The College Portfolio Problem*

S. Nageeb Ali   Ran I. Shorrer †

July 30, 2021

Abstract

A college applicant faces the following risky choice: she applies to a portfolio of colleges while being uncertain about which school would admit her. Admissions decisions are correlated insofar as being rejected by a lower ranked school may imply that she is rejected by a higher ranked school. We show that solutions to this decision problem involve applying to a combination of *reach*, *match*, and *safety* schools. When application costs decrease, a college applicant broadens the range of schools to which she applies by including both those that are more selective and those that are safer options.

*We thank Kehinde Ajayi, Hector Chade, David Dillenberger, Simone Galperti, Guillaume Haeringer, Jacob Leshno, Ross Rheingans-Yoo, Uzi Segal, Denis Shishkin, Lones Smith, Joel Sobel, Andy Skrzypacz, and Rakesh Vohra for their suggestions. 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.

Contents

1 Introduction

Ann faces a risky choice with high stakes: she is applying to colleges. There are many colleges but she can apply only to a few. Ann, like most applicants, is uncertain about which colleges would admit her. She is advised to apply to a broad range of colleges including (i) some that she is very excited about even if her chances are low (*reaches*), (ii) some where she is a good *match*, and (iii) some *safety* schools where she has high odds of being admitted. This strategy hedges not only against the idiosyncratic risk of each college application but also the correlation across colleges: being rejected from one reach school may imply her odds are low at another reach school, but may not be all that informative about her prospects at a safety school. Correlation often looms large in the mind of applicants and is a reason that applicants diversify their portfolio rather than apply to a large number of reaches.

Why might admissions prospects be correlated? In the US college and graduate school admissions process, applicants are uncertain about admissions criteria, the overall pool of applicants, and where one stands in that pool. In other contexts, the administrative process of admissions explicitly correlates admissions decisions. Some admissions processes require that applicants first submit applications, then take a common exam, and then admissions decisions are based on those exam scores. Each school has a minimum score for admissions and an applicant at the time of application is uncertain of the score she will obtain. This process is or has been used (until recently) for high school and college admissions in a number of countries, including China, Ghana, Kenya, Mexico, Turkey, and the UK.[1] Centralized school choice algorithms also correlate admissions prospects if a common lottery is used to break priority ties for over-demanded seats, as is done in Amsterdam and NYC.

We study the college portfolio problem when admissions prospects are highly correlated. Which colleges should Ann apply to? How does the composition of her portfolio depend on the number of colleges to which she applies? If she can apply to more colleges, should she expand the range to which she applies? Or should she be more aggressive and add only more selective colleges?

This decision problem has an interesting structure. Ann first chooses a portfolio of gambles, each of which has a single prize ("acceptance"). These gambles award prizes with a specific correlation structure, corresponding to our discussion above. Ann chooses her favorite prize and there is no intrinsic value to multiple prizes. We use an elementary and graphical approach that reduces this decision problem to a "coverage" problem. We characterize features of the optimal solution and show that it matches the advice of diversifying one's portfolio by applying to colleges at various levels of selectivity. We show that if applications become less costly and Ann can apply to more colleges, she broadens this range further by including both more selective and safer schools.

Our results use two elementary principles of optimal portfolios. The first, formalized in Lemma 1, is that Ann applies less aggressively if she is less optimistic about her chances in a likelihood-ratio sense, namely lower relative odds of admissions at higher ranked colleges. We call this a *bad news* effect. The second, formalized in Lemma 2, is that Ann applies more aggressively if she is more

[1] Ajayi (2013) and Ajayi and Sidibe (2020) study the process for high school admissions in Ghana and Lucas and Mbiti (2012) study that in Kenya. In the UK, a university applicant applies to at most five universities, receives admissions offers contingent on a score in a subsequent examination, and of these contingent admissions offers, can mark one as a preferred choice and another as an insurance option before she learns her score. See Broecke (2012) for details and analysis.

risk loving. This risk loving effect applies, for example, if Ann's outside option increases in value. Although these two effects appear distinct, we show that each is the "dual" of the other.

Our main result, Theorem 1, is on the value of diversification. We study how Ann's optimal portfolio changes if she can apply to more colleges or faces reduced application costs. We show that she uses this greater flexibility to diversify her portfolio further, applying to colleges at various levels of selectivity, including more reach and safety schools. We prove this result by thinking through Ann's portfolio choice using the two principles described above. If Ann could add more safety schools to a portfolio, these schools serve as backup options if she is rejected by more selective colleges. Having these backup options makes her more risk loving when choosing other colleges, and by the risk loving effect, we know that she is then more aggressive in those choices. Conversely, if Ann adds more reaches to her portfolio, she has to think about backup options if she is rejected by those reaches. Being rejected is bad news, and the bad news effect implies that she should then be less aggressive in her other choices. Thus, the two effects combined push Ann to diversify her portfolio.

The prediction that candidates should apply to diverse portfolios appears consistent with practice. Ajayi (2013) and Pallais (2015) find that applicants do indeed diversify their portfolios when they can apply to more schools. The College Board in their advice to applicants, as seen in How to Strengthen Your College List, suggests that "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." Our work rationalizes this practice of diversification.

While it may appear natural for applicants to rationally diversify their portfolio, this prediction is not shared by existing models. Most models of the college portfolio problem build on the pioneering work of Chade and Smith (2006), who formulate and study the class of "simultaneous search" problems. They assume that an applicant's admissions prospects across schools are stochastically independent conditional on all information available to her. They characterize the optimal solution via a greedy algorithm that adds schools to one's portfolio on the basis of which school adds the highest marginal benefit over the existing portfolio. In their model, an applicant does not value safety schools, and their logic suggests that if application costs decrease, applicants should expand their portfolios upwards by applying only to more selective schools. This framework has been extended in a number of directions as well as taken to empirical evidence, maintaining this assumption of stochastic independence.[2]

While we view stochastic independence to be plausible for many settings, there are three reasons to study the correlated case that we do here. First, as articulated above, admissions prospects are indeed highly correlated in a number of settings, and are often perceived to be so even in decentralized US College and graduate school admissions. Second, as Chade and Smith (2006) show, an applicant has no motive to diversify when admissions prospects are independent. These predictions appear at odds with the advice that applicants receive, the choices they make to apply to safety schools, and how applicants expand their portfolio upwards and downwards when application costs

[2]See, for example, Card and Krueger (2005), Galenianos and Kircher (2009), Kircher (2009), Chade, Lewis, and Smith (2014), Fu (2014), Olszewski and Vohra (2016), Larroucau and Rios (2018), and Walters (2018).

decrease. Third, the correlated case that we study admits an elementary geometric approach, which in conjunction with existing approaches may be useful to better understand college portfolio choices.

Similar to us, Avery and Levin (2010) model correlation in their study of early admissions programs. In their model, colleges value ability and a match-specific enthusiasm with the applicant. Colleges assess ability in a common way, which generates correlation in admissions decisions. Avery and Levin focus on how applicants may use early admissions to signal their enthusiasm and colleges benefit from having early admissions programs. We too study the role of early admissions programs in Section 6.1, focusing on its role in helping an applicant search for the right college (and abstracting from the role of signaling). We show that *early decision* programs offer limited value to Ann in our model because she is committed to attending the college to which she applies early. By contrast, *early action* programs are almost equivalent to doubling the number of applications that she can submit: because Ann is not committed to attending the school to which she applies early, she can use its acceptance and rejection decision to direct her future search.

Outline of Paper: Section 2 illustrates our results through a simple example. We describe our model in Section 3. We study how optimal portfolios respond to beliefs and risk sensitivity in Section 4. We prove our main result in Section 5. Section 6.1 studies early admissions policies and Section 6.2 offers a detailed comparison to Chade and Smith (2006). Section 7 concludes. All proofs are in Appendices.

2 Example

Suppose that Ann can apply to a subset of four colleges, $\{1, 2, 3, 4\}$. The colleges are ordered in terms of her preferences so that College i is her i^{th} favorite. She can attend at most one college and she prefers attending any of these colleges to her outside option of not attending college at all.

Admissions decisions are based on a score s that Ann obtains. That score may reflect how her application compares to the pool of all applicants, or correspond to her performance on a common examination, or be the realization of a common lottery used to break priority ties. Each college i has its own score threshold τ_i. If Ann applies to College i, her application is accepted if and only if her score weakly exceeds τ_i. Of the schools that accept her, Ann attends her favorite college. If all her applications are rejected, she obtains her outside option.

For simplicity, suppose that s is drawn uniformly from $[0, 1]$, and each college's threshold τ_i is in $[0, 1]$. Table 1 summarizes for each college, Ann's utility from attending the college, its score threshold, and the admissions probability. The payoff of her outside option (of not attending college) equals 0.

	College 1	**College 2**	**College 3**	**College 4**
Utility (u_i)	1	0.45	0.25	0.1
Score Threshold (τ_i)	0.78	0.5	0.125	0.25
Admissions Probability	0.22	0.5	0.875	0.75

Table 1: Utilities, score thresholds, and admission probabilities for each college.

Ann's favorite college, College 1, has the highest threshold, and we consider this to be her "reach"

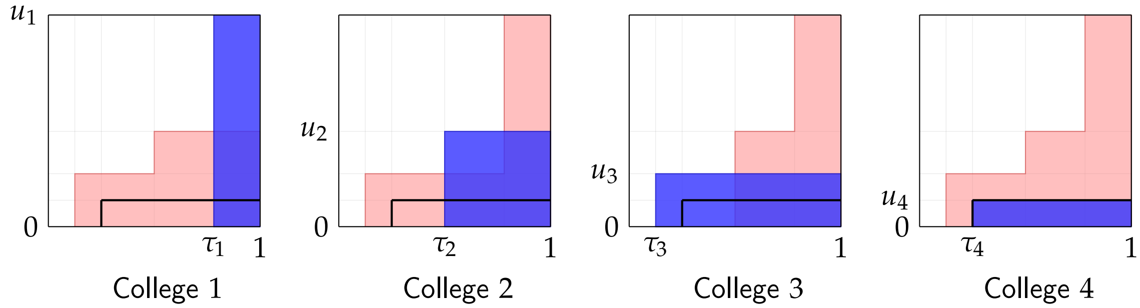

Figure 1: Single-College Portfolios. Each figure depicts a single-college portfolio with scores and thresholds on the horizontal axis and utility on the vertical axis. If Ann applies to College i, her application is accepted whenever her score exceeds τ_i, and if accepted, she obtains a utility of u_i. The area of the blue rectangle is therefore the expected utility of the portfolio.

school. College 2 has a lower threshold but a lower desirability, and therefore, we refer to it as her "match" school. College 3 is a "safety" school for Ann. Finally, College 4 has a higher threshold than College 3 but offers less utility to her; we therefore refer to it as a "dominated" school.

Optimal Single-College Portfolio: If Ann can apply to only one college, she does so to a college that maximizes $(1-\tau_i)u_i$. Based on the numbers above, Ann's optimal single-college portfolio is to apply to her match college, College 2.

Although this logic is obvious, it is useful for our subsequent analysis to see it through the lens of a coverage problem. Figure 1 shows the expected utility associated with each of the four single-college portfolios. We depict each college application as a rectangle in score-utility space: Ann is accepted by College i if her score weakly exceeds τ_i, and then she obtains utility u_i. The expected utility of a portfolio corresponds to the area of that rectangle. An optimal portfolio maximizes that area, or equivalently, coverage of the unit square.

Optimal Two-College Portfolio: Now suppose that Ann can apply to two colleges. Being able to apply to a second college affects Ann's reasoning in two ways. First, if both applications are accepted, she attends her preferred college. Second, her admissions prospects are correlated across the two colleges in her portfolio, and now there is a question as to how Ann responds to that correlation.

The optimal two-college portfolio is $\{1,3\}$. Thus, she is applying to her reach school, College 1, with College 3 as a backup. The reader may wonder why she no longer applies to College 2 and instead expands her range both upwards and downwards rather than in a single direction. We use calculations and pictures to explain why.

Suppose Ann compares portfolios $\{1,2\}$ and $\{1,3\}$. For each portfolio, if her score exceeds 0.78, she is accepted by College 1, and chooses to attend that college. So these portfolios generate identical (ex post) payoffs if her score exceeds 0.78. But what if her score is below 0.78? With portfolio $\{1,2\}$, Ann is accepted by College 2 if her score exceeds 0.5. Conditional on being rejected by College 1, her conditional expected payoff is

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

By contrast, if she chooses the portfolio $\{1,3\}$, she is admitted into College 3 so long as her score

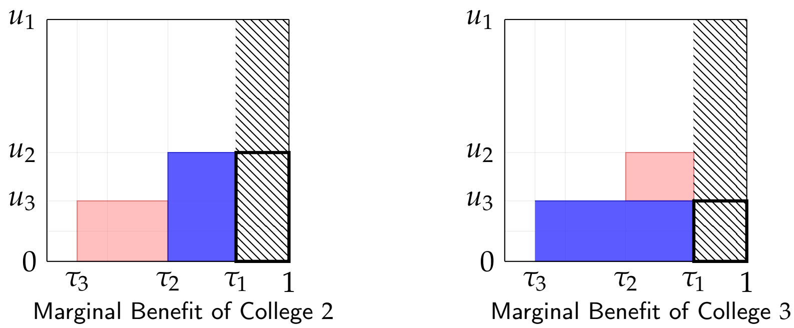

Figure 2: Comparing portfolios $\{1, 2\}$ and $\{1, 3\}$. Each figure depicts the marginal benefit of adding a college to College 1. The dashed area depicts regions where Ann is admitted to College 1, the area framed in black is the overlap between College i and College 1, and the blue area is what remains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

exceeds 0.125. Therefore, her conditional expected payoff is

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

Thus, we see that the portfolio $\{1, 3\}$ is better than $\{1, 2\}$.

Ann's deliberations here are reminiscent of pivot calculations in auctions and elections: her choice between portfolios $\{1, 2\}$ and $\{1, 3\}$ matters *only* if she is rejected by College 1. In that case, College 3 is a better backup option because being rejected by College 1 is worse news for her prospects at College 2 than at College 3.

This logic is also seen in Figure 2. If she chooses portfolios $\{1, 2\}$ or $\{1, 3\}$, Ann enrolls in College 1 if she is accepted by that college. Therefore she obtains the area of the rectangle from College 1 in each case. Thus, in comparing the two portfolios, the choice depends on how much area the other college in that portfolio adds when Ann is rejected by College 1. We see that removing the parts of all rectangles with scores above τ_1 takes away a larger part of College 2's rectangle than that of College 3 making the latter relatively more attractive. Hence, the portfolio $\{1, 3\}$ attains a higher coverage.

How about if Ann compares portfolio $\{1, 3\}$ to a less risky portfolio $\{2, 3\}$? If Ann's score is below the threshold for College 3, then both $\{1, 3\}$ and $\{2, 3\}$ would result in her obtaining her outside option, and so this contingency does not affect her decision. So let's imagine that her score is above the threshold for College 3. Conditioning on this event, if Ann chooses portfolio $\{1, 3\}$, she obtains

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

By contrast, if she chooses portfolio $\{2, 3\}$, her conditional expected payoff 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. For Ann's choice to matter, her score must be sufficiently high that she is accepted by at least College 3. Once Ann conditions on that event, College 3 effectively becomes her

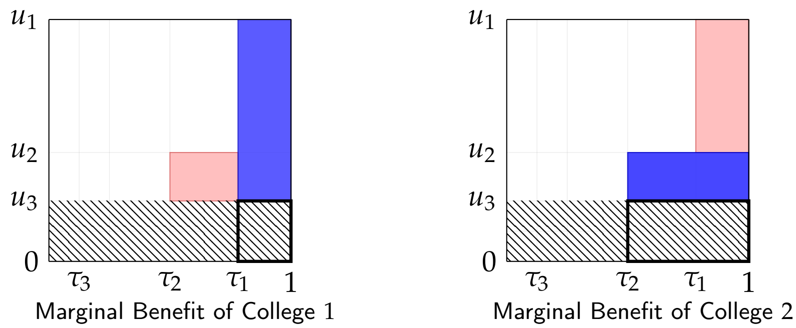

Figure 3: Comparing portfolios $\{1,3\}$ and $\{2,3\}$. Each figure shows the benefit of adding a college to College 3. Because Ann is accepted by College 3 whenever she is accepted by 1 or 2, Ann values each college relative to an outside option of u_3. The dashed area removes outcomes with lower utility. The solid blue area is that of the rectangle above u_3.

outside option. We see this reflected in both (3) and (4), where she obtains a payoff of u_3 whenever she is rejected by a better college. Because u_3 is strictly positive, this is an increase in the value of Ann's outside option. So Ann should ask herself which of College 1 and 2 is better to apply to given that College 3 is her outside option. Having an increased outside option makes Ann more "risk loving" (in a sense that we formalize later) and willing to apply more aggressively. This logic can be seen in Figure 3, where obtaining an outside option of u_3 removes a larger part of College 2's rectangle, making College 1 relatively more attractive. As the figure shows, the increased outside option removes a bottom slice of each rectangle, but makes College 1 relatively more attractive by removing a larger part of College 2's rectangle.

Figures 2 and 3 appear to look at the problem differently. In comparing $\{1,2\}$ and $\{1,3\}$, we described a *bad news* effect that leads Ann to be less aggressive and choose 3 as her backup option. By contrast, in comparing $\{2,3\}$ and $\{1,3\}$, we interpreted matters in terms of an *increased outside option* that leads Ann to be more aggressive and apply to College 1. We show that, all else equal, (i) bad news lead to less aggressive portfolios, and (ii) increased outside options lead to more aggressive portfolios. While these two effects appear distinct, they are in fact mirror images of each other. We show how in Figure 4: transposing a figure around the off-diagonal connecting $(1,0)$ and $(0,1)$, we see that the choice between two portfolios where one has the bad news effect can be mapped to a choice in which one has the increased outside options effect (and vice versa). This transposition does not change the collection of rectangles that cover the biggest area, so showing that bad news lead

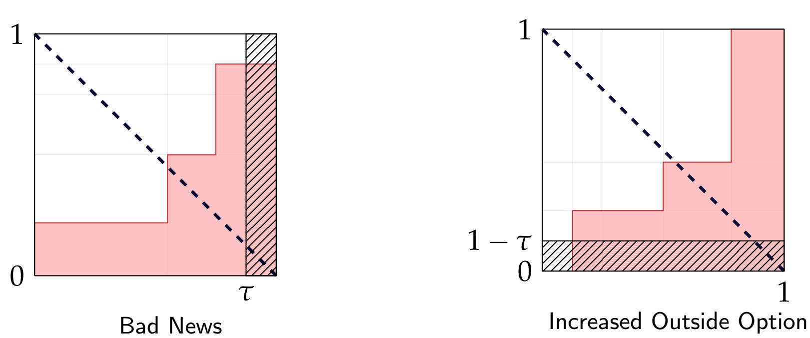

Figure 4: The left figure illustrates the effect of giving Ann the bad news that her score is below τ. Transposing the figure over the $-45°$ degree line yields an isometric coverage problem, where Ann's outside option increases from 0 to $1-\tau$.

to less aggressive portfolios is equivalent to showing that increased outside options lead to more aggressive portfolios.

So we see that the optimal two-college portfolio, $\{1,3\}$, expands both upwards and downwards relative to the optimal single-college portfolio. Our main result shows that this is a general property: the optimal $(k+1)$-college portfolio expands both upwards and downwards relative to the optimal k-college portfolio. The logic combines the two effects described above. By the bad news effect, adding an additional college at the top makes Ann less aggressive in choosing the remaining k colleges. But also, by the increased outside option effect, adding an additional college at the bottom makes Ann more aggressive in optimally choosing the remaining k colleges. Therefore, whenever application costs reduce and Ann can therefore apply to more colleges, she pursues a more diversified portfolio.

3 Model

Ann is applying to college. There is a finite set of colleges $C \equiv \{1, 2, ..., n\}$ with generic element i. Ann's utility from attending College i is u_i. Colleges are ordered in terms of Ann's preferences so that $u_1 \geqslant u_2 \geqslant \cdots \geqslant u_n$.

Here is how the process works. Ann first chooses a portfolio P of colleges, where P is a subset of C. After this portfolio choice, she obtains a score s in $[0,1]$. Each college i has a *minimal score threshold* $\tau_i \in [0,1]$ such that if Ann applies to college i, her application is accepted if and only if $s \geqslant \tau_i$. Ann can attend at most one college and chooses among the colleges that accept her application. If she does not attend any college, she obtains her *outside option*, which generates a payoff of $\overline{u}$. The *utility assessment* $U \equiv (\overline{u}; u_1, \ldots, u_n)$ specifies utilities associated with all of the possible outcomes.

When choosing her portfolio, Ann is uncertain about her score, and her beliefs are represented by a cumulative distribution function F. Given a distribution F, we denote its left-continuous version as $F^-(x) \equiv \lim_{y \uparrow x} F(y)$. Since Ann chooses her favorite college among those that both accept her and improve upon her outside option, her expected value of a portfolio P is

$$V(P, U, F) \equiv \int_0^1 \max_{\{i \in P : s \geqslant \tau_i\}} \max\{u_i, \overline{u}\} dF. \tag{5}$$

This expression shows that Ann's score matters only to the extent that it affects colleges' acceptance and rejection decisions. Thus, we can summarize distribution F by the vector $(F^-(\tau_1), \ldots, F^-(\tau_n))$. Any two distributions that have the same vector will generate identical values for all portfolios.

Applying to colleges is costly. Ann faces application costs that depends on the number of colleges to which she applies: if she chooses portfolio P, the total cost is $\phi(|P|)$. These application costs may reflect potentially exogenous limits to how many colleges Ann can apply, the monetary fees associated with applications, and the time spent completing the application process and learning about colleges. We normalize costs so that $\phi(0) = 0$ and assume that ϕ is non-decreasing. Special cases of this setting include (i) those where Ann can apply to a certain number of colleges for free but cannot apply to more than that number as well as (ii) those where each application fee is the same. Ann chooses a portfolio based on its value and its associated application costs: her net utility from portfolio P is $V(P, U, F) - \phi(|P|)$.

We simplify our exposition through several assumptions. First, we eliminate *dominated colleges* and *replicas*. College j *dominates* College i if $u_j \geqslant u_i$ and $\tau_j \leqslant \tau_i$ with at least one inequality being strict. If College j is in the portfolio, College i adds no value, and if College j is not in the portfolio, replacing College i by College j increases Ann's expected payoff. Therefore, there always exists an optimal portfolio that excludes dominated colleges. Colleges i and j are *replicas* if $u_j = u_i$ and $\tau_j = \tau_i$. Ann does not gain from applying to both colleges i and j, since she chooses at most one of them, and if she is accepted by one college, then she is accepted by the other. We prune away dominated colleges and replicas.

Assumption 1. *For any two colleges i and j, if $i < j$ then $u_i > u_j$ and $\tau_i > \tau_j$.*

Given Assumption 1, we say that College i is *higher ranked* than College j if $i < j$; higher ranked colleges are strictly more selective and more desirable to Ann.

Our second assumption restricts attention to beliefs such that the selectivity of colleges has implications for Ann's beliefs about the likelihood of acceptance.

Assumption 2. *Ann believes that less selective colleges are more likely to accept her: we consider distributions F such that if $i < j$, then $F^{-}(\tau_i) = F^{-}(\tau_j)$ implies that $F^{-}(\tau_i) \in \{0, 1\}$.*

Assumption 2 means that if Ann believes that she is equally likely to be accepted by College i and a less selective College j, then she believes that she is guaranteed to be accepted or rejected by each college; otherwise, she ascribes strictly positive probability to College j accepting her application and College i rejecting it. A special case of Assumption 2 is a distribution F that is full support on $[0, 1]$.

Third, we assume that Ann does not make applications that she views to be redundant. In principle, if the marginal cost of applications is 0, an optimal portfolio could include applications to colleges that she knows would reject her or that she would surely reject in favor of her outside option. Such applications add no value to Ann, and so there exist optimal portfolios that exclude such redundant applications. We focus on portfolios in which no application is redundant.

Assumption 3. *Ann does not apply to colleges that would reject her or that she would surely reject: the set of feasible portfolios is $\mathcal{P}(U, F) \equiv \{P \subseteq C : \text{for all } \tilde{P} \subsetneq P,\ V(\tilde{P}, U, F) < V(P, U, F)\}$.*

Assumption 3 asserts that Ann chooses from portfolios that are minimal in terms of set inclusion: for any portfolio in the feasible set, removing a college from that portfolio reduces her value.

We impose a genericity assumption that rules out indifference.

Assumption 4. *Every pair of feasible portfolios delivers distinct values: $V(P, U, F) \neq V(\tilde{P}, U, F)$ for distinct P and $\tilde{P}$ in $\mathcal{P}(U, F)$.*

As mentioned in our discussion after (5), V can be summarized as a function of two finite-dimensional vectors U and $(F^{-}(\tau_1), \ldots, F^{-}(\tau_n))$. The assumption fails for a non-generic set of parameters. In those cases, our results apply with the caveat that we are specifying *an* optimal solution.

Finally, we assume that given her outside option and her beliefs, applying to college is neither worthless nor hopeless for Ann.

Assumption 5. *We consider (U, F) such that there exists at least one College i such that $u_i > \overline{u}$ and $F(\tau_i) < 1$.*

Discussion: The modeling assumption of a score may be motivated in at least three ways. First, it may represent some component of Ann's ability (relative to the application pool) that is observable to colleges but not to her, as modeled in Avery and Levin (2010). Second, it may be an explicit test score that is used to determine admissions, as is done in college or high school admissions in many parts of the world. Third, it may be the realization of a common lottery used to break ties as has been done in centralized matching algorithms that use "single tie-breaking" rules.

Although we formulate the problem in terms of a score and thresholds, it is not necessary. Suppose Ann's (marginal) beliefs about college admissions prospects are represented by $(\alpha_1, \ldots, \alpha_n)$, and Ann believes that if her application is rejected by College i, then it is rejected by every higher ranked college. This would map to our model with $F^{-}(\tau_i) = 1 - \alpha_i$, and trivially, every vector can be mapped to a score distribution and thresholds.

Finally, we represent the problem as specifying a *set* of options, but an equivalent formulation is that of a *rank-ordered list* where Ann explicitly ranks the colleges in the set and is automatically enrolled in the college that she ranks highest among those that accept her.

4 Elementary Properties of College Portfolio Solutions

Here we study how Ann's portfolio choice is influenced by her beliefs or risk attitudes. We show that Ann applies less aggressively when she obtains bad news (in the sense of the likelihood ratio dominance order). We also show that she applies more aggressively if she becomes more "risk loving" in a sense that we formalize. While these two effects may appear to be distinct, we show that each is the mathematical dual of the other. These elementary results generalize the bad news and increased outside option effects that we illustrated in Section 2, and are at the core of our subsequent analysis.

We obtain these properties by considering a special case of Ann's problem. Suppose that Ann can apply to at most k colleges. Her problem then is

$$P^*(k, U, F) \equiv \underset{\{P \in \mathcal{P}(U,F): |P| \leqslant k\}}{\operatorname{argmax}} V(P, U, F).$$

One can think of this problem as Ann incurring application costs of 0 if her portfolio includes k or fewer applications, and a cost of infinity otherwise. But, as we show, this special case is germane for general costs since Ann's decisionmaking can be thought of as identifying the optimal portfolio of k items for each k, and then choosing the utility-maximizing value of k.

Let us formalize what it means for one portfolio to be *more aggressive* than another, adapting the definition of Chade and Smith (2006). For a portfolio P, let $P^{(i)}$ be the i^{th} *highest ranked* element of the portfolio. By Assumption 1, higher ranked colleges are both more attractive and more selective.

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 $\succeq_A$.

A more aggressive portfolio targets higher ranked schools. Case (a) applies if the portfolios being compared have the same number of applications. Here, we use the standard vector dominance order, stipulating that the i^{th} best college in P is higher ranked than that in $\tilde{P}$. The other two cases extend this definition to portfolios that differ in the number of applications.[3] Case (b) applies if P has fewer applications than $\tilde{P}$. Here, we say that portfolio P is more aggressive than $\tilde{P}$ if P is more aggressive (according to case (a)) than the portfolio that has the top $|P|$ items of $\tilde{P}$. One can then view $\tilde{P}$ as taking a portfolio that is already less aggressive than P by case (a) and expanding downwards to include colleges that are even less selective. Case (c) applies if P includes more applications than $\tilde{P}$. Here, the condition stipulates that even after omitting the top $|P| - |\tilde{P}|$ best colleges from P, we still have a portfolio that is more aggressive than $\tilde{P}$. One can then view P as taking a portfolio that is already more aggressive than $\tilde{P}$ and expanding upwards to include colleges that are even more selective.

4.1 How Beliefs Affect Portfolio Choice

Consider the following hypothetical scenario: Ann is about to submit her applications, but right before she does so, she obtains bad news that makes her more pessimistic about her admissions prospects. How should this influence her portfolio choice?

To answer this question, let us first formalize bad news. Given a cumulative distribution function G, let $\mu(i, G)$ denote the probability that College i is the best college that would accept an application from Ann. Setting $\tau_0 = 1$, notice that College i is the best college that would accept Ann's application if her score is in the interval $[\tau_i, \tau_{i-1})$. Therefore, $\mu(i, G) \equiv G^-(\tau_{i-1}) - G^-(\tau_i)$ where $G^-(\cdot)$ is the left-continuous version of G. We say that one distribution has bad news relative to another if it has relatively lower odds of obtaining entry into more selective colleges.

Definition 2. *Distribution H has **bad news** relative to G if for every College i and less selective College j,*

$$\mu(j, G)\mu(i, H) \leqslant \mu(i, G)\mu(j, H). \tag{6}$$

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

Definition 2 adapts the standard definition of the likelihood ratio dominance order to our setting: re-arranging (6) implies that distribution G has a relatively higher likelihood ratio of Ann being admitted into College i versus the less selective College j than distribution H. Definition 2 is implied by score distributions being ranked by the standard definition of the likelihood-ratio dominance order; namely that if G and H have densities, then $g(s)h(s') \leqslant g(s')h(s)$ for all $s' > s$. The order $\succeq_{LR}$ is less stringent because it compares the distributions only at the score thresholds for admissions.[4]

[3]Comparing portfolios of different cardinality, as in cases (b) and (c), is necessary for our analysis because Ann may obtain information that rules out or guarantees admissions at certain colleges. In both cases, by Assumption 3, this news may shrink the size of her chosen portfolio.

[4]Because of this distinction, G need not first-order stochastically dominate H. For example, returning to the colleges and cutoffs of Section 2, suppose that $G(s)$ puts probability $\frac{1}{2}$ on scores $\{0, 1\}$ whereas H is uniform on $[0, 1]$. Even though $G \succeq_{LR} H$, neither distribution first-order stochastically dominates the other.

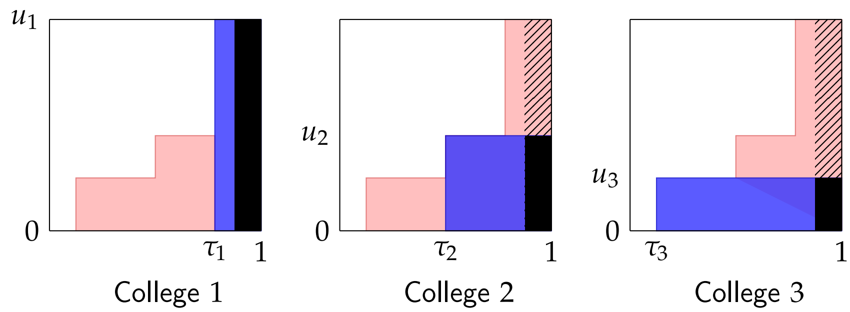

Figure 5: Illustrating Lemma 1. We depict the value of single-college portfolios, as in Figure 1, but where Ann learns that her score is lower than 0.875. Bad news reduces the area of the rectangle corresponding to each college, and this reduction is the black region of each rectangle. The rectangle of College 1 incurs a greater reduction than that of College 2, which in turn, incurs a greater reduction than that of College 3.

Lemma 1. *Bad news leads to a less aggressive portfolio:*

$$G \succeq_{LR} H \Rightarrow P^*(k, U, G) \succeq_A P^*(k, U, H).$$

Lemma 1 shows that shifting Ann's beliefs downwards (in the sense of likelihood ratios) induces her to apply less aggressively. We call this *the bad news effect*. We have already seen it at play in the example in Section 2. Therein we noted that the optimal single-college portfolio (based on Ann's prior) was College 2. What Equations (1) and (2) show is that if Ann were to (hypothetically) obtain the bad news that her score is below 0.78, the optimal single-college portfolio for her posterior belief shifts down to College 3. Lemma 1 shows that being less aggressive in response to bad news is a general property of Ann's decisionmaking.

Although the argument for Lemma 1 is involved, we can describe the key idea. Consider the special case where H is a right truncation of G: all that Ann learns is that her score is below some threshold τ so that $H(s) = G(s|s \leqslant \tau)$. We see in Figure 5 that this form of bad news removes more area from rectangles that correspond to more selective schools, and this makes less selective schools relatively more attractive (in terms of the area they cover). This logic suggests that with bad news, Ann should be less aggressive at least in terms of the most selective college to include in her portfolio. Stepping outside of this special case, the general ranking from the likelihood ratio dominance order does not correspond to removing slices from the right of each rectangle. But it nevertheless has the property that more area is removed from the right than from the left. This intuition suggests that Ann should be less aggressive in terms of the most selective college to include in her portfolio.[5]

What about the other colleges in the portfolio? Once we have shown that Ann is less aggressive in terms of her top choice, this downward force propagates to her choice of other colleges to include in her portfolio. The reason is that Ann chooses these remaining colleges optimally assuming that she is rejected by the top-ranked college in her portfolio, and these posterior beliefs remain LR-ordered. To see why, suppose that College i is the most selective college in her optimal k-portfolio when her

[5]This intuition is partial because Ann may conceivably choose to be more aggressive in her second or third top choice when her beliefs are lowered in the sense of Definition 2. This would then create a countervailing force that could conceivably push her to be more aggressive in her top choice. Addressing this potential complication is a significant component of our proof where we show that indeed Ann's top choice is less aggressive when her beliefs are lowered.

distribution is G and College j is that when her distribution is H. Suppose that, as argued above, College i is more selective than College j, i.e., $\tau_i \geqslant \tau_j$. In each portfolio, Ann chooses the remaining colleges optimally conditional on being rejected by the top-ranked college in the portfolio (since her choice matters only when rejected by the top-ranked college in the portfolio). But because $G \succeq_{LR} H$ and $\tau_i \geqslant \tau_j$, it follows that Ann's posterior beliefs conditional on rejection from each top choice are also LR-ordered:

$$G(\cdot|s < \tau_i) \succeq_{LR} H(\cdot|s < \tau_j).$$

Therefore, by induction, the remainder of Ann's optimal portfolio when her beliefs are described by G is more aggressive than that when her beliefs are described by H.

4.2 How Risk Attitudes Affect Portfolio Choice

Now consider an alternative hypothetical scenario: 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 change influence her portfolio choice?

We think of this more broadly as a change in Ann's *risk sensitivity*. Improving her outside option makes her more willing to take risks since being rejected by all of the colleges in her portfolio is less costly. We formalize this idea by adapting standard notions of risk sensitivity.

Definition 3. *Utility assessment $U' \equiv (\overline{u}'; u'_1, \ldots, u'_n)$ is more **risk loving** than utility assessment $U \equiv (\overline{u}; u_1, \ldots, u_n)$ if there exists a convex non-decreasing transformation $\nu : \Re \to \Re$ such that for every College i,*

$$\max\{u'_i, \overline{u}'\} = \nu(\max\{u_i, \overline{u}\}). \tag{7}$$

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

Definition 3 adapts a standard notion of being risk-loving to Ann having an outside option. If in U and U', each college were to be better than the outside option, and the outside options were identical, Definition 3 is the usual Arrow–Pratt definition of being risk loving.

Lemma 2. *Being more risk loving leads to a more aggressive portfolio:*

$$U' \succeq_{RL} U \Rightarrow P^*(k, U', G) \succeq_A P^*(k, U, G).$$

Lemma 2 shows that being more risk loving induces Ann to apply more aggressively. We have already seen a version of this exhibited in our example in Section 2 where we discussed how having a better outside option induces Ann to shoot for College 1 rather than College 2. An increase in Ann's outside option naturally makes her more risk loving: keeping all of Ann's utilities from each college the same but increasing her outside option from $\overline{u}$ to $\overline{u}'$ implies that when thinking about her portfolio, the utility of being admitted to College i is $\max\{u_i, \overline{u}'\} = \max\{\max\{u_i, \overline{u}\}, \overline{u}'\}$, which reflects a convex non-decreasing transformation of $\max\{u_i, \overline{u}\}$. Therefore, the increased outside option effect that we alluded to in Section 2 is a special case of the risk-loving effect established in Lemma 2.[6]

We prove Lemma 2 by using a "duality" between risk sensitivity and beliefs, which we alluded to in Section 2. To see the connection, let us consider two utility assessments U and U' where U' is more risk loving, and a fixed probability distribution F. Let us normalize each outside option to 0 and each utility to be a number in $[0, 1]$, reflecting the proportional gain over the outside option relative to the gain from Ann's favorite college.[7] As illustrated in Figure 4 on p. 6, we can transpose this problem over the off-diagonal of the unit square (the -45 degree line) to flip utilities into probabilities and probabilities into utilities. Let $T[U]$ and $T[U']$ be distributions that are the transposed versions of U and U' respectively. We show that

$$U' \succeq_{RL} U \Rightarrow T[U'] \preceq_{LR} T[U].$$

In other words, if U' is more risk loving than U, then the distribution $T[U']$ has bad news relative to the distribution $T[U]$.[8] Lemma 1 shows that bad news induces Ann to choose a less aggressive portfolio, which implies that in the isometric transposed problem, she chooses a more aggressive portfolio when she is more risk loving.

5 Optimal Diversification

In this section, we prove our main result that Ann should diversify her portfolio when application costs reduce. We do this in two steps. First, we show that if she can apply to more colleges, she should expand the range of colleges to which she can apply in a particular way. Second, it then follows that she pursues a more diversified portfolio when the marginal costs of each application is lowered.

Given a portfolio P whose cardinality weakly exceeds k, let $\lceil P \rceil^k$ and $\lfloor P \rfloor^k$ denote the set of the best and worst k colleges in P. If the cardinality of P is less than k, then $\lceil P \rceil^k = \lfloor P \rfloor^k = P$. Using this notation, we state our result.

Theorem 1. *Larger optimal portfolios are more diverse: if $\tilde{k} > k$, then*

$$\lceil P^*(\tilde{k}, U, F) \rceil^k \succeq_A P^*(k, U, F), \tag{8}$$

$$\lfloor P^*(\tilde{k}, U, F) \rfloor^k \preceq_A P^*(k, U, F). \tag{9}$$

Theorem 1 asserts that being able to apply to more colleges induces Ann to be both more aggressive at the top and less aggressive at the bottom. This implies that Ann is "spreading out" her pool of applications. We illustrate this property in Figure 6 where we see how Ann's portfolio depends on

[6] This increased outside option effect may be germane to the role of unequal outside options in segregation. If colleges represent *public schools* whose admissions is determined by a centralized examination or matching procedure, the outside option may reflect Ann's value from attending a *private school* outside that system. Lemma 2 implies that if Ann and Bob have the same preferences and beliefs about their admissions prospects, but differ in that Bob is richer and more easily able to attend a private school, then he will apply more aggressively. Unequal outside options therefore lead to segregation whereby poorer students apply less aggressively and therefore are more likely to attend less selective schools.

[7] For utility $U = (\overline{u}; u_1, \dots, u_n)$, we use the normalization $\left(0; 1, \dots, \frac{u_n - \overline{u}}{u_1 - \overline{u}}\right)$.

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

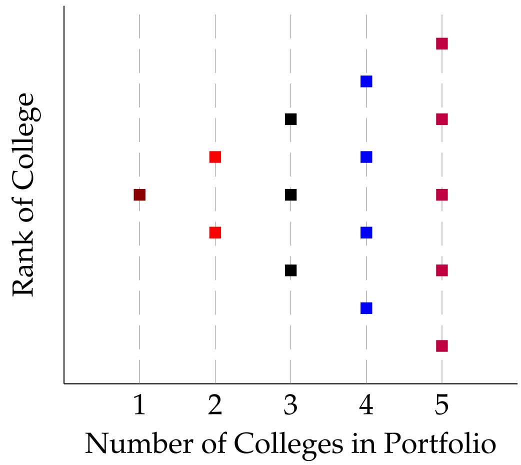

Figure 6: Illustrating Theorem 1. Ann expands the ranges of colleges to which she applies both upwards and downwards when she can apply to more colleges. The horizontal axis shows the number of colleges in a portfolio and the vertical axis shows the rank / selectivity / desirability of colleges.

the number of colleges to which she applies. When she can apply to more colleges, she expands this range by including both higher and lower ranked colleges.

To further interpret Theorem 1, suppose that $\tilde{k} = k + 1$ so that we are studying how Ann's optimal portfolio choice expands when she can apply to just one more college. The ordering shown in (8) says that the k best colleges in Ann's optimal $(k+1)$-portfolio is more aggressive than her optimal k-portfolio. So apart from her least preferred college in her optimal $(k+1)$-portfolio, all of Ann's colleges have shifted up relative to the optimal k-portfolio. The ordering shown in (9) pushes in the opposite direction: *except* for the top ranked alternative in the optimal $(k+1)$-portfolio, Ann's optimal $(k+1)$-portfolio is *less* aggressive than her optimal k-portfolio.

How can these two conclusions be mutually consistent? The only way is that (i) Ann's highest ranked college in her optimal $(k+1)$-portfolio is more of a reach than her highest ranked college in her optimal k-portfolio, (ii) Ann's lowest ranked college in her optimal $(k+1)$-portfolio is more of a safety than her lowest ranked college in her optimal k-portfolio, and (iii) the remaining colleges "refine" the grid from her optimal k-portfolio.

Although we have illustrated this when Ann can apply to only one more college, the idea generalizes to when she can apply to $\tilde{k} - k$ more colleges. As we see in Figure 6, as one moves from the optimal k-portfolio to the optimal $\tilde{k}$-portfolio where $\tilde{k} > k$, the highest ranked k items in the optimal $\tilde{k}$-portfolio are higher ranked than the optimal k-portfolio. Analogously, the lowest ranked k items in the optimal $\tilde{k}$-portfolio are lower ranked than the optimal k-portfolio. Ann uses greater flexibility to diversify her portfolio in both directions.

We prove Theorem 1 by combining ideas from Lemmas 1 and 2. Let us argue why the result is true for $\tilde{k} = k + 1$ because the general analysis will then follow by induction. Ann's selection of an optimal $(k+1)$-portfolio can be viewed as reflecting a two-stage optimization process: first, she picks which college is the highest ranked college in her portfolio (her "first choice"), and second, an optimal portfolio of k backup options. Notice that these k backup options matter *only if* she is rejected by her first choice, and so she should choose this portfolio of k backup options conditioning on that event. But conditioning on that event generates posterior beliefs that are likelihood-ratio dominated by her prior. The bad-news effect (formalized in Lemma 1) implies that the optimal portfolio of k backup options must be less aggressive than the optimal k-portfolio she would have chosen based on her prior alone, which corresponds to (9).

To see why (8) must also be true, she could have viewed the selection of an optimal $(k+1)$-portfolio through the lens of a different two-stage optimization process: first, she picks which college is the lowest ranked in her portfolio (her "safety school"), and second, an optimal portfolio of k colleges that are more selective ("improvements"). Suppose she chooses College i to be her safety school; by Assumption 3, going to this college generates a utility of u_i that exceeds that of her outside option $\overline{u}$. Her choice of the other k colleges to include in her portfolio matters only if her score guarantees that she is at least admitted into College i because otherwise she obtains $\overline{u}$ regardless of her choice. Conditioning on this event—namely $s \geqslant \tau_i$—implies that if she is rejected by the remaining k colleges, she obtains u_i. Choosing these remaining k colleges is therefore isomorphic to a modified optimal k-portfolio problem in which her outside option is now u_i. Because this modified problem has a better outside option than that of the original problem, the increased outside option effect (formalized in Lemma 2) implies that the optimal portfolio of k improvements over College i is more aggressive than the optimal k-portfolio.

These results have implications for Ann's behavior when her application costs reduce for a general cost function. Recall that $\phi(|P|)$ denotes the total application cost incurred for Ann to have the portfolio P. The marginal cost of the m^{th} application is $\phi(m) - \phi(m-1)$. Marginal costs are lower in cost function $\tilde{\phi}$ relative to ϕ if the marginal cost for each m is lower in $\tilde{\phi}$ relative to ϕ. The following describes how reductions in application fees influence Ann's behavior.

Corollary 1. *If marginal costs are reduced, Ann applies to a larger and more diverse portfolio: she applies to more colleges, and by Theorem 1, she is more aggressive in her selection of the highest ranked college to include in her portfolio and less aggressive in her selection of the lowest ranked college to include in her portfolio.*

These results illustrate why applicants should diversify their portfolios when they can apply to more colleges. In practice, it has been seen that reductions to application costs or allowing students to apply to more colleges does expand the range of colleges to which they apply (Ajayi, 2013; Pallais, 2015).[9] Our model also predicts that Ann may apply to a safety school: in our example in Section 2, College 3 is not only less selective than Colleges 1 and 2 but is also dominated by each in terms of ex ante expected utility as a single college. Nevertheless Ann includes College 3 in her optimal two-college portfolio as her backup if she is rejected by College 1.[10] Our analysis matches the common wisdom that when applicants can apply to several colleges, they should include some reaches, some matches, and some safety schools.

[9]Moreover, an important friction that applicants face is informational, and one way to reduce application costs is by providing applicants with information. Hoxby and Avery (2013) find that providing low-income applicants with information about selective colleges also induces them to apply to more selective colleges.

[10]As Chade and Smith (2006) show, this would not happen if college admissions decisions were stochastically independent.

6 Other Results

6.1 Early Admissions Policies

Many colleges have early admissions programs: applicants can send in a single early application and obtain a decision prior to applying to other colleges. These programs are either *early decision* where a student is committed to enroll if her application is accepted or *early action* where a student obtains an admissions decision but is not committed to enroll if her application is accepted. In this section, we study how this affects Ann's portfolio choice.

We consider a stylized setting: Ann can apply to a single college ahead of time, and then following that application, she can apply to k colleges in the regular admissions process. To focus on the effect most relevant to our work, we assume that each college has the same score threshold for both early and regular decisions. We first see what Ann would do if every early admissions program is early decision, namely if she is accepted by the college to which she applies early, she is committed to enrolling in that college.

Theorem 2. *If Ann can apply to a single college early decision and to k colleges in the regular cycle, she obtains the same distribution over outcomes as applying to $(k+1)$ colleges simultaneously:*

- *she applies early to the highest ranked college in $P^*(k+1, U, F)$;*
- *if rejected by that college, she applies to the remaining colleges in that portfolio.*

Theorem 2 illustrates that early decision offers limited informational value to Ann. Here is why: suppose Ann chooses to apply early decision to College i. She enrolls in that college if she is accepted. If she is rejected, she infers that her score is lower than τ_i. Therefore, her posterior belief for the regular cycle is $F(\cdot|s < \tau_i)$, and she would choose the remaining k colleges that are optimal for those beliefs. But notice that this is isomorphic to her choice when formulating her optimal $(k+1)$-portfolio: she chooses the highest ranked college to which she applies and the best k colleges conditional on being rejected by that college.

By contrast, early action programs convey significant advantages.

Theorem 3. *If Ann can apply to a single college early action and to k colleges in the regular cycle, she obtains the same distribution over outcomes as applying to $(2k+1)$ colleges simultaneously:*

- *she applies early to the $(k+1)^{th}$ ranked college in $P^*(2k+1, U, F)$;*
- *if accepted by that college, she applies to the (at most) k colleges in that portfolio that are ranked higher;*
- *if rejected by that college, she applies to the (at most) k colleges in that portfolio that are ranked lower.*

The logic of Theorem 3 is that because Ann is not committed to enrolling in a college that admits her early, she can use its admissions decision to direct her future search. Suppose she applies early action to College i. If she is rejected, matters are identical to that of an early decision program where she learns the bad news that she is rejected. By contrast, if she is accepted, she can apply to k more colleges knowing that if she is rejected by those k colleges, she has an outside option of attending College i. The increased outside option effect means that she should choose the remaining k colleges aggressively with this outside option. These deliberations are isomorphic to the choice that Ann would make when choosing her optimal $(2k+1)$-portfolio: that problem can be broken down into optimally choosing which is the best college for the median position in the portfolio, followed by the selection of k colleges above and below that median college.

Our comparison of early action and early decision illustrates a broader point: when college prospects are correlated, early action allows an applicant to direct her search and so she may not wish to apply to any of her k favorite colleges under early action. By contrast, with an early decision program, the fact that Ann is committed to enrolling in a college that admits her early implies that she cannot use this to help guide her applications process.

We should mention a caveat to our analysis: we omit the idea that applying early offers a signal to a college of an applicant's interest and can therefore boost admissions prospects. This effect has been modeled by Avery and Levin (2010) (and also discussed by Avery, Fairbanks, and Zeckhauser 2009) as a rationale for colleges to have early admissions programs. To see how this admissions boost would affect our results, suppose that if Ann applies early to College i, she is accepted if her score exceeds $\tau_i - b$, where $b > 0$. For early decision processes, if this boost is moderate, then it does not change the college to which Ann applies to early. Now if Ann is rejected, she learns that her score is below $\tau_i - b$, which is worse news than without the boost. Therefore, the remaining k colleges to which she applies are less aggressive than without the boost. For an early action program, the admissions boost has a similar effect if she is rejected by the college to which she applies early without having an effect on the colleges to which she applies if she is accepted (since the admissions boost does not interact with the increased outside option effect). Thus, if applying early boosts one's admissions prospects moderately, it may induce applicants to apply less aggressively than they would otherwise.[11]

6.2 A Comparison with Chade and Smith (2006)

Chade and Smith (2006) offer the first analysis of simultaneous search. In this section, we highlight how our work complements their study and the new forces that it introduces.

They study the case of independent admissions probabilities.[12] Each college is characterized by a pair (u_i, α_i), which denote the payoff of attending College i and the probability with which an application from Ann is accepted by College i, independently of whether her application would be accepted by other colleges. Suppose the value of the outside option, $\overline{u}$, is 0. For a portfolio P, we order its elements in descending order on the basis of the payoff of attendance so that $P^{(i)}$ is the i^{th} best element (in terms of payoffs). Then the value of a portfolio P in their setting is

$$V^{CS}(P) = \sum_{i}^{|P|} \alpha_{P(i)} u_{P(i)} \prod_{j}^{i-1} (1 - \alpha_{P(j)}).$$

[11] A second, more subtle, effect is on the degree to which colleges can infer an applicant's interest from her choice to apply early. Theorem 2 shows that for an early decision program, a college that receives an early application would know that the applicant is applying early to her top choice. By contrast, for an early action program, a college that receives an early application would know that it is not this applicant's first choice.

[12] Their results extend to slight degrees of correlation sufficiently close to stochastic independence.

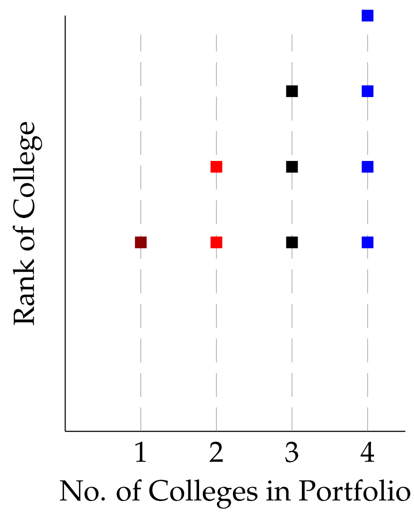

Figure 7: Illustrating optimal portfolios in Chade and Smith (2006). Ann expands the ranges of colleges to which she applies only upwards when she can apply to more colleges.

The idea is that Ann attends the i^{th} best college in portfolio P (obtaining utility $u_{P(i)}$) if:

- she is accepted by that college, which occurs with probability $\alpha_{P(i)}$, and
- she is rejected by all better colleges, which occurs with probability $\prod_j^{i-1}(1-\alpha_{P(j)})$.

They derive a number of elegant results in this setting. We summarize those that most easily compare with ours. Their Theorem 2 argues, as they interpret, that applicants should not apply to safety schools: the optimal k-portfolio is more aggressive than the portfolio of the k items that maximize $\alpha_i u_i$.[13] Their Theorem 1 identifies a “marginal improvement algorithm" that generates the optimal k-portfolio, and shows that the optimal k-portfolio is nested within the optimal $(k+1)$-portfolio. They suggest that this nesting should take the form of expanding the set only in an upwards direction (i.e., more aggressively), and we formalize this logic in our Supplementary Appendix. Figure 7 illustrates this form of nesting. Therefore we see that an applicant would not use greater flexibility to diversify her portfolio, in contrast to the predictions of our model.

We can use our decomposition of choice behavior into the bad news and increased outside option effects to explain why the correlation structure of our model generates a preference for diversification while independent admissions probabilities would lead someone to expand only upwards. As we discussed in Section 5, the bad news effect of being rejected by top schools motivates Ann to expand downwards and the increased outside option effect of applying to more colleges at the bottom induces Ann to expand upwards. If admission decisions are stochastically independent, then there is no bad news in being rejected by one’s top ranked schools and so the first effect disappears. But the second effect remains, and hence as depicted in Figure 7, an applicant would expand only upwards.

These results show the relationship of our work to Chade and Smith (2006). Their work introduces the idea that one obtains the prize of attending a college only if one is rejected by all better colleges in one’s portfolio. They suggest (and we agree) that the independent admissions framework

[13] Section 2 shows that this property does not hold in our model: the optimal two-college portfolio in our model includes College 3 even though $\alpha_3 u_3 < \min\{\alpha_1 u_1, \alpha_2 u_2\}$.

depicts settings in which an applicant knows her true caliber (and where she lies on the distribution of applicants) and colleges obtain conditionally i.i.d. signals of her caliber. But it generates conclusions about the rationale of safety schools and diversification that are difficult to reconcile with the observed practice of applicants and the advice commonly given to them. Moreover, in some settings, college admissions processes are either highly correlated by design or perceived to be so.

7 Discussion

This paper offers an elementary treatment of a stylized college portfolio problem. A student applies to a set of colleges knowing that her admissions probabilities are correlated in a specific way across colleges. This correlation may arise because (i) her admissions prospects are determined by a common examination (as in many countries), or (ii) different colleges use very similar scoring methods and the student is uncertain about the exact criteria, or (iii) a common lottery is used to break priority ties in a centralized algorithm. We show that the student pursues a diversified portfolio in which she applies to a mix of reach, match, and safety schools. Reducing application costs induces her to diversify her portfolio even further. This pattern appears to match diversification strategies that applicants pursue in practice, and how applicants have behaved when application fees have decreased.

We have interpreted this portfolio problem in the context of college applications but it may be relevant to other settings too. Firms often face a choice between multiple marketing or organizational strategies and can adopt only one. These strategies may be correlated in higher return strategies succeed only if medium return strategies would also succeed. Such a firm may wish to do small-scale experiments on a diversified set of strategies before picking one. Closer to home, an early-stage researcher or PhD student may choose a set of research problems to work on simultaneously but will have time to write only one. Some projects are low risk and low return whereas others are high risk and high return. Success in these problems may depend on the researcher's ability, and early-stage researchers may be uncertain about their ability, which introduces correlation into the perceived chances of success. Our analysis suggests that such researchers benefit from starting problems across the spectrum of difficulty, and pushing to completion the best project that bears fruit.

Although we do not view this kind of decision problem to be special, we have not seen a general treatment. A close parallel is the study of rank-dependent expected utility (henceforth RDEU) introduced by Quiggin (1982). Therein, a decisionmaker uses a probability weighting function that is rank-dependent. In our model, the value of a portfolio can also be seen as weighting utilities by changes in a rank-dependent cumulative probability. To see how, suppose that the outside option equals 0. The value of a portfolio P then is

$$\sum_{i=1}^{|P|} u_{P^{(i)}} \left(F^{-}(\tau_{P^{(i-1)}}) - F^{-}(\tau_{P^{(i)}}) \right),$$

where $P^{(i)}$ denotes the i^{th} best college and we set $F^{-}(\tau_{P^{0}}) = 1$. Notice that Ann obtains the utility from the i^{th} ranked item with a change in a rank-dependent cumulative probability. In formalizing the connection between LR-dominance and being risk loving, we rewrite the above expression as

$$\sum_{i=1}^{|P|}(1 - F^{-}(\tau_{P(i)}))\,(u_{P(i)} - u_{P(i+1)})\,,$$

where we set $u_{P|P|+1} = 0$. In making this switch, we are effectively treating the probability $(1 - F^{-}(\tau_{P(i)}))$ as the utility and $(u_{P(i)} - u_{P(i+1)})$ as a change in probability (if utilities are suitably normalized). The similarity in functional form with that of RDEU models suggests a similar duality between probability and utility may be useful for that setting. And indeed, Yaari (1987) and Segal (1989, 1993) approach RDEU models with that focus and also study the induced coverage problem.

It would be interesting to compare this simultaneous portfolio problem with a fully sequential procedure similar to Weitzman (1979) but in which prospects are correlated as we have modeled them. Our analysis of early action programs suggests that with sequential search, Ann may sample an option not merely because she might choose it but also because it offers information that guides her future search. It may be fruitful to study how Ann would optimally sample and search.

References

AJAYI, K. (2013): "School choice and educational mobility: Lessons from secondary school applications in Ghana," Working paper.

AJAYI, K. AND M. SIDIBE (2020): "An empirical analysis of school choice under uncertainty," Working paper.

AVERY, C., A. FAIRBANKS, AND R. J. ZECKHAUSER (2009): *The Early Admissions Game*, Harvard University Press.

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.

CARD, D. AND A. B. KRUEGER (2005): "Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas," *ILR Review*, 58, 416–434.

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.

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

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

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.

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

LARROUCAU, T. AND I. RIOS (2018): "Demand estimation in the college admissions problem: Should we assume truth telling?" Working paper.

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.

Quiggin, J. (1982): "A theory of anticipated utility," *Journal of economic behavior & organization*, 3, 323–343.

Segal, U. (1989): "Anticipated utility: A measure representation approach," *Annals of Operations Research*, 19, 359–373.

——— (1993): "The measure representation: A correction," *Journal of Risk and Uncertainty*, 6, 99–107.

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.

Yaari, M. E. (1987): "The dual theory of choice under risk," *Econometrica*, 55, 95–115.

A Main Appendix

A.1 Proof of Lemma 1 on p. 10

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

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

Our assumptions guarantee that $|P^*(k, U, G)|$ is then of the form $\{j, j+1, \ldots, j+m\}$ for some $m < k-1$.[14] Furthermore, $\mu(i, G) = 0$ for each $i < j$. Since $G \geq_{LR} H$, we have that $\mu(i, H) = 0$ for each $i < j$. Furthermore, if $\mu(i, H) > 0$ for some $i < j+m$, then $\mu(i', H) > 0$ for each $i' \in \{i, \ldots, j+m\}$. Hence, $P^*(k, U, G) \geq_A P^*(k, U, H)$, as the latter portfolio either consists of a bottom part of the former portfolio – potentially with some additional colleges in $\{j+m+1, \ldots, n\}$ – or of k colleges that include a subset of the colleges in the former portfolio in addition to colleges in $\{j+m+1, \ldots, n\}$.

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

As in the previous case, our assumptions guarantee that $|P^*(k, U, H)|$ takes the form $\{j, j+1, \ldots, j+m\}$ for some $m < k-1$, and that $\mu(i, H) = 0$ for each $i > j+m$ (otherwise adding i to the portfolio would be beneficial). Since $G \geq_{LR} H$, we have $\mu(i, G) = 0$ for each $i > j+m$. Thus, by the same argument as above, $P^*(k, U, G) \geq_A P^*(k, U, H)$.

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

We prove that $P^*(k, U, G) \geq_A P^*(k, U, H)$ by induction on k.

Base Step: $(k = 1)$

[14] By Assumption 5 the portfolio is nonempty, by Assumption 3 all the colleges on the portfolio strictly contribute to its value, and by Assumption 2 they must form a consecutive interval.

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

$$V(\{i\}, U, H) - \overline{u} = \left(1 - H^{-}(\tau_i)\right)\,(u_i - \overline{u}) \geqslant \left(1 - H^{-}\left(\tau_j\right)\right)(u_j - \overline{u}) = V(\{j\}, U, H) - \overline{u}. \tag{10}$$

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

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

where the first line follows by definition, the second follows from $G \succeq_{LR} H$, and the third follows by definition. Multiplying (10) with the inequality above yields

$$\left(1 - G^{-}(\tau_i)\right)\left(1 - H^{-}\left(\tau_j\right)\right)\left(1 - H^{-}(\tau_i)\right)(u_i - \overline{u}) \geqslant \left(1 - H^{-}(\tau_i)\right)\left(1 - G^{-}\left(\tau_j\right)\right)\left(1 - H^{-}\left(\tau_j\right)\right)(u_j - \overline{u}).$$

We note that $(1 - H^{-}(\tau_i)) > 0$ (otherwise the empty portfolio is optimal, contradicting Assumption 5), and that $\left(1 - H^{-}\left(\tau_j\right)\right) \geqslant (1 - H^{-}(\tau_i))$, since $\tau_j < \tau_i$. Hence, we can divide both sides of the inequality by the positive term $\left(1 - H^{-}\left(\tau_j\right)\right)(1 - H^{-}(\tau_i))$ and obtain

$$\left(1 - G^{-}(\tau_i)\right)(u_i - \overline{u}) \geqslant \left(1 - G^{-}\left(\tau_j\right)\right)(u_j - \overline{u}),$$

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

Inductive Step: $(k > 1)$

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 $P^{*(1)}(k, U, G) \leqslant P^{*(1)}(k, U, H)$ then the result follows from the inductive hypothesis. This follows from two observations. First,

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

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,

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

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

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

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

We now create a chain of portfolios $Q_0, \dots, Q_m$. 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 $P \equiv P^*(k,U,G)$ and $R \equiv P^*(k,U,H)$. Let $Q_0, Q_1, \dots, Q_m$ be portfolios such that $Q_i = \{R^{(1)}, \dots, R^{(i)}, P^{(i+1)}, \dots, P^{(k)}\}$. In other words, Q_i selects the top i colleges from portfolio R and the remaining $k-i$ colleges from P.[15] Observe that for each $j \leqslant i \leqslant m$ we have $Q_i^{(j)} = R^{(j)}$ and $Q_j \succeq_A Q_{j-1}$. It also follows that $Q_m \succeq_A P$ and $Q_m \succeq_A R$.

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 $i \leqslant m$,

$$V(Q_i,U,D) - V(Q_{i-1},U,D) = \left(D^-(\tau_{R^{i-1}}) - D^-(\tau_{R^i})\right)(u_{R^i} - u_{P^i}) - \left(D^-(\tau_{R^i}) - D^-(\tau_{P^i})\right)(u_{P^i} - u_{P^{i+1}}).$$[16]

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

$$V(Q_j,U,G) - V(Q_0,U,G) \leqslant 0. \tag{11}$$

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

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

where the inequality is (11) and the equality is computation. Therefore,

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

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

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

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

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

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

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

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

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

which implies

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

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 $W^l_D := (D^-(\tau_{R^{l-1}}) - D^-(\tau_{R^l}))\,(u_{R^l} - u_{P^l})$ and $L^l_D := (D^-(\tau_{R^l}) - D^-(\tau_{P^l}))\,(u_{P^l} - u_{P^{l+1}})$. W^l_D (respectively L^l_D) 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 $Q^{(l)}_{l-1} = P^l$ but not for $Q^{(l)}_l = R^l$). Additionally, denote $q_{R^l} := \frac{\mu(R^l, H)}{\mu(R^l, G)}$.[17]

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

$$\sum_{l=1}^{j} L^l_H = L^j_H + \sum_{l=1}^{j-1}\left(L^l_H - W^l_H + W^l_H\right) = L^j_H + \sum_{l=1}^{j-1} W^l_H + \sum_{l=1}^{j-1}\left(L^l_H - W^l_H\right) \\ \geqslant q_{R^j} L^j_G + \sum_{l=1}^{j-1} W^l_H + \sum_{l=1}^{j-1} q_{R^l}\left(L^l_G - W^l_G\right),$$

where the inequality follows from $W^l_H / W^l_G = (H^-(\tau_{R^{l-1}}) - H^-(\tau_{R^l}))\,/\,(G^-(\tau_{R^{l-1}}) - G^-(\tau_{R^l}))$ being bounded above by q_{R^l}, and $L^l_H / L^l_G = (H^-(\tau_{R^l}) - H^-(\tau_{P^l}))\,/\,(G^-(\tau_{R^l}) - G^-(\tau_{P^l}))$ being bounded be-

[17] 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).

low by q_{R^l} for all $l \leqslant m$. By rewriting the RHS of the above inequality, it then follows that

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

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

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

contradicting the optimality of $P^*(k,U,H) \equiv R$. 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 $\tau_{R^m} > \tau_{P^m}$ by Assumption 1). It must be strictly positive by Assumption 4. □

A.2 Proof of Lemma 2 on p. 12

Proof. Our argument proceeds in multiple steps.

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

For utility assessment $U \equiv (\overline{u}; u_1, \ldots, u_n)$, we denote its normalization by

$$N[U] \equiv (0; \tilde{u}_1, \ldots, \tilde{u}_n) \text{ where } \tilde{u}_i \equiv \max\left\{\frac{u_i - \overline{u}}{u_1 - \overline{u}}, 0\right\}. \tag{12}$$

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 $G^-(\tau_{P0}) = 1$, we can write the value of a portfolio as

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

where the first equality is the definition of a portfolio's value (Equation 5 on p. 7), the second equality is calculating the integral, the third equality is algebra, the fourth equality substitutes (12), 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:

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

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, $\overline{u}$, 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 $\lambda_j \leqslant \lambda_i$, 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.[18]

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, $U \equiv (0; u_1, \ldots, u_n)$, let

[18]It suffices to consider a discrete probability distribution that puts positive probability only on scores in $\{\tau_1, \ldots, \tau_n, 0\}$.

$w(U) \equiv (u_1, \ldots, u_n)$ 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 $U \equiv (0; u_1, \ldots, u_n)$ 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

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

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

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.**

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

$$V\left(P, U, G\right) = V(\mathcal{T}[P], U_T(G), G_T(U)). \tag{13}$$

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

$$V\left(P, U, G\right) = u_{P(1)} \underbrace{\left(1 - G^-(\tau_{P(1)})\right)}_{\text{Accepted by } P^{(1)}} + \sum_{i=2}^{k} u_{P(i)} \underbrace{\left((1 - G^-(\tau_{P(i)})) - (1 - G^-(\tau_{P(i-1)}))\right)}_{\text{Accepted by } P^{(i)} \text{ but not by } P^{(i-1)}}. \tag{14}$$

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. Re-arranging the RHS of the above expression yields

$$\left(\sum_{i=1}^{k-1}(1-G^{-}(\tau_{P(i)}))\,(u_{P(i)}-u_{P(i+1)})\right)+(1-G^{-}(\tau_{P(k)}))u_{P(k)}. \tag{15}$$

Let $\hat{U} \equiv U_T(G)$ be the transposition of the distribution and $\hat{G} \equiv G_T(U)$ be a transposition of the utilities. Finally, let $\hat{P} \equiv \mathcal{T}[P]$ 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 (15) 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 (14) is equal to $V\left(\hat{P},\hat{U},\hat{G}\right)$.

Step 4: **Being more risk loving implies bad news in the transposed problem.**[19]

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

$$U' \succeq_{RL} U \Rightarrow G_T(U') \preceq_{LR} G_T(U).$$

Observe that $U' \succeq_{RL} U$ implies that there exists a convex nondecreasing function $\nu : \Re \to \Re$ 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 $j > i$,

$$(u'_{i+1}-u'_i)(u_{j+1}-u_j) \geqslant (u_{i+1}-u_i)(u'_{j+1}-u'_j). \tag{16}$$

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 $u_{j+1} = u_j$ then $u'_{j+1} = u'_j$ and both sides are equal to 0. Otherwise, (16) 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 (16) to argue that $G_T(U') \preceq_{LR} G_T(U)$. Let $G_T(U') = \tilde{H}$ and $G_T(U) = \tilde{G}$. Note that

$$\mu(i,\tilde{H}) = \tilde{H}^{-}(\tau_{i-1}) - \tilde{H}^{-}(\tau_i) = u'_{n+1-i} - u'_{n+2-i}, \tag{17}$$

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

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

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 (17), the second uses (16), and the third uses (17). Therefore, we see that $G_T(U') \preceq_{LR} G_T(U)$.

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

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

$$U' \succeq_{RL} U \Rightarrow G_T(U') \preceq_{LR} G_T(U).$$

Therefore, by Lemma 1, for every k,

$$P^*(k, U_T(G), G_T(U)) \succeq_A P^*(k, U_T(G), G_T(U')). \tag{18}$$

It then follows that

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

where the equalities follow from Step 3 and the ordering follows from (18). □

A.3 Proof of Theorem 1 on p. 13

Proof. We first prove the result for $P^*(k+1, U, F)$ and $P^*(k, U, F)$.

We observe that for all k, $|P^*(k, U, F)| \leqslant |P^*(k+1, U, F)|$. If $|P^*(k, U, F)| = |P^*(k+1, U, F)|$, then $|P^*(k+1, U, F)| \leqslant k$, and therefore $P^*(k, U, F) = P^*(k+1, U, F)$, in which case we are done.

Now suppose that $|P^*(k+1, U, F)| = k+1$. Let College i denote the highest ranked college in $P^*(k+1, U, F)$ and let $G(s) = F(s | s < \tau_i)$ be Ann's belief about her score conditional on being rejected by College i. Observe that $G \preceq_{LR} F$.

Since Ann should choose the colleges in $P^*(k+1, U, F)$ other than College i assuming she is rejected from College i, it follows that

$$P^*(k+1, U, F) = \{i\} \bigcup P^*(k, U, G). \tag{19}$$

Therefore,

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

where the equality follows from (19) and the ordering follows from $G \preceq_{LR} F$ and Lemma 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 $U' \succeq_{RL} U$. 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

$$P^*(k+1, U, F) = \{j\} \bigcup P^*(k, U', F). \tag{20}$$

Therefore,

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

where the equality follows from (20) and the ordering follows from $U' \succeq_{RL} U$ and Lemma 2.

Therefore, we have shown that Theorem 1 holds for $\tilde{k} = k+1$. The argument follows for general k and $\tilde{k} > k$ by induction, using the transitivity of $\succeq_A$. □

A.4 Proofs for Section 6.1

Proof of Theorem 2 on p. 16. Ann can apply to a single college ahead of time, and may apply to an additional k colleges if she is rejected by that college. A strategy for Ann involves no more than $k+1$ schools. Therefore, an optimal strategy cannot achieve more utility than applying to all these colleges simultaneously and choosing the best one. Hence, the strategy specified in the statement must be optimal since it achieves the highest utility among all all portfolios of up to $k+1$ schools. □

Proof of Theorem 3 on p. 16. Ann can apply to a single college ahead of time, and may apply to an additional k colleges if she is rejected by that college and to (potentially different) k colleges if she is accepted by that college. Altogether, a strategy for Ann involves no more than $2k+1$ schools. Therefore, an optimal strategy cannot achieve more utility than applying to all these colleges simultaneously and choosing the best one. Hence, the strategy specified in the statement must be optimal since it achieves the highest utility among all all portfolios of up to $2k+1$ schools. □

B Supplementary Appendix (Not for Publication)

Below, we provide the analysis supporting our discussion in Section 6.2. We show that when admissions probabilities are independent, Ann should expand her portfolio only upwards. Chade and Smith (2006) suggest this result under certain conditions, and using the logic they outline, we prove it more generally.

So as to be self-contained, the set of college types $C = \{1, \ldots, n\}$ comprises n colleges. Being admitted into a college ot type i generates utility u_i. 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 also remove Assumptions 1 and 2 and relax Assumption 4 to require uniqueness up to replacing replicas.

We begin with a preliminary lemma that considers a setting with independent admissions probabilities and stochastic outside options. Analyzing this setting facilitates our analysis of the independent simultaneous search problem studied by Chade and Smith (2006).

Lemma 3. *Let $\{\tilde{u}_j\}_{j=1}^r$ be independent random variables taking the value L_j with probability β_j and 0 otherwise, and let $(u_1, \ldots, u_n)$ denote the student's utilities from attending colleges of each types. Then, for each $r' \leqslant r$, if the student has access to the stochastic outside option $\max_{j \leqslant r'} \tilde{u}_j$, then there exists a (deterministic) utility assessment $(\mathbb{E}[\max_{j \leqslant r'} \tilde{u}_j], v_1^{r'}, \ldots, v_n^{r'})$ that attributes the same expected utility for each portfolio. Furthermore, $(\mathbb{E}[\max_{j \leqslant r'} \tilde{u}_j], v_1^{r'}, \ldots, v_n^{r'}) \geq_{RL} (\mathbb{E}[\max_{j \leqslant r'-1} \tilde{u}_j], v_1^{r'-1}, \ldots, v_n^{r'-1})$.*

Proof. Denote by $G^{r'}$ the CDF of $\max_{j \leqslant r'} \tilde{u}_j$. A direct calculation shows that the profile

$$v_i^{r'} = \beta^{r'}(u_i) = u_i + \int_{u_i}^{\infty} 1 - G^{r'}(z) dz$$

attributes the same expected utility for each portfolio. This formulation illustrates the idea that the student only benefits from admission to school i if the realized outside option is lower than u_i, and that when its realization is lower, the student only benefits from the marginal improvement over the outside option.

Next, we denote

$$\phi(x) := \begin{cases} \int_0^{\infty} 1 - G^{r'}(z) dz & \text{if } x \leqslant \int_0^{\infty} 1 - G^{r'-1}(z) dz \\ x + \int_{\text{inv}\beta^{r'-1}(x)}^{\infty} G^{r'-1}(z) - G^{r'}(z) dz & \text{else;} \end{cases}$$

and[20] note that

$$v_i^{r'} = \phi\left(v_i^{r'-1}\right).$$

[20]The inverse of $\beta^{r'-1}(\cdot)$ exists for values greater than $\int_0^{\infty} 1 - G^{r'-1}(z) dz$ since $\beta^{r'-1}(\cdot)$ is increasing for values greater than $\int_0^{\infty} 1 - G^{r'-1}(z) dz$.

By the Leibniz rule and the implicit function theorem, for values of x greater than $\int_0^\infty 1 - G^{r'-1}(z)dz$, we have

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

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

Theorem A.1. *If each school has m replicas, then for each $k < m$, $P(k+1, U, I) \succeq_A P(k, U, I)$.*

Proof. We prove the stronger claim that if $P^{(1)}(k, U, I)$ has a replica that is not included in $P(k, U, I)$, then $P(k+1, U, I) \succeq_A P(k, U, I)$. Chade and Smith (2006) have shown that there exists an optimal portfolio of size $k+1$, $P(k+1, U, I)$, such that $P(k+1, U, I) = P(k, U, I) \cup \{x\}$, unless we are in the trivial case that $P(k+1, U, I) = P(k, U, I)$. Let x denote a college such that $P(k, U, I) \cup \{x\}$ is an optimal size $k+1$ portfolio. Let y denote a replica of $P^{(1)}(k, U, I)$ that is not included in $P(k, U, I)$. By the weak axiom of revealed preferences $P(k, U, I)$ continues to be optimal if we constrain Ann to include $P(k, U, I) \setminus \{P^{(1)}(k, U, I)\}$ in her portfolio and to only choose additional schools from the set $\{x, y, P^{(1)}(k, U, I)\}$. We can equivalently think of the constrained problem as the problem of choosing a singleton portfolio from $\{x, y, P^{(1)}(k, U, I)\}$ with a stochastic outside option distributed as the utility from the portfolio $P(k, U, I) \setminus \{P^{(1)}(k, U, I)\}$ (we always assume that stochastic outside options are independent from schools admissions decisions). Since $P^{(1)}(k, U, I) \in P(k, U, I)$ and y is a replica of $P^{(1)}(k, U, I)$ that does not belong to $P(k, U, I)$, we have that $\{y\}$, like $\{P^{(1)}(k, U, I)\}$, is an optimal size-1 portfolio in this problem. By the weak axiom of revealed preferences, it is also optimal when the set of available schools is only $\{x, y\}$.

The argument above shows that $\{x\}$ is an optimal size-1 portfolio from the menu $\{x, y\}$ with an outside option that is distributed as the utility from the portfolio $P(k, U, I)$. By Lemma 3, Ann is more risk loving when making this latter choice. Since she is choosing a size-1 portfolios in both cases (from the menu $\{x, y\}$), the correlation structure between colleges' admissions decisions is irrelevant, and so by Lemma 2 and the definition of x, $\{x\} \succ_A \{y\}$, and so $P(k+1, U, I) \succ_A P(k, U, I)$. □