1} - z)\right) u_{\ell+1} (F(\tau_{\underline{m}} - z))^{\bar{k}} dz. \tag{23}$$

The expression (23) offers a lower bound because it stipulates that Ann benefits from the additional application to a college of type $(\ell+1)$ only if she is rejected by all other colleges, including colleges that are less selective, and assuming that all colleges of type $\{\ell+1, \ldots, \underline{m}\}$ use the threshold of college $\underline{m}$. Our argument below shows that (23) exceeds (22) for sufficiently large portfolios.

Let $\kappa := \sum_{i=1}^{\ell} k_i - 1$ be the number of applications submitted to schools of type $1, \ldots, \ell$, excluding the marginal application to a school of type ℓ. Define $\hat{F}_\kappa(z) := \left[\left(\prod_{i=1}^{\ell-1} (F(\tau_i - z))^{k_i} \right) (F(\tau_\ell - z))^{k_\ell - 1} \right]^{1/\kappa}$ as the geometric mean. Observe that for each $z \in \mathbb{R}$, $F(\tau_\ell - z) \leqslant \hat{F}_\kappa(z) \leqslant F(\tau_1 - z)$. Furthermore, $\hat{F}_\kappa$ is strictly decreasing and continuous, and therefore invertible. Hence, for each $x \in [0,1]$,

$$\tau_\ell - F^{-1}(x) \leqslant \hat{F}_k^{-1}(x) \leqslant \tau_1 - F^{-1}(x). \tag{24}$$

Using a change-of-variables ($x = \hat{F}_\kappa(z)$), we can rewrite (22) as

$$-\int_0^1 g(\hat{F}_\kappa^{-1}(x)) x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{1}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx \tag{25}$$

and the lower bound (23)

$$-\int_0^1 g(\hat{F}_\kappa^{-1}(x)) x^\kappa \left(1 - F(\tau_{\ell+1} - \hat{F}_\kappa^{-1}(x))\right) u_{\ell+1} F^{\bar{k}}(\tau_{\underline{m}} - \hat{F}_\kappa^{-1}(x)) \frac{-1}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx. \tag{26}$$

Our argument proceeds by showing that the ratio of (26) to (25) approaches infinity, which implies that the marginal benefit of applying to a replica of College $(\ell+1)$ is significantly higher than applying to a replica of College ℓ.

To this end, we consider a lower bound for (26). By (24), as x approaches 1 we have that $\hat{F}_\kappa^{-1}(x)$ approaches $-\infty$. Hence, the positive expressions $1 - F(\tau_{\ell+1} - \hat{F}_\kappa^{-1}(x))$ and $1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))$ both approach 0 as x approaches 1. Therefore, for any $M > 0$, there exists $0 < \varepsilon_M < 1/4$ such that in the interval $(1-\varepsilon_M, 1)$ we have

$$1 - F(\tau_{\ell+1} - \hat{F}_\kappa^{-1}(x)) = 1 - F\left((\tau_{\ell+1} - \tau_\ell) + \tau_\ell - \hat{F}_\kappa^{-1}(x)\right) > M\left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right).$$ [47]

Furthermore, using (24) and the monotonicity of F and $\hat{F}_\kappa$, we get that for any $x \geqslant 1 - \varepsilon_M$,

$$F^{\bar{k}}(\tau_{\underline{m}} - \tau_1 + F^{-1}(1-\varepsilon_M)) \leqslant F^{\bar{k}}(\tau_{\underline{m}} - \hat{F}_\kappa^{-1}(1-\varepsilon_M)) \leqslant F^{\bar{k}}(\tau_{\underline{m}} - \hat{F}_\kappa^{-1}(x)).$$

Restricting the domain of integration in (26) to $[1-\varepsilon_M, 1]$ yields the following lower bound on (26)

$$\int_{1-\varepsilon_M}^{1} g(\hat{F}_\kappa^{-1}(x))x^\kappa \left(1 - F(\tau_{\ell+1} - \hat{F}_\kappa^{-1}(x))\right) u_{\ell+1} F^{\bar{k}}(\tau_{\underline{m}} - \hat{F}_\kappa^{-1}(x)) \frac{-1}{\hat{F}_\kappa'\left(\hat{F}_\kappa^{-1}(x)\right)} dx$$

which, using the inequalities we developed for the domain $[1-\varepsilon_M, 1]$, provides us with the following relaxed lower bound on (26):

$$-\int_{1-\varepsilon_M}^{1} g(\hat{F}_\kappa^{-1}(x))x^\kappa M\left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_{\ell+1} F^{\bar{k}}(\tau_{\underline{m}} - \tau_1 + F^{-1}(1-\varepsilon_M)) \frac{1}{\hat{F}_\kappa'\left(\hat{F}_\kappa^{-1}(x)\right)} dx. \tag{27}$$

We now consider an upper bound for (25). Our approach uses a concentration argument in which we show that in large portfolios, when conditions on her marginal application being relevant—i.e., many weakly more desirable schools have rejected her—then her beliefs are concentrated on very low values of the common score.

First, we consider the integral from (25) restricted to low beliefs ($x \geqslant 1/2$). Observe that

$$\int_{\frac{1}{2}}^{1} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}_\kappa'\left(\hat{F}_\kappa^{-1}(x)\right)} dx > \left(\frac{3}{4}\right)^\kappa \int_{\frac{3}{4}}^{1} \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}_\kappa'\left(\hat{F}_\kappa^{-1}(x)\right)} dx.$$

Additionally, by (24) we have

$$\left(\frac{3}{4}\right)^\kappa \int_{\frac{3}{4}}^{1} \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}_\kappa'\left(\hat{F}_\kappa^{-1}(x)\right)} dx > \left(\frac{3}{4}\right)^\kappa \int_{\frac{3}{4}}^{1} \left(1 - F(F^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}_\kappa'\left(\hat{F}_\kappa^{-1}(x)\right)} dx.$$

[47] To see this, write $\Delta = \tau_\ell - \tau_{\ell+1} > 0$ and $z = \tau_\ell - \hat{F}_\kappa^{-1}(x)$, and observe that $(1 - F(z-\Delta)) / (1 - F(z))$ approaches infinity as z approaches infinity (by Lemma 4).

Similarly, we obtain

$$\int_0^{\frac{1}{2}} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx < \left(\frac{1}{2}\right)^\kappa \int_0^{\frac{1}{2}} \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx$$

and by (24)

$$\left(\frac{1}{2}\right)^\kappa \int_0^{\frac{1}{2}} \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx < \left(\frac{1}{2}\right)^\kappa \int_0^{\frac{1}{2}} \left(1 - F(\tau_\ell - \tau_1 + F^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx.$$

Combining these inequalities, we obtain

$$\frac{\int_{\frac{1}{2}}^1 x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx}{\int_0^{\frac{1}{2}} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx} > \frac{(3/4)^\kappa}{(1/2)^\kappa} \frac{\int_{\frac{3}{4}}^1 \left(1 - F(F^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx}{\int_0^{\frac{1}{2}} \left(1 - F(\tau_\ell - \tau_1 + F^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx}.$$

As the RHS above approaches infinity as $m \to \infty$,[48] we obtain that for sufficiently large m,

$$2\int_{\frac{1}{2}}^1 x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx > \int_0^1 x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx.$$

Observe that we can write the LHS above as

$$2\int_{\frac{1}{2}}^{1-\varepsilon_M} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx + 2\int_{1-\varepsilon_M}^1 x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx.$$

As the expression $\frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)}$ is uniformly bounded above by some $L > 0$ in the domain $[\frac{1}{2}, 1-\varepsilon_M]$, it follows that

$$\begin{aligned}\int_{\frac{1}{2}}^{1-\varepsilon_M} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx &\leqslant L u_\ell \int_{\frac{1}{2}}^{1-\varepsilon_M} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) dx \\ &\leqslant \frac{(1-\varepsilon_M)^{\kappa+1}}{\kappa+1} L u_\ell.\end{aligned}$$

Similarly, the expression $\left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)}$ is uniformly bounded below by some $\Delta > 0$

48 The right hand side grows to infinity since $(\frac{3}{4}/\frac{1}{2})^\kappa$ approaches infinity, while the ratio of integrals is bounded below by a positive number. To see this, note that the integrals depend on m only through the vector $(k_1/\kappa, k_2/\kappa, \ldots, (k_l-1)/\kappa)$, and the values of these integrals are continuous in this vector. Thus, they attain a minimum and a maximum in the ℓ-simplex.

in the domain $[1-\varepsilon_M, 1-\varepsilon_M/2]$. Thus,

$$\int_{1-\varepsilon_M}^{1} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx > \int_{1-\varepsilon_M}^{1-\frac{1}{2}\varepsilon_M} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx$$
$$\geqslant \frac{\varepsilon_M \Delta u_\ell}{2}(1-\varepsilon_M)^\kappa.$$

Since for sufficiently large values of κ we have

$$\frac{\varepsilon_M \Delta u_\ell}{2}(1-\varepsilon_M)^\kappa \geqslant \frac{(1-\varepsilon_M)^{\kappa+1}}{\kappa+1} L u_\ell,$$

we obtain that for sufficiently large values of k

$$\int_{\frac{1}{2}}^{1} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx < 2\int_{1-\varepsilon_M}^{1} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx.$$

The above offers the following upper bound on the marginal benefit of the k_ℓ-th application to a school of type ℓ:

$$4\int_{1-\varepsilon_M}^{1} x^\kappa \left(1 - F(\tau_\ell - \hat{F}_\kappa^{-1}(x))\right) u_\ell \frac{-g(\hat{F}_\kappa^{-1}(x))}{\hat{F}'_\kappa\left(\hat{F}_\kappa^{-1}(x)\right)} dx. \tag{28}$$

To complete the proof, observe that taking the ratio of (27) to the (28), we obtain

$$\frac{M u_{\ell+1} F^B(\tau_{\underline{m}} - \tau_1 + F^{-1}(1-\varepsilon_M))}{4u_\ell}. \tag{29}$$

The numerator of (29) increases to infinity as M increases to infinity, and the denominator is independent of M. Thus, for sufficiently large values of M, (29) is greater than 1. As the numerator is a lower bound for the marginal benefit of sending the k_ℓ-th application to a college of type $(\ell+1)$ and the denominator is an upper bound of the marginal benefit of sending it to a college of type ℓ, we obtain a strict improvement, contradicting that portfolio $P(k)$ is optimal. □

F.4 Proof of Proposition 5

We write the payoff from a two-college portfolio in which College i is the top choice and College j is the backup:

$$\mathcal{V}(ij,\rho) := \underbrace{(1-\Phi(\tau_i))\, u_i}_{\text{College } i \text{ accepts}} + \underbrace{\left(\Phi(\tau_i) - \Phi_2(\tau_i,\tau_j,\rho)\right) u_j}_{\text{College } i \text{ rejects, College } j \text{ accepts}}.$$

Let ij be the optimal pair under correlation ρ and $i'j'$ be that under correlation $\rho' > \rho$, where in each case, we write the top choice first.

Suppose towards a contradiction that $i < i' < j' < j$. By definition, $\mathcal{V}(ij,\rho) \geqslant \mathcal{V}(ij',\rho)$ and

$\mathcal{V}(i'j',\rho') \geqslant \mathcal{V}(i'j,\rho')$. Re-arranging these inequalities yields

$$\frac{\Phi(\tau_{i'}) - \Phi_2(\tau_{i'},\tau_j,\rho')}{\Phi(\tau_{i'}) - \Phi_2(\tau_{i'},\tau_{j'},\rho')} \leqslant \frac{u_{j'}}{u_j} \leqslant \frac{\Phi(\tau_i) - \Phi_2(\tau_i,\tau_j,\rho)}{\Phi(\tau_i) - \Phi_2(\tau_i,\tau_{j'},\rho)}. \tag{30}$$

By Lemma 5,

$$\frac{\Phi(\tau_i) - \Phi_2(\tau_i,\tau_j,\rho)}{\Phi(\tau_i) - \Phi_2(\tau_i,\tau_{j'},\rho)} < \frac{\Phi(\tau_i) - \Phi_2(\tau_i,\tau_j,\rho')}{\Phi(\tau_i) - \Phi_2(\tau_i,\tau_{j'},\rho')}.$$

Thus, to reach a contradiction to (30), it suffices to show that

$$\frac{\Phi(\tau_i) - \Phi_2(\tau_i,\tau_j,\rho')}{\Phi(\tau_i) - \Phi_2(\tau_i,\tau_{j'},\rho')} \leqslant \frac{\Phi(\tau_{i'}) - \Phi_2(\tau_{i'},\tau_j,\rho')}{\Phi(\tau_{i'}) - \Phi_2(\tau_{i'},\tau_{j'},\rho')}.$$

For $\tau \geqslant \tau_{j'}$, let

$$\Lambda(\tau) := \frac{\Phi(\tau) - \Phi_2(\tau,\tau_j,\rho')}{\Phi(\tau) - \Phi_2(\tau,\tau_{j'},\rho')}$$

denote the relative odds under correlation ρ' of being accepted by College j, rather than j', and rejected by a College whose threshold is τ. As College j is less selective than College j', observe that $\Lambda(\tau)$ is greater than 1. We show that $\frac{\mathrm{d}\Lambda}{\mathrm{d}\tau} < 0$ for any $\tau > \tau_{j'}$; intuitively, the bad news effect from being rejected by the top choice attenuates as the top choice becomes more selective. Observe that

$$\Lambda(\tau) = \frac{\int_{-\infty}^{\tau}\left(1-\Phi\left(\frac{\tau_j-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)dx}{\int_{-\infty}^{\tau}\left(1-\Phi\left(\frac{\tau_{j'}-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)dx}.$$

Taking the derivative with respect to τ, $\frac{\mathrm{d}\Lambda}{\mathrm{d}\tau}$ then has the same sign as

$$\left(1-\Phi\left(\frac{\tau_j-\rho\tau}{\sqrt{1-\rho^2}}\right)\right)\phi(\tau)\int_{-\infty}^{\tau}\left(1-\Phi\left(\frac{\tau_{j'}-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)dx -$$
$$\left(1-\Phi\left(\frac{\tau_{j'}-\rho\tau}{\sqrt{1-\rho^2}}\right)\right)\phi(\tau)\int_{-\infty}^{\tau}\left(1-\Phi\left(\frac{\tau_j-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)dx,$$

which in turn has the same sign as

$$\frac{\int_{-\infty}^{\tau}\left(1-\Phi\left(\frac{\tau_{j'}-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)dx}{\int_{-\infty}^{\tau}\left(1-\Phi\left(\frac{\tau_j-\rho x}{\sqrt{1-\rho^2}}\right)\right)\phi(x)dx} - \frac{\left(1-\Phi\left(\frac{\tau_{j'}-\rho\tau}{\sqrt{1-\rho^2}}\right)\right)\phi(\tau)}{\left(1-\Phi\left(\frac{\tau_j-\rho\tau}{\sqrt{1-\rho^2}}\right)\right)\phi(\tau)}$$

which is negative by Equation (19). □

F.5 Effect of Increasing Weight on Common Component: Examples

This subsection illustrates how increasing the weight on the common component, ρ, can induce the optimal two-college portfolio to become less or more aggressive.

Example 1. Let $C = \{1, 2, 3, 4\}$, $u_1 = 1.0099$, $u_2 = 1$, $u_3 = 0.5$, and $u_4 = 0.2$, and admission thresholds implicitly defined by $\Phi(\tau_1) = 65/81$, $\Phi(\tau_2) = 0.8$, $\Phi(\tau_3) = 0.5$, and $\Phi(\tau_4) = 0.01$. If $\rho = 1$, the optimal 2-college portfolio is $\{2, 4\}$ whereas if $\rho = 0$, the optimal 2-college portfolio is $\{1, 3\}$.

Example 2. Let $C = \{1, 2, 3\}$, $u_1 = 1$, $u_2 = 0.5$, and $u_3 = 0.48$, and admission thresholds implicitly defined by $\Phi(\tau_1) = 0.99$, $\Phi(\tau_2) = 0.5$, $\Phi(\tau_3) = 0.49$. If $\rho = 1$ the optimal 2-college portfolio is $\{1, 2\}$ whereas if $\rho = 0$, the optimal 2-college portfolio is $\{2, 3\}$.

G Results for Section 4.2

Proof of Proposition 6. First, we show that the utility from the optimal $(2^k - 1)$-portfolio is an upper bound for that achieved by the optimal k-strategy. Any k-strategy details one project to attempt first (corresponding to the "start here" label in Figure 9), two projects to attempt next (in case of success and failure), and generally 2^j projects in the j-th step. Thus, any strategy can attempt at most $2^k - 1$ projects. Since the optimal $(2^k - 1)$-portfolio chooses the best such set of projects, it guarantees at least as much utility.[49]

We next show that there exists a k-strategy that attains this upper bound (and is therefore optimal). Ann first attempts the median project in the optimal $(2^k - 1)$-portfolio. If it succeeds, she does not gain from attempting any lower-ranked projects; similarly, if it fails, she has no reason to attempt any higher-ranked project. Based on this observation, Ann attempts the median project among the top $2^{k-1} - 1$ projects of the optimal $(2^k - 1)$-portfolio if her first attempt succeeds (the first blue arrow in Figure 9) and the median project among the bottom $2^{k-1} - 1$ projects if her first attempt fails (the first red arrow in Figure 9). Generally, in each Step j she attempts the median project among the remaining $2^{k-j+1} - 1$ relevant projects. In this way, she is guaranteed to choose the same project as if she attempted all projects in the optimal $(2^k - 1)$-portfolio simultaneously. □

This analysis compares the optimal k-strategy with optimal k-portfolio. Alternatively, one might be interested in sequential search with a constant marginal cost of each attempt. We highlight some contrasts below using an example.

Example 3. Let $C = \{1, 2, 3\}$ and assume that $\mathbf{s}$ is distributed uniformly on $[0, 1]$, with $\tau_1 = 0.9$, $\tau_2 = 0.5$, and $\tau_3 = 0$. Furthermore, let $u_1 = 1.1$, $u_2 = 0.5$, and $u_3 = 0.1$. We assume a constant marginal cost of each attempt, i.e., $\phi(x) := cx$ for some $c > 0$. We will consider two specific values of c: $c_1 = 0.051$ and $c_2 = 0.01$.

We begin by considering Ann's dynamic search strategy. First, we observe that Project 2 is so attractive that she will not stop searching without trying Project 2 unless she is successful with Project 1. There are other strategies that can be easily ruled out. For example, strategies where

[49] We note that this upper bound is independent of the correlation in project outcomes.

Ann tries Project 3 and, if successful, then tries Project 1. Ann can save search costs by first trying Project 1, and only trying Project 3 in case of failure.

The values of c we consider are low enough that Ann is willing to try Project 1 even after success in Project 2, and is willing to try Project 3 even after failure in Project 2. This leaves us with two reasonable strategies: *Top to Bottom* (first try Project 1, Project 2 if failure, Project 3 if that too fails), or Middle Out (first try Project 2, Project 1 if success, Project 3 if failure). The expected cost from *Top to Bottom* is $c \times (1 + 0.9 + 0.5) = 2.4c$. The cost for *Middle Out* is $2c$; either way Ann attempts two projects. Hence, this latter strategy is optimal.

Let us now compare the static portfolio to the dynamic strategy. For $c = 0.051$, the optimal static portfolio is $\{1, 2\}$. This is weakly (and sometime strictly) more aggressive than the set of colleges searched by Ann (either $\{1, 2\}$ or $\{2, 3\}$). For $c = 0.01$, the optimal static portfolio includes *all* three colleges but Ann only searches two in the dynamic setup.

These results contrast with those from independent success: Chade and Smith (2006) show that agents stop searching after the first success and that sequential search, relative to the simultaneous problem, opts for more aggressive projects.

References

Ajayi, Kehinde F. 2024. "School choice and educational mobility: Lessons from secondary school applications in Ghana." *Journal of Human Resources*, 59(4): 1207–1243.

Ajayi, Kehinde F, Willa H Friedman, and Adrienne M Lucas. 2020. "When information is not enough: Evidence from a centralized school choice system." Working Paper 27887, National Bureau of Economic Research.

Chade, Hector, and Lones Smith. 2006. "Simultaneous search." *Econometrica*, 74(5): 1293–1307.

Karlin, Samuel, and Yosef Rinott. 1980. "Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions." *Journal of Multivariate Analysis*, 10(4): 467–498.

Lucas, Adrienne M., and Isaac M. Mbiti. 2012. "The determinants and consequences of school choice errors in Kenya." *American Economic Review*, 102(3): 283–288.

Plackett, R. L. 1954. "A Reduction Formula for Normal Multivariate Integrals." *Biometrika*, 41(3/4): 351–360.