Hedging in the College Application Problem*

S. Nageeb Ali   Ran I. Shorrer †

March 27, 2023

Abstract

An applicant applies to a portfolio of colleges while being uncertain about her prospects. Admissions decisions are correlated insofar as rejection from one college makes her pessimistic about her odds at other colleges. We develop a simultaneous search framework that accounts for this correlation. An applicant hedges by applying to a broad range of colleges, including safety schools. Those who have more confidence or better outside options apply to higher-ranked schools. If application costs decrease, the applicant broadens her optimal portfolio. We find an algorithm that yields the optimal solution in polynomial time.

*This paper supersedes and builds on Shorrer (2019). We have benefited from suggestions from Kehinde Ajayi, Hector Chade, Jacques Cremer, David Dillenberger, Simone Galperti, Guillaume Haeringer, Jacob Leshno, Ross Rheingans-Yoo, Uzi Segal, Denis Shishkin, Lones Smith, Joel Sobel, Rakesh Vohra, and various seminar and conference audiences. Salvador Candelas, Nilufer Gok, and Xiao Lin provided excellent research assistance. Ran Shorrer was supported by the United States–Israel Binational Science Foundation (BSF Grant 2016015).

†Pennsylvania State University. Emails: nageeb@psu.edu and rshorrer@gmail.com.

1 Introduction

College applicants face a risky choice with high stakes: they can apply to only a few colleges and are uncertain about which will admit them. To hedge their bets, many applicants apply to a broad range of colleges, including some they consider *reaches*, *matches*, and *safeties*. This strategy is generally recommended by guidance counselors and advisers. For example, in the context of US college admissions, the College Board suggests, "Before you start your applications, strengthen your list to include three reach colleges, two match colleges, and one safety college to ensure you apply to a balanced list of schools that match your academic abilities."[1] Diversifying one's application portfolio in this way hedges against not only the idiosyncratic risks of each college application but also the potential correlation across college decisions.

Why might admissions be correlated? In some contexts, correlation is an element of the design. For example, in many countries, applicants submit applications before taking a centralized exam that determines admissions decisions across schools. This process has been or is currently used for high school and college admissions in China, Ghana, Kenya, Mexico, Turkey, and the United Kingdom, to name just a few examples. Correlation also arises in centralized matching systems that use *single tie-breaking*, i.e., a common lottery that is used across schools to break priority ties for overdemanded seats. Public school seats in Amsterdam and New York City are allocated by this procedure.

In contexts such as the decentralized process of US undergraduate and graduate admissions, one would posit a different source of correlation. Here, an applicant is likely uncertain about her relative caliber: that is, where she stands in the pool of applicants. She may also worry that aspects of her application—e.g., the strength of her letters of recommendation, the quality of her application essay, the rarity of her extracurricular activities—will affect her admissions prospects across schools.

This paper studies how an applicant accounts for this correlation in the application process. Which schools should she apply to? Should she be aggressive by applying exclusively to high-ranked colleges, gambling that at least one accepts? Or should she hedge by also applying to safety schools?

To answer these questions, we build on the canonical framework of *simultaneous search* developed by Chade and Smith (2006). They model a decisionmaker who chooses a portfolio of gambles. Each gamble has a single prize ("acceptance") and once these gambles mature, the decisionmaker chooses her favorite prize. Although this analysis has been extended in a number of different directions,[2] the literature generally treats all gambles as stochastically independent. Hence, rejection by one school does not influence one's belief about the likelihood of admission elsewhere.

While we view independence to be the natural starting point, this assumption precludes scenarios like those described above. Our approach instead invokes the notion of a "common score" that determines admissions decisions across colleges. Each college has a known threshold and accepts an applicant only if her score clears its threshold. The applicant chooses her portfolio prior to learning her score. Therefore, she anticipates that if her application is rejected by, say, College i, then it will also be rejected by any college that has a weakly higher threshold.

[1] See https://secure-media.collegeboard.org/CollegePlanning/media/pdf/BigFuture-Strengthen-Your-College-List.pdf.
[2] See, for example, Galenianos and Kircher (2009), Chade, Lewis, and Smith (2014), and Olszewski and Vohra (2016).

In the context of our motivating examples, this common score may be the test result from the centralized examination that applicants are required to take after choosing their portfolio. Or it may be the lottery number that is used to break priority ties after one has submitted one's rank-order list in a single tie-breaking matching procedure. It may also simply reflect the applicant's standing in the pool of applicants or the strength of her letters of recommendation, analogous to the uncertainty that an applicant faces in decentralized undergraduate and graduate admissions.

The common-score formulation is tractable and has a graphical characterization. The optimal portfolio problem can be represented as a "coverage" problem in which each portfolio is depicted as a collection of rectangles in the unit square in which one dimension represents probability and the other utility. We use this graphical approach to deduce properties of the optimal solution.

In order to describe these properties, let us first classify colleges. The college that an applicant would pick if she could apply to only one is referred to as her "match": this school maximizes her expected payoff given its acceptance rate and her payoff from attending it. A "safety school" is less selective than the match: it has a higher acceptance rate but its reward is sufficiently low that given a single shot, the applicant would prefer to aim higher.[3] A "reach" is more selective than the match: its reward may be high but the applicant would not target it exclusively given its low acceptance rate.

Our main finding is that if the applicant applies to several colleges, she hedges by applying to *both* safety schools and reaches. The rationale for aiming high is that in a portfolio of multiple colleges, the applicant has backup options that make her more willing to gamble in her top choice. But she also aims low in choosing those backup options. The reason is that backup options matter only if she is rejected by her top choice and, by the Sure-Thing Principle, she should condition on this information. Since being rejected by her top choice constitutes bad news about her common score, the applicant is motivated to choose backup options less selective than her match.

We obtain these conclusions by characterizing how the optimal portfolio shifts with an applicant's beliefs and risk sensitivity. Theorem 1 shows that the optimal solution exhibits a *bad-news* effect wherein the decisionmaker applies to colleges less aggressively when she is less optimistic about her score (in a likelihood-ratio sense). Theorem 2 finds that the applicant applies more aggressively if she becomes more risk loving in an Arrow-Pratt sense. Combining the two effects, Theorems 3 and 4 show that if an applicant can apply to more colleges—say, if application costs fall—she chooses a more dispersed portfolio that expands both upwards and downwards in terms of college selectivity.[4]

These predictions diverge from those of the leading framework. Chade and Smith (2006) show that given independent admissions probabilities, the optimal portfolio is aggressive and precludes applications to safety schools: in particular, their Theorem 2 asserts that the optimal portfolio of k colleges is more aggressive than the k best single-college portfolios. When each college has many replicas, optimal portfolios exclude any college that is less selective than the single college in the optimal single-college portfolio.

Because applications to safety schools cannot be rationalized with independent admissions, Chade,

[3] Arguably, this notion of a "match" is conservative as an applicant might choose a college with a high acceptance rate if she can apply to only one. The point is that a safety school is even less selective than that.

[4] Pallais (2015) shows that portfolios become more dispersed in decentralized college admissions when application costs fall. Ajayi (2022) shows the same occurs in centralized admissions if the applicant is allowed to apply to more colleges.

Eeckhout, and Smith (2017) in their survey suggest

> *By the same token, "safety schools" can only be understood if acceptances are not independent....This remains a challenging but important research avenue.*

Against this backdrop, one contribution of this paper is that it posits a tractable model of correlation in which an applicant hedges by including safety schools in her optimal portfolio.

Beyond identifying properties of the optimal portfolio, we find an algorithm that delivers the optimal portfolio in polynomial time. Our approach treats simultaneous search as a recursive problem: an applicant decomposes the problem of identifying an optimal portfolio of k colleges into a two-stage process during which she first chooses the top-ranked program to include in her portfolio ("her first choice") and then chooses the optimal *continuation* of $(k-1)$ colleges if she is rejected by that first choice. In principle, this recursive problem could suffer from the curse of dimensionality where the continuation has to condition on all colleges that have rejected one's application. A property of the common-score formulation is that a sufficient statistic for the "history" of colleges that have rejected one's application thus far is the least selective college that has done so. Using this property, in the main text we describe an algorithm that achieves the optimal portfolio of k colleges in about n^3 steps; in the online appendix, we use the bad-news effect to speed up the algorithm to about $n^2 \log n$ steps.[5] We also show that this approach achieves the optimal solution in settings outside our baseline model, in particular if portfolios have to satisfy "tier constraints" that limit the number of schools in each tier. This generality is potentially useful as a number of countries feature such constraints.

We study three extensions of the baseline framework. The first extension allows the applicant to be uncertain about the threshold used by each college. We find that so long as the relative ranking of selectivity across colleges remains unchanged, the applicant chooses the same portfolio as if each college's threshold were known and set to its "certainty equivalent." The second extension departs more significantly by bridging the independent-success and common-score models. In this extension, a student obtains for each college a college-specific score that is the weighted average of independent and common components. Putting all the weight on the common component results in our baseline framework, whereas doing the same on the independent component leads to the canonical simultaneous-search model. Although the mixed model is considerably more challenging to analyze, we show that a motive for diversification manifests in two ways. First, when there is any weight on the common component, an applicant will apply to many colleges in both the most and least selective tiers if she is choosing a large portfolio. By contrast, when there is no weight on the common component, the applicant will apply almost exclusively to colleges in the most selective tier. Second, in studying two-college portfolios, we show that increasing correlation never makes the optimal portfolio less diverse in the sense of having a less aggressive first choice and a more aggressive

[5] Chade and Smith (2006) show that given independent admissions, the optimal portfolio is reached through a *marginal improvement algorithm* that selects colleges on the basis of their marginal benefit; this algorithm takes about n^2 steps. It finds the optimal portfolio of k colleges by choosing, at the first stage, the college with the highest expected value; at stage $j \leqslant k$, it chooses the college that adds the highest marginal benefit to the optimal portfolio of $(j-1)$ colleges constructed so far. This algorithm is myopic in that a college is added to the optimal portfolio without accounting for the other colleges that may be added later. A property that this algorithm requires is that the optimal portfolio of k colleges nests that of $(k-1)$ colleges, which holds in their setting. This property fails here.

second choice. The third extension models sequential search in the common-score setting and shows that it results in *informationally directed* search.

Let us put these results in context. Although the common-score approach may appear stark, admissions processes that use a common score appear worldwide. Here are a few examples:

- The UK employs a centralized university admissions process in which each student applies to at most five universities prior to taking the A-level examinations. Universities respond to most of these applications with a rejection or a provisional acceptance that specifies a threshold that the applicant's A-level score must clear in order for her to be admitted.[6] The applicant responds to these offers prior to learning her A-level score by selecting a first choice and a backup in case her score does not clear the threshold of that first choice, i.e., a two-college portfolio.

- In Kenya, more than a million students each year take the *Kenya Certificate of Primary Education* (KCPE), which serves as the entrance examination for secondary schools.[7] Students submit their choices when registering for the KCPE: that is, prior to taking the exam or learning their scores. Students are permitted to list up to six schools. Schools admit students in priority of their examination score until they run out of spots. Although schools do not have predetermined score thresholds, the relative selectivity of schools remains highly stable over time (Lucas and Mbiti, 2012), and our analysis in Section 5.1 allows for that uncertainty.

- A similar system is used in Ghana where students choose a portfolio of six senior high schools and then take a standardized common examination. Ajayi and Sidibe (2022) highlight that this system matches several hundred thousands of students with more than two thousand schools every year, "making it de facto one of the largest matching systems in the world."

Analyses of portfolio choices for these settings have been limited, partly because of the considerations that correlation raises. In her study of the admissions process in Ghana, Ajayi (2022) estimates preferences through the revealed-preference argument that an applicant must prefer a school in her portfolio to all others that are equally selective. Ajayi and Sidibe (2022) model the common-score approach directly and, building on our work, implement the algorithm that we propose in this paper.

A common score also arises in centralized matching procedures that involve single tie-breaking, which is used to allocate seats in public schools in Amsterdam and New York City. These matching procedures involve deferred acceptance and constrain the applicant to a maximal number of schools (e.g., twelve for New York City) in her rank-order list. The constraint implies that the mechanism is no longer strategyproof. Single tie-breaking then results in a portfolio choice identical to that in this paper: the optimal rank-order list simply lists schools in the optimal portfolio in the order of the applicant's preference.[8] In this spirit, Idoux (2023) uses a simultaneous-search approach to study applicant behavior in New York City, where she assumes that applicants account for the correlation

[6] Only about ten percent of applications receive an unconditional acceptance (Broecke, 2012).

[7] See https://tinyurl.com/2x3h2cd4.

[8] Deducing the optimal rank-order list in, say, the "Boston Mechanism," where one stands a better chance at schools that one ranks higher is more challenging. In that context, Calsamiglia, Fu, and Güell (2020) use a recursive approach to find optimal rank-order lists that they then apply to school-choice data in Barcelona.

across admissions decisions. However, she assumes that applicants follow a behavioral heuristic so as to avoid the computational burden of comparing all rank-order lists. The algorithms that we develop would deduce the optimal rank-order list while also sidestepping this challenge.[9]

When it comes to decentralized college admissions, as in the US, a more suitable model may be that of Section 5.2: namely, where each college uses the combination of a common score and a match-specific term. Such models feature in leading theoretical analyses of the college admissions process, although the focus is largely on the decisions of colleges. Avery and Levin (2010) model how a college may wish to structure its early admissions program when it cares about both applicants' *caliber* (the common score) and *fit* (the match-specific term). Che and Koh (2016) study how competitive pressures may induce a college to assign more weight to its match-specific term given the choices of other colleges. Both of these studies abstract from the portfolio choice problem by assuming that there are only two colleges and each applicant applies to both. Chade, Lewis, and Smith (2014) model the decisions of colleges and applicants jointly, restricting attention to two colleges. Most of their analysis assumes independence but they also discuss settings in which students are uncertain about their caliber and colleges observe either perfectly correlated signals (as in our baseline framework) or affiliated noisy signals (analogous to the analysis in Section 5.2). Our analysis complements this prior literature by offering a detailed study of the application problem and the motive to hedge.[10]

An empirical literature estimates structural models of school and college admissions, mostly under the assumption that applicants view admission decisions as being stochastically independent; see, for example, Howell (2010), Fu (2014), and Walters (2018).[11] Kapor (2020) flexibly models a correlation structure allowing for a common caliber observed by all colleges (but not applicants) and an idiosyncratic match-specific term. His structural estimates indicate that applicants are imperfectly informed about their caliber and therefore perceive correlation across admissions decisions. Our extension in Section 5.2 hews to the correlation structure of his model; we show that some weight on the common component motivates applicants to apply to safety schools. This highlights the benefits of modeling correlation flexibly: if one estimates preferences assuming stochastic independence in settings where an applicant actually treats admissions as correlated, one might incorrectly infer from applications to safety schools that the applicant genuinely prefers those schools.

Our results also contribute to the active discussion of segregation in colleges and universities. Hoxby and Avery (2013) study why elite colleges are missing "low-income high achievers." They find that lower-income students are far less likely to apply to selective colleges than higher-income

[9] She highlights that if applicants rank 12 of 60 programs, there are more than 10^{20} lists to consider. Our algorithm in the main text would narrow this down to $\approx 43,200$ lists and that in the online appendix would consider $\approx 4,320$. Our results also identify how choices from Idoux's heuristic diverge from the rational benchmark. The heuristic stipulates that when the applicant decides whether to add a school to her ranked list, subject to a cost, she envisions that she will incur no further costs. The applicant therefore naively assumes she will add all other schools to the list. Because this overstates the number of backup options she will have, the applicant effectively becomes more risk loving. Theorem 2 implies that the resulting list is then more aggressive than that of the rational benchmark.

[10] Simultaneous search also features in the study of labor markets (Galenianos and Kircher, 2009; Kircher, 2009) where firms post wage offers and workers apply to several firms simultaneously. Here too, the literature assumes that applications are accepted independently across firms. Embedding our portfolio problem in this market setting may shed light on labor markets in which workers perceive hiring decisions as being informative about their future prospects.

[11] In some cases, such as Fu, Guo, Smith, and Sorensen (2022), the authors suggest that the model ought to allow for correlation but they assume independence for computational tractability.

students with the same characteristics; Ajayi (2022) and Ajayi and Sidibe (2022) note similar effects in their study of secondary-school applications in Ghana. Our work suggests two potential mechanisms. Theorem 1 concludes that those who are more pessimistic about their caliber apply less aggressively. This interpretation is in line with Hoxby and Avery's focus on how many low-income high achievers attend school districts that do not support selective public high schools and are unlikely to have teachers who have attended a selective college. Theorem 2 points to a second contributing mechanism: even if all applicants have the same beliefs, those with better outside options apply more aggressively. For example, if higher-income students can afford private schools outside the public school system, they may target better public schools. We show in the online appendix that this risk-loving effect emerges also in the independent-admissions framework. Hence, the general point stands: unequal outside options may exacerbate segregation through applicants' portfolio choice.[12]

2 Example

We illustrate our approach using a simple example. Ann can apply to a subset of colleges $\{1,2,3\}$, ordered so that College i is her i^{th} favorite. Admissions decisions are based on a *common score* s, which may reflect her standing in the pool of all applicants or her performance on an examination taken after she has chosen her portfolio. If Ann applies to College i, her application is accepted if her score exceeds its threshold, τ_i. The timing is as follows: (i) Ann chooses a portfolio; (ii) Ann's score is realized, resulting in college admissions decisions; and (iii) Ann chooses a college among those that accept her application. She obtains her outside option if all her applications are rejected.

For simplicity, suppose that s is distributed uniformly on $[0,1]$, and each threshold is in $[0,1]$. Table 1 summarizes for each college Ann's utility from attending that college, its threshold, and its (marginal) acceptance rate. The utility of her outside option is 0.

	College 1	**College 2**	**College 3**
Utility (u_i)	1	0.45	0.25
Score Threshold (τ_i)	0.78	0.5	0.125
Acceptance Rate	0.22	0.5	0.875

Table 1

Optimal Single-College Portfolio: Given a single shot, Ann chooses the college that maximizes $(1-\tau_i)u_i$. Based on the numbers above, this is College 2. Although this problem is trivial, depicting it graphically is useful for the subsequent analysis. Figure 1 depicts the expected utility of each single-college portfolio as a coverage problem: each college is represented as a rectangle in a score-utility space whose horizontal edge depicts its acceptance rate and vertical edge depicts Ann's utility from attending it. The optimal portfolio corresponds to the rectangle that covers the most area.

Optimal Two-College Portfolio: Now suppose Ann can apply to two colleges; we establish that the optimal such portfolio is $\{1,3\}$.

[12]These results complement recent work by Calsamiglia, Martínez-Mora, and Miralles (2021) and Akbarpour, Kapor, Neilson, van Dijk, and Zimmerman (2022) who note similar effects in manipulable centralized matching procedures.

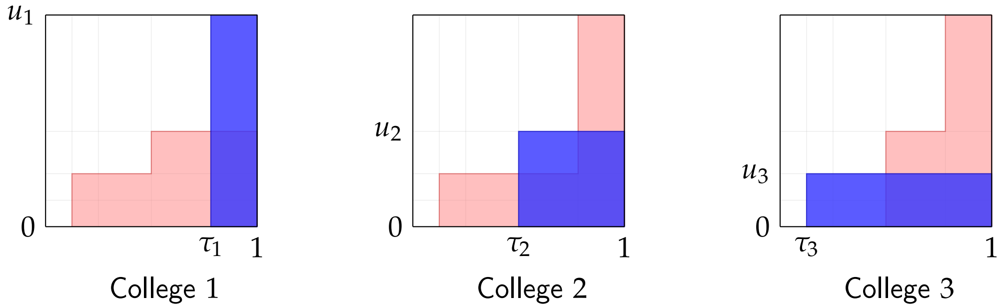

Figure 1: Single-college portfolios. Each figure plots the scores and thresholds on the horizontal axis and utility on the vertical axis. The area of the blue rectangle is the expected utility of that single-college portfolio.

To explain why, we first compare portfolios $\{1,2\}$ and $\{1,3\}$. For either portfolio, if Ann is accepted by College 1, she will attend that college and then obtain the same payoff. By the Sure-Thing Principle, her optimal choice between these two portfolios should condition on her being rejected by College 1, i.e., obtaining a score less than its threshold of 0.78. Portfolio $\{1,2\}$ then delivers a conditional expected payoff of

$$Pr(s \geqslant 0.5|s < 0.78) \times u_2 \approx 0.16. \tag{1}$$

By contrast, if Ann chooses portfolio $\{1,3\}$, her conditional expected payoff is

$$Pr(s \geqslant 0.125|s < 0.78) \times u_3 \approx 0.21. \tag{2}$$

We dub this the *bad-news* effect: although College 2 is better ex ante, College 3 is the better backup option once Ann conditions on the bad news that she has been rejected by College 1.

This *bad-news* effect is reflected in the coverage problem depicted in Figure 2. For portfolios $\{1,2\}$ and $\{1,3\}$, Ann enrolls in College 1 if her score exceeds τ_1. She therefore obtains the area of the rectangle associated with College 1 in either case. The choice then hinges on how much area the other college adds, i.e., the area added when she is rejected by College 1. Removing the parts of all rectangles with scores above τ_1 takes away a larger slice of College 2's rectangle than it does of

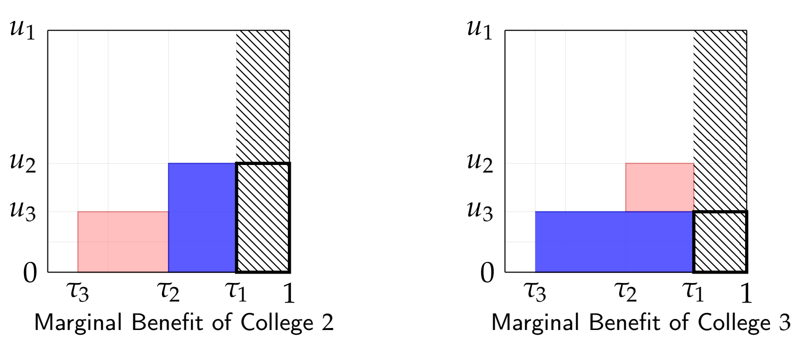

Figure 2: Comparing portfolios $\{1,2\}$ and $\{1,3\}$. Each figure shows the marginal benefit of adding a college to College 1. The dashed area shows regions where Ann is admitted to College 1, the area framed in black is the overlap between College i and College 1, and the blue area is what remains after removing the overlap.

College 3's, thereby making the latter more attractive.[13]

We turn to the choice between portfolios $\{1,3\}$ and $\{2,3\}$. If her score is so low that College 3 rejects her, she obtains her outside option in either case. So Ann should condition on her score clearing the threshold for College 3. In that event, portfolio $\{1,3\}$ delivers an expected payoff of

$$u_3 + \underbrace{Pr(s \geqslant 0.78 | s \geqslant 0.125)}_{\text{Accepted by College 1}} \times (u_1 - u_3) \approx 0.42. \tag{3}$$

This expression reflects Ann's certainty that she accrues the payoff of College 3 for sure; with some chance, she also obtains the additional bump that comes from being accepted by College 1. By contrast, the conditional expected payoff of portfolio $\{2,3\}$ is

$$u_3 + \underbrace{Pr(s \geqslant 0.5 | s \geqslant 0.125)}_{\text{Accepted by College 2}} \times (u_2 - u_3) \approx 0.35. \tag{4}$$

Therefore, Ann is better off choosing portfolio $\{1,3\}$.

Let us interpret why. Ann's choice between $\{1,3\}$ and $\{2,3\}$ matters only if she is accepted by at least College 3. Once she conditions on that event, College 3 is effectively her outside option. We see this reflected in (3) and (4), where she obtains a payoff of u_3 whenever she is rejected by a better college. As $u_3 > 0$, this outside option is better than that of the original problem. This better outside option makes Ann more "risk loving" (in an Arrow-Pratt sense) and that induces her to apply more aggressively. This logic is reflected in Figure 3, where obtaining the outside option of u_3 removes a bottom slice of the rectangles associated with Colleges 1 and 2 but as a larger part of College 2's rectangle is removed, College 1 becomes relatively more attractive.

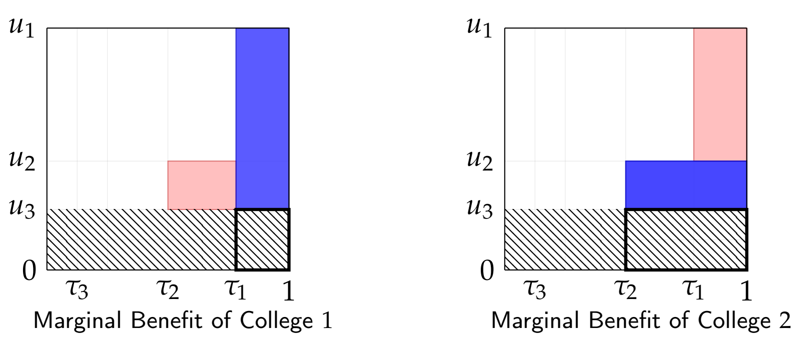

Figure 3: Comparing portfolios $\{1,3\}$ and $\{2,3\}$. Each figure shows the benefit of adding a college to College 3. As the choice matters only if Ann is at least accepted by College 3, she effectively has an outside option of u_3. The solid blue area shows what remains of each rectangle after removing those with utility lower than u_3.

In comparing $\{1,2\}$ and $\{1,3\}$, we documented a *bad-news* effect that induces Ann to be less aggressive and choose College 3 as her backup option if she is rejected by College 1. By contrast, in comparing $\{2,3\}$ and $\{1,3\}$, we described a *risk-loving* effect that leads Ann to be more aggressive

[13] Equivalently, the expected payoff of a portfolio equals the total area covered by the union of its rectangles. Because one does not double count the intersection of rectangles in $\{1,2\}$ (or $\{1,3\}$), the other college is relevant only in the region where Ann is rejected by College 1.

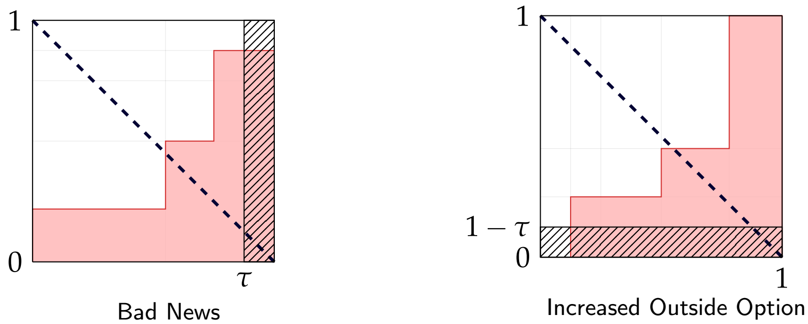

Figure 4: The left figure shows the effect of the bad news that Ann's score is below τ. Transposing each rectangle over the $-45°$ line yields an isometric coverage problem in which her outside option increases from 0 to $(1-\tau)$.

and apply to College 1 once she has College 3 as a backup. We show generally that (i) bad news lead to less aggressive portfolios (Theorem 1), and (ii) becoming more risk-loving leads to more aggressive portfolios (Theorem 2). Although the two effects appear distinct, each is a mirror image of the other. We depict this duality in Figure 4: as the expected payoff of a portfolio is equal to the total area covered by the union of the corresponding rectangles, the comparison of areas does not change if one transposes each shape around the off-diagonal.

We also see that the optimal two-college portfolio $\{1,3\}$ expands in both directions relative to the optimal single-college portfolio. Theorem 3 shows that this expansion obtains generally: of the optimal $(k+1)$-portfolio, the top k colleges are more aggressive than the optimal k-portfolio and the bottom k colleges are less aggressive than the optimal k-portfolio. Hence, the larger portfolio is more spread out. This intuition combines the two aforementioned forces: (i) by the bad-news effect, having an additional college at the top makes Ann less aggressive in choosing the other k colleges as these are relevant only if she is rejected by that top college; (ii) but by the risk-loving effect, an additional college at the bottom offers her an outside option that makes her more aggressive in optimally choosing the other k colleges.

To see how these predictions contrast with those of Chade and Smith (2006), suppose each college has the same acceptance rate as in Table 1 but that admission decisions are stochastically independent across colleges. Their Theorem 2 implies that the optimal two-college portfolio is $\{1,2\}$; moreover, the optimal $(k+1)$-portfolio nests the optimal k-portfolio and, if replicas of each college are available, always expands upwards. Thus, regardless of the number of applications Ann makes, she *never* applies to College 3 (or a replica thereof). This exemplifies their conclusion that Ann does not apply to a "safety school": that is, a college less selective than the optimal single-college portfolio.

3 Model

Ann applies to colleges in $C := \{1,2,...,n\}$. The application process works as follows. Ann first chooses a portfolio $P \subseteq C$ of colleges. After this choice is made, a score s in $[0,1]$ is realized. College i has a *score threshold* $\tau_i \in [0,1]$ such that it accepts Ann's application if $s \geqslant \tau_i$. Once admissions decisions are made, Ann chooses her outside option or a college that accepts her. Ann's utility from attending College i is u_i and that from her outside option is u_o. The *utility assessment*

$U := (u_o; u_1, \ldots, u_n)$ specifies utilities based on these outcomes.

Ann's (ex ante) beliefs about her score are represented by the cumulative distribution function F. Without loss of generality, we assume that F is atomless and has a strictly positive density on an interval support. Given Ann's assessment and beliefs, the expected value of portfolio P is

$$V(P, U, F) := \int_0^1 \max_{\{i \in P : s \geqslant \tau_i\} \cup \{o\}} u_i \, dF. \tag{5}$$

Thus, in choosing her optimal portfolio, Ann anticipates that once admission decisions have been made, she will choose between her outside option and her favorite among all colleges that admit her.

We make a few assumptions to simplify the exposition.

As Ann does not benefit from applying to a college dominated by another—in terms of desirability or selectivity—or to replicas of the same college, we rule out such colleges. Hence, we order colleges so that if $i < j$, then College i is more desirable ($u_i > u_j$) and selective ($\tau_i > \tau_j$).

We also focus on colleges that improve upon Ann's outside option and to which she stands a chance of being accepted. These are the *rationalizable colleges*, $C^*(U, F) := \{i \in C : u_i > u_o, F(\tau_i) < 1\}$. We assume that at least one college is rationalizable.

Finally, we rule out ties. We assume that no two portfolios in $C^*(U, F)$ achieve the same value: that is, for every distinct pair P and $\tilde{P}$ in $C^*(U, F)$, $V(P, U, F) \neq V(\tilde{P}, U, F)$. Given any belief F, this assumption holds for almost all assessments.

We now write the value of a portfolio explicitly. For portfolio P, denote its i^{th} highest-ranked element by $P^{(i)}$. Setting $F(\tau_{P^0}) := 1$,

$$V(P, U, F) = \sum_{i=1}^{|P|} \left(F(\tau_{P^{(i-1)}}) - F(\tau_{P^{(i)}})\right) \max\{u_{P^{(i)}}, u_o\}. \tag{6}$$

Thus, the expected value of a portfolio weighs the utility of attending College $P^{(i)}$ by the probability that $P^{(i)}$ is the highest-ranked college in portfolio P that admits Ann.

The value of a portfolio must be balanced against its cost. Ann's cost depends on the number of applications she makes: the cost schedule is $\phi : \mathbb{N} \to \mathbb{R}_+ \cup \{\infty\}$ where $\phi(|P|)$ is the total cost of portfolio P. We assume that ϕ is weakly increasing. A portfolio is optimal if it solves

$$\max_{P \subseteq C} \left[V(P, U, F) - \phi(|P|)\right]. \tag{OPT}$$

To solve (OPT), one may identify the optimal portfolio of k colleges for each k and then choose k optimally. The (unique) solution to the first stage of this optimization problem is

$$P^*(k, U, F) := \operatorname*{argmax}_{\{P \subseteq C^*(U,F) : |P| \leqslant k\}} V(P, U, F).$$

We call this object the optimal k-portfolio. Our results assess how that portfolio shifts with Ann's prior (Theorem 1), her assessment (Theorem 2), and the number of applications k (Theorem 3). To describe these shifts, we order portfolios by their aggressiveness and dispersion, formalized below.

Definition 1. *For non-empty portfolios P and $\tilde{P}$, P is **more aggressive** than $\tilde{P}$ if any of the following is true:*

(a) $|P| = |\tilde{P}|$ *and for every* $i = 1, \ldots, |P|$, $P^{(i)} \leqslant \tilde{P}^{(i)}$.

(b) $|P| < |\tilde{P}|$ *and for every* $i = 1, \ldots, |P|$, $P^{(i)} \leqslant \tilde{P}^{(i)}$.

(c) $|P| > |\tilde{P}|$ *and for every* $i = 1, \ldots, |\tilde{P}|$, $P^{(i+|P|-|\tilde{P}|)} \leqslant \tilde{P}^{(i)}$.

We denote the aggressiveness order by $\geqslant_A$.

A more aggressive portfolio targets more selective schools. Case (a) compares portfolios of the same cardinality by the standard vector dominance order, stipulating that the i^{th} best college of P is higher ranked than that of $\tilde{P}$. The other cases extend this definition to portfolios of different cardinality. Case (b) applies if P has fewer colleges than $\tilde{P}$; P is deemed more aggressive if, according to case (a), P is more aggressive than the portfolio comprising the best $|P|$ colleges in $\tilde{P}$. If P has more colleges, case (c) asserts that P is more aggressive if the portfolio comprising its $|\tilde{P}|$-least selective colleges is nevertheless more aggressive than $\tilde{P}$. We note that $\geqslant_A$ is a transitive but incomplete order.

We use this notion of aggressiveness to formalize when one portfolio is more *dispersed* than another. Let $\lceil P \rceil^k$ and $\lfloor P \rfloor^k$, respectively, denote the k most and least selective colleges in portfolio P.

Definition 2. *Portfolio P is **more dispersed** than $\tilde{P}$ if all of the following are true:*

(a) *Portfolio P has more colleges:* $|P| \geqslant |\tilde{P}|$.

(b) *The $|\tilde{P}|$-most selective colleges in P is more aggressive than $\tilde{P}$:* $\lceil P \rceil^{|\tilde{P}|} \geqslant_A \tilde{P}$.

(c) *The $|\tilde{P}|$-least selective colleges in P is less aggressive than $\tilde{P}$:* $\lfloor P \rfloor^{|\tilde{P}|} \leqslant_A \tilde{P}$.

We denote the dispersion order by $\geqslant_D$.

Informally, a larger portfolio P is more dispersed than $\tilde{P}$ if it expands in both directions, to wit, including colleges higher and lower ranked than those in $\tilde{P}$. Suppose that $\tilde{P}$ has k colleges and P has more. Definition 2(b) stipulates that the k most selective colleges in P constitute a portfolio more aggressive than $\tilde{P}$. The notion of more aggressive that we use is that of Definition 1(a): the highest-ranked college of $\tilde{P}$ is less selective than that of P, the second-highest ranked college of $\tilde{P}$ is less selective than that of P, and so on and so forth until all colleges in $\tilde{P}$ have been exhausted. This ranking implies that portfolio P unambiguously expands upwards to include more selective colleges. Simultaneously, Definition 2(c) stipulates that the portfolio comprising the k least selective colleges in P is less aggressive than $\tilde{P}$. Thus, the lowest ranked college of $\tilde{P}$ is *more selective* than that of P, the second-lowest ranked college of $\tilde{P}$ is also more selective than that of P, and so on and so forth until all colleges in $\tilde{P}$ have been exhausted. Hence, portfolio P also unambiguously expands downwards relative to portfolio $\tilde{P}$. Because the larger portfolio unambiguously spreads out in both directions, we view it as being more dispersed. This order, too, is transitive but incomplete.

A Comparison To Independent Simultaneous Search: Prior to analyzing our setting, we compare it to the standard setup. For this comparison, we consider a different but equivalent representation of the common-score problem. Suppose the acceptance rate of College i is α_i; moreover, if Ann's application is rejected by College i, then it would also be rejected by all higher-ranked colleges. This formulation is isomorphic to our setup with $F(\tau_i) = 1 - \alpha_i$. We can then rewrite (6) as

$$\sum_{i=1}^{|P|} \underbrace{\left(\alpha_{P(i)} - \alpha_{P(i-1)}\right)}_{\substack{\text{Accepted by } P^{(i)}, \\ \text{Rejected by better colleges in P}}} \max\{u_{P(i)}, u_o\}, \tag{7}$$

where we normalize $\alpha_{P(0)} = 0$.

We contrast this value with that of Chade and Smith (2006), where admissions decisions are independent across colleges. The value of a portfolio then is

$$V^I(P, U, \alpha) := \sum_{i}^{|P|} \underbrace{\left(\alpha_{P(i)} \prod_{j}^{i-1} (1 - \alpha_{P(j)})\right)}_{\substack{\text{Accepted by } P^{(i)}, \\ \text{Rejected by better colleges in P}}} \max\{u_{P(i)}, u_o\}. \tag{8}$$

In this expression, rejection from colleges better than $P^{(i)}$ does not affect Ann's probability of acceptance to college $P^{(i)}$. By contrast, in the common-score formulation of Equation (7), Ann puts lower odds on being accepted by $P^{(i)}$ if she is rejected by better colleges.

The working paper version of Chade and Smith (2006) anticipates the formulation of Equation (7) and identifies it as a setting in which their marginal improvement algorithm does not reach the optimal portfolio. They assert that "*It is an exciting open problem to find an algorithm that works efficiently in these cases: future research beckons.*" Below we describe properties of the optimal portfolio and an efficient algorithm that computes it.

4 Optimal Portfolios

4.1 How Beliefs Affect Portfolio Choice

Suppose Ann is about to submit her applications, but right before doing so, she obtains bad news that makes her pessimistic about her prospects. How should this influence her portfolio?

Let us first formalize bad news. For an atomless CDF G, the probability that College i is the best college that would admit Ann is $\mu(i, G) := G(\tau_{i-1}) - G(\tau_i)$, where we set $\tau_0 = 1$.

Definition 3. *Distribution H has* ***bad news*** *relative to G if for every College i and less selective College j,*

$$\mu(j, G)\mu(i, H) \leqslant \mu(i, G)\mu(j, H). \tag{9}$$

In such a case, we write $G \succcurlyeq_{LR} H$.

Definition 3 adapts the standard definition of the likelihood ratio dominance order to our setting: rearranging (9) implies that relative to distribution H, distribution G has a higher likelihood ratio of College i being the best college that accepts Ann versus any less selective College j.[14]

[14]Definition 3 is implied by the standard likelihood-ratio dominance order but is weaker as it imposes restrictions on the

Theorem 1. *Bad news leads to a less aggressive portfolio:* $G \succcurlyeq_{LR} H \Rightarrow P^*(k, U, G) \succcurlyeq_A P^*(k, U, H)$.

We offer some intuition for this result. Suppose that, as a special case of bad news, H is a right truncation of G: $H(s) = G(s|s \leqslant \tau)$. This is tantamount to Ann's learning that her score is below τ. This news reduces Ann's odds of admission to each school, but as depicted in Figure 2, it has a relatively stronger effect at more selective schools. Although bad news can go beyond truncating the prior, it shares the property that more area is removed from the right, thereby disadvantaging more selective schools. The logic then *suggests* that Ann ought to be less aggressive at least in terms of the most selective college she decides to include in her portfolio. It is merely suggestive, however, because Ann may conceivably choose a more aggressive second or third choice, which leads to a countervailing force on the optimal first choice. A significant component of the proof addresses this possibility and shows that indeed, if $G \succcurlyeq_{LR} H$, the top choice in the optimal k-portfolio under distribution G is higher ranked than that under H.

Once this property is established, an inductive argument shows that this force propagates to all subsequent choices. Suppose Ann chooses Colleges i and j as her top choices under distributions G and H, respectively. We argue that the following holds:

$$G(s|s \leqslant \tau_i) \succcurlyeq_{LR} H(s|s \leqslant \tau_i) \succcurlyeq_{LR} H(s|s \leqslant \tau_j).$$

The first inequality states that because $G \succcurlyeq_{LR} H$, rejection from the same top choice (College i) results in posterior beliefs that remain LR-ordered. The second inequality identifies an amplification in which, because Ann has chosen a less selective college as her top choice with distribution H, its rejection conveys *even worse* news. As the second choice of an optimal k-portfolio given beliefs G and H, respectively, corresponds to the first choice of an optimal $(k-1)$-portfolio given beliefs $G(s|s \leqslant \tau_i)$ and $H(s|s \leqslant \tau_j)$, respectively, it follows from the inductive hypothesis that the second choice under distribution G is also higher ranked. By induction, this property holds for all subsequent choices.

4.2 How Risk Attitudes Affect Portfolio Choice

Now suppose that, right before applying, Ann obtains a better job with her high school diploma so that the value of her outside option increases. How does this influence her portfolio? We formalize this as a change in Ann's *risk sensitivity*, in the Arrow-Pratt sense.

Definition 4. *Assessment* $U' := (u'_o; u'_1, \ldots, u'_n)$ *is more* ***risk loving*** *than* $U := (u_o; u_1, \ldots, u_n)$ *if there exists a convex non-decreasing transformation* $\nu : \mathbb{R} \to \mathbb{R}$ *such that for every* $i \in \mathcal{C} \cup \{o\}$,

$$\max\{u'_i, u'_o\} = \nu(\max\{u_i, u_o\}). \tag{10}$$

In such a case, we write $U' \succcurlyeq_{RL} U$.

Definition 4 extends the standard notion of risk loving to allow for an outside option. A special case of a more risk-loving utility assessment is one in which the utility for each college remains the same but that of her outside option increases, which we considered in Section 2.

CDF only at the thresholds.

Theorem 2. *Risk-love leads to a more aggressive portfolio:* $U' \succcurlyeq_{RL} U \Rightarrow P^*(k, U', G) \succcurlyeq_A P^*(k, U, G)$.

We prove Theorem 2 by exploiting a duality between risk sensitivity and beliefs. Fix a probability distribution F and two assessments U and U' where each outside option is normalized to 0 and each utility is a number in $[0, 1]$.[15] As shown in Figure 4 on p. 9, we can transpose this problem around the off-diagonal of the unit square to flip utilities and admissions probabilities. Denoting $T[U]$ and $T[U']$ as the transposed versions of U and U', respectively, we show

$$U' \succcurlyeq_{RL} U \Rightarrow T[U'] \preccurlyeq_{LR} T[U].$$

In other words, if U' is more risk loving than U, then its transposition $T[U']$ has bad news relative to the transposition $T[U]$.[16] Theorem 1 implies that Ann applies less aggressively given the beliefs $T[U']$. As transposition flips the labels of colleges, it follows that Ann applies more aggressively given the more risk-loving assessment U'.

Theorem 2 has implications for how unequal outside options may exacerbate segregation. Say the schools represent *public schools* and the outside option is one's value of attending a *private school* that is outside the system. It follows from Theorem 2 that if Ann and Bob have identical preferences and beliefs but differ in that Bob can afford private school while Ann cannot, then Bob will apply more aggressively. For example, Bob may target the flagship public university or better schools in the school district whereas Ann may opt for less selective options. Because students with weaker outside options apply less aggressively, they are more likely to be matched with less selective schools.

4.3 Larger Portfolios Are More Dispersed

We now study how the optimal portfolio changes if Ann can apply to more colleges or if application costs fall. We first show that compared to optimal portfolios of a smaller size, those of a larger size are more dispersed, as per Definition 2. In other words, Ann expands both upwards and downwards.

Theorem 3. *Larger optimal portfolios are more dispersed:* $k' \geqslant k \Rightarrow P^*(k', U, F) \succcurlyeq_D P^*(k, U, F)$.

Figure 5 illustrates this dispersion: the top k items in the optimal k'-portfolio are more aggressive than the optimal k-portfolio and that in turn is more aggressive than the bottom k items in the optimal k'-portfolio. As seen in Section 2, the expansion upwards and downwards can be strict. Indeed, this will be the case if the set of colleges is "rich" in that the gaps in selectivity and payoffs across colleges are small. In such a case, it follows that in any optimal portfolio of two or more colleges, Ann applies to a college lower ranked (and less selective) than the college in her optimal single-college portfolio.

The logic for Theorem 3 invokes both the bad-news and risk-loving effects. Let us compare the optimal $(k+1)$-portfolio to the optimal k-portfolio; the general case then follows by induction. Observe that the optimal $(k+1)$-portfolio can be obtained through a two-stage process where Ann first picks the highest ranked college for the portfolio (her "first choice") and then picks k backup options

[15]For utility $U = (u_o; u_1, \ldots, u_n)$, we use the normalization $\left(0; 1, \ldots, \frac{u_n - u_o}{u_1 - u_o}\right)$.

[16]The converse is also true but we prove and use only one direction of this implication.

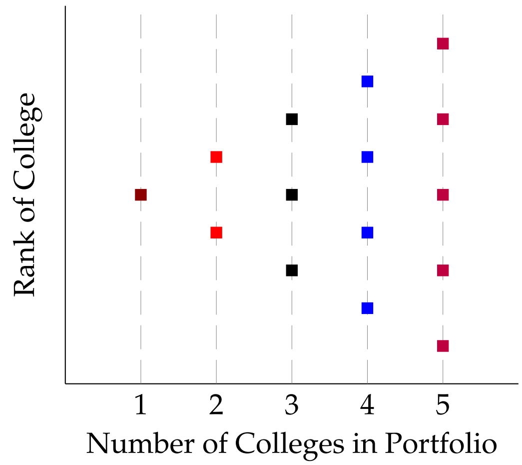

Figure 5: Illustrating Theorem 3. The optimal k-portfolio is more dispersed for higher values of k.

optimally. These backups matter *only if* she is rejected by the first choice; hence, they must be optimal conditional on that event. But being rejected by her first choice is bad news. The ordering in Definition 2(c) then follows from Theorem 1: the k optimally chosen backups must be less aggressive than the optimal k choices based on the prior. This is the force for including safer colleges.

To see why the larger portfolio also moves up, note that the optimal $(k+1)$-portfolio is also obtained by first selecting a college to be the lowest ranked in her portfolio (her ultimate "safety") and then picking k colleges that improve on that safety. Suppose College i is the safety. Because Ann applies only to a college that improves on the outside option, its utility u_i exceeds that of her outside option u_o. Ann's choice of k improvements matters only if she is at least admitted to College i; otherwise, she obtains u_o regardless of her choice. Conditioning on this event—namely $s \geqslant \tau_i$—implies that if she is rejected by k improvements, she obtains u_i. Optimally choosing k improvements is therefore equivalent to the modified optimal k-portfolio problem with an outside option of u_i. As this modified problem has a better outside option, Theorem 2 implies that the optimal portfolio of k improvements is more aggressive than the optimal k-portfolio, establishing Definition 2(b).

This greater-dispersion property implies that larger portfolios expand both upwards and downwards. Chade and Smith (2006, Section 5.2) show in a two-college setting that if admission decisions are stochastically independent, the optimal portfolio expands *only* upwards. To facilitate comparison, we generalize their conclusion in Theorem A.1 in the online appendix. Specifically, we show that if admissions are independent and each of n colleges has many replicas, then the optimal $(k+1)$-portfolio nests the optimal k-portfolio and is more aggressive (as per Definition 1(c)).[17] An implication is that Ann applies to only those colleges that are higher ranked and more selective than the college in her optimal single-college portfolio; she does not apply to safety schools.

One can reconcile these disparate predictions using the bad-news and risk-loving effects. With independent admissions, rejection at her top choice is uninformative about Ann's prospects at her backup options, thereby nullifying the bad-news effect. Ann then has no motive to choose less aggressive backup options. But the risk-loving effect remains unabated, as more backup options offer

[17] For the common-score framework, there is no loss of generality in omitting replicas. For the independent-admissions framework, replicas are valuable and hence we allow for them in this analysis.

more insurance. Hence Ann expands only upwards when she can apply to more colleges.

Theorem 3 compares optimal portfolios of different cardinalities, but it also has implications for how the optimal portfolio varies more generally with the cost schedule. We say that cost schedule ϕ' has **lower marginal costs** than cost schedule ϕ if for each $k \in \mathbb{N}$, $\phi'(k+1) - \phi'(k) \leqslant \phi(k+1) - \phi(k)$. Denote the optimal portfolio for a generic cost schedule ϕ by $P(\phi) := \text{argmax}_{P \subseteq C} [V(P) - \phi(|P|)]$.[18]

Theorem 4. *Suppose the cost schedule ϕ' has lower marginal costs than ϕ. Then the optimal portfolio for cost schedule ϕ' is more dispersed than that for ϕ: $P(\phi') \geqslant_D P(\phi)$.*

Theorem 4 is a corollary of the earlier result: lower marginal costs induce Ann to apply to more colleges and Theorem 3 then implies that the optimal portfolio is more dispersed.

4.4 An Algorithm That Obtains the Optimal Portfolio

Finding the optimal portfolio is complex because the value of adding a college to one's portfolio depends on the other colleges in one's portfolio. Solving this combinatorial optimization problem by exhaustively checking all 2^n portfolios—or even all k-portfolios for a given k—is computationally infeasible for large n. Chade and Smith (2006) show that such computations are unnecessary if admissions decisions are independent: a greedy algorithm reaches the optimal portfolio in about n^2 steps. A property that this algorithm relies upon is that the optimal $(k+1)$-portfolio nests the optimal k-portfolio, which holds in their setting but fails in ours (as illustrated in Section 2).

We instead turn to dynamic programming. We describe an efficient algorithm that delivers the optimal portfolio in about n^3 steps, and in the online appendix, we use the bad-news effect to speed it up to $n^2 \log n$ steps.[19] Conveniently, this approach also reaches the optimal portfolio in settings in which the applicant faces a "tier" constraint that limit her to a certain number of schools in each tier. Such constraints features in several countries including Ghana and Kenya.

For a set of colleges, $\mathcal{C}'$, denote Ann's beliefs conditional on being rejected by every college in $\mathcal{C}'$ by $F_{\mathcal{C}'} := F(\cdot | s < \tau_i \text{ for every } i \in \mathcal{C}')$; if $\mathcal{C}'$ is empty, then $F_{\mathcal{C}'}$ coincides with F. Given a belief $\tilde{F}$, we denote the value of portfolio P by $V(P, \tilde{F})$ and the optimal k-portfolio by $P^*(k, \tilde{F})$; these expressions omit the utility assessment U as that is held constant.

Consider the scenario in which all that Ann learns is that she would be rejected by all of the colleges in $\mathcal{C}'$ (were she to apply). Her optimal k-portfolio would then solve

$$V\left(P^*(k, F_{\mathcal{C}'}), F_{\mathcal{C}'}\right) := \max_{j \in \mathcal{C}} \left\{ \underbrace{(1 - F_{\mathcal{C}'}(\tau_j))u_j}_{\text{Accepted by College } j} + F_{\mathcal{C}'}(\tau_j) \underbrace{V\left(P^*(k-1, F_{\mathcal{C}' \cup \{j\}}), F_{\mathcal{C}' \cup \{j\}}\right)}_{\text{Continuation after rejection from College } j} \right\}. \tag{11}$$

This finds the optimal k-portfolio using a two-stage process in which Ann optimally chooses a college to be the most selective in her k-portfolio—College j above—anticipating that if she is rejected, the

[18] As utility assessments and beliefs are fixed, we suppress those arguments.

[19] For a given k, the running time to find the optimal k-portfolio is kn^2. For illustrative purposes, suppose the applicant were to apply to four colleges from a set of eighty, as in Fu et al. (2022). There are over 1.5 million four-college portfolios; of these, the algorithm here compares $\approx 25,000$ portfolios, and the faster one in the online appendix compares $\approx 2,240$.

Let $\mathcal{C}^\dagger := \mathcal{C} \cup \{0\}$, where 0 is a fictitious college that rejects all applications (i.e., $\tau_0 = 1$).

Step 1. For each $i \in \mathcal{C}^\dagger$, find the optimal single-college portfolio following rejection from i, i.e., $P^*(1, U, F_{\{i\}})$. This is the *optimal single-college continuation* following i, denoted by $C(i, 1)$.

Step 2. For each $i \in \mathcal{C}^\dagger$, find the optimal two-college portfolio following rejection from i. By Equation (12), this solves for

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

This finds the optimal College j anticipating that if j also rejects, one follows with continuation $C(j, 1)$ identified in Step 1. The *optimal two-college continuation* following i is $C(i, 2) := \{j^*\} \cup C(j^*, 1)$.

...

Step k. If $k \leqslant n$: For each $i \in \mathcal{C}^\dagger$, find the optimal k-college portfolio following rejection from i,

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

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

If $k = n$: Terminate.

Output: This algorithm outputs the optimal k-college portfolio, $C(0, k)$, for every $k \leqslant n$.

Figure 6: An algorithm that achieves the optimal portfolio in $O(n^3)$ steps.

continuation is the optimal $(k-1)$-portfolio for her posterior beliefs following rejection, $F_{\mathcal{C}' \cup \{j\}}$. This formulation treats simultaneous search as a dynamic programming problem in which the relevant "history" is the set of colleges that have rejected Ann.

Equation (11) could be subject to the curse of dimensionality in which the continuation problem being solved depends on $\mathcal{C}'$, leading to 2^n feasible histories. But given the common-score formulation, once Ann knows that she is being rejected by all colleges in $\mathcal{C}'$, she will apply only to colleges less selective than the least selective college in $\mathcal{C}'$. It then follows that $F_{\mathcal{C}' \cup \{j\}} = F_{\{j\}}$. Hence, we can simplify the Bellman equation to

$$V\left(P^*(k, F_{\mathcal{C}'}), F_{\mathcal{C}'}\right) := \max_{j \in \mathcal{C} \text{ s.t. } \{j\} \preccurlyeq_A \mathcal{C}'} \left\{ (1 - F_{\mathcal{C}'}(\tau_j))u_j + F_{\mathcal{C}'}(\tau_j) V\left(P^*(k-1, F_{\{i\}}), F_{\{j\}}\right) \right\}. \tag{12}$$

The continuation problem now depends exclusively on the College j that is chosen here and not on $\mathcal{C}'$. Therefore, it suffices to solve only continuation problems following rejection from single-college portfolios. As this substantially reduces the number of histories that need to be considered, the optimal portfolio problem can be solved efficiently.

Figure 6 describes an algorithm that does so in about n^3 steps. Invoking Equation (12), it recursively computes the optimal portfolio of k colleges following every rejection history (including the empty set), beginning with $k = 1$. The algorithm has n steps, and at each step, it performs at most n comparisons for $(n+1)$ options, leading to a total number of computation steps that is $O(n^3)$. These computations identify the optimal k-portfolio for each k. One can then compare these portfolios and optimally choose k, accounting for costs, without increasing the running time.

We generalize this algorithm in the online appendix. First, we use the bad-news effect (Theorem 1) to reduce the running time to $O(n^2 \log n)$. Second, we show that the algorithm can accommodate "tier constraints" that compel an applicant to apply to a certain number of schools in each tier. Third, we find that a virtually identical algorithm obtains the optimal portfolio in the independent-admissions framework. Although the Marginal Improvement Algorithm of Chade and Smith (2006) is more efficient if costs depend only on the number of colleges, the algorithm we develop can also accommodate tier constraints. Hence, it potentially expands the scope of their study. Moreover, it illustrates that virtually the same algorithm works for both the independent-admissions and common-score models.

5 Extensions

5.1 Uncertainty About Thresholds

The baseline analysis presumes that Ann is uncertain about her standing but knows perfectly the thresholds used by colleges. In practice, applicants are likely uncertain about both.[20] Continuity implies that the optimal portfolio remains strictly optimal for small perturbations. Herein, we document that a stronger property holds: 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.

To this end, 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 $(\tilde{\boldsymbol{\tau}}_{\boldsymbol{1}}, \ldots, \tilde{\boldsymbol{\tau}}_{\boldsymbol{n}})$ 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 $F(\tau_i^{CE}) = \mathbb{E}[F(\tilde{\boldsymbol{\tau}}_{\boldsymbol{i}})]$. 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.[21]

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 (2022) document that the relative selectivity of schools is extremely stable in the context of school admissions in Kenya and Ghana, respectively.

Theorem 5. *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. Using the formulation in Equation (6) on p.

[20] The UK is an exception insofar as provisional acceptance offers specify a threshold that the A-level must clear.

[21] A certainty equivalent exists as F is continuous and increasing on its interval support.

10, observe that the expected value of a portfolio P with uncertain thresholds is

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

The expression on the right-hand side coincides with the expected value of a portfolio P under belief F and using the (deterministic) certainty-equivalent threshold for each school.[22]

5.2 College-Specific Scores with a Common Component

Here, we posit a model with a flexible correlation structure to bridge the common- and independent-score approaches. The setting involves a college obtaining a college-specific score that is a weighted mix of a common component and an idiosyncratic college-specific term, analogous to the correlation structures in Avery and Levin (2010), Che and Koh (2016) and Kapor (2020).

In this setting, it is essential to allow colleges to have replicas as an applicant may like to apply to multiple replicas of the same college (since rejection at one replica need not imply rejection at the other). Thus, there are n college types $1, 2, \ldots, n$, and there are R colleges of each type. The set of colleges $\mathcal{C}$ comprises these nR colleges and a generic college is denoted by c. Each college c obtains its own score $s_c := \rho s + \left(\sqrt{1-\rho^2}\right)\varepsilon_c$, where s is a common component that affects all schools' assessments, ε_c is a college-specific term, and $\rho \in [0,1]$ denotes the weight on the common score. For analytical tractability, we assume that ε_c's and s are independent standard normal variables; our specification guarantees that s_c is also standard normal. Each college has a threshold $\tau_c \in \mathbb{R}$. Colleges of the same type have identical score thresholds and offer the same utility. We refer to colleges of type i as copies or replicas of College i and maintain the assumptions that lower-indexed types are more desirable and more selective; that is, $u_i > u_j$ and $\tau_i > \tau_j$ if $i < j$.

Observe that setting $\rho = 1$ reduces to our baseline model whereas setting $\rho = 0$ reduces to the independent-admissions model of Chade and Smith (2006). At intermediate values of ρ, rejection by a college is bad news for Ann's prospects at higher-ranked colleges but she nevertheless has a chance. To interpret this setting, the common component may be the student's caliber and the school-specific term ε_c is a college-specific match quality.

Although a general analysis of this setup is beyond our scope, we obtain results that speak to the implications of correlation. First, so long as there is some weight on the common component, large portfolios include colleges of both the most selective and the least selective type. By contrast, if there is no weight on the common component, large portfolios almost exclusively target the most selective type. Second, in the context of two-college portfolios, increasing the weight on the common component never pushes an applicant to a less dispersed portfolio.

Dispersion in Large Portfolios: We first study large portfolios. Suppose that there are infinitely many colleges of each type. As a benchmark, consider the case of independent admissions.

[22] 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.

Theorem 6. *Suppose colleges place no weight on the common component ($\rho = 0$). Then large portfolios target almost exclusively only the most selective colleges: 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.*

Theorem 6 reproduces the "upward-expansion" conclusion of Chade and Smith (2006). The logic is that once Ann has applied to sufficiently many other colleges, she effectively has a high outside option. This makes her risk loving and willing to gamble on only the most selective colleges.

This is a knife-edge conclusion. If colleges place even slight weight on the common component, Ann has an incentive to hedge by applying to safety schools. To that end, let College $\underline{m}$ be the lowest-ranked college whose utility exceeds that of Ann's outside option u_o.

Theorem 7. *Suppose colleges place some weight on the common component ($\rho > 0$). Then large portfolios must feature arbitrarily many copies of the most and least selective colleges: for every integer M, there exists k such that for all $k' > k$, the optimal k'-portfolio has at least M copies of both Colleges 1 and $\underline{m}$.*

The logic for why Ann applies to the best college is already in Theorem 6, so we focus on why she also targets the least selective rationalizable college. If she targets only replicas of College 1, the marginal benefit of an additional application obtains only if all those applications fail. Such rampant failure is bad news about the common component of her scores as it would be unlikely to stem completely from idiosyncratic college-specific terms. This force makes her defensive and leads her to apply to at least some less selective colleges. Even for low weights on the common component, this bad-news effect accumulates so that she applies to arbitrarily many copies of College $\underline{m}$.

Thus, the motive to hedge and apply to safety schools generalizes beyond the baseline model of Section 3. If colleges value both a common component—which could reflect recommendation letters, application essays, and one's overall standing in the pool—as well as match-specific qualities, an applicant may seek out colleges that have low selectivity even if they result in low utility.

Increasing Weight on the Common Component: We now study how the optimal portfolio changes with changes in ρ. Our specification ensures that changing ρ affects only the correlation in admissions decisions across colleges and not the (marginal) acceptance rate of any single college. So as to make this comparison meaningful, we focus on portfolios of two colleges each.[23]

Increasing the weight on the common component can increase the dispersion of a portfolio because of the forces described in the baseline model. However, it need not always be the case. For example, increasing ρ could lead to a more aggressive first choice, which dampens the bad-news effect and thereby induces a more aggressive backup. Symmetrically, it could also lead to a less aggressive backup, which makes Ann less risk loving and opt for a less aggressive first choice.

However, what *cannot happen* is that increases in ρ lead the optimal portfolio to become *less dispersed*. Say that a portfolio of two colleges $\{i', j'\}$ is less dispersed than $\{i, j\}$ if $i < i' < j' < j$.[24]

Theorem 8. *If $\rho' > \rho$, the optimal portfolio with weight ρ' is not less dispersed than that with weight ρ.*

[23] Apart from tractability, we restrict our attention to two colleges because with $k \geqslant 3$ colleges, there is no straightforward way to discuss one k-college portfolio as more or less dispersed than another.

[24] To simplify notation, here we abstract from copies, which allows us to identify each college with its index.

Here is the central idea. Suppose to the contrary that the optimal portfolio at weight ρ', $\{i', j'\}$, were less dispersed than that, $\{i, j\}$, at weight ρ. Since College i' is less selective than College i, being rejected by i' is worse news about the common component. Thus, even if the weight on the common component were ρ, the applicant would already be inclined to choose a less selective backup. Increasing the weight on the common component to ρ' amplifies this motive because the bad news then has even more significant implications for her prospects. While the underlying logic is intuitive, the proof is intricate because the information conveyed by a rejection from the first college takes the form of an inequality. Hence, the posterior belief following that rejection has no simple closed form. We sidestep this issue by using approaches from Plackett (1954) that represent the change in distribution from changes in ρ through a heat equation.

5.3 Informationally Directed Search

In this section, we study a sequential variant with the common-score structure and compare it to the portfolio problem. We use the same setup as in Section 3, but for interpretation, we call the colleges $\{1, \ldots, n\}$ "projects." Each project i has a threshold τ_i, and Ann's attempt at that project succeeds only if her score s exceeds the threshold. As before, if $i < j$, then project i is harder ($\tau_i > \tau_j$) and more rewarding ($u_i > u_j$). We interpret s as Ann's ability and projects as varying in difficulty such that if she fails at one project, she is bound to fail at any harder project. In contrast to the college portfolio problem, Ann tries projects sequentially, observing whether an attempt succeeds before deciding which project to try next. Of the projects where her attempts succeed, Ann can take to completion and "consume" only one. But search is with recall as she can go back to a project she attempted before. Finally, if all of her attempts fail, she obtains her outside option u_o.

In this problem, past outcomes are informative about future attempts and one need not stop with a success. Success at one project instead spurs Ann towards more ambitious projects whereas failure lowers her ambition. Because Ann is forward-looking, she takes this into account in choosing where to start. To see this logic, suppose that Ann faces a cost function in which she can attempt up to k projects for free but no more. We call the optimal strategy in this setting her optimal k-strategy and compare it to the optimal portfolios of simultaneous search.

Theorem 9. *The optimal k-strategy achieves the same value as the optimal* $(2^k - 1)$*-portfolio.*

Theorem 9 shows that sequential search offers significant gains: searching k projects sequentially results in the same payoff as simultaneous search with $(2^k - 1)$ projects. We illustrate why in Figure 7. Suppose Ann can try at most three projects. She first attempts the median project of the optimal (simultaneous) seven-portfolio. If that succeeds, she aims higher and attempts the second-best project in that portfolio, whereas if it fails, she shifts to the second-worst project in that portfolio. Further successes and failures direct her search locally so that she achieves the same value as she would with the optimal seven-portfolio.

This strategy exemplifies features of informationally directed search. The optimal strategy would be qualitatively different were the success at each project independent of those at other projects. In that case, nothing would be learned from past attempts. Chade and Smith (2006) study that Weitzman problem with a constant marginal cost. They show that the optimal strategy stops at the first

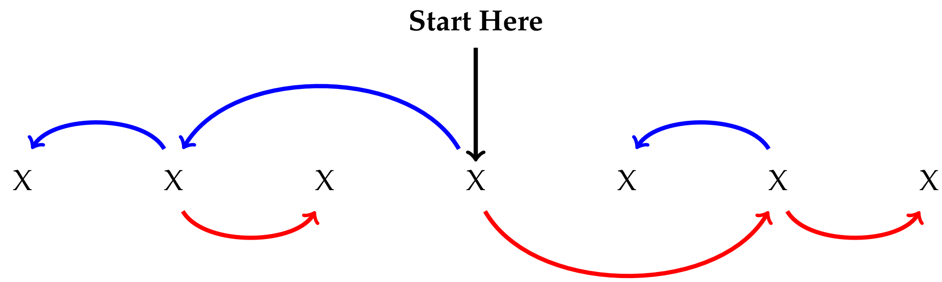

Figure 7: The optimal 3-strategy. Projects in the optimal 7-portfolio are ordered in descending order of difficulty. Blue and red arrows show what Ann tries following success and failure respectively.

success and is more aggressive and broader than the optimal simultaneous portfolio. In the online appendix, we show that these properties may fail here because our searcher has the additional motive to start with the project that best guides her future search. We have not seen this particular process of informationally directed search studied in prior work; it would be interesting to obtain more general predictions and compare those to precedents such as Callander (2011) and Urgun and Yariv (2023).

References

AJAYI, K. (2022): "School choice and educational mobility: Lessons from secondary school applications in Ghana," *Journal of Human Resources*.

AJAYI, K. AND M. SIDIBE (2022): "School Shoice under Imperfect Information," Working paper.

AJAYI, K. F., W. H. FRIEDMAN, AND A. M. LUCAS (2020): "When information is not enough: Evidence from a centralized school choice system," Tech. rep., National Bureau of Economic Research.

AKBARPOUR, M., A. KAPOR, C. NEILSON, W. VAN DIJK, AND S. ZIMMERMAN (2022): "Centralized School choice with unequal outside options," *Journal of Public Economics*, 210, 104644.

AVERY, C. AND J. LEVIN (2010): "Early admissions at selective colleges," *American Economic Review*, 100, 2125–2156.

BROECKE, S. (2012): "University selectivity and earnings: Evidence from UK data on applications and admissions to university," *Economics of Education Review*, 31, 96–107.

CALLANDER, S. (2011): "Searching and learning by trial and error," *American Economic Review*, 101, 2277–2308.

CALSAMIGLIA, C., C. FU, AND M. GÜELL (2020): "Structural estimation of a model of school choices: The Boston mechanism vs. its alternatives," *Journal of Political Economy*, 128, 642–680.

CALSAMIGLIA, C., F. MARTÍNEZ-MORA, AND A. MIRALLES (2021): "School Choice Design, Risk Aversion and Cardinal Segregation," *The Economic Journal*, 131, 1081–1104.

CHADE, H., J. EECKHOUT, AND L. SMITH (2017): "Sorting through Search and Matching Models in Economics," *Journal of Economic Literature*, 55, 493–544.

CHADE, H., G. LEWIS, AND L. SMITH (2014): "Student portfolios and the college admissions problem," *Review of Economic Studies*, 81, 971–1002.

CHADE, H. AND L. SMITH (2006): "Simultaneous search," *Econometrica*, 74, 1293–1307.

CHE, Y.-K. AND Y. KOH (2016): "Decentralized college admissions," *Journal of Political Economy*, 124, 1295–1338.

FU, C. (2014): "Equilibrium tuition, applications, admissions, and enrollment in the college market," *Journal of Political Economy*, 122, 225–281.

FU, C., J. GUO, A. J. SMITH, AND A. SORENSEN (2022): “Students’ heterogeneous preferences and the uneven spatial distribution of colleges,” *Journal of Monetary Economics*, 129, 49–64.

GALENIANOS, M. AND P. KIRCHER (2009): “Directed search with multiple job applications,” *Journal of Economic Theory*, 144, 445–471.

HOWELL, J. S. (2010): “Assessing the impact of eliminating affirmative action in higher education,” *Journal of Labor Economics*, 28, 113–166.

HOXBY, C. AND C. AVERY (2013): “The Missing" One-Offs": The Hidden Supply of High-Achieving, Low-Income Students,” *Brookings Papers on Economic Activity*, 1–50.

IDOUX, C. (2023): “Integrating New York City Schools: The Role of Admission Criteria and Family Preferences,” Working paper.

KAPOR, A. (2020): “Distributional effects of race-blind affirmative action,” Working Paper.

KIRCHER, P. (2009): “Efficiency of simultaneous search,” *Journal of Political Economy*, 117, 861–913.

LUCAS, A. M. AND I. M. MBITI (2012): “The determinants and consequences of school choice errors in Kenya,” *American Economic Review*, 102, 283–288.

OLSZEWSKI, W. AND R. VOHRA (2016): “Simultaneous selection,” *Discrete Applied Mathematics*, 200, 161–169.

PALLAIS, A. (2015): “Small differences that matter: Mistakes in applying to college,” *Journal of Labor Economics*, 33, 493–520.

PLACKETT, R. L. (1954): “A Reduction Formula for Normal Multivariate Integrals,” *Biometrika*, 41, 351–360.

SHORRER, R. I. (2019): “Simultaneous Search: Beyond Independent Success,” *EC’19: Proceedings of the 2019 ACM Conference on Economics and Computation*.

URGUN, C. AND L. YARIV (2023): “Contiguous Search: Exploration and Ambition on Uncharted Terrain,” Working Paper.

WALTERS, C. R. (2018): “The demand for effective charter schools,” *Journal of Political Economy*, 126, 2179–2223.

WEITZMAN, M. L. (1979): “Optimal search for the best alternative,” *Econometrica*, 47, 641–654.

A Main Appendix

A.1 Proof of Theorem 1 on p. 13

Proof. The proof of this argument involves three cases, and a double induction argument.

Case 1: $|\mathbf{P}^*(\mathbf{k}, \mathbf{U}, \mathbf{G})| < \mathbf{k}$.

Let j^* be the least selective school on $P^*(k, U, G)$. Our assumptions guarantee that $\mu(j^*, G) > 0$. Since $G \succcurlyeq_{LR} H$ this implies that for any $i < j^*$ such that $\mu(i, G) = 0$ we also have that $\mu(i, H) = 0$. Furthermore, for any $i, j \in P^*(k, U, G)$ such that $i < j$, if $\mu(j, H) = 0$ then $\mu(i, H) = 0$ (by the definition of LR-dominance since $\mu(i, G) > 0$ and $\mu(j, G) > 0$ as they are on the portfolio). Therefore, if $|P^*(k, U, H)| = k$ then it consists of a subset of the colleges in $P^*(k, U, G)$ and some colleges lower ranked than j^*. Otherwise, $|P^*(k, U, H)| < k$ in which case it consists of all colleges in $P^*(k, U, G)$ that are less selective than some i^* in addition to some colleges less selective than j^*. In either case, $P^*(k, U, G) \succcurlyeq_{LR} P^*(k, U, H)$.

Case 2: $|\mathbf{P}^*(\mathbf{k}, \mathbf{U}, \mathbf{H})| < |\mathbf{P}^*(\mathbf{k}, \mathbf{U}, \mathbf{G})| = \mathbf{k}$.

Let i^* be the most aggressive college on $P^*(k, U, H)$. Our assumptions guarantee that $\mu(i^*, H) > 0$. Since $G \succcurlyeq_{LR} H$ this implies that for any $i^* < j$ such that $\mu(j, H) = 0$ we also have that $\mu(j, G) = 0$. Furthermore, for any $i, j \in P^*(k, U, H)$ such that $i < j$, if $\mu(i, G) = 0$ then $\mu(j, G) = 0$ (by the definition of LR-dominance since $\mu(i, H) > 0$ and $\mu(j, H) > 0$ as they are on the portfolio). Therefore, since $|P^*(k, U, G)| = k$, it consists of a subset of the colleges in $P^*(k, U, H)$ and some colleges more selective than i^*, and so $P^*(k, U, G) \succcurlyeq_{LR} P^*(k, U, H)$.

Case 3: $|\mathbf{P}^*(\mathbf{k}, \mathbf{U}, \mathbf{H})| = |\mathbf{P}^*(\mathbf{k}, \mathbf{U}, \mathbf{G})| = \mathbf{k}$

We prove that $P^*(k, U, G) \succcurlyeq_{A} P^*(k, U, H)$ by induction on k.

Base Step: $(k = 1)$

If $P\left(k, U, H\right) = \{i\}$ then because i is the uniquely optimal single-college portfolio, it follows that for every College j,

$$V(\{i\}, U, H) - u_o = \left(1 - H\left(\tau_i\right)\right)\,\left(u_i - u_o\right) \geqslant \left(1 - H\left(\tau_j\right)\right)\left(u_j - u_o\right) = V(\{j\}, U, H) - u_o. \tag{13}$$

Consider a College j that is lower ranked than i. Therefore, $j > i$. Observe that

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

where the first line follows by definition, the second follows from $G \succcurlyeq_{LR} H$, and the third follows by definition. Multiplying (13) with the inequality above yields

$$\left(1-G\left(\tau_i\right)\right)\left(1-H\left(\tau_j\right)\right)\left(1-H\left(\tau_i\right)\right)\left(u_i-u_o\right) \geqslant \left(1-H\left(\tau_i\right)\right)\left(1-G\left(\tau_j\right)\right)\left(1-H\left(\tau_j\right)\right)\left(u_j-u_o\right).$$

We note that $\left(1-H\left(\tau_i\right)\right) > 0$ (otherwise the empty portfolio is optimal, contradicting our assumptions, and that $\left(1-H\left(\tau_j\right)\right) \geqslant \left(1-H\left(\tau_i\right)\right)$, since $\tau_j < \tau_i$. Hence, we can divide both sides of the inequality by the positive term $\left(1-H\left(\tau_j\right)\right)\left(1-H\left(\tau_i\right)\right)$ and obtain

$$\left(1-G\left(\tau_i\right)\right)\left(u_i-u_o\right) \geqslant \left(1-G\left(\tau_j\right)\right)\left(u_j-u_o\right),$$

and therefore $P^*(k,U,G) \neq \{j\}$ for any $j > i$. Hence, $P^*(k,U,G)$ must be at least as aggressive as $\{i\}$.

<u>Inductive Step: $(k > 1)$</u>

Suppose that the statement holds for all portfolio sizes strictly smaller than k. We show that this implies that it also holds for portfolios of size k.

We begin by noting that if $P^{*(1)}\left(k,U,G\right) \leqslant P^{*(1)}\left(k,U,H\right)$ then the result follows from the inductive hypothesis. This follows from two observations. First,

$$P^*\left(k,U,G\right) = \left\{P^{*(1)}\left(k,U,G\right)\right\} \bigcup P^*\left(k-1,\bar{u},G\left(\cdot \mid s < \tau_{P^{*(1)}(k,U,G)}\right)\right).$$

The reason is that a student applying to P attends $P^{(1)}$ whenever accepted to $P^{(1)}$, and so the rest of her portfolio must be optimal conditional on being rejected from $P^{(1)}$. More generally,

$$P^*\left(k,U,G\right) = \left\{P^{*(1)}\left(k,U,G\right),\ldots,P^{*(j)}\left(k,U,G\right)\right\} \bigcup P^*\left(k-j,\bar{u},G\left(\cdot \mid s < \tau_{P^{*(j)}(k,U,G)}\right)\right).$$

Second, if $\tau \geqslant \tau'$ then $G\left(\cdot \mid s < \tau\right) \succcurlyeq_{LR} H\left(\cdot \mid s < \tau'\right)$. This follow by the transitivity of $\succcurlyeq_{LR}$, since $G\left(\cdot \mid s < \tau\right) \succcurlyeq_{LR} G\left(\cdot \mid s < \tau'\right)$ and $G\left(\cdot \mid s < \tau'\right) \succcurlyeq_{LR} H\left(\cdot \mid s < \tau'\right)$.

Therefore, it suffices to show that $P^{*(1)}\left(k,U,G\right) \leqslant P^{*(1)}\left(k,U,H\right)$, which is what we do in the remainder of this proof.

Suppose otherwise. Let $m \geqslant 1$ be the maximal index with $P^{*(m)}\left(k,U,G\right) > P^{*(m)}\left(k,U,H\right)$. By the inductive hypothesis and the two observations above, $P^{*(j)}\left(k,U,G\right) \leqslant P^{*(j)}\left(k,U,H\right)$ for all $j > m$ (this condition may be vacuous if $m = k$), and $P^{*(j)}\left(k,U,G\right) > P^{*(j)}\left(k,U,H\right)$ for all $j \leqslant m$.

We now create a chain of portfolios $Q_0,\ldots,Q_m$. To simplify notation, for a general portfolio Q, we sometimes write Q^i instead of $Q^{(i)}$ to denote the i^{th} ranked college. To simplify notation, we denote $P := P^*(k,U,G)$ and $R := P^*(k,U,H)$. Let $Q_0,Q_1,\ldots,Q_m$ be portfolios such that $Q_i = \left\{R^{(1)},\ldots,R^{(i)},P^{(i+1)},\ldots,P^{(k)}\right\}$. In other words, Q_i selects the top i colleges from portfolio R and the remaining $k-i$ colleges from P.[25] Observe that for each $j \leqslant i \leqslant m$ we have $Q_i^{(j)} = R^{(j)}$ and $Q_j \succcurlyeq_A Q_{j-1}$. It also follows that $Q_m \succcurlyeq_A P$ and $Q_m \succcurlyeq_A R$.

[25] Since $|P| = |R| = k$, the definition of m guarantees that for each $i \leqslant m$ we have $|Q_i| = k$.

For $D \in \{G, H\}$ and $j \leqslant m$ we have

$$V(Q_j, U, D) - V(Q_0, U, D) = \sum_{i=1}^{j} V(Q_i, U, D) - V(Q_{i-1}, U, D).$$

Next, we note that for $i \leqslant m$,

$$V(Q_i, U, D) - V(Q_{i-1}, U, D) = (D(\tau_{R^{i-1}}) - D(\tau_{R^i}))\,(u_{R^i} - u_{P^i}) - (D(\tau_{R^i}) - D(\tau_{P^i}))\,(u_{P^i} - u_{P^{i+1}}).$$[26]

Since $Q_0 = P := P^*(k, U, G)$, we have

$$V(Q_j, U, G) - V(Q_0, U, G) \leqslant 0. \tag{14}$$

We prove by induction on j that this implies

$$V(Q_j, U, H) - V(Q_0, U, H) \leqslant 0.$$

Base Step ($j = 1$): Observe that

$$0 \leqslant V(Q_1, U, G) - V(Q_0, U, G) = (1 - G(\tau_{R^1}))\,(u_{R^1} - u_{P^1}) - (G(\tau_{R^1}) - G(\tau_{P^1}))\,(u_{P^1} - u_{P^2}),$$

where the inequality is (14) and the equality is computation. Therefore,

$$(1 - G(\tau_{R^1}))\,(u_{R^1} - u_{P^1}) \leqslant (G(\tau_{R^1}) - G(\tau_{P^1}))\,(u_{P^1} - u_{P^2}).$$

Additionally, since $G \geqslant_{LR} H$

$$(G(\tau_{R^1}) - G(\tau_{P^1}))\,(1 - H(\tau_{R^1})) \leqslant (1 - G(\tau_{R^1}))\,(H(\tau_{R^1}) - H(\tau_{P^1}))\,.$$

Since all terms in the two inequalities above are nonnegative, we can multiply them to obtain

$$\begin{aligned}&(G(\tau_{R^1}) - G(\tau_{P^1}))\,(1 - H(\tau_{R^1}))\,(1 - G(\tau_{R^1}))\,(u_{R^1} - u_{P^1}) \leqslant \\ &(1 - G(\tau_{R^1}))\,(H(\tau_{R^1}) - H(\tau_{P^1}))\,(G(\tau_{R^1}) - G(\tau_{P^1}))\,(u_{P^1} - u_{P^2}).\end{aligned}$$

Since R^1 is in $P^*(k, u, H)$ we have $(1 - H(\tau_{R^1})) > 0$ and so $(1 - G(\tau_{R^1})) > 0$ since $G >_{LR} H$. Additionally, $(G(\tau_{R^1}) - G(\tau_{P^1})) > 0$, as otherwise $V(Q_1, U, G) > V(P, U, G)$, contradicting the optimality of P. We can therefore divide both sides of the inequality by $(1 - G(\tau_{R^1}))\,(G(\tau_{R^1}) - G(\tau_{P^1}))$ and obtain

$$(1 - H(\tau_{R^1}))\,(u_{R^1} - u_{P^1}) \leqslant (H(\tau_{R^1}) - H(\tau_{P^1}))\,(u_{P^1} - u_{P^2})$$

which implies

$$V(Q_1, U, H) - V(Q_0, U, H) = (1 - H(\tau_{R^1}))\,(u_{R^1} - u_{P^1}) - (H(\tau_{R^1}) - H(\tau_{P^1})) \cdot (u_{P^1} - u_{P^2}) \leqslant 0.$$

[26]We use the notational convention that for a portfolio Q of size k, $u_{Q^{k+1}} = u_o$ and $\tau_{Q^0} = 1$.

Inductive Step ($j > 1$): We assume that the statement holds for all $j' < j$. The following notation will be useful. For $D \in \{G, H\}$ and $l \leqslant m$, denote $W_D^l := (D(\tau_{R^{l-1}}) - D(\tau_{R^l}))\,(u_{R^l} - u_{P^l})$ and $L_D^l := (D(\tau_{R^l}) - D(\tau_{P^l}))\,(u_{P^l} - u_{P^{l+1}})$. W_D^l (respectively L_D^l) represent the gains (losses) for an agent whose beliefs are given by D from the changing her portfolio from Q_{l-1} to Q_l (she gains if she ends up attending the $Q_l^{(l)} = R^l$ and she loses if her score suffices for $Q_{l-1}^{(l)} = P^l$ but not for $Q_l^{(l)} = R^l$). Additionally, denote $q_{R^l} := \frac{\mu(R^l,H)}{\mu(R^l,G)}$.[27]

Because Q_0 is optimal under distribution G, we know that $\sum_{l=1}^{j} L_G^l \geqslant \sum_{l=1}^{j} W_G^l$ for all $j \leqslant m$. Observe that

$$\begin{aligned}\sum_{l=1}^{j} L_H^l = L_H^j + \sum_{l=1}^{j-1}\left(L_H^l - W_H^l + W_H^l\right) &= L_H^j + \sum_{l=1}^{j-1} W_H^l + \sum_{l=1}^{j-1}\left(L_H^l - W_H^l\right)\\ &\geqslant q_{R^j} L_G^j + \sum_{l=1}^{j-1} W_H^l + \sum_{l=1}^{j-1} q_{R^l}\left(L_G^l - W_G^l\right),\end{aligned}$$

where the inequality follows from $W_H^l / W_G^l = (H(\tau_{R^{l-1}}) - H(\tau_{R^l}))\,/\,(G(\tau_{R^{l-1}}) - G(\tau_{R^l}))$ being bounded above by q_{R^l}, and $L_H^l / L_G^l = (H(\tau_{R^l}) - H(\tau_{P^l}))\,/\,(G(\tau_{R^l}) - G(\tau_{P^l}))$ being bounded below by q_{R^l} for all $l \leqslant m$. By rewriting the RHS of the above inequality, it then follows that

$$\begin{aligned}\sum_{l=1}^{j} L_H^l &\geqslant q_{R^j} L_G^j + \sum_{l=1}^{j-1} W_H^l + \sum_{l=1}^{j-1} q_{R^j}\left(L_G^l - W_G^l\right) + \sum_{l=1}^{j-1}\sum_{b=l}^{j-1}(q_{R^b} - q_{R^{b+1}})\left(L_G^l - W_G^l\right)\\ &= q_{R^j} L_G^j + \sum_{l=1}^{j-1} W_H^l + \sum_{l=1}^{j-1} q_{R^j}\left(L_G^l - W_G^l\right) + \sum_{b=1}^{j-1}(q_{R^b} - q_{R^{b+1}})\sum_{l=1}^{b}\left(L_G^l - W_G^l\right)\\ &= q_{R^j} L_G^j + \sum_{l=1}^{j-1} W_H^l + \sum_{l=1}^{j-1} q_{R^j}\left(L_G^l - W_G^l\right) + \sum_{b=1}^{j-1}(q_{R^b} - q_{R^{b+1}})\,(V(Q_0, U, G) - V(Q_b, U, G))\\ &\geqslant q_{R^j} L_G^j + \sum_{l=1}^{j-1} W_H^l + \sum_{l=1}^{j-1} q_{R^j}\left(L_G^l - W_G^l\right) = \sum_{l=1}^{j-1} W_H^l + \sum_{l=1}^{j} q_{R^j}\left(L_G^l - W_G^l\right) + q_{R^j} W_G^j\\ &= q_{R^j}\left(V(Q_0, U, G) - V(Q_j, U, G)\right) + \sum_{l=1}^{j-1} W_H^l + q_{R^j} W_G^j \geqslant \sum_{l=1}^{j-1} W_H^l + q_{R^j} W_G^j \geqslant \sum_{l=1}^{j} W_H^l,\end{aligned}$$

where the second inequality follows from q_{R^b} being increasing in b, and $V(Q_0, U, G) \geqslant V(Q_b, U, G)$. The third inequality uses the optimality of $V(Q_0, U, G)$ again, and that q_{R^j} is nonnegative. The fourth and final inequality follows from W_H^l / W_G^l being bounded above by q_{R^l}.

To complete the proof, we note that

$$\begin{aligned}&V\left(\{P^1, P^2, \ldots, P^m, R^{m+1}, \ldots, R^k\}, U, H\right) - V(R, U, H) =\\ &\sum_{l=1}^{m}\left(L_H^l - W_H^l\right) + (H(\tau_{R^m}) - H(\tau_{P^m}))\,(u_{P^{m+1}} - u_{R^{m+1}}) > 0\end{aligned}$$

[27] This expression is well defined as $\mu(R^l, G) > 0$ since otherwise $R = P^*(k, U, G)$ is not optimal (as it would be beneficial for the agent to replace R^l with a more aggressive option) or not minimal (the agent can drop R^l from her portfolio).

contradicting the optimality of $P^*(k,U,H) := R$. The expression is nonnegative since we have shown above that the first term is nonnegative, and the second expression is the product of two nonnegative terms ($u_{P^{m+1}} \geqslant u_{R^{m+1}}$ by the definition of m, and $H(\tau_{R^m}) \geqslant H(\tau_{P^m})$ since $u_{R^m} > u_{P^m}$ and so $\tau_{R^m} > \tau_{P^m}$). Since ties are ruled out, it must be strictly positive. □

A.2 Proof of Theorem 2 on p. 14

Proof. Our argument proceeds in multiple steps.

Step 1: **There is no loss of generality in normalizing utilities.**

For utility assessment $U := (u_o; u_1, \ldots, u_n)$, we denote its normalization by

$$N[U] := (0; \tilde{u}_1, \ldots, \tilde{u}_n) \text{ where } \tilde{u}_i := \max\left\{\frac{u_i - u_o}{u_1 - u_o}, 0\right\}. \tag{15}$$

We argue that there is no loss of generality in making this normalization. Consider a utility assessment U, distribution G, and a portfolio P. Recall that $P^{(i)}$ denotes the i^{th} best item in the portfolio. Using $G(\tau_{P^0}) = 1$, we can write the value of a portfolio as

$$\begin{aligned}
V(P,U,G) &= \int_0^1 \max_{\{i \in P: s \geqslant \tau_i\}} \max\{u_i, u_o\} dG \\
&= \sum_{i=1}^{|P|} \max\{u_o, u_{P^{(i)}}\} \left(G(\tau_{P^{(i-1)}}) - G(\tau_{P^{(i)}})\right) \\
&= u_o + \sum_{i=1}^{|P|} \max\{0, u_{P^{(i)}} - u_o\} \left(G(\tau_{P^{(i-1)}}) - G(\tau_{P^{(i)}})\right) \\
&= u_o + (u_1 - u_o) \sum_{i=1}^{|P|} \tilde{u}_{P^{(i)}} \left(G(\tau_{P^{(i-1)}}) - G(\tau_{P^{(i)}})\right) \\
&= u_o + (u_1 - u_o) V(P, N[U], G),
\end{aligned}$$

where the first equality is the definition of a portfolio's value (Equation 5 on p. 10), the second equality is calculating the integral, the third equality is algebra, the fourth equality substitutes (15), and the final equality uses the definition of $V(P, N[U], G)$. Hence, normalizing utilities does not affect the relative ranking of the value of portfolios: for portfolios P and P',

$$V(P,U,G) \geqslant V(P',U,G) \Leftrightarrow V(P,N[U],G) \geqslant V(P',N[U],G).$$

The normalization also maintains the same risk-loving order:

$$U' \succcurlyeq_{RL} U \Leftrightarrow N[U'] \succcurlyeq_{RL} N[U].$$

In light of the above, we restrict attention to normalized utility assessments. Note that in a normalized utility assessment, U, the utility of each College i, u_i, is in $[0,1]$, the utility of the best college, u_1, is equal to 1, and the value of the outside option, u_o, equals 0.

Step 2: **We define the transposition.**

In this step, we show how to transpose utilities and probabilities, formalizing the idea of Figure 4. Conceptually, the transposition does two flips. First, for each college, it flips the utility and the *acceptance probability* so that a college with a high utility and low acceptance probability flips into being a college with low utility and high acceptance probability, (ii) to order colleges descending in the order of Ann's preferences, we also flip the order of colleges.

Given a distribution of scores G, let $\lambda(G) = (G(\tau_1), \ldots, G(\tau_n))$ denote a vector of rejection probabilities for the n colleges. Analogously, given a vector $\lambda \in [0,1]^n$ where $\lambda_j \leqslant \lambda_i$ for $j \geqslant i$, let $\pi(\cdot, \lambda)$ be any CDF such that $\pi(\tau_i, \lambda) = \lambda_i$. In other words, $\pi(\cdot, \lambda)$ selects a probability distribution on scores that generates the vector of rejection probabilities λ.[28]

We transpose utilities into probabilities as follows. For a vector $x \in [0,1]^n$, let $T[x]$ be the *transposed vector* $(1 - x_n, \ldots, 1 - x_1)$. Given a normalized utility assessment, $U := (0; u_1, \ldots, u_n)$, let $w(U) := (u_1, \ldots, u_n)$ represent the utility of prizes. For a normalized utility assessment U, we define the *transposed distribution of scores* $G_T(U)$ as the CDF G such that for every score s,

$$G(s) = \pi\left(s, T[w(U)]\right).$$

Therefore, $G_T(U)$ is a score distribution that generates the vector of rejection probabilities $(1-u_n, \ldots, 1-u_1)$.

Analogously, let us transpose probabilities into utilities. Given a distribution of scores G, we say that a utility assessment $U := (0; u_1, \ldots, u_n)$ is its transposition $U_T(G)$ if

$$w(U) = T[\lambda(G)].$$

In other words, in $U_T(G)$, the value of the outside option is 0, and the utility of being accepted by Colleges 1 through n is respectively $1 - G(\tau_n), \ldots, 1 - G(\tau_1)$.

Notice that transposition flips the order of colleges. The transposition maps the i^{th} ranked college, College i, that offers utility u_i and rejection probability $G(\tau_i)$ to the college ranked $(n+1-i)^{th}$ with utility $1 - G(\tau_i)$ and rejection probability $1 - u_i$. Accordingly, we map portfolios in the original problem to those in the transposed problem using the operator $\mathcal{T} : 2^N \to 2^N$ where

$$\mathcal{T}[P] := \{i \in C : n + 1 - i \in P\}.$$

Because transposition flips the order of colleges, it follows that if $P \succcurlyeq_A \tilde{P}$, then $\mathcal{T}[P] \preccurlyeq_A \mathcal{T}[\tilde{P}]$.

We also note that if a utility assessment U and distribution G satisfy our assumptions, so does the transposed model with utility assessment $U_T(G)$ and distribution $G_T(U)$.

Step 3: **Transposition leads to an isomorphic problem.**

[28] It suffices to consider probability distribution functions that are constant on intervals of the form (τ_i, τ_{i+1}).

In this step, we prove that transposing utilities and acceptance probabilities leads to a problem that is isomorphic to the original problem. Specifically, we show that

$$V(P, U, G) = V(\mathcal{T}[P], U_T(G), G_T(U)). \tag{16}$$

Therefore for every k, $P^*(k, U, G) = \mathcal{T}[P^*(k, U_T(G), G_T(U))]$.

Consider a normalized utility profile U and a distribution G. Consider a portfolio P where as usual, $P^{(i)}$ denotes the i^{th} ranked item in the portfolio P. Let $|P| = k$. Observe that we can write the value of a portfolio as

$$V(P, U, G) = u_{P^{(1)}} \underbrace{(1 - G(\tau_{P^{(1)}}))}_{\text{Accepted by } P^{(1)}} + \sum_{i=2}^{k} u_{P^{(i)}} \underbrace{((1 - G(\tau_{P^{(i)}})) - (1 - G(\tau_{P^{(i-1)}})))}_{\text{Accepted by } P^{(i)} \text{ but not by } P^{(i-1)}}. \tag{17}$$

The above expression computes the value of a portfolio based on Ann obtaining the payoff of College $P^{(i)}$ if she is accepted by that college but rejected by every higher ranked college in portfolio P. Rearranging the RHS of the above expression yields

$$\left(\sum_{i=1}^{k-1} (1 - G(\tau_{P^{(i)}}))\,(u_{P^{(i)}} - u_{P^{(i+1)}})\right) + (1 - G(\tau_{P^{(k)}}))u_{P^{(k)}}. \tag{18}$$

Let $\hat{U} := U_T(G)$ be the transposition of the distribution and $\hat{G} := G_T(U)$ be a transposition of the utilities. Finally, let $\hat{P} := \mathcal{T}[P]$ be the "transposed" portfolio. Observe that by construction, $|\hat{P}| = k$, $\hat{u}_{\hat{P}^{(i)}} = 1 - G(\tau_{P^{(k+1-i)}})$, and $\hat{G}(\tau_{\hat{P}^{(i)}}) = 1 - u_{P^{(k+1-i)}}$. These substitutions in (18) yield

$$\hat{u}_{\hat{P}^{(1)}}(1 - \hat{G}(\tau_{\hat{P}^{(1)}})) + \sum_{i=2}^{k} \hat{u}_{\hat{P}^{(i)}} \left((1 - \hat{G}(\tau_{\hat{P}^{(i)}})) - (1 - \hat{G}(\tau_{\hat{P}^{(i-1)}}))\right),$$

which by comparison to (17) is equal to $V(\hat{P}, \hat{U}, \hat{G})$.

Step 4: **Being more risk loving implies bad news in the transposed problem.**[29]

We show that for normalized utility assessments U and U',

$$U' \succcurlyeq_{RL} U \Rightarrow G_T(U') \preccurlyeq_{LR} G_T(U).$$

Observe that $U' \succcurlyeq_{RL} U$ implies that there exists a convex nondecreasing function $\nu : \mathbb{R} \to \mathbb{R}$ such that $\nu(0) = 0$, $\nu(1) = 1$, and for every $i \in \{1, \ldots, N\}$, $u'_i = \nu(u_i)$. We argue that for every i and $j > i$,

$$(u'_{i+1} - u'_i)(u_{j+1} - u_j) \geqslant (u_{i+1} - u_i)(u'_{j+1} - u'_j). \tag{19}$$

To see why the above inequality holds, consider the following two cases. First, if $u_{i+1} = u_i$ then the right hand side is equal to 0 and since both sides are nonnegative we are done. Second, if $u_{j+1} = u_j$

[29]The converse is also true, but we do not use that direction.

then $u'_{j+1} = u'_j$ and both sides are equal to 0. Otherwise, (19) can be rewritten as

$$\frac{u'_{i+1} - u'_i}{u_{i+1} - u_i} \geqslant \frac{u'_{j+1} - u'_j}{u_{j+1} - u_j}$$

or equivalently

$$\frac{\nu\left(u_{i+1}\right) - \nu\left(u_i\right)}{u_{i+1} - u_i} \geqslant \frac{\nu\left(u_{j+1}\right) - \nu\left(u_j\right)}{u_{j+1} - u_j}$$

which follows from the convexity of ν.

We use (19) to argue that $G_T(U') \preccurlyeq_{LR} G_T(U)$. Let $G_T(U') = \tilde{H}$ and $G_T(U) = \tilde{G}$. Note that

$$\mu(i, \tilde{H}) = \tilde{H}(\tau_{i-1}) - \tilde{H}(\tau_i) = u'_{n+1-i} - u'_{n+2-i}, \tag{20}$$

where the first equality is the definition of $\mu(i, H)$ and the second follows from $H = G_T(U)$. Therefore, it follows that for $i < j$,

$$\begin{aligned}
\mu(j, \tilde{G})\mu(i, \tilde{H}) &= (u_{n+1-j} - u_{n+2-j})(u'_{n+1-i} - u'_{n+2-i}) \\
&\leqslant (u'_{n+1-j} - u'_{n+2-j})(u_{n+1-i} - u_{n+2-i}) \\
&= \mu(i, \tilde{G})\mu(j, \tilde{H}),
\end{aligned}$$

where the first equality uses (20), the second uses (19), and the third uses (20). Therefore, we see that $G_T(U') \preccurlyeq_{LR} G_T(U)$.

Step 5: **We now combine these steps to complete the proof.**

Suppose that $U' \succcurlyeq_{RL} U$. As noted in Step 1, it is without loss of generality to treat U and U' as normalized utility assessments. For a distribution G, let $P' = P^*(k, U', G)$ and $P = P^*(k, U, G)$. By Step 4,

$$U' \succcurlyeq_{RL} U \Rightarrow G_T(U') \preccurlyeq_{LR} G_T(U).$$

Therefore, by Theorem 1, for every k,

$$P^*(k, U_T(G), G_T(U)) \succcurlyeq_A P^*(k, U_T(G), G_T(U')). \tag{21}$$

It then follows that

$$P^*(k, U, G) = \mathcal{T}[P^*(k, U_T(G), G_T(U))] \preccurlyeq_A \mathcal{T}[P^*(k, U_T(G), G_T(U'))] = P^*(k, U', G),$$

where the equalities follow from Step 3 and the ordering follows from (21). $\square$

A.3 Proof of Theorem 3 on p. 14

Proof. Consider any $k \geqslant 1$ and suppose $\tilde{k} = k + 1$. We consider two cases below.

The first case applies if $|P^*(k+1, U, F)| \leqslant k$; i.e., the optimal $(k+1)$-portfolio has no more than k colleges. Then $P^*(k, U, F) = P^*(k+1, U, F)$, which trivially implies the desired conclusion.

The second case applies if $|P^*(k+1, U, F)| = k+1$; i.e., the optimal $(k+1)$-portfolio has $(k+1)$ colleges. Denote the highest ranked college in $P^*(k+1, U, F)$ by College i and let $G(s) = F(s|s < \tau_i)$ be Ann's belief about her score conditional on being rejected by that college. Observe that $G \preccurlyeq_{LR} F$. Because Ann chooses the other colleges in $P^*(k+1, U, F)$ assuming she is rejected from College i, it follows that

$$P^*(k+1, U, F) = \{i\} \cup P^*(k, U, G). \tag{22}$$

Hence,

$$\lfloor P^*(k+1, U, F) \rfloor^k = P^*(k, U, G) \preccurlyeq_A P^*(k, U, F),$$

where the equality follows from (22) and the ordering follows from $G \preccurlyeq_{LR} F$ and Theorem 1.

Similarly, let College j be the lowest ranked college in $P^*(k+1, U, F)$. Let U' be a utility assessment that is identical to U but the outside option is u_j. Note that $U' \succcurlyeq_{RL} U$. Since Ann choose all colleges other than j in the portfolio assuming that she is accepted by College j (since it is her lowest ranked college), it follows that

$$P^*(k+1, U, F) = \{j\} \cup P^*(k, U', F). \tag{23}$$

Therefore,

$$\lceil P^*(k+1, U, F) \rceil^k = P^*(k, U', F) \succcurlyeq_A P^*(k, U, F),$$

where the equality follows from (23) and the ordering follows from $U' \succcurlyeq_{RL} U$ and Theorem 2.

Thus, we have shown that Theorem 3 holds for every $k \geqslant 1$ and $\tilde{k} = k + 1$. The argument for general $\tilde{k} > k$ then follows from the transitivity of $\succcurlyeq_A$. □

Online Appendices (Not for Publication)

The online appendices are organized as follows:

- Appendix B describes how to speed up the algorithm using Theorem 1, accommodate tier constraints, and how this algorithm can solve for the optimal portfolio also when admission decisions are stochastically independent.
- Appendix C collects some new results for the simultaneous search framework with independent admissions.
- Appendix D collects all examples and proofs for Section 5.2.
- Appendix E collects all examples and proofs for Section 5.3.

B Algorithms for Solving for the Optimal Portfolio

B.1 Speeding up the Algorithm Using Theorem 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 Theorem 1 to reduce the number of computation steps required for each step of the algorithm to $O(n \log n)$. Therefore, $O(n^2 \log n)$ 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 $i_1 = \texttt{median}(\mathcal{C}^\dagger)$ (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:

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

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

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

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 Theorem 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 $i_{2.2} = \texttt{median}(\{i_1+1,\dots,n\})$, find the optimal k-college portfolio following rejection from $i_{2.2}$:

$$j^*_{2.2} := \underset{j\in\mathcal{C}\text{ s.t. }\{j\}\preccurlyeq_A\{j^*_1\}}{\operatorname{argmax}} \left\{(1-F_{i_{2.2}}(\tau_j))u_j + F_{i_{2.2}}(\tau_j)V(C(j,k-1),F_{\{j\}})\right\}.$$

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 Theorem 1. The *optimal k-college continuation* following $i_{2.2}$ is then $C(i_{2.2},k) := \{j^*_{2.2}\} \cup C(j^*_{2.2},k-1)$.

$\vdots$

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

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

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

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

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 Theorem 1. The *optimal k-college continuation* following $i_{m.2}$ is then $C(i_{m.2},k) := \{j^*_{m.2}\} \cup C(j^*_{m.2},k-1)$.

$\vdots$

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

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

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 Theorem 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 Theorem 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 as 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 $O(n^2 \log n)$.

B.2 Accommodating Tier Constraints

Some education systems impose constraints not only on the number of colleges to which the student can apply but also on the composition of the portfolio. For example, in Kenya, applicants to secondary schools are restricted to rank two national schools, two provincial schools, and two district schools (Lucas and Mbiti, 2012). Similarly, in recent years, 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 6 can be extended to accommodate 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 $(i, j) \in C^\dagger \times \{0, \ldots m\}$ (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.

C Additional Results for Stochastically Independent Admissions

In this Appendix, we obtain and collect some new results for stochastically independent admissions, i.e., the simultaneous search framework modeled by 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 admissions probabilities, and unlike our framework, "replicas" are valuable for an 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 index copies). Similarly, colleges that are less desirable and more selective than another college are not ruled out.[30] Finally, to accommodate replicas, we require uniqueness only up to replacing replicas.

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

C.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. Theorem A.1 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.[31] We contrast this with the baseline framework in which the outside option u_o is deterministic.

Lemma 1 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 assessment in the sense of Definition 4.

Lemma 1. *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 assessment $V_r := (v_0^r; v_1^r, \ldots, v_n^r)$ (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

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

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 $\beta^r(u_i) = u_i + \int_{u_i}^\infty (1 - G^r(z)) dz$. This setup establishes the first part of Lemma 1: a direct calculation shows that this utility assessment 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

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

where we use the fact that the inverse of $\beta^r(\cdot)$ exists for values greater than $\int_0^\infty 1 - G^r(z) dz$ since $\beta^r(\cdot)$

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

is increasing for values greater than $\int_0^\infty 1 - G^r(z)dz$. 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

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

Since the step function $G^{r'}/G^r$ is non-decreasing from 0 to 1, and since ψ is constant for values of x lower than $\int_0^\infty 1 - G^r(z)dz$, this implies that ψ is convex. $\square$

Theorem A.1. *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.[32] 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 1, 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 Theorem 2 and the definition of x that $\{x\} \succcurlyeq_A \{y\}$, and so $P(k+1) \succcurlyeq_A P(k)$.[33] $\square$

C.2 The Risk-Loving Effect with Stochastically Independent Admissions

In Section 1, we claim that a result parallel to Theorem 2 holds even if admission decisions are stochastically independent. An implication of this result is that in Chade and Smith (2006), like

[32]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.

[33]We can invoke Theorem 2 because Ann is choosing a single-college portfolio in both cases and hence the correlation structure between colleges' admissions decisions is irrelevant.

our common-score framework, unequal outside options lead to more aggressive applications and therefore segregation in the composition of schools.

Theorem A.2. *Risk-love leads to a more aggressive portfolio:* $U' \succcurlyeq_{RL} U \Rightarrow P^*(k, U', \alpha) \succcurlyeq_A P^*(k, U, \alpha)$.

Lemma 2. *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 assessments* V *and* V' *such that* $\overline{V} = \overline{V'} = 0$ *and* $V' \succcurlyeq_{RL} V$.

Proof. Direct calculation shows that

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

and

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

are profiles as required by the statement. To see that $V' \succcurlyeq_{RL} V$ 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 V to V' (in particualr, $\beta(0) = 0$). 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 Theorem A.2. The proof proceeds by induction on k. The case of $k = 1$ follows from Theorem 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 2 (the rest of each portfolio is the optimal size $k-1$ portfolio from $C \setminus \{P^{*(i)}(k, U', \alpha)\}$ 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

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

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 2, she would choose a portfolio of $k-1$ colleges that is more aggressive than $\lceil P^*(k, U, \alpha) \rceil^{k-1}$ (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.[34]

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 Theorem 2). □

C.3 Algorithm for Solving for the Optimal Portfolio

Here, we adapt the algorithm of Section 4.4 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 assessment 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

$$V\left(P^*(k,U,\alpha,i)\right) := \max_{j \in \mathcal{C} \text{ s.t. } \{i\} \succ_A \{j\}} \left\{\alpha_j u_j + (1-\alpha_j) V\left(P^*(k-1,U,\alpha,j)\right)\right\}. \tag{24}$$

Using this dynamic program, one can run a routine analogous to Figure 6: 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 B 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 B.2, our approach can expand the scope of their analysis by accommodating tier constraints, unlike that algorithm.

D Examples and Proofs for Section 5.2

D.1 Proof of Theorem 6 on p. 19

Lemma 3. *The number of replicas of College 1 on the optimal k-portfolio approaches infinity as k increases to infinity.*

Proof. Since the parameters of the problem are fixed throughout the proof, for each k, we denote the optimal k-portfolio by $P(k)$.

Toward contradiction, suppose that $\liminf |\{x \in P(k) \mid x \text{ is a replica of College } 1\}| = m < \infty$. Then, there exists an increasing sequence of portfolio sizes such that the number of replicas of College 1 on the optimal portfolio is at most m. 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 $l > 1$ be the lowest such index.

[34]This follows since the corresponding profiles V from Lemma 2 are $\succcurlyeq_{RL}$-ranked.

Consider a subsequence of portfolio sizes such that the number of replicas of College l increases to infinity and the composition of higher ranked colleges is constant (this is possible as, by construction, the number of such colleges on the optimal portfolio is uniformly bounded across all portfolio sizes in the sequence). For sufficiently large portfolio, admission to a replica of College l is nearly guaranteed (even when conditioning on rejection from all better schools).[35]

By contrast, rejection from all applications to Colleges $1, \ldots, (l-1)$ occurs with probability greater than $K > 0$ (which does not depend on the size of the portfolio, since we selected portfolios with the same composition of applications to these colleges). Additionally, the probability of admission to College 1 conditional on rejections from all replicas of Colleges $1, \ldots, (l-1)$ is bounded below by $\Delta > 0$ (which does not depend on the size of the portfolio).[36] This is a contradiction, because the implication is that for sufficiently large portfolio sizes, Ann strictly prefers to replace a replica of College l with a replica of College 1 (contradicting the optimality of $P(k)$). □

Proof of Theorem 6. By Lemma 3 there exists k such that a replica of College 1 appears on the optimal k-portfolio, $P(k)$. By Theorem A.1, for any $k' > k$ the $P(k')$ consists of the same colleges as $P(k')$ in addition to $k' - k$ replicas of College 1. □

D.2 Proof of Theorem 7 on p. 20

Proof. That the number of copies of College 1 approaches infinity as the size of the portfolio increases follows from Lemma 3. We now prove that the number of copies of College $\underline{m}$ also approaches infinity as the size of the portfolio increases. The case $\underline{m} = 1$ is degenerate, so we henceforth assume that $\underline{m} \neq 1$. To simplify notation, we denote by 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. We also normalize the outside option to zero throughout this proof (without loss of generality).

Given an increasing sequence of portfolio sizes, Let $l < \underline{m}$ denote the largest index such that $\limsup |\{x \in P^*(k) \mid x \text{ is a replica of College } l\}| = \infty$ (where we suppress the dependence of l on the sequence). Toward contradiction, assume that there exists an increasing sequence of portfolio sizes such that $l \neq \underline{m}$. Then, there is no loss in assuming that each portfolio in the sequence include an application to a copy of College l. Additionally, the number of copies of colleges $l+1, \ldots, \underline{m}$ is uniformly bounded by some $B < \infty$.

For a given portfolio size k, let $k_1, k_2, \ldots k_{l-1}$ denote the number applications to copies of Colleges $1, 2, \ldots, (l-1)$, respectively. Furthermore, denote by k_l the number of applications to copies of College l minus 1 (i.e., excluding the marginal application to college l). In what follows, we will show that, for sufficiently large portfolios, replacing the $k_l + 1$-st application to College l with an application to College $(l+1)$ will be strictly beneficial to Ann (contradicting the optimality of the portfolios).

[35] E.g., find a s^* low enough that $\Pr\{s > s^* \mid \text{rejections from Colleges } 1, \ldots, (l-1)\}$ is above $\sqrt{q}$ and a large enough number of applications to College l that conditional on s^* admissions occurs with probability $\sqrt{q}$ (and thus conditional on any $s > s^*$ admission occurs at least with probability $\sqrt{q}$).

[36] To see this, note that there is a positive probability that the common component of Ann's score $\rho s \in [\tau_l, \tau_l + 1]$ but, for all of the independent draws for replicas of Colleges $1, \ldots, (l-1)$ are such that $\left(\sqrt{1-\rho^2}\right)\varepsilon_c < -1$.

We proceed by conducting a cost-benefit analysis for replacing the $k_l + 1$-st application to College l with an application to College $(l+1)$. The benefit from the $k_l + 1$-st application to College l is bounded above by

$$\int_{-\infty}^{\infty} g(z) F^{k_1}(\tau_1 - z) \times \cdots \times F^{k_l}(\tau_l - z) \left(1 - F(\tau_l - z)\right) u_l dz.$$

This expression is an upper bound on the loss from forgoing the marginal application to a copy of College l. It uses the fact that the marginal application to College l is only beneficial if Ann is rejected by all other weakly preferred colleges, and ignores the fact that if Ann is rejected by all these colleges, she may still get into some college she likes more than the outside option.

The marginal benefit from an application to school $l+1$ is bounded below by

$$\int_{-\infty}^{\infty} g(z) F^{k_1}(\tau_1 - z) \times \cdots \times F^{k_l}(\tau_l - z) \left(1 - F(\tau_{l+1} - z)\right) u_{l+1} F^B(\tau_{\underline{m}} - z) dz$$

In this expression, we assumed a "worst case" where Ann only benefits from the application to College $l+1$ if she is rejected by all other colleges she likes weakly less (taking the maximal number of such applications, B, and assuming they are all to the least selective rationalizable college, $\underline{m}$).

Denote $\hat{F}_k(z) := (F^{k_1}(\tau_1 - z) \times \cdots \times F^{k_l}(\tau_l - z))^{1/\sum k_i}$. A change of variables ($x = \hat{F}_k(z)$) yields the upper bound

$$\int_0^1 g(\hat{F}_k^{-1}(x)) x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{1}{\hat{F}_k' \left(\hat{F}_k^{-1}(x)\right)} dx \tag{25}$$

and the lower bound

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

We will complete the proof by showing that the ratio between the marginal benefit and the marginal cost approaches infinity as portfolio sizes increase. The proof proceeds in several steps.

Step 0. Since F is strictly increasing, by the definition of $\hat{F}_k$ we have that $\hat{F}_k$ is strictly increasing (thus invertable) and

$$\tau_l - F^{-1}(x) \leqslant \hat{F}_k^{-1}(x) \leqslant \tau_1 - F^{-1}(x).$$

Step 1. By Step 0, as x approaches 1 we have that $\hat{F}_k^{-1}(x)$ approaches $-\infty$. Hence, the positive expressions $1 - F(\tau_{l+1} - \hat{F}_k^{-1}(x))$ and $1 - F(\tau_l - \hat{F}_k^{-1}(x))$ 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

$$1 - F(\tau_{l+1} - \hat{F}_k^{-1}(x)) = 1 - F\left((\tau_{l+1} - \tau_l) + \tau_l - \hat{F}_k^{-1}(x)\right) > M\left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right).$$ [37]

[37] To see this, write $\Delta = \tau_l - \tau_{l+1} > 0$ and $z = \tau_l - \hat{F}_k^{-1}(x)$, and observe that $1 - F(z - \Delta)/1 - F(z)$ approaches infinity as z approaches infinity.

Furthermore, using Step 0 and the monotonicity of F, we get that for any $x \geqslant 1-\varepsilon_M$

$$F^B(\tau_{\underline{m}} - \tau_1 + F^{-1}(1-\varepsilon_M)) \leqslant F^B(\tau_{\underline{m}} - \hat{F}_k^{-1}(x)).$$

Since

$$\int_{1-\varepsilon_M}^{1} g(\hat{F}_k^{-1}(x))x^k \left(1 - F(\tau_{l+1} - \hat{F}_k^{-1}(x))\right) u_{l+1} F^B(\tau_{\underline{m}} - \hat{F}_k^{-1}(x)) \frac{1}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx$$

is clearly a lower bound on Equation (26), the above inequalities allow us to provide a relaxed lower bound on the benefit of including a copy of College (l+1):

$$\int_{1-\varepsilon_M}^{1} g(\hat{F}_k^{-1}(x))x^k M \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_{l+1} F^B(\tau_{\underline{m}} - \tau_1 + F^{-1}(1-\varepsilon_M)) \frac{1}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx. \tag{27}$$

Step 2. We now derive a relaxed upper bound on the marginal benefit from an application to a copy of College l. Intuitively, we will show that in large portfolios (after many rejections from other colleges that are weakly more desirable), Ann's beliefs are concentrated on pessimistic values.

First, consider the integral from Equation (25) restricted to low beliefs ($x \geqslant 1/2$). We have

$$\int_{\frac{1}{2}}^{1} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx > \frac{3}{4}^k \int_{\frac{3}{4}}^{1} \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx$$

where we have used the fact that the integrand is positive everywhere. Additionally, by Step 0 we have

$$\frac{3}{4}^k \int_{\frac{3}{4}}^{1} \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx > \frac{3}{4}^k \int_{\frac{3}{4}}^{1} \left(1 - F(F^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx.$$

Similarly, focusing on high beliefs, we get

$$\int_{0}^{\frac{1}{2}} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx < \frac{1}{2}^k \int_{0}^{\frac{1}{2}} \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx$$

and by Step 0

$$\frac{1}{2}^k \int_{0}^{\frac{1}{2}} \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx < \frac{1}{2}^k \int_{0}^{\frac{1}{2}} \left(1 - F(\tau_l - \tau_1 + F^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx.$$

Altogether we get

$$\frac{\int_{\frac{1}{2}}^{1} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'(\hat{F}_k^{-1}(x))} dx}{\int_{0}^{\frac{1}{2}} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'(\hat{F}_k^{-1}(x))} dx} > \frac{(3/4)^k}{(1/2)^k} \frac{\int_{\frac{3}{4}}^{1} \left(1 - F(F^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'(\hat{F}_k^{-1}(x))} dx}{\int_{0}^{\frac{1}{2}} \left(1 - F(\tau_l - \tau_1 + F^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'(\hat{F}_k^{-1}(x))} dx}.$$

Since the right hand side approaches infinity as k grows to infinity,[38] we get that for sufficiently large k

$$2\int_{\frac{1}{2}}^{1} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx > \int_0^1 x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx.$$

Step 3. We now further relax the upper bound from Step 2. The intuition, again, is that beliefs on the common component are concentrated on lowest values (this time, $1 - \varepsilon_M < x$).

$$\int_{\frac{1}{2}}^{1} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx =$$

$$\int_{\frac{1}{2}}^{1-\varepsilon_M} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx + \int_{1-\varepsilon_M}^{1} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx.$$

The expression $\frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)}$ is uniformly bounded above by some $L > 0$ in the domain $[\frac{1}{2}, 1 - \varepsilon_M]$.[39] Thus,

$$\int_{\frac{1}{2}}^{1-\varepsilon_M} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx \leqslant L u_l \int_{\frac{1}{2}}^{1-\varepsilon_M} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) dx \leqslant \frac{(1-\varepsilon_M)^{k+1}}{k+1} L u_l.$$

Similarly, the expression $\left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)}$ is uniformly bounded below by some $\Delta > 0$ in the domain $[1 - \varepsilon_M, 1 - \varepsilon_M/2]$. Thus,

$$\int_{1-\varepsilon_M}^{1} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx > \int_{1-\varepsilon_M}^{1-\frac{1}{2}\varepsilon_M} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx$$

$$> \frac{\varepsilon_M \Delta u_l}{2} (1 - \varepsilon_M)^k.$$

Altogether we get that for sufficiently large values of k

$$\int_{\frac{1}{2}}^{1} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx < 2\int_{1-\varepsilon_M}^{1} x^k \left(1 - F(\tau_l - \hat{F}_k^{-1}(x))\right) u_l \frac{g(\hat{F}_k^{-1}(x))}{\hat{F}_k'\left(\hat{F}_k^{-1}(x)\right)} dx. \tag{28}$$

[38] The right hand side grows to infinity since $(\frac{3}{4}/\frac{1}{2})^k$ approaches infinity, while the ratio of integrals is bounded below by a positive number. To see this, note that the integrals depend on k only through the vector $(k_1/k, k_2/k, \ldots, k_l/k)$, and the values of these integrals are continuous in this vector. Thus, they attain a minimum and a maximum in the l-simplex (both must be positive, since the integrand is positive). (An alternative approach is to take a subsequence of portfolios such that $(k_1/k, k_2/k, \ldots, k_l/k)$ converge.)

[39] See Footnote 38.

Step 4. To complete the proof we show that the ratio between the relaxed lower bound of Equation (27) and the relaxed upper bound of Equation (28) is greater than 1, establishing that for sufficiently large portfolios on the sequence replacing the marginal application to a copy of College l with an application to a copy of College $(l+1)$ would be beneficial (a contradiction).

The ratio is equal to $\frac{Mu_{l+1}F^B(\tau_{\underline{m}}-\tau_1+F^{-1}(1-\varepsilon_M))}{4u_l}$. For a sufficiently large choice of M , this expression is arbitrarily large (all terms in the numerator are weakly increasing in M) and, in particular, it is greater than 1. $\square$

D.3 Effect of Increasing Weight on Common Component

This subsection analyzes the effect of increasing the weight on the common component, ρ, on two-college portfolios. We begin with an example that illustrates how moving from $\rho = 0$ to $\rho = 1$, the optimal portfolio becomes strictly less aggressive.

Example A.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 $\Phi(\tau_4) = 0.01$. When $\rho = 1$, the optimal 2-college portfolio is $\{2,4\}$. By contrast, when $\rho = 0$, the optimal 2-college portfolio is $\{1,3\}$.

The example below shows that the opposite effect can also arise: moving from $\rho = 0$ to $\rho = 1$ leads to a more aggressive portfolio.

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

We now provide a proof of Theorem 8. We denote by Φ (resp., ϕ) the CDF (resp., PDF) of the standard (univariate) normal distribution and by $\Phi_2(\cdot,\cdot,\rho)$ (resp., $\phi_2(\cdot,\cdot,\rho)$) the CDF (resp., PDF) of the standard bivariate normal distribution with correlation ρ. The proof uses two lemmas.

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

Proof. Observe that

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

The first term of the product is the ratio of (nonzero) probabilities, and the term in parenthesis is the difference between two inverse Mills ratios, which are known to be increasing. $\square$

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

Proof. Let $B := \Phi_2(\tau_i, \infty, \rho) = \Phi(\tau_i)$, the probability of rejection from school i (which is constant across all correlation levels), $g(\rho) := \Phi_2(\tau_i, \tau_j, \rho)$, 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 $\dot{H} > 0$.[40]

[40]For consistency with the literature, we represent the derivative with respect to ρ as a dot.

To begin with, note that $\dot{H} :=_{sgn} \dot{g}(B-f) - \dot{f}(B-g)$, and so

$$\dot{H} > 0 \iff \frac{(B-f)}{(B-g)} > \frac{\dot{f}}{\dot{g}} \ .$$

It is well known that conditional of $s_i = x$ 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,

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

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

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

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.

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 get:

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

as required. □

Proof of Theorem 8. Fixing the rest of the parameters of the problem, let $P(\overline{\rho})$ denote the optimal size-2 portfolio under correlation $\overline{\rho}$. Toward contradiction, assume that $\rho' > \rho$ but $P(\rho') := \{i', j'\}$ is less dispersed than $P(\rho) := \{i, j\}$, i.e., $i < i' < j' < j$. Since $P(\rho)$ is optimal under ρ, $j = P^{(2)}(\rho)$ must be optimal conditional on being rejected from i under the correlation ρ. Hence

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

Similarly, since $P(\rho')$ is optimal under ρ',

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

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 get a contradiction, it suffices to show that

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

Let $\kappa(\tau) := \frac{\Phi(\tau) - \Phi_2(\tau, \tau_j, \rho')}{\Phi(\tau) - \Phi_2(\tau, \tau_{j'}, \rho')}$. We will show that $\kappa'(\tau) < 0$ for any $\tau > \tau_{j'}$ (i.e., that the bad news effect from being rejected from the first school is stronger the less selective it is).

By Fubini's theorem,

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

So $\kappa'(\tau)$ 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 has the same sign as

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

By Lemma 4 the first term is bounded below by the value of the integrands at τ (we called this quantity λ in the proof of Lemma 5), which is exactly equal to the second term (with opposite sign). This establishes that $\kappa'(\tau) > 0$ as required. □

E Examples and Proofs for Section 5.3

Proof of Theorem 9 on p. 21. 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 7), 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.[41]

We next show that in our setting, 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 7) and the median project among the bottom $2^{k-1} - 1$ projects if her first attempt fails (the first red arrow in Figure 7). 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. □

Example A.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 application, 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 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 the independent setting. In that setting, Chade and Smith (2006) show that agents stop searching after the first success. Moreover, if attempts do not succeed, the agent searches more than they would in the simultaneous problem and goes for more aggressive choices.

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