Simultaneous Search: Beyond Independent Successes*

Ran I. Shorrer
Penn State

April 8, 2019

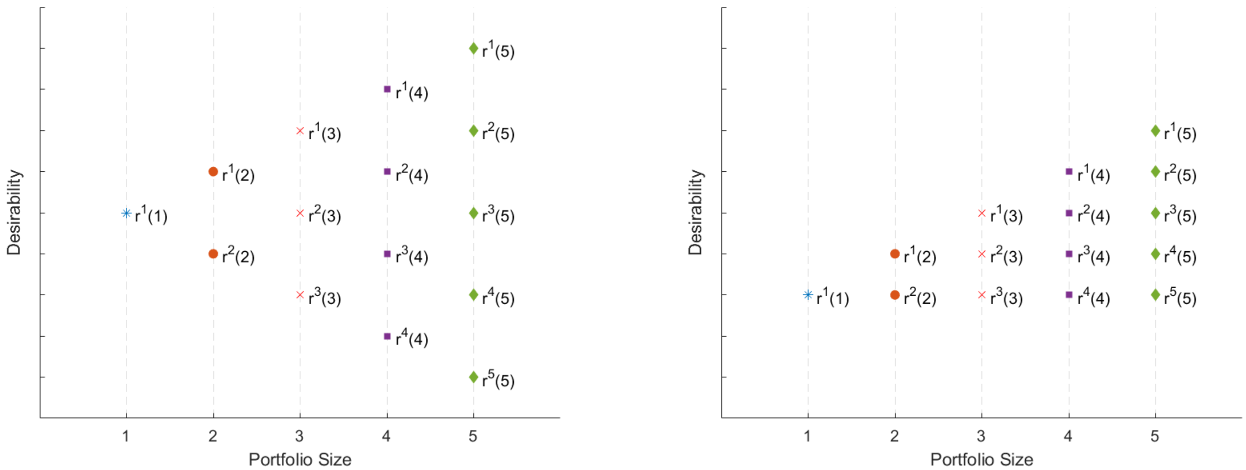
Abstract

A key decision commonly faced by students is how to optimally choose their portfolio of college applications. Students are often advised to apply to a combination of "reach," "match," and "safety" schools. Empirically, when reductions in the cost of application permit students to apply to more schools, they expand the range of selectivity of schools to which they apply both upwards and downwards. However, this ubiquitous practice of diversification is difficult to reconcile with existing theoretical analyses of search decisions. To solve this, I develop a generalized framework for simultaneous search problems in which students' optimal behavior generates these patterns. My framework generates many predictions that are consistent with empirical findings on school choice. The important departure that I make from existing models is that colleges' admissions decisions are not independent.

When applying to schools and colleges, a key decision commonly faced by students is how to optimally choose their portfolio of schools. In many settings, large numbers of schools are available, but due to costs or constraints students apply only to a few, often without perfect information about how the school will respond to their application. Determining which subset of programs to apply to—balancing the desire to attend sought-after programs with the need to hedge—forms a critical part of the decision problem, with these decisions deeply affecting the final outcomes in the market. To achieve this balance, students are often advised to apply to a combination of "reach," "match," and "safety" schools (Avery, Howell and Page, 2014). In practice, it is also seen that when reductions in application costs permit students to apply to more schools, they expand the range of schools to which they apply *both* upwards and downwards, including safer and more selective schools (Ajayi, 2011; Pallais, 2015).

Although it is well-understood that applying to a diverse set of colleges is crucial for students' success and carries significant implications, this practice of diversification is difficult to reconcile

*I am especially grateful to Kehinde Ajayi who introduced me to this problem. I also thank Mohammed Akbarpour, Nageeb Ali, Nick Arnosti, Kalyan Chatterjee, Oren Danieli, Henrique de Oliveira, Pieter Gautier, Michael Gechter, Ed Green, Nima Haghpanah, Avinatan Hassidim, Sándor Sóvágó , Kala Krishna, Vijay Krishna, Tomás Larroucau, Jacob Leshno, Dilip Mookherjee, Amanda Pallais, Alex Rees-Jones, Ignacio Rios, Assaf Romm, Karl Schurter, Shouyong Shi, Andy Skrzypacz, and Ricky Vohra. Nilufer Gok and Xiao Lin provided excelent research assistance. This research was supported by a grant from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel.



(a) Perfectly alligned admission decisions

(b) Independent admission decisions

Figure 1: Optimal portfolios under independent and perfectly aligned admission decisions

Panel 1a Illustrates my result that when admission decisions are perfectly aligned, optimal portfolios of larger sizes span a wider range of options and refine the grid. Panel 1b illustrates the findings of Chade and Smith (2006) (which I slightly strengthen using my general framework): i) the optimal portfolio of size $k+1$ nests the optimal portfolio of size k (Theorem 4), and ii) an option that appears in the optimal portfolio of size $k+1$ but not in the optimal portfolio of size k is more aggressive than any option in the optimal portfolio of size k (when each option has close substitutes, Theorem 5).

with existing theoretical analyses of search decisions. In a leading analysis of simultaneous search, Chade and Smith (2006) show that students should not apply to "safety" schools. Furthermore, they show that if application costs decrease, students should expand the range of schools only in an upward direction of applying to schools of higher selectivity. This lack of consistency between theory and evidence carries important implications for inferences that rely on such theoretical models (see Example 1).

In this paper, I develop a new analysis of simultaneous search in which the optimal behavior involves diversification and expanding the range of schools when application costs decrease. The important departure that I make from existing models—used for both theoretical and empirical studies—is that colleges' admission decisions are not necessarily independent.[1] In the general environment that I consider, admission decisions at some schools convey information about the probability of acceptance at other schools (e.g., finding out that one is rejected at MIT is bad news with respect to the probability of acceptance at Harvard). In the leading application, admissions decisions are perfectly aligned—they are based on a common index, with different schools having

[1]Apart from Chade and Smith (2006), examples of other studies that make the common assumption that colleges' admission decisions are independent conditional on the information known to the student include Card and Krueger (2005), Chade, Lewis and Smith (2014), and Fu (2014).

different bars for admission. I show that such correlation in admission decisions leads the model to predict the behaviors described above (Figure 1).

Example 1. A school district makes admissions decisions based on a single lottery, where higher numbers get priority. There is a unit mass of students, half of whom reside in the East Neighborhood, and the others in the West Neighborhood. After applying to schools, students in both neighborhoods draw lottery numbers uniformly from the interval $[0, 1]$.

In each location, x, there are two schools. A good school, g_x, with capacity $\frac{1}{4}$, and a bad school, b_x, with unlimited capacity. Students prefer schools that are nearer to their place of residence, and they prefer good schools. Specifically, the utility derived by student s from attending school m is given by $u_s(x) := \beta \mathbb{1}_{m \text{ is local for } s} + \gamma \mathbb{1}_{m \text{ is good}}$, where β and γ are both greater than zero, and where the utility from the outside option is normalized to zero. Students can only apply to two schools. The above facts are all commonly known.

In equilibrium, each student applies to both of the schools in her neighborhood. To see this, note that under this profile of strategies, if a student can be admitted to the good school in the other neighborhood (if her lottery draw is greater than one half), she will also be admitted to the good school in her own neighborhood, which she prefers. Consequentially, she prefers to apply only to the good local school and to use her second application to guarantee admission to the bad local school, in case she is not admitted to the good school.

Next, imagine an econometrician who believes that schools admission decisions are independent. Except for this fact, the econometrician's model is correctly specified, and he believes that students have rational expectations and that they choose their applications portfolio optimally. Observing students' applications and admission probabilities this econometrician concludes that

$$\frac{1}{2} \cdot (\beta + \gamma) + \frac{1}{2} \cdot 1 \cdot \beta \geq \frac{1}{2} \cdot (\beta + \gamma) + \frac{1}{2} \cdot \frac{1}{2} \cdot \gamma$$

where the left hand side represents his (correct) perception of students' expected utility from the portfolios they choose and the right hand side is what he believes to be the expected utility from applying to both of the good schools (see Agarwal and Somaini, 2018). This inequality implies that parameters in the identified set satisfy the inequality

$$2\beta \geq \gamma.$$

Thus, if students attribute high relative importance to quality (e.g., $\beta = 1$, $\gamma = 100$) this preference will not be reflected in the econometrician's estimates. When considering the possibility of increasing good schools' capacities, a policy maker that relies on the ecomometricain's estimates may make misguided choices. □

Correlation in admission prospects is ubiquitous. For example, admissions decisions are perfectly aligned in school districts in the U.S. and around the world where a single lottery is used to break priority ties in oversubscribed schools (e.g., De Haan et al., 2015; Abdulkadiroğlu et al., 2017*b*). Admissions decisions are perfectly aligned when a centralized admission exam determines admissions— a common practice for high-school and college admissions across the world. Even in the absence of a centralized admissions exam, (perceived) correlation in admission decisions is present when college applicants think that colleges are seeking students with certain characteristics, but applicants face uncertainty about where they stand relative to others.[2] The evidence outlined in the first paragraph suggest that this is the case in the American college admissions market. Such correlation also exists in labor markets where common factors affect all firms' hiring decisions.

I derive the results outlined above, as well as other important comparative statics that match the findings of earlier empirical papers, by solving a broad class of *simultaneous search* problems—portfolio choice problems in which an agent chooses a portfolio of stochastic options, but only consumes the best realized one. These problems have many applications (e.g., in consumer search (De Los Santos, Hortaçsu and Wildenbeest, 2012), labor markets (Galenianos and Kircher, 2009), and industrial organization Wong (2014)), but to fix ideas, I will only refer to agents as students. In the class of simultaneous search problems that I study, portfolios can be represented by a Rank-Order List (ROL) where the probability that the k-th-ranked option on the ROL is consumed depends only on higher-ranked options. The leading examples are: *i)* a student applies to k colleges and attends the most-preferred college that admits her, and *ii)* a student submits an ROL to a centralized mechanism where this property holds. Many popular allocation mechanisms feature this structure, including variants of deferred acceptance, top trading cycles, and Boston mechanism.[3]

I show that the portfolio choice problems in this class can be solved using dynamic programming – a concept similar to backwards induction. The key idea is that the optimal continuation (or suffix) of an ROL can be calculated by conditioning on the event that the agent will be rejected by all the options that are ranked higher on the ROL (prefix). When the portfolio choice problem satisfies an additional condition, which assures that the optimal suffix is identical for many prefixes, the *exact* optimal ROL can be found "quickly," in running time polynomial in the number of options.[4] In the absence of such a solution, empirical studies had to resort to approximations of the optimal portfolio (e.g., Ajayi and Sidibe, 2015).

My framework has many potential applications. In this paper, I apply it to study the cases where

[2] Many studies of application patterns to American colleges use SAT test-score sending as a proxy for application (e.g., Card and Krueger, 2005). An important institutional detail is that *before* taking the exam applicants choose up to four schools that will receive their scores (they have the option to add schools to the list before and after the exam for a fee). This feature suggests that students, especially first-time test takers, may be facing substantial uncertainty about their strength as applicants when determining their application portfolio.

[3] This condition rules out, for example, mechanisms that maximize the number of acceptable matches, as under such mechanisms ranking an undersubscribed school last reduces the probability of assignment to an oversubscribed school ranked higher.

[4] In the context of centralized school choice, Casalmiglia, Fu and Güell (2014) prove this claim and use it to calculate optimal ROLs in their empirical study of school choice in Barcelona, where a variant of the Immediate Acceptance (Boston) mechanism is in place. The independent admissions environment is studied by Chade and Smith (2006) who show that it is solved by a greedy algorithm. Specifically, they show that for any k, an optimal portfolio of size $k+1$ nests an optimal portfolio of size k. Chade and Smith (2005; 2006) provide some extensions, and so do Olszewski and Vohra (2016) who take a different approach to the problem. With the exception of the environment with independent admissions decisions, the applications I study do not fall under these conditions as well.

admissions decisions are perfectly aligned and where they are conditionally independent, as well as to cases where the market is cleared centrally using variants of the deferred acceptance (DA, Gale and Shapley, 1962), immediate acceptance (Boston), or top trading cycles (TTC, Abdulkadiroğlu and Sönmez, 2003) algorithms. In all of these cases I show that risk tolerance leads students to apply more aggressively.[5]

A corollary of the results on risk aversion is that allocation mechanisms that are commonly used in centralized school choice, such as constrained deferred acceptance (Haeringer and Klijn, 2009; Calsamiglia, Haeringer and Klijn, 2010) and variants of immediate acceptance, give applicants with better outside options (e.g., private schools) incentives to apply more aggressively.[6] This finding suggests that school segregation may occur *endogenously* when such school allocation mechanisms are in place, even when all applicants share the same preferences, in spite of the mechanism treating all applicants symmetrically (cf. Hastings, Kane and Staiger, 2009; Oosterbeek, Sóvágó and van der Klaauw, 2019).

This result provides intuition for the upward expansion of optimal portfolios under both independent and perfectly aligned admission decisions. The optimal portfolio of size $k+1$ can be thought of as the optimal portfolio of size k where the student's outside option is the (stochastic) $k+1$-st ranked school (rather than nothing, in the case of choosing the optimal size-k portfolio).

I also show that, in all of the applications I study, students with more optimistic beliefs about their admission chances (to be precisely defined) apply more aggressively. This explains the divergence in predictions for the case where admission decisions are independent conditional on the information known to the student and where they are highly correlated. Recall that admission to less desirable schools only matters when a student is rejected by the most desirable school in his portfolio. Thus, one can think of the optimal portfolio of size $k+1$ as consisting of the most desirable school in this portfolio and the optimal size-k portfolio conditional on being rejected from the most desirable school in the optimal portfolio of size $k+1$. Since, when admission decisions are aligned, news of rejection lead to lower beliefs, this continuation portfolio is less aggressive than the optimal portfolio of size k. By contrast, when admissions decisions are independent, a rejection by the first choice carries no further information about admissions chances.

These predictions are in line with empirical studies that find that improved outside options lead applicants to apply more aggressively under variants of deferred acceptance (Krishna, Lychagin and Robles, 2015; Istomin, Krishna and Lychagin, 2018), and of immediate acceptance (Calsamiglia and Güell, 2014); that applicants with optimistic (subjective) beliefs apply more aggressively than ap-

[5] In Section 7, I discuss how my approach can be used to analyze college admissions markets with early admissions. Avery and Levin (2010) study a special case of this model, and find that colleges can lure strong cautious students away from higher-ranked schools by adopting an early decision policy (Avery, Fairbanks and Zeckhauser, 2009). Poor outside options is one reason for students to be cautious (See Kim, 2010).

[6] Calsamiglia, Martínez-Mora and Miralles (2015) study a model where school quality is determined endogenously as a function of the quality of the peers in the school. They show that in equilibrium outside options have a similar effect when immediate acceptance mechanisms are in place. Akbarpour and van Dijk (2018) find that when immediate acceptance is in place, students with better outside options are over represented in the best school. Delacroix and Shi (2006) derive similar results in a model of on-the-job search, where workers differ in their current jobs, which they can hold on to.

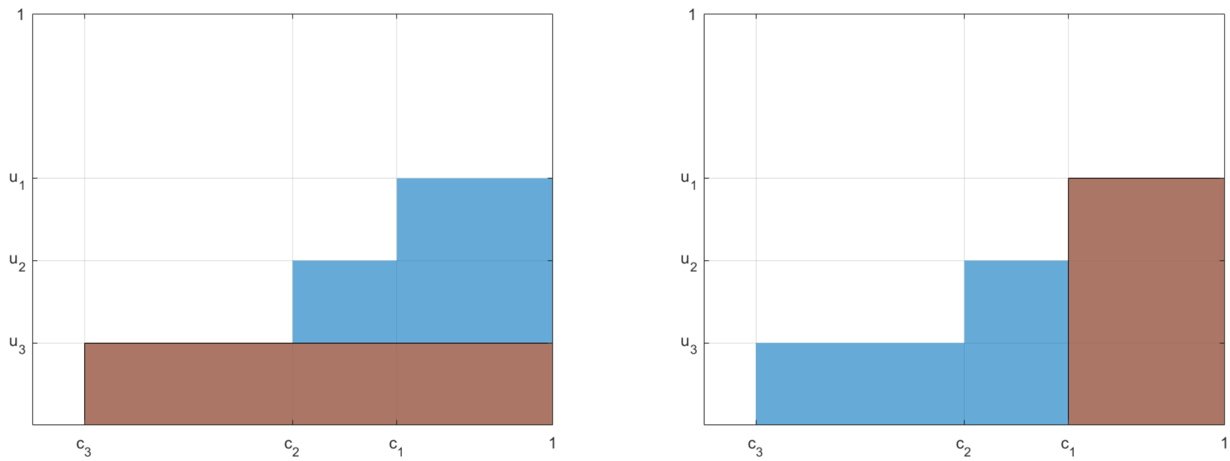
plicants with pessimistic (subjective) beliefs under variants of deferred acceptance (Bobba and Frisancho, 2014), and of immediate acceptance (Kapor, Neilson and Zimmerman, 2018; Pan, 2019); and that applicants with weak relative preferences for their most-preferred schools are more likely to rank another school first under a variant of immediate acceptance (i.e., applicants whose preferences were relatively concave applied less aggressively, Wu and Zhong, 2014; Kapor, Neilson and Zimmerman, 2018).

Finally, I highlight a potential tradeoff between efficiency and the desire to protect unsophisticated agents that arises when school districts choose whether oversubscribed schools use a single lottery or *independent* (school-specific) lotteries to break priority ties between applicants. I show that identifying the optimal portfolio in decentralized admissions or under constrained deferred-acceptance mechanisms is simpler in two well-defined senses when priority ties are broken with multiple independent lotteries rather than with a single lottery.[7] The tension arises since efficiency may be improved by the use of a single lottery (e.g., Ashlagi, Nikzad and Romm, 2015). The TTC algorithm introduces correlation in the set of feasible schools a student can be admitted to (Che and Tercieux, 2017; Leshno and Lo, 2017). For this reason, all of the results on the perfectly aligned admissions environment apply to markets cleared by TTC, regardless of the randomization method used to break priority ties. Specifically, regardless of the randomization method, identifying the optimal portfolio under constrained TTC is hard in both of the above-mentioned senses.

Perfectly aligned admissions. The application I focus on is that of identifying the optimal portfolio when decision makers are vertically ranked but they face uncertainty about their strength relative to others. While I derive most of the results using my framework, I also derive some results by relating this specific portfolio choice problem to the well-studied NP-hard max-coverage problem (Hochbaum, 1996). As illustrated in Figures 2 and 3, I show that there is a correspondence between colleges and rectangles in $\mathbb{R}^2$, where a college that yields utility u_i and whose admission cutoff is at c_i is associated with the rectangle that has vertices at $(c_i, 0)$ and $(1, u_i)$. The expected utility from a portfolio of k colleges is equal to the area covered by the k corresponding rectangles (I assume, without loss of generality, that scores are distributed uniformly on $[0, 1]$. Thus, the probability of admission to college i is equal to the distance between c_i and 1. For details, see Section 2). This representation allows me to provide a clear graphical intuition for many of my results.

I show that the optimal application portfolio in such environments can be described as *"a combination of 'reach,' 'match,' and 'safety' schools,"* as recommended by the College Board (Avery, Howell and Page, 2014). By contrast, with conditionally independent admission decisions utility maximization *"precludes 'safety schools'"* (Chade and Smith, 2006). Moreover, when a reduction in the cost of application leads a student to apply to more colleges, she optimally applies to a wider range of schools in terms of selectivity and desirability. This finding reflects the empirical findings of Pallais (2015) in (decentralized) American college admissions and those of Ajayi (2011) in (centralized) Ghanaian high-school choice.

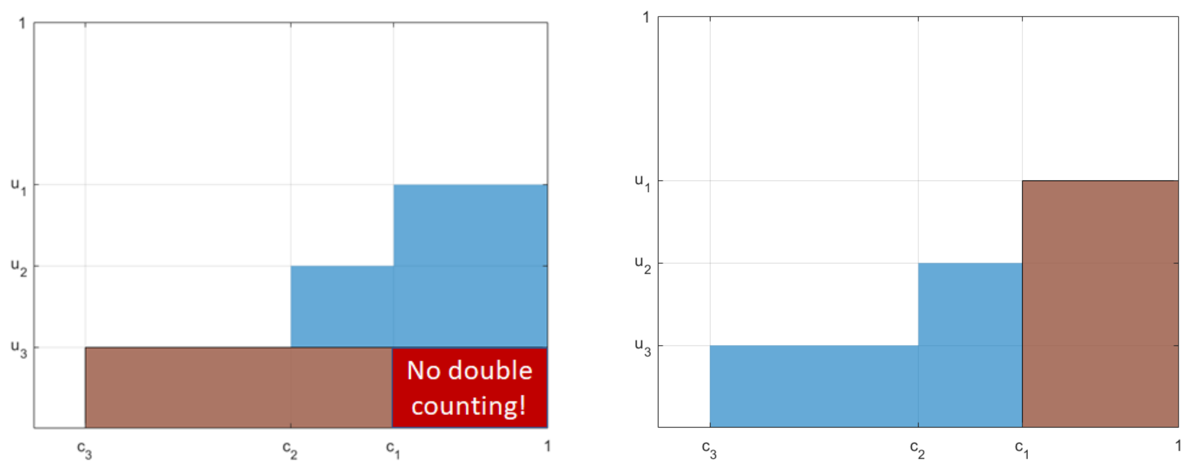
[7] Lucas and Mbiti (2012) and Ajayi and Sidibe (2015) find that a large fraction of the applicants in Kenya and in Ghana submit a suboptimal ROL that indicates that they find the "single-lottery" nature of the environment difficult.



(a) "safety" school

(b) "reach" school

Figure 2: The utility from different size-1 portfolios
This figure assumes that the distribution of scores is uniform. As a result, the probability of acceptance to school i is equal to $1 - c_i$, the probability of passing this school's admission cutoff. This quantity is equal to the distance between c_i and 1 on the horizontal axis. Thus, the shaded area represents the expected utility from a portfolio consisting of a single school (school 3 in Panel 2a and school 1 in Panel 2b) – the probability of admission, $1-c_i$, times the utility from attending, u_i. School 3 is a "safety" school – it has a high probability of admission, but yields a low utility from attending. School 1 is a "reach" school – it has a low probability of admission, but yields a high utility from attending.



(a) “safety” school

(b) “reach” school

Figure 3: The utility from a size-2 portfolio

This figure illustrates the expected utility from applying to both college 1 and college 3. It highlights the fact that a student can only attend a single college. As a result, when she attends the most desirable college on her portfolio to which she is admitted, and does not derive further utility from being also admitted to the less desirable college. This is illustrated by the red area, missing from the “safety school” rectangle (labeled “No double counting!”). This rectangle corresponds to cases where the student is admitted to both schools, and thus attends the more preferred “reach” school.

The graphical intuition for the incentive to diversify is that the rectangles corresponding to schools of similar selectivity (similar cutoffs) have a large overlap, so choosing several of them does not increase the covered area substantially relative to choosing just one of them.[8] For example, in the portfolio choice problem depicted in Figure 2, each rectangle covers approximately the same area, but the one corresponding to college 2 is slightly larger. As a result, the optimal size-1 portfolio consists of college 2, but the optimal size-2 portfolio consists of colleges 1 and 3, as the large overlap of other rectangles with the rectangle corresponding to college 2 dwarfs the size advantage of this rectangle.

As mentioned above, I show that when there is a constraint on the size of the portfolio, higher beliefs about one's chances of success in the sense of the monotone likelihood ratio property (MLRP) lead to a more aggressive portfolio choice, in the sense that for all i, the i-th-ranked school on one ROL is more aggressive than the i-th-ranked school on another. One implication of this finding is that students can "fall through the cracks": better students, who expect to receive higher exam scores, are sometimes more likely to be unassigned. When the marginal cost of application is positive and finite, the number of applications is nonmonotonic in beliefs. Students with higher beliefs may apply to more or to fewer colleges, further compromising the assortativeness of the match.[9] Figure 4 illustrates how news that one's score is below a certain number (which induces MLRP-lower beliefs) disproportionately decreases the expected utility from applications to more selective schools.

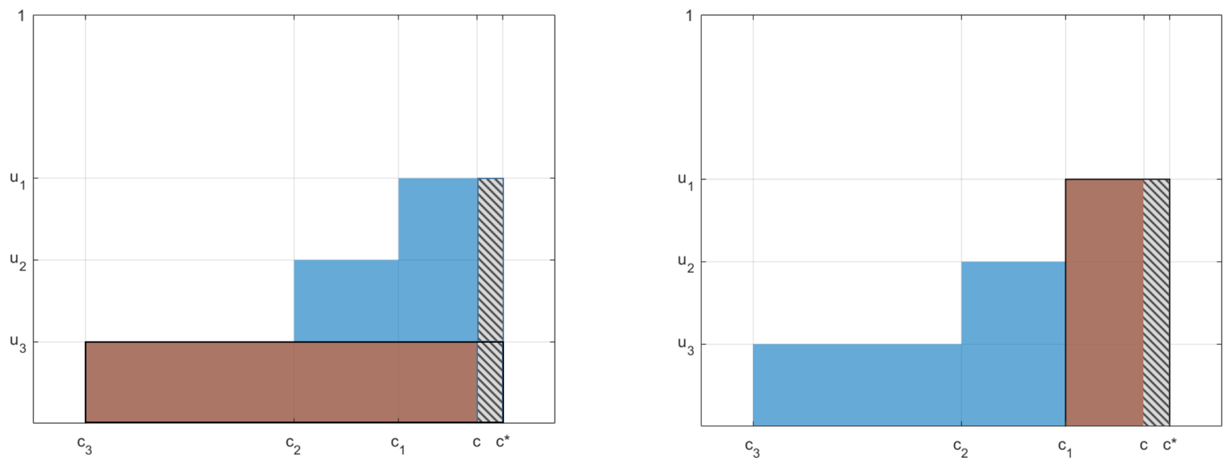
To prove that risk tolerance leads to more aggressive portfolios, I exploit an interesting duality between portfolio choice problems, where increasing risk aversion in the primal problem translates to increasing beliefs in the dual problem. The special case corresponding to an improved deterministic outside option is illustrated in Figure 5. This duality also establishes that this particular choice problem can also be solved by starting from the top of the ROL, instead of from the bottom.

Organization of the paper. In the next section, I introduce the model and the class of portfolio choice problems that this paper addresses. Section 2 analyzes the leading application: the case of perfectly aligned admissions. Section 3 analyzes the "classic" independent admissions case. I derive novel comparative statics, and show how my approach simplifies the proofs of established results. Section 4 compare different tie-breaking rules in terms of their strategic simplicity. Sections 5 and 6 cover the top-trading-cycles and Boston algorithms. Section 7 reviews other applications. Section 8 discusses the results. Throughout, omitted proofs a relegated to the appendix.

Although in many of the applications I study there exists, generically, a unique optimal portfolio, I choose to treat the general case. The reason is that my results hold in equilibrium in large markets, and there is no reason to think that agents have unique optimal portfolios in equilibrium. This

[8] The economic intuition resembles that of Athey and Ellison (2011) in the context of position auctions with consumer search. Athey and Ellison (2011) study the cases where the probabilities that websites meet the needs of a consumer are independent, and where similar sellers are likely to meet the same needs. They find that in the latter case the optimal auction incentivizes product variety.

[9] Nagypal (2004) derives similar comparative statics in a model with a continuum of colleges whose selectivity varies continuously, where students' beliefs belong to certain families.



(a) "safety" school

(b) "reach" school

Figure 4: The effect of news of rejection

The figure illustrates the effect of negative news on one's score on the expected utility from size-1 portfolios. Specifically, it shows that being rejected by the option with cutoff c, which is less selective than c^*, affects the expected utility from (the more selective) college 1 (Panel 4b) more than it affects the expected utility from (the less selective) college 3 (Panel 4a).

comes at the cost that statements and proofs are slightly more cumbersome.

1 Model

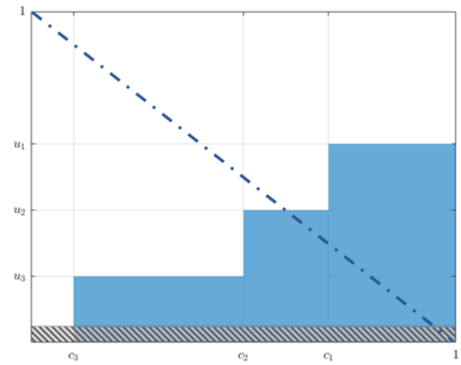
There is a set, X, of stochastic *options* (hereafter *schools* or *colleges*). Unless otherwise mentioned, the set is finite and the options are indexed by the integers 1 to N. Agents (hereafter *students* or *applicants*) are expected-utility maximizers. They can attend at most one school, and they know the utility they will derive from placement in each school.[10] I envision that the number of students is large, and consider the problem of a single student. I assume, without loss of generality, that lower integers (schools) are (weakly) more desired by this student. The utility from being unassigned is normalized to zero.

A Rank-Order List (ROL) is an ordered list of schools. I focus on cases where application portfolios can be summarized by an ROL. The two leading examples are: 1) a student applies to a subset of X and chooses the best school that admits her,[11] and 2) a student submits an ROL of alternatives in X to a centralized mechanism that determines her assignment. For an ROL, r, let r^l denote the l-th-highest-ranked school on r.

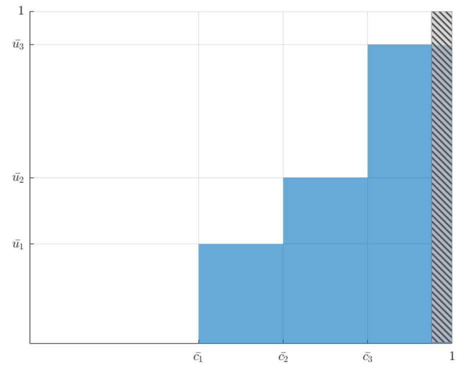
The cost associated with a portfolio of size k is $C(k)$, where $C(0) = 0$ and $C(\cdot)$ is increasing. I sometimes further assume that $C(\cdot)$ is convex. Special cases include constrained choice ($C(x) = 0$

[10] The assumption that searchers know their utility from prizes is common in the search literature, but there are some exceptions (e.g., Immorlica et al., 2018; Albrecht, Gautier and Vroman, 2006).

[11] In addition to the subset of X, my definition of a portfolio includes, a strict ranking based on which options are chosen in case of multiple successes.



(a) primal problem



(b) dual problem

Figure 5: The effect of an improved outside option
Panel 5a illustrates the effect of an improvement to an agent's outside option. It shows that less selective schools are disproportionately affected (the reduction in the area they cover is disproportionately large). Panel 5b illustrates the dual problem. It corresponds to the same coverage problem, but the graph is transposed over the line connecting $(1, 0)$ and $(0, 1)$. In the transposed figure, the improved outside option looks like the negative news from Figure 4. Hence, the optimal coverage in the dual problem becomes less aggressive, which corresponds to a more aggressive portfolio in the primal problem. Detailed arguments appear in the body of the paper.

if $x \leq k$, and $C(x) = \infty$ otherwise) and constant marginal cost ($C(x) = cx$).

I focus on environments that satisfy Condition 1. This condition assures that optimal portfolios can be identified using dynamic programming (which is similar to backward induction).

Condition 1. Each ROL, r, represents a portfolio such that the probability of assignment to r^{k+1} conditional on not being assigned to higher-ranked alternatives $(r^1, ..., r^k)$ is independent of lower-ranked alternatives $(r^{k+2}, r^{k+3}...)$.

Observation 1. In environments that satisfy Condition *1*, the optimal portfolios of a given size can be found using dynamic programming.

Proof. The condition assures that the choice of lower-ranked alternatives on an ROL does not affect the outcomes with respect to higher-ranked alternatives. Thus, to identify an optimal ROL (portfolio), the applicant can decide on the best alternative for the k-th rank assuming that she was not assigned to any higher-ranked alternative, while taking into account the best course of action conditional on not being assigned to this alternative. □

When an additional condition is satisfied, the dynamic program can be solved quickly. This is particularly useful for for empirical analysis of application portfolios.

Definition. Denote by H^k the set of all possible vectors $h^k = (r^1, r^2..., r^{k-1})$.

Condition 2. There exists $U \in \mathbb{N}$, such that for every instance with N options, for every k, and for every i, there exists a partition of H^k, $\{H_t^k\}_{t=1}^{N^U+U}$ such that the probability of a student being assigned to i, conditional on being rejected by all higher-ranked alternatives and on using the same continuation list starting from k, is constant for each partition element.

The two conditions are satisfied, for example, by the serial dictatorship environment studied in Abdulkadiroğlu et al. (2017*b*), which is a special case of the perfectly aligned admissions environment studied in Section 2. In this setting, the Most Informative Disqualification (MID) summarizes the information that is relevant for the purpose of decision making. The MID is the most forgiving cutoff the applicant failed to pass if the k-th-ranked alternative becomes relevant, which is a sufficient statistic for updating one's beliefs about her priority score.

Observation 2. When a mechanism satisfies Conditions *1* and *2*, the optimal ROL of a given length can be found in a number of steps polynomial in N.

Proof. Condition 2 assures that there are only a polynomial number of histories to consider in every step of the dynamic program, and the number of steps is at most N. □

2 Simultaneous Search by Ranked Applicants

I now turn to the decision problem of main interest: simultaneous search by vertically differentiated applicants who face uncertainty about their market position. A motivating example is school choice with a centralized entrance exam (or central lottery) where admission to each school depends on passing a school-specific score cutoff.[12] In such markets, a student gets to attend the best school to which she applied whose admission cutoff she surpassed. It is often the case that students are required to make their application decision not knowing their score on the admission exam.

Without loss of generality, I make the following assumptions: 1) the decision maker's belief about her admission score is an atomless distribution over all possible scores. Furthermore, I assume that scores are distributed uniformly between 0 and 1 (otherwise, apply the probability distribution transform to all scores, including the above-mentioned cutoffs). 2) Every school is preferred to the outside option whose value is normalized to 0.

2.1 The Portfolio Choice Problem

Denote school i's admission cutoff by c_i. The portfolio choice problem can be summarized by the following equation:

$$\max_r \sum_{1\le i\le |r|} \max\left\{0, \min_{0\le j<i}\{c_{r^j}\} - c_{r^i}\right\} u(r^i) - C(|r|),$$

[12] In the absence of aggregate uncertainty (i.e., correlation in the distribution of types in the population), when schools offer the same number of seats and the market is large, it is safe to assume that admission cutoffs are stable from one year to the next (Azevedo and Leshno, 2016).

where $c_{r^0} \equiv 1$, and $\min_{j<i}\{c_{r^j}\}$ is the MID. A first observation is that it is never beneficial to list an option lower than a (weakly) less selective (i.e., lower cutoff) option.

A first observation is that when looking for an optimal portfolio, there is no loss in ignoring schools such that $i > j$ and $c_j \leq c_i$ for some j. Graphically, these are colleges whose rectangles in the corresponding coverage problem are covered by another rectangle. To keep the notation simple, I assume that all such options have been removed from the menu of options, and there are $M \leq N$ schools, relabeled as $1, 2, ...M$, such that $c_i > c_j \iff u_i > u_j$. Still, in what follows I use N, not M, to denote the number of options.

Lemma 1. *There always exists an optimal ROL that is strictly decreasing in selectivity and in desirability (i.e., lower-ranked options have higher indices and strictly lower admission cutoffs).*

Proof. An ROL that is inconsistent with one's preferences (i.e., not monotonic in indices) is weakly dominated (e.g., Haeringer and Klijn, 2009). The result follows from this fact and the discussion in the paragraph above.

□

The portfolio choice problem can thus be simplified to

$$\max_{r|c_{r^{i-1}} > c_{r^i}} \sum_{1 \leq i \leq |r|} (c_{r^{i-1}} - c_{r^i})\, u(r^i) - C(|r|).$$

2.1.1 Unconstrained Choice

I begin by considering the case of unconstrained choice ($C \equiv 0$). In this case, agents clearly have a weakly dominant strategy of applying everywhere and attending the best school that accepts them (i.e., using an ROL of all the alternatives in order of preference). The utility an agent derives from this strategy is equal to the area under a step function, as illustrated in Figure 6.

2.1.2 Constrained Choice

In a constrained choice problem, the applicant may apply only to a limited number of schools. To gain intuition, I start by considering the special case of constrained choice where the constraint is extreme: the agent can choose at most one school to apply to ($C(x) = 0$ if $x \leq 1$, and $C(x) = \infty$ otherwise). In this case, the optimal strategy is clear: choose the school that maximizes $(1 - c_i)\, u(i)$. Graphically, this is the largest rectangle under the step function whose lower-right corner is the coordinate $(1, 0)$, as illustrated in Figure 2.

The agent faces a trade-off between high ex-post utility and high ex-ante admission probability. This is illustrated by comparing the different portfolios depicted in Figure 2. The rectangle depicted in Panel 2a represents an application to a "safety school," school 3, with a high admission probability and low ex-post utility from consumption. The rectangle depicted in Panel 2b represents an application to the more selective school 1, which yields a higher ex-post utility from consumption, but entails more risk in terms of the probability of admission.

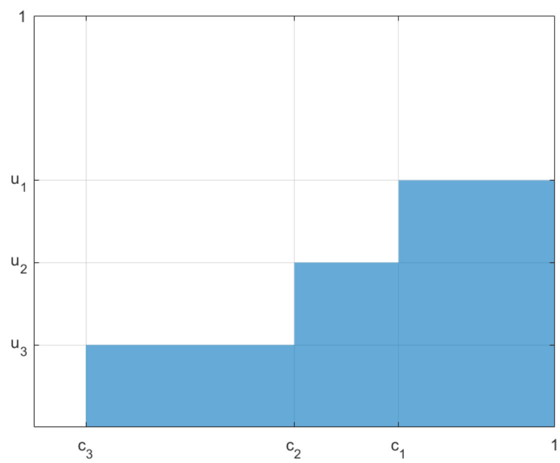


Figure 6: The utility from the optimal unconstrained ROL

This figure assumes that the distribution of scores is uniform. As a result, the probability of acceptance to school i is equal to $1 - c_i$, the probability of passing this school's admission cutoff. This quantity is equal to the distance between c_i and 1 on the horizontal axis. Thus, the shaded area represents the expected utility from a portfolio consisting of all schools, where the student gets to attend the best school whose cutoff she surpassed.

Next, we consider the optimal ROL of length k (which corresponds to the solution in the case where $C(x) = 0$ if $x \leq k$, and $C(x) = \infty$ otherwise). Graphically, the problem corresponds to the coverage problem (e.g., Hochbaum, 1996) of identifying k rectangles as above whose union covers the largest area. The challenge is that the rectangles overlap, but their intersections should not be double counted.

To this end, I propose the Probabilistically Sophisticated Algorithm (PSA), which uses dynamic programming to solve for the optimal ROL. I denote by $r(k, c)$ an optimal ROL of size k when the score is uniformly distributed in $[0, c]$, and let $v(k, c)$ denote the corresponding expected utility. Furthermore, $r_F(k, c)$ and $v_F(k, c)$ are similarly defined, where the score is distributed according to F conditional on it being lower than c. I use the convention that $v(0, \cdot) = c_{r(0,\cdot)} \equiv 0$.

The success of the algorithm hinges on the following lemma.

Lemma 2. *For all $k \in \mathbb{N}$, $r^1(k+1, c) = \underset{i|c_i \leq c}{\arg\max} \frac{1}{c} \{(c - c_i)u_i + c_i v(k, c_i)\}$ and for all $k+1 \geq j > 1$, $r^j(k+1, c) = r^{j-1}(k, c_{r^1(k+1,c)})$.*[13] *More generally, for all $k \in \mathbb{N}$, and for any distribution of scores, F, $r^1_F(k+1, c) = \underset{i|c_i \leq c}{\arg\max} \frac{1}{F(c)} \{(F(c) - F(c_i))u_i + F(c_i)v_F(k, c_i)\}$, and, for all $k+1 \geq j > 1$, $r^j_F(k+1, c) = r^{j-1}_F(k, c_{r^1_F(k+1,c)})$.*

Algorithm 1 Probabilistically Sophisticated Algorithm (PSA)

Step 1 For all possible values of r^{k-1} calculate $r^1(1, c_{r^{k-1}})$ and $v(1, r^{k-1})$.

Step j<k For all possible values of r^{k-j} calculate $r^1(j, c_{r^{k-j}})$ and $v(j, r^{k-j})$, as in Lemma 2.

Step k Calculate $r^1(k, 1)$, and set $r^1 = r^1(k, 1)$, $r^i = r^1(k+1-i, c_{r^{i-1}})$.

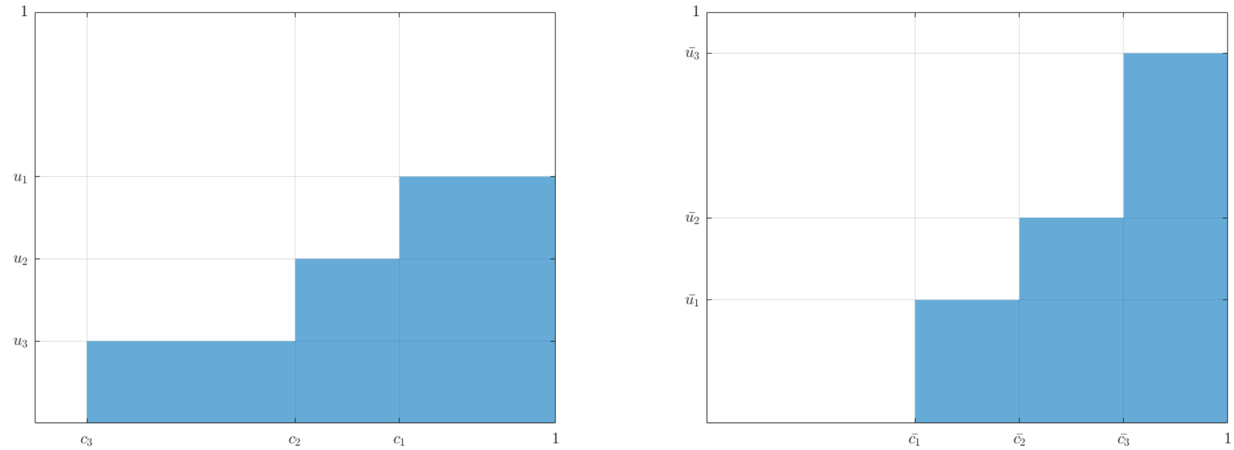
Remark 1. For ease of exposition, the algorithm does not explicitly treat cases where the agent ranks alternatives inconsistently with her preferences. These cases are (weakly) suboptimal (Lemma 1). For the same reason, I also assume that there are more than k acceptable alternatives ($k < N$).

2.1.3 Costly Choice

Each step of the PSA requires no more than N^2 steps. By setting $k = N$ it is possible to find the optimal portfolio of any size in less than N^3 steps. When portfolio costs differ in size, one can solve for the optimal portfolio of each size and then choose the best one by accounting for costs.

When the cost is convex, one can leverage the fact that the marginal benefit from increasing k is decreasing. Although intuitive, the decreasing marginal benefit property is not at all trivial due to the possibility that optimal portfolios of varying sizes are not nested. Indeed, Example 4 provides an instance of the coverage problem by rectangles where this is not the case. The example does not, however, possess the same structure as the problems studied in this section as the rectangles

[13] The set $\{i|c_i \leq c\}$may be empty. In this case, I use the the convention that the arg max is the empty set and the ROL will have an empty enrty. This does not occur on the run of the algorithm (it will repeat the same option). In any case, such choices are inconsequential for the agent's assignment.



(a) primal problem

(b) dual problem

Figure 7: A problem and its dual problem
Panel 7a depicts the (primal) choice problem that was depicted in Figure 6 ($c_3 = .1$, $c_2 = .5$, $c_1 = .7$; $u_3 = .3$, $u_2 = .4$, $u_1 = .6$). Panel 7b depicts its dual choice problem ($\bar{c}_3 = .7$, $\bar{c}_2 = .6$, $\bar{c}_1 = .4$; $\bar{u}_3 = .9$, $\bar{u}_2 = .5$, $\bar{u}_1 = .3$).

do not share a common bottom-right corner. I prove that the marginal benefit from increasing k is decreasing later (Theorem 2), as the proof invokes Theorem 1.

2.2 The Dual Choice Problem

Before I start, it may be beneficial to give some intuition for the transformation I am about to present. The main insight of this section is that if one transposes the coverage problem depicted in Figure 6 (i.e., the portfolio choice problem) over the diagonal connecting $(1, 0)$ and $(0, 1)$, one is looking at an equivalent coverage problem, and the only difference is that higher indices correspond to higher rectangles.

Formally, consider the (primal) problem defined by $(c_i, u_i)_{i=1}^N$, the uniform distribution over the unit interval, and some cost function. First, note that there is no loss of generality from normalizing utilities by multiplying each u_i and the cost function by $\frac{1}{\max\{u_i\}}$. This normalization assures that the agent derives a utility of 1 from the alternative he likes best. Next, for all i, let

$$\bar{c}_i = 1 - u_i$$

and

$$\bar{u}_i = 1 - c_i.$$

Since cutoffs lie in the unit interval, $\bar{u}_i \in [0, 1]$ for all $1 \leq i \leq N$, and, thanks to the normalization of utilities, $\bar{c}_i \in [0, 1]$ for all $1 \leq i \leq N$. The problem defined by $(\bar{c}_i, \bar{u}_i)_{i=1}^N$ — with the same cost function and the uniform distribution — is the dual choice problem of the primal problem.

Definition. Given an ROL, r, of length k, let $\overset{\circlearrowright}{r}$ be the ROL such that $\overset{\circlearrowright}{r_i} = r_{k-i}$ for all i.

In words, $\overset{\circlearrowright}{r}$ is the ROL r turned upside down.

Proposition 1. *ROL r is a solution of a portfolio choice problem iff $\overset{\circlearrowright}{r}$ is a solution to the dual problem.*

Proof. Note first that in the dual problem, lower index options are less desirable. Thus, if the optimal portfolio consists of the same alternatives, it is optimal to rank them in reverse order (since truthfulness is optimal).

To see that the same alternatives should be ranked in both problems, note that the corresponding coverage problem is isometric (the only difference is that the plane has been transposed). □

Remark 2. I previously assumed that alternatives that are more selective and less desirable have been removed from the menu. Without this assumption, the statement would not be correct, since the applicant might be able to rank undesirable alternatives in irrelevant positions on an optimal ROL. This issue can arise only when applicants are effectively unconstrained (i.e., when the marginal cost of applying to an additional college at some optimal portfolio is zero). In what follows, I assume that applicants are effectively constrained.

Remark 3. The proposition suggests that solving the dual problem "backwards" is tantamount to solving the primal problem "forwards." The appendix contains a forward-looking version of the PSA (using the notation of the primal problem), which is sometimes more convenient.

2.3 Comparative Statics

2.3.1 Beliefs

I now consider how the optimal portfolio of size k changes with the decision maker's beliefs. To this end, I dispense with the assumption that the distribution of scores is uniform.

Definition. Given two portfolios of size k, r and r', we say that r *at least as aggressive as* r' if $r^j \leq r'^j$ (r^j is at least as desirable as r'^j) for all $1 \leq j \leq k$. Furthermore, r *more aggressive* than r' if r is at least as aggressive as r' but r' is not at least as aggressive as r.

Proposition 2. *Let s and b be agents with identical preferences, but s has more optimistic beliefs about her score in the sense of MLRP. Then for all $k \leq N$, there exists an s-optimal size-k portfolio that is at least as aggressive as any b-optimal size-k portfolio.*

Proof. The proof is by induction on k. The case of $k = 1$ is obvious, because the MLRP shift to beliefs implies that admission chances to programs that yield higher ex-post utility are disproportionately higher.

Assume that the assertion is correct for all $l < k \leq N$. By Lemma 2 and the inductive hypothesis, if $r_s^1 \leq r_b^1$ we are done, since s's beliefs conditional on her being rejected from r_s^1, which is more selective than r_b^1, MLRP-dominate s's beliefs conditional on being rejected by r_b^1, and these beliefs MLRP-dominate b's beliefs conditional on b being rejected from r_b^1 by assumption.

To complete the proof, I rule out the case of $r_s^1 > r_b^1$. Without loss of generality, assume that b's beliefs are distributed uniformly on the unit interval and that s's beliefs are given by the increasing probability distribution function f, and assume that $r_s^1 > r_b^1$. Let j be the minimal index such that $r_s^j \leq r_b^j$ (if such index exists). Then $r_s^i \leq r_b^i$ for all $i \in \{j, ..., k\}$ by the inductive hypothesis and the argument in the previous paragraph.

A possible deviation for b is to drop r_b^1 from her portfolio, and to add the highest option on r_s not on r_b, while continuing to rank colleges according to her true preferences (note that $r_s^1 = r_b^2$ is not excluded). Denote the resulting portfolio by $\bar{r}_b$. This deviation will cause her to lose $x > 0$ utils if the realized score is above $c_{r_b^1}$, and to gain $y > 0$ utils if the realized score is above the cutoff of the new addition, $c_{r_s^n}$, but lower than the cutoffs of all higher-ranked options (i.e., lower than $c_{r_b^m} := \min\left\{c_{\bar{r}_b^i} | \bar{r}_b^i < r_s^n\right\}$). Since these higher-ranked options are all ranked lower than r_b^1, their cutoffs are all lower than $c_{r_b^1}$. Since r_b is chosen by b over $\bar{r}_b$, $A_1 := \left(1 - c_{r_b^1}\right) \cdot x$ is greater than $B_1 := \left(c_{r_b^m} - c_{r_s^n}\right) \cdot y$, where $\left(1 - c_{r_b^1}\right)$ and $\left(c_{r_b^m} - c_{r_s^n}\right)$ are the probabilities that the relevant events are realized under the assumption that beliefs are uniform.

If $\bar{r}_b^i \neq r_s^i$ for some $i < j$, the argument can be iterated. This time we get $A_1 + A_2 \geq B_1 + B_2$, where A_2 is the probability that the second change causes a utility loss times the magnitude of the loss, and B_2 is the probability of a gain times the magnitude of the gain (both are the incremental gains/losses after the first change having taken place). Clearly, scores contributing to B_2 are all lower than those contributing to A_2.

In general, for all $m < j$, we get $\sum_{i=1}^{m} A_i \geq \sum_{i=1}^{m} B_i$ and scores that contribute to B_m are all lower than those that contribute to A_m. Of note, the inequalities do not necessarily hold for each i separately. That is, B_2 may be higher than A_2, but the difference must be smaller than the difference between A_1 and B_1.

Now, consider s's deviation of changing her ROL to equal r_b^i for all $i < j$. For simplicity, assume first that $r_s^j = r_b^j$. The gain/loss score ranges are the same and so are the magnitudes of gain/loss, but the probabilities that the realized score is in these ranges change. Denote by A_i^f and B_i^f the corresponding quantities for a, and let f_i denote the value of f at the lowest score in the range associated with A_i (and A_i^f). This score is at least as high as the highest score in the range associated with B_i (and B_i^f); thus, by MLRP, we get $A_i^f \geq f_i A_i$ and $f_i B_i \geq B_i^f$. Hence,

$$A_1^f \geq f_1 A_1 \geq f_1 B_1 \geq B_1^f,$$

and so

$$\begin{aligned} A_1^f + A_2^f = B_1^f + (A_1^f - B_1^f) + A_2^f \geq f_1 B_1 + f_1(A_1 - B_1) + f_2 A_2 \geq f_1 B_1 + f_2(A_1 - B_1) + f_2 A_2 = \\ = f_1 B_1 + f_2(A_1 - B_1 + A_2) \geq f_1 B_1 + f_2 B_2 \geq B_1^f + B_2^f, \end{aligned}$$

The inequalities use the fact that f_i is decreasing in i (by MLRP, since A_i's critical scores are decreasing by construction) and that $A_1 \geq B_1$, and later that $A_1 + A_2 \geq B_1 + B_2$.

Generally,

$$\begin{aligned}
A_1^f + A_2^f + ... + A_m^f &= B_1^f + B_2^f + ... + B_{m-1}^f + A_m^f + \sum_{i<m}(A_i^f - B_i^f) \geq \\
&\geq B_1^f + B_2^f + ... + B_{m-1}^f + f_m A_m + \sum_{i<m} f_i(A_i - B_i) = \\
&= B_1^f + B_2^f + ... + B_{m-1}^f + f_m A_m + f_m \sum_{i<m}(A_i - B_i) + \sum_{i<m}(f_i - f_{i+1}) \cdot \sum_{j \leq i}(A_i - B_i) \geq \\
&\geq B_1^f + B_2^f + ... + B_{m-1}^f + f_m \left(A_m + \sum_{i<m}(A_i - B_i) \right) \geq \\
&\geq B_1^f + B_2^f + ... + B_{m-1}^f + f_m B_m \geq B_1^f + B_2^f + ... + B_m^f,
\end{aligned}$$

but this means that a's deviation is weakly profitable.[14]

Finally, if $r_s^j \neq r_b^j$, it must be that $r_b^j > r_s^j$ (by the definition of j). This means that B_{j-1} is actually smaller from the perspective of s (as her outside option is more desirable). This only increases s's incentive to use the above deviation.

Thus, it is optimal for s to use the first $j-1$ entries on b's ROL. Repeating this process, enumerating over all b-optimal ROLs completes the proof (since s's ROL becomes more aggressive in each iteration).

□

"Falling through the Cracks"

It is clear that better students, who have higher expectations about their desirability to schools, are sometimes more likely to be assigned to some college. I now show that the opposite is also possible. Namely, in some cases better students are more likely to remain unmatched, due to search frictions.

Proposition 3. *For all $k < N$, there exist pairs of agents s and b with identical preferences, but s's beliefs about her score are more optimistic in the sense of MLRP, such that under their respective optimal size-k portfolios s is more likely than b to be unassigned.*

Proof. See Appendix. □

2.3.2 Relaxation of Constraints

What happens to the optimal portfolio when the constraint on the number of schools is relaxed?[15] To answer this question, the following lemma will be useful.

[14] I have used the fact that f_i's are decreasing, and so the expressions $(f_i - f_{i+1}) \cdot \sum_{j \leq i}(A_i - B_i)$ are a product of positive numbers.

[15] There are two ways to address this question. Here, I take the partial equilibrium approach, by considering a large market and relaxing the constraint for an individual agent. In a separate project, I consider the equilibrium implications of relaxing the constraint for all agents, which affects admission cutoffs.

Lemma 3. *For all k, $c' < c \implies r(k,c)$ is at least as aggressive as $r(k,c')$.*

Proof. Follows from Proposition 2. $\square$

Theorem 1. *Optimal portfolios of larger sizes span a wider range of options in terms of desirability (selectivity), and they refine the grid. Formally, for all k, $r^1(k+1,1) \geq r^1(k,1) \geq r^2(k+1,1) \geq r^2(k,1) ... \geq r^{k+1}(k+1,1)$.*

Theorem 1 echoes the findings of Ajayi (2011) and Pallais (2015). Example 4.1 in Section 4.1 illustrates that all the inequalities in the statement of Theorem 1 can hold strictly.

Proof. For all $j \leq k$, $r^{j+1}(k+1,c) = r^j(k, c_{r^1(k+1,c)})$. Since $c_{r^1(k+1,1)} \leq 1$ we have $r^j(k, c_{r^1(k+1,1)}) \leq r^j(k,1)$. Thus, $r^j(k,1) \geq r^{j+1}(k+1,1)$ for all $j \leq k$.

The same argument applied to the dual problem implies that $\overset{\circlearrowright}{r}{}^{j}(k,1) \geq \overset{\circlearrowright}{r}{}^{j+1}(k+1,1)$ for all $j \leq k$, and thus $r^{k-j}(k,1) \leq r^{k-j}(k+1,1)$ for all $j \leq k$, by Proposition 1. $\square$

To gain intuition, note that when an option is ranked highly, the probability that the next-ranked option will be relevant is scaled down additively (that is, the probability of assignment to the next-highest option is decrease by the same amount, regardless of the identity of this option). This means that the attractiveness of schools with higher ex-post value (high u_i) decreases more when another school is ranked above them, or when the school above them is replaced with a lower-cutoff school (as illustrated in Figure 4). Symmetrically, when the constraint on the portfolio size is relaxed the continuation value is higher, making aggressive gambles more appealing. Intuitively, these forces give agents incentives to widen the range of selectivity of their applications. In Section 3 I show that only the latter force applies when admissions are independent conditional on the information available to the applicant.

I now prove that the marginal benefit from relaxing the constraint on the size of the portfolio is decreasing. The proof invokes Theorem 1 and uses a technique similar to the one developed in Wong (2014).

Theorem 2. *The marginal benefit from increasing the size of the portfolio is decreasing. Formally, the function $MB(k) = v(k,1) - v(k-1,1)$ is decreasing in k.*

Proof. For brevity, denote by $r(k) := r(k,1)$ an optimal portfolio of size k. Then

$$v(k+1) = U(r(k+1)) = \sum_{i=1}^{k+1} \left(c_{r^{i-1}(k+1)} - c_{r^i(k+1)}\right) u(r^i(k+1)),$$

where $U(\cdot)$ represent the expected utility from a portfolio, and similarly

$$v(k-1) = U(r(k-1)) = \sum_{i=1}^{k-1} \left(c_{r^{i-1}(k-1)} - c_{r^i(k-1)}\right) u(r^i(k-1)).$$

I will identify two (potentially suboptimal) size-k portfolios, p and p', such that

$$v(k+1)+v(k-1)\leq U(p)+U(p').$$

By the definition of $v(k)$ this will imply that

$$v(k+1)+v(k-1)\leq 2v(k)$$

and hence

$$v(k+1)-v(k)\leq v(k)-v(k-1)$$

as required.

To construct the portfolios p and p', note that it is possible to list all the $2k$ entries of $r(k+1)$ and $r(k-1)$, such that entries are sorted according to decreasing order of desirability of the corresponding school and for all $i\in\{1,2,...k-1\}$ the entry $r^i(k+1)$ appears before $r^i(k-1)$ on the sorted list. The last requirement is possible to satisfy by Theorem 1.

Since $r(k+1)$ has two more entries than $r(k-1)$, there must exist a minimal m such that $r^m(k+1)$ and $r^{m+1}(k+1)$ appear consecutively on the sorted list. Let p be the ROL that is equal to $r(k+1)$ up to the m-th entry followed by the last $k-m$ entries of $r(k-1)$, and p' be the ROL consisting of the first $m-1$ entries of $r(k-1)$ followed by the last $k+1-m$ entries of $r(k+1)$. Then

$$v(k+1)+v(k-1)-u(p)-u(p')=\left(c_{r^m(k+1)}-c_{r^{m-1}(k-1)}\right)\left(u(r^{m+1}(k+1))-u(r^m(k-1))\right)\leq 0,$$

where the inequality follows by the following observations. First, $c_{r^m(k+1)}\leq c_{r^{m-1}(k-1)}$ because $r^{m-1}(k-1)$ appears before $r^m(k+1)$ on the sorted list (by the minimality of m), and thus $r^{m-1}(k-1)$ is weakly more desirable. If the inequality did not hold, $r^m(k+1)$ would have been a dominated choice (by $r^{m-1}(k-1)$), in contradiction to the optimality of $r(k+1)$. Second, $u(r^{m+1}(k+1))\geq u(r^m(k-1))$, again by the definition of m. This completes the proof. $\square$

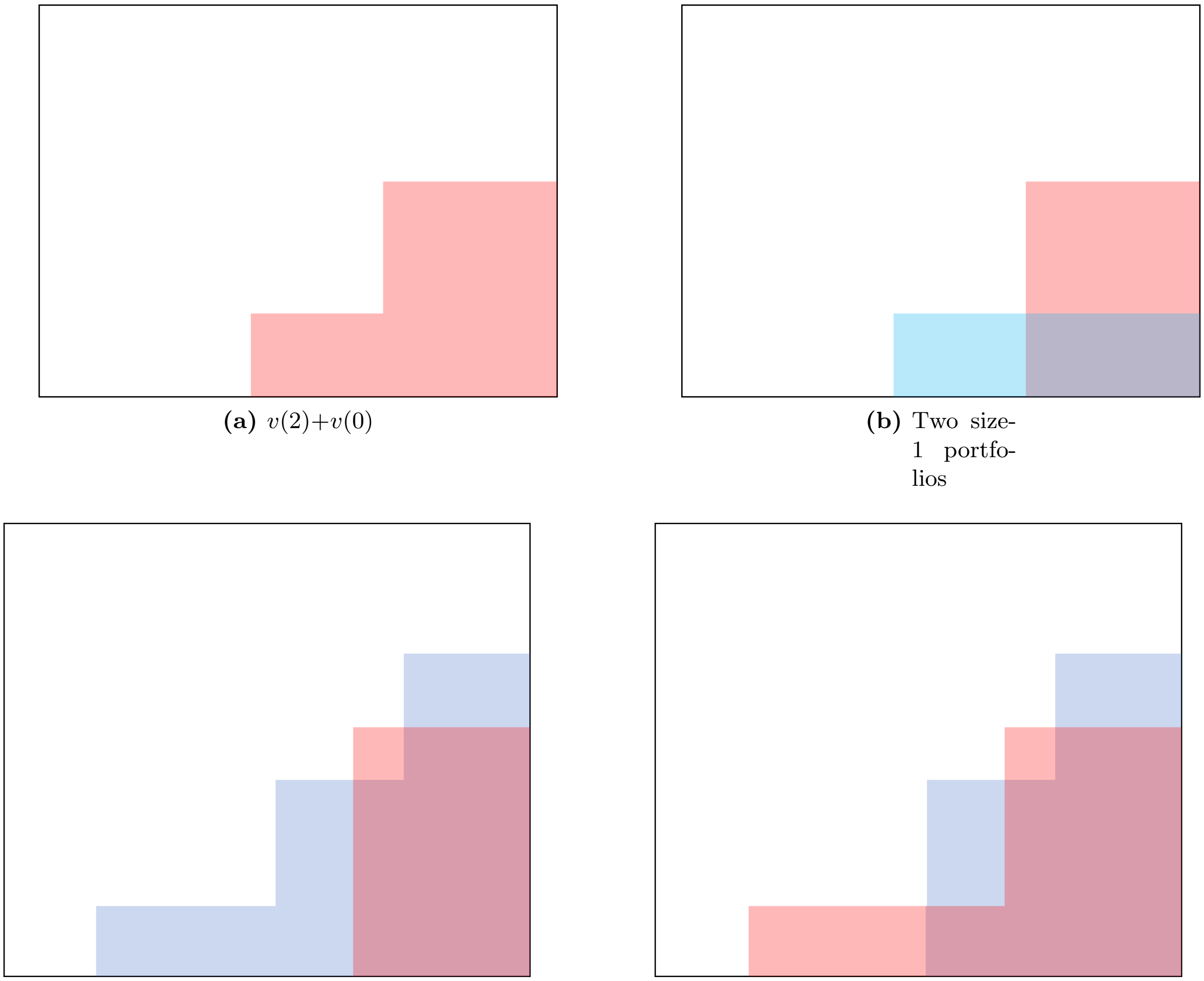
2.3.3 Costly Portfolios

In this section I consider the case that applying to an additional option is costly but the size of the portfolio is not constrained. Specifically, I concentrate on the case of constant marginal cost. I show that the combination of uncertainty and portfolio cost may lead to further failures in the assortativeness of the match.

Proposition 4. *Let the marginal cost of application be $c>0$. Then, there exist pairs of agents s and b with identical preferences, but s's beliefs about her score are more optimistic in the sense of MLRP, such that:*

1) s's optimal portfolio is larger than b's,

2) s's optimal portfolio is smaller than b's,



(a) $v(2)+v(0)$

(b) Two size-1 portfolios

(c) $v(3)+v(1)$

(d) Two size-2 portfolios

Figure 8: Illustration of the proof of Theorem 2

Panels 8a and 8b show two size-1 portfolios whose covered areas sum to more than the area covered by the optimal size-2 portfolio (plus zero, the coverage of the optimal size-0 portfolio). Panels 8c and 8d show two size-2 portfolios whose covered areas sum to more than the sum of areas covered by the optimal portfolios of size 1 and 3. In each panel, each of the portfolios is represented by a different color (pink or light blue). The darker areas indicate areas that are double counted when summing the areas covered by the two portfolios as they are covered by both. The figures highlight that the union of the two portfolios cover the same area, but in the cases of same size portfolios there is larger overlap – i.e., a larger area is double counted.

3) s's optimal portfolio is identical to b's.

Proof. See Appendix. □

Remark 4. At points of indifference between a large and a small portfolio, the larger one spans a wider range of options in terms of selectivity (by Theorem 1). Hence, in cases where a marginal improvement in beliefs leads to a reduction in the portfolio size assortativeness is compromised.

Remark 5. I focus on the simple case of constant marginal cost to reassure the reader that the result is not driven by the cost function. The proof can easily be generalized to cover other cost functions.

2.3.4 Risk Aversion

How does risk aversion affect the optimal portfolio? Adopting the definition of Coles and Shorrer (2014), I say that agent b is *more risk averse* than s if there exists a concave nondecreasing function, ϕ, such that $\phi(0) = 0$ and for any school, i, the utility that b derives from attending it, u_i^b, is equal to $\phi(u_i^s)$. In other words, the two agents share the same ordinal preferences, but the relative marginal benefit from attending a preferred school is smaller for the more risk averse agent. The assumption that $\phi(0) = 0$ is just a convenient normalization, as the concavity of ϕ is unaffected by positive affine transformations. This normalization holds the value of the outside option fixed.

Proposition 5. *Let s and b be agents with identical beliefs and ordinal preferences, but b is more risk averse than s. Then for all k, there exists an s-optimal size-k portfolio that is* at least as aggressive as *any b-optimal size-k portfolio.*

Proof. What are the implications of b being more risk averse than s for the dual decision problems? To begin with, note that $\bar{u}_i^s = \bar{u}_i^b$ for all i, as these depend on values of c_i exclusively, and these are equal for both students.

Next, note that if b is more risk averse than s, there must exist a concave transformation, ϕ, such that $\phi(0) = 0$ and $u_i^b = \phi(u_i^s)$. Set the convention that $c_0 \equiv 1$ and $u_{N+1} \equiv 0$, and observe that since ϕ is concave and increasing, it follows that for all $N \geq i > 1$,

$$\frac{\phi(u_{i-1}^s) - \phi(u_i^s)}{u_{i-1}^s - u_i^s} \geq \frac{\phi(u_i^s) - \phi(u_{i+1}^s)}{u_i^s - u_{i+1}^s}.$$

Adding and reducing 1 everywhere yields

$$\frac{[1 - \phi(u_i^s)] - \left[1 - \phi(u_{i-1}^s)\right]}{[1 - u_i^s] - \left[1 - u_{i-1}^s\right]} \geq \frac{\left[1 - \phi(u_{i+1}^s)\right] - [1 - \phi(u_i^s)]}{\left[1 - u_{i+1}^s\right] - [1 - u_i^s]},$$

which can be rewritten as

$$\frac{\bar{c}_{i-1}^b - \bar{c}_i^b}{\bar{c}_{i-1}^s - \bar{c}_i^s} \geq \frac{\bar{c}_i^b - \bar{c}_{i+1}^b}{\bar{c}_i^s - \bar{c}_{i+1}^s}.$$

Now, note that given a set of options ordered by their selectivity, the only thing that matters for the purposes of decision making is the probabilities of the score realizations between each pair of cutoffs (and not the cardinal value of the cutoff, which is sufficient when the distribution is assumed to be uniform). Based on this observation, and on the above inequality, there are choice problems equivalent to $(\bar{c}_i^s, \bar{u}_i^s)_{i=1}^N$ and $\left(\bar{c}_i^b, \bar{u}_i^b\right)_{i=1}^N$ such that $\bar{c}_i^s = \bar{c}_i^b$ for all i, $\bar{u}_i$'s are unchanged (and thus equal), and the beliefs of b are more optimistic in the sense of MLRP.[16] Thus, the optimal portfolio for b in the dual problem is at least as aggressive as that of s by Proposition 2, which means that the optimal portfolio of s in the primal problem is at least as aggressive as that of b by Proposition 1. □

2.3.5 Outside Options

How do agents' outside options affect their optimal portfolios? To answer this question, I dispense with the normalization of the value of the outside option to 0. Instead, I compare the optimal behavior of individuals who are identical, except that they have different access to outside options. This may occur, for example, when a centralized school-choice mechanism allocates seats in public schools, but families differ in their ability to pay for a private school, or when students have different default assignments in case they are not placed by the mechanism (as is the case in New Orleans; Gross, DeArmond and Denice, 2015).

Theorem 3. *Let s and b be agents who are identical, except that b's outside option MLRP-dominates s's outside option (and both outside options are independent of the score). Then for all k, there exists a b-optimal size-k portfolio that is at least as aggressive as any s-optimal size-k portfolio.*

Proof. Let $\boldsymbol{o_x}$ denote the (stochastic) outside option of agent x. Write $\hat{u}_x(z) = \mathrm{E}\left[\max\left\{o_x, u_x(z)\right\}\right]$, where 0 stands for no assignment, and $u_x(0) = 0$. Note that the optimal portfolio for agent x solves

$$\max_{r} \sum_{1\leq i\leq |r|} \max\left\{0, \min_{j<i}\{c_{r^j}\} - c_{r^i}\right\} \hat{u}_x(r^i) + \max\left\{0, \min_{r^j\in r}\{c_{r^j}\}\right\} \mathrm{E}\left[o_x\right] - C(|r|),$$

where the "new" term corresponds to the outside option that was normalized to 0 previously (equivalently, deduct from each $\hat{u}_x$ the expected value to x from consuming her outside option, $\mathrm{E}\left[o_x\right]$, and use the previous formulation).

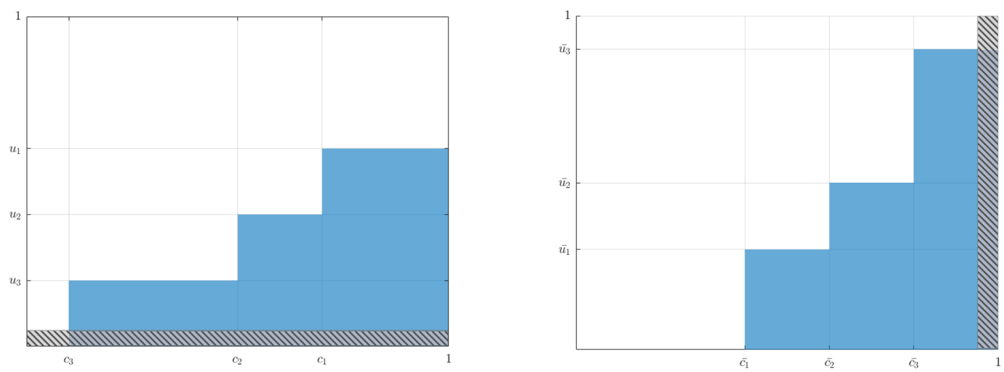
Denote the CDF of the random variable $\boldsymbol{o_x}$ by F_x, and note that

$$\hat{u}_x(z) = u_x(z) + \int_{u_x(z)}^{\infty} [1 - F_x(s)]ds,$$

[16] For example, hold $\bar{c}_i^s$'s fixed and let the distribution of scores for b be constant $\left(\frac{\bar{c}_i^b - \bar{c}_{i+1}^b}{\bar{c}_i^s - \bar{c}_{i+1}^s}\right)$ between consecutive pairs of cutoffs.

thus, b is less risk averse than s if o_b MLRP-dominates o_s.[17] $\square$

Corollary 1. *Let s and b be agents who are identical, except that b's (deterministic) outside option is more desirable than s's. Then for all k, there exists a b-optimal size-k portfolio that is at least as aggressive as any s-optimal size-k portfolio.*



(a) **(b)**

Figure 9: Illustration of a proof of Corollary 1
Panel 9a shows the effect of an increase in the outside option (the gray striped area is removed from the coverage problem). Panel 9b presents the corresponding change in the dual problem. The effect on the dual problem is similar to the one illustrated in Figure 4, suggesting an alternative proof for Corollary 1 (by Proposition 1 and Lemma 3).

3 Simultaneous Search with Independent Successes

The class of simultaneous search problems that I study includes as a special case the environment where colleges' admission decisions are independent conditional on the information available to each student, and students attend the most-preferred college that admits them. This is the case, for example, when oversubscribed schools use *independent* (school-specific) lotteries to break priority ties between applicants and schools' priorities are determined based on observable characteristics such as neighborhood of residence (e.g., Abdulkadiroğlu, Pathak and Roth, 2009; Dobbie and Fryer, 2011). This environment is studied by Chade and Smith (2006) who show that a "greedy" Marginal Improvement Algorithm (MIA) identifies the optimal portfolio. Put simply, they show that the optimal portfolio of size k, $r(k)$, is included in the optimal portfolio of size $k+1$, $r(k+1)$. Moreover, Chade and Smith (2006) use their solution to derive some comparative statics.

[17] When the outside option is smooth, $\hat{u}_b(z) = u_b(z) + \int_{u_b(z)}^{\infty} [1 - F_b(s)]ds = u_s(z) + \int_{u_s(z)\equiv u_b(z)}^{\infty} [1 - F_b(s)]ds = \hat{u}_s(z) + \int_{u_s(z)}^{\infty} [1 - F_b(s)] - [1 - F_s(s)]ds$. Thus, $\frac{\mathrm{d}\hat{u}_b}{\mathrm{d}\hat{u}_s(z)} = 1 + F_b(u_s(z) - F_s(u_s(z))$ and $\frac{\mathrm{d}^2\hat{u}_b}{\mathrm{d}^2\hat{u}_s(z)} = f_b(u_s(z) - f_s(u_s(z))$. When outside options are not smooth, there are simple ways to "spread" them without crossing any c_i, making the distributions smooth, without changing any parameter relevant to portfolio choice.

In this section, I provide short and transparent proofs for several of their results using my approach, and derive additional, novel comparative statics.

Theorem 4. *(Chade and Smith, 2006) For each k, the MIA identifies an optimal portfolio of size k.*

Proof. It is clear that Condition 1 is satisfied. Hence, the optimal portfolio can be represented as an ROL and can be solved backwards.

The proof is by induction on k. The case of $k = 1$ is trivial. For larger values of k, there are three cases to consider. First, if $r^1(k)$, the first-ranked option on the optimal ROL of size k, does not appear on (some) $r(k-1)$, then the optimal behavior in the continuation problem is $r(k-1)$. This follows from independence of irrelevant alternatives (IIA), since with independent successes, rejection from one school is not at all informative about chances of success with other schools, and the menu of available options following a rejection from $r^1(k)$ contains all the options in $r(k-1)$ and is contained in the menu from which $r(k-1)$ is chosen.

Second, $r^1(k) = r^1(k-1)$ (for some$r(k-1)$) . The statement holds by induction as the optimal continuations (after being rejected by $r^1(k) = r^1(k-1)$) are the optimal portfolios of size $(k-1)$ and $(k-2)$ from a smaller menu: the original menu of schools excluding all $r^1(k)$.

Third, $r^1(k) = r^j(k-1)$ for some $j > 1$ (for all $r(k-1)$). Hence, $r^1(k-1) < r^1(k)$ (or there is indifference, and a rearrangement of the list would yield the same expected utility and be covered by the second case). Since $|r(k)| > |r(k-1)|$, there exists an option, i, on $r(k)$ that does not appear on $r(k-1)$. In this case, i cannot be $r^1(k)$. Consider the two decision problems of choosing the optimal portfolio of size $k-1$ from the menu of options that does not include i where the outside option is equal to i) 0, and ii) the (stochastic) option i. The solutions of these problems correspond to $r(k-1)$ and to the remainder of $r(k)$ (i.e, $r(k)$ excluding $\{i\}$). But by Theorem $8'$ (which is stated and proved below), applicants with better outside options apply more aggressively, and so it must be that there is another $r(k)$ with $r^1(k) \leq r^1(k-1)$.

□

The proof makes strong use of the fact that no information is conveyed by rejections, and so the optimal continuation does not depend on the "past" or the particular "prefix," other than to note that these choices are no longer available. The heart of the argument is best illustrated by the special case of $k = 2$. In this case, either the highest-ranked school on the optimal ROL of size 2 is the optimal portfolio of size 1, or the optimal portfolio of size 1 is available if the agent is rejected from his first-ranked school, and is thus optimal (by IIA, as rejection by the first-ranked school conveys no information about admission chances in other schools).

The fact that optimal portfolios are nested also implies that $v(k)$, the expected utility from an optimal size-k portfolio, is concave. The k-th option that was added to the ROL could have been added in the $k-1$-st step, and its marginal value would have been even higher (either because the probability that it is relevant is higher, in case the $k-1$ addition to the ROL is more desirable, or because the continuation value is lower, in case the $k-1$ addition to the ROL is less desirable).

Lemma 4. *(Chade and Smith, 2006)The marginal benefit from increasing the size of the portfolio is decreasing.*

Corollary 2. *(Chade and Smith, 2006) If $C(\cdot)$ is convex, the greedy algorithm identifies an optimal portfolio.*

Remark 6. The MIA identifies the optimal portfolio quickly. To see that a dynamic program can do the same, consider the portfolio choice problem where if the agent ranks a lower-index (more desirable) option lower than another option, his chances of admission to that school are 0. This change does not alter the solution, as optimal portfolios are monotonic. And the altered problem satisfies Condition 2.

I now formalize another observation made by Chade and Smith (2006), which they term the upward diversity of optimal portfolios. To make the result transparent, I assume that there are many colleges of each type, where a type is defined by (u_c, p_c), i.e., the cardinal utility from attending the college and the probability of admission. There may also be infinitely many college types (as in the labor market models of Galenianos and Kircher, 2009 and Kircher, 2009).

Theorem 5. *The optimal portfolios are upwardly diverse in the sense that the decision maker derives higher utility from consuming options that belong to larger portfolios only. Formally, if a college c is in $r(k+1)$ but not in $r(k)$, then $u_c \geq u_{c'}$ for any other college c' in $r(k)$.*

Proof. First note that in the current setting if $u_c \geq u_{c'}$ and $p_c \geq p_{c'}$ with one inequality being strict, then colleges of the same type as c' will never appear on any optimal ROL as they are dominated by colleges of the same type as c, and I assumed that there are arbitrarily many colleges of each type. Next, note that the beliefs of an applicant about admission probabilities to a college do not change if the applicant learns she is rejected by other colleges (due to independence). Thus, the decision maker is facing the same problem of finding the optimal portfolio, with the only difference being that the length of the ROL to be chosen is shorter by one (as there are many copies of c). That is, $r(k) = \left(r^1(k), r(k-1)\right)$.[18] Denoting by $v(r)$ the expected utility from the ROL r, I note that $r^1(k)$ is the solution to the following problem:

$$r^1(k) = \operatorname*{argmax}_c p_c u_c + (1-p_c)v(r(k-1)).$$

Since $v(r(k-1))$ is increasing in k, and the same types of colleges are available for any k (since there are plenty), it is clear that higher u (and lower p) colleges become increasingly attractive as k grows large. To see this, note that the problem of choosing an optimal size-1 portfolio is identical to the problem where agents are vertically differentiated (correlation in admission chances only matters when portfolios have more than one school). Thus, the choice of $r^1(k)$ is identical to the choice of an optimal size-1 portfolio with an outside option of $v(k-1)$. The result therefore follows from the fact that $v(k-1)$ is increasing in k and from Corollary 1.[19] $\square$

[18] I slightly abuse notation to avoid introducing notation for concatenation; $\left(r^1(k), r(k-1)\right)$ should be read as the concatenation of $r^1(k)$ and $r(k-1)$.

[19] For an alternative proof, note that, by induction, the outside option $r(k-1)$ MLRP-dominates the outside option $r(k-2)$, and apply Theorem 3.

The proof shows that in every step, the MIA adds a (weakly) more selective college that yields higher utility if attended. The only reason that this does not always happen in the environment studied by Chade and Smith (2006) is that sometimes there are no perfect substitutes for colleges already on the portfolio.

Graphically, the problem faced by the decision maker can be described as follows. The choice of the last college is the rectangle with the larger area under the curve (among options that were not ranked previously). For the j-th-before-last school, the graph should be transformed so that u_c is replaced by $u_c + \frac{1-p_c}{p_c} \cdot v(r(j-1)) = u_c + (\frac{1}{p_c} - 1) \cdot v(r(j-1))$. The second part of the expression is increasing in j, and increasingly so the lower p_c is. Thus lower p_c (higher u_c) options become more attractive as the constraint on the length of the ROL is relaxed.

This graphical approach can be used to prove another theorem of Chade and Smith (2006), namely, that the optimal portfolio of size k is more aggressive than the portfolio consisting of the k schools that yield the highest expected utility (on their own).

Theorem 6. *(Chade and Smith, 2006) The best portfolio of size k is at least as aggressive as the k most attractive singleton portfolios.*

I present an alternative proof.

Proof. The proof proceeds by induction. The case of $k = 1$ is obvious. For $k > 1$, note that the k-th-ranked option on the optimal ROL, $r^k(k)$, is one of the k most attractive singletons. The $k - 1$ higher-ranked schools are the optimal portfolio of size $k - 1$ from the menu $X \backslash r^k(k)$ with the stochastic outside option $r^k(k)$. By Theorem 8′ (which is stated and proved below), applicants with better outside options apply more aggressively. Hence, this portfolio is more aggressive than the one chosen from the same menu $(X \backslash r^k(k))$ with the outside option equal to zero. By induction, this last portfolio is more aggressive than the $k - 1$ most attractive singletons in $X \backslash r^k(k)$. This completes the proof, as these singletons, together with $r^k(k)$, are the k most attractive singleton portfolios in the original problem. □

Remark 7. As noted by Chade and Smith (2006), Theorem 6 essentially rules out "safety schools": an applicant never applies to a school for its high admission rates unless this is justified by the application yielding high expected utility on its own. This result stands in sharp contrast to our findings in the case of vertically differentiated applicants (for a counterexample, see Section 4.1).

3.1 Risk Aversion

Recall the assumption that lower index options are more desirable. For simplicity, assume that no option is realized with certainty.[20] As before, given two portfolios of the same size, r and r', we say that r is *at least as aggressive* as r' if $r^i \leq r'^i$ for all i.

[20] Without the assumption that no option is realized with certainty, the definition an ROL being at least as aggressive will have to account for cases where a desirable option is realized with certainty, deeming the rest of the ROL irrelevant. This can only occur when agents are effectively unconstrained. While accounting for this possibility will complicate the definitions, the result will continue to hold.

Proposition 6. *Let s and b be agents with identical beliefs and ordinal preferences, but b is more risk averse than s. Then for all k, there exists an s-optimal size-k portfolio that is at least as aggressive as any b-optimal size-k portfolio.*

Proof. Let r_s and r_b be optimal ROLs of size k of agents s and b respectively. Assume that r_s is not at least as aggressive as r_b. I construct an s-optimal ROL that is at least as aggressive as both r_s and r_b. Iterating this process, enumerating over all optimal r_b's, yields the desired s-optimal ROL.

Since r_s is not at least as aggressive as r_b, there exists $i \leq k$ such that $r_s^i > r_b^i$. Let i^* be the maximal such index. For each agent, $x \in \{s, b\}$, normalize her utility function (using a different normalization) so that her expected utility from the portfolio r_x conditional on being rejected from the first i^* options on her ROL (i.e., the continuation value at i^*) is equal to 0, and her utility from consuming $r_b^{i^*}$ is equal to 1 (that is, both agents assign to b's i^*-th choice a value of 1).[21]

Since i^* is the maximal index with the property $r_s^i > r_b^i$, and it is suboptimal to rank options in a way that is inconsistent with true preferences (thus, $r_b^j \geq r_s^j > r_s^{i^*}$ for all $j > i^*$), $r_s^{i^*}$ does not appear on r_b (otherwise, there would be a greater index with $r_s^i > r_b^i$).[22] Hence, had $r_b^{i^*}$ been changed to $r_s^{i^*}$, b's (original) optimal continuation would still be available.

Denote by p the probability of acceptance to $r_b^{i^*}$ and by p' the probability of acceptance to $r_s^{i^*}$. Since b chose $r_b^{i^*}$ over $r_s^{i^*}$ we have that

$$p = p \cdot 1 + (1-p) \cdot 0 \geq p' \cdot u_b(r_s^{i^*}) + (1-p') \cdot 0 = p' \cdot u_b(r_s^{i^*}),$$

where the zeros stand for the expected utility from the optimal continuation, which is still available after the deviation. Similarly, since s chose $r_s^{i^*}$ and his optimal continuation over $r_b^{i^*}$and the same continuation (which may be suboptimal),[23] we have that

$$p' \cdot u_s(r_s^{i^*}) = p' \cdot u_s(r_s^{i^*}) + (1-p') \cdot 0 \geq p \cdot 1 + (1-p) \cdot 0 = p.$$

These two inequalities together imply that

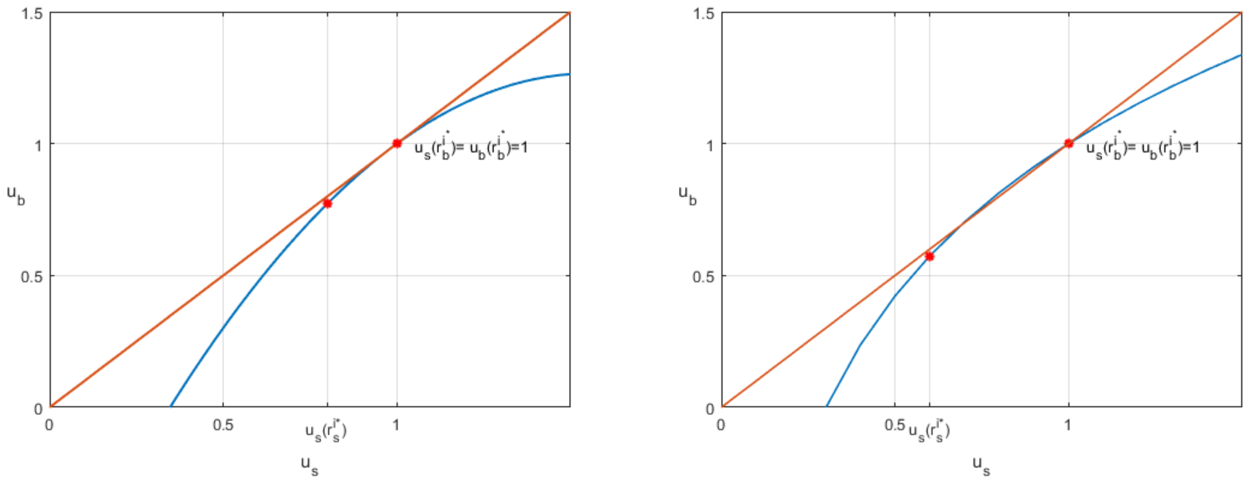
$$u_s(r_s^{i^*}) \geq \frac{p}{p'} \geq u_b(r_s^{i^*}).$$

As Figure 10 illustrates, this inequality implies that for any school $j > r_s^{i^*}$ (i.e., schools that are less desirable than $r_s^{i^*}$) we have $u_s(j) \geq u_b(j)$ and $u_s(0) \geq u_b(0)$, since there is a concave nondecreasing function ϕ such that $u_b(\cdot) = \phi \circ u_s(\cdot)$, and $u_b(r_b^{i^*}) = u_s(r_b^{i^*}) > u_s(r_s^{i^*})$. Thus, by using r_b's suffix, s is assured of at least as high a expected utility (namely, at least 0). But this is what s gets from his optimal suffix, and so this suffix is optimal, and all the inequalities hold with equality.

[21] This normalization is possible since ranking consistently with one's true preferences is optimal; thus, the expected value from b's continuation portfolio is the average of utilities of options less preferred than $r_b^{i^*}$, and s's continuation is such an average over options less preferred than $r_s^{i^*}$, which, by the definition of i^*, is less preferred than $r_b^{i^*}$.

[22] For simplicity, but without loss, I always assume that both agents break ties in favor of lower-index options.

[23] It is possible that $r_b^{i^*}$ belongs to r_s. I treat this case in the Appendix using an argument analogous to the one used in the body of the paper.



(a) **(b)**

Figure 10: The case of $u_s(r_s^{i^*}) > u_b(r_s^{i^*})$

The figure plots u_b as a function of u_s, after both normalizations (blue curve). The red curve is the identity map, representing the identity $u_s \equiv u_s$. The figure illustrates that if $u_s(r_s^{i^*}) > u_b(r_s^{i^*})$, u_b must be steeper than u_s to the left of $u_s(r_s^{i^*})$, and hence lower than u_s to the left of $u_s(r_s^{i^*})$. In the case of equality (not illustrated), the function must be weakly lower for similar reasons.

To complete the proof, there are two cases to consider. First, if i^* is equal to k, then the normalization has the outside option valued at 0, and the problem of choosing the k-th-ranked school is no different from the case with vertically differentiated students (correlation in admission successes has no implications for the optimal choice of portfolios of size 1). Thus, changing $r_s^{i^*}$ to $r_b^{i^*}$ does not compromise optimality, and results in an ROL that is at least as aggressive as r_s.

Otherwise, $i^* < k$, which means that $u_s(j) = u_b(j)$ for all $j \geq r_s^{i^*}$ and for the outside option. But then the same argument implies that changing $r_s^{i^*}$ to $r_b^{i^*}$ does not compromise optimality, and results in an ROL that is at least as aggressive as r_s. Repeating the argument using the resulting ROL (fewer than k times) results in an s-optimal ROL that is at least as aggressive as both r_s and r_b. Finally, repeating this process using the resulting ROL as r_s and enumerating over all b-optimal ROLs completes the proof. $\square$

3.2 Beliefs

Recall that lower indices options are assumed more desirable.

Theorem 7. *(Chade and Smith, 2006) Assume that s and b are two agents who are identical, except that b's admission chances at each option, $\beta_1, \beta_2, ..., \beta_N$, dominate s's chances, $\alpha_1, \alpha_2, ..., \alpha_N$, in the sense that $\beta_i \geq \alpha_i$ for all i, and $\frac{\beta_i}{\alpha_i} \geq \frac{\beta_{i+1}}{\alpha_{i+1}}$. Then there exists a b-optimal size-k portfolio that is at least as aggressive as any s-optimal size-k portfolio.*

Proof. By way of contradiction, let k be the minimal size of ROL for which one can find a counterexample. For $k = 1$, if $j^* \in \arg\max \beta_i u_i$ and $i^* \in \arg\max \alpha_i u_i$ then if $j^* > i^*$ we have $\frac{u_{i^*}}{u_{j^*}} \geq 1$

and hence

$$\frac{\beta_{i^*}}{\beta_{j^*}} \cdot \frac{u_{i^*}}{u_{j^*}} \geq \frac{\alpha_{i^*}}{\alpha_{j^*}} \cdot \frac{u_{i^*}}{u_{j^*}} \geq 1,$$

where the first inequality follows from β dominating α and the second follows from the optimality of i^* for s. But this means that i^* is also optimal for b. This argument is analogous to the one used in the model with vertically differentiated applicants.

For general $k > 1$, there are three cases to consider. First, if $r_b^1(k) = r_s^1(k)$, remove this option from the set of available options, X, and note that the optimal continuation is the optimal portfolio of size $k-1$ in the resulting problem. But this decision problem contradicts the minimality of k.

Second, if $r_b^1(k) < r_s^1(k)$, then, since ranking consistently with one's preferences is optimal, either $r_b^2(k) \leq r_s^1(k)$ and a fortiori $r_b^2(k) < r_s^2(k)$, or $r_b^j(k) \leq r_s^j(k)$ for all $1 < j \leq k$ by IIA and the minimality of k.[24] In the first case, either one can iterate the argument until a stage where $r_b^{j+1}(k) > r_s^j(k)$ (then use the minimality of k and IIA), or one gets all the way to $r_b^k(k) \leq r_s^{k-1}(k)$, which means r_b is at least as aggressive as r_s, in contradiction to the instance being a counterexample.

Finally, I claim that if $r_b^1(k) > r_s^1(k)$ then there exists another b-optimal size-k ROL that is at least as aggressive as both r_b and r_s (and hence there exists one such ROL that is at least as aggressive as all s-optimal size-k ROLs). Otherwise, let $\hat{\imath}$ be the maximal index, i, such that $r_b^i(k) > r_s^i(k)$. I claim that s's expected utility increases if she changes $r_s^{\hat{\imath}}$ to $r_b^{\hat{\imath}}$. First, note that s does not rank $r_b^{\hat{\imath}}$ on her ROL. This holds since, if $\hat{\imath} < k$, then $r_s^{\hat{\imath}} < r_b^{\hat{\imath}} < r_b^{\hat{\imath}+1} \leq r_s^{\hat{\imath}+1}$, where the middle inequality is due to the optimality of ROLs that are consistent with preferences, and the other inequalities follow from the definition of $\hat{\imath}$.

Next, note that b is free to imitate s's suffix after $\hat{\imath}$ and b has better chances with each option. Thus, b's continuation value after the $\hat{\imath}$-th rank is (weakly) higher. Holding the rest of the ROL fixed, at the history that agent x was rejected by all his $\left(\hat{\imath}-1\right)$-first-ranked options (the only relevant one for the purpose of deciding on the $\hat{\imath}$-th rank) agent x must be choosing optimally. It is not difficult to see that this problem has the same solution as the problem of choosing the optimal portfolio of size 1 from the set of options $\left\{r_s^{\hat{\imath}}, r_b^{\hat{\imath}}\right\}$ where agent x's outside option is $V_x^{\hat{\imath}}$, which is the expected value for x from her optimal portfolio conditional on being rejected by her first $\hat{\imath}$-ranked options. But for portfolios of size 1, there is no difference between the environment with independent admissions that is studied here and the one studied in Section 2. The result therefore follows from Theorem 3. □

3.3 Outside Options

Theorem 8. *Let s and b be agents who are identical, except that b's outside option MLRP-dominates s's outside option (and both outside options are independent of the realization of options in X). Then for all k, there exists a b-optimal size-k portfolio that is at least as aggressive as any s-optimal size-k portfolio.*

[24] For the full argument, consider the choice of a portfolio of size $k-1$ by the two agents from the menu of all options except for the two options they rank first. IIA applies since $r_b^2(k) > r_s^1(k)$ (and $r_s^2(k) > r_s^1(k)$).

Proof. The proof is identical to the proof of Theorem 3, with the only difference being that Proposition 6 is invoked instead of Proposition 5. □

Corollary 3. *Let s and b be agents who are identical, except that b has a more desirable (deterministic) outside option than s. Then for all k, there exists a b-optimal size-k portfolio that is at least as aggressive as any s-optimal size-k portfolio.*

4 Strategic Simplicity and Tie-Breaking Rules

School districts around the world determine school assignment based on coarse priorities (i.e., many students share the same priority for the same school). In order to respect schools' capacity constraints, lotteries are often used to break ties in priority. There are two common approaches to breaking priority ties in oversubscribed schools: a single common lottery or multiple independent lotteries. A large literature studies different aspects of this design choice (e.g., Abdulkadiroğlu, Pathak and Roth, 2009; Arnosti, 2016; Ashlagi, Nikzad and Romm, 2015; Ashlagi and Nikzad, 2016; De Haan et al., 2015; Schmelzer, 2016).

This section connects the literature on priority-tie-breaking in school choice to the discussion on protecting unsophisticated agents (e.g., Rees-Jones, 2017; Hassidim et al., 2017), by highlighting an aspect that, to the best of my knowledge, has not received attention in the literature. Namely, how difficult it is for humans to identify the optimal portfolio of applications? I provide two formal senses in which identifying the optimal portfolio is easier when ties are broken using multiple lotteries.

4.1 The "Greedy" Approach

Theorem 4 (due to Chade and Smith, 2006) shows that the "greedy" marginal improvement algorithm identifies the optimal portfolio when ties are broken using multiple lotteries (and more generally, when admission chances are independent conditional on the information available to the applicant). The following example shows that this result does not extend to the case of a single lottery (and, more generally, to vertically differentiated applicants). To be sure, the example also shows that the "reverse greedy" approach also fails. A greedy algorithm identifies an optimal portfolio of any size only when the optimal portfolios are nested. In the example, the optimal portfolios of size 1 and of size 2 are disjoint.

Example 2. Consider an environment with three colleges such that $c_1 = \frac{3}{4}$, $c_2 = \frac{1}{2}$, and $c_3 = 0$, and $u_1 = 4$, $u_2 = 2.01$, and $u_3 = 1$. The values were selected so that singleton portfolios yield the same expected utility (i.e., $(1 - c_i) \cdot u_i = 1$), except that 2 yields a slightly higher expected utility. It is thus clear that the best singleton portfolio consists of college 2.

Next we consider the optimal portfolio of size 2. Denote by v_{ij} the value from the two-school portfolio $\{i, j\}$. Then

$$v_{12} = \frac{1}{4} \cdot 4 + (\frac{1}{2} - \frac{1}{4}) \cdot 2.01 \approx \frac{6}{4},$$

$$v_{13} = \frac{1}{4} \cdot 4 + (1 - \frac{1}{4}) \cdot 1 = \frac{7}{4},$$

$$v_{23} = \frac{1}{2} \cdot 2.01 + (1 - \frac{1}{2}) \cdot 1 \approx \frac{6}{4}.$$

Hence, the optimal portfolio of size 2 consists of colleges 1 and 3, and does not include college 2. A greedy algorithm will only achieve a fraction of approximately $\frac{6}{7}$ of the expected utility from the optimal portfolio.

Ajayi and Sidibe (2015) use a greedy approach to approximate the optimal portfolio in their empirical study of school choice in Ghana. To get a satisfactory approximation, they enhance the simple algorithm in various ways. The following proposition provides an explanation of their success: the baseline that they chose, the greedy algorithm, cannot perform too poorly. Even before introducing the modifications of Ajayi and Sidibe (2015), a greedy algorithm is assured to achieve at least 63% of the expected utility from the optimal portfolio.

Proposition 7. *For each k, the greedy algorithm achieves at least a $(1 - \frac{1}{e})$-fraction of the expected utility from the optimal size-k portfolio.*

Lemma 5. *The expected utility from a portfolio is a nonnegative, monotonic submodular function.*

Proof. The graphical representation of the problem shows clearly that the constrained portfolio choice problem is a coverage problem, and coverage problems are a famous example of such functions (e.g., Hochbaum, 1996). □

Proof. (of the proposition) Follows from Lemma 5 and well-known results in approximation (e.g., Hochbaum, 1996).[25] □

4.2 Uniqueness of Local Optima

The following proposition illustrates that when ties are broken using multiple lotteries the global optimum is the unique local optimum, in the sense that for every suboptimal portfolio there is another portfolio that differs by at most one school which yields higher expected utility. I then show that this result does not hold when ties are broken using a single lottery.

Our findings in this section have practical implications for demand estimation using reports to centralized mechanisms. The "one-shot swaps" result below was first proved by Larroucau and Rios (2018) in order to show that there is no loss of information by restricting attention to inequalities that follow from "one-shot swaps" in their empirical study of Chilean college admissions. My "multiplicity of local optima" result shows that this is not the case when admissions chances are correlated.

[25] The statement and proof of the results I use are nicely exposited in the Math∩Programing blog at https://jeremykun.com/2014/07/07/when-greedy-algorithms-are-good-enough-submodularity-and-the-1-1e-approximation/.

Proposition 8. *("One-shot Swaps," Larroucau and Rios, 2018) Assume that ties are broken using multiple lotteries, and let r' be a suboptimal portfolio of length k. Then there exists another portfolio, r'', that yields higher utility, such that the sets of schools on r' and r'' differ by at most one school.*

Proof. The proof includes cases with a stochastic outside option, and proceeds by induction. The case of $k = 1$ is trivial. For $k > 1$, let r' be a suboptimal portfolio of length k. If r' is not ordered according to true preferences, sort it according to true preferences. If sorting does not suffice to achieve an improvement, there are two cases to consider.

First, if there exists a school, $i \in X$, that appears on both r' and $r(k)$, an optimal ROL of length k, then remove it from both ROLs. Note that the ROL resulting from $r(k)$ is an optimal length-$(k-1)$ ROL from the menu $X\backslash\{i\}$ with the stochastic outside option i, and the ROL resulting from r' is suboptimal in the same decision problem. Hence, the result follows from the inductive hypothesis.

Second, if the set of schools on r' is disconnected from all optimal portfolios of size k, then the best singleton portfolio does not appear on r' (as it appears on optimal portfolios of size k by Theorem 4). But this means that replacing the k-th (last) school on r' with this school will (weakly) increase the expected utility. If the increase is only weak, do not make the replacement and, instead, try replacing the $(k-1)$-st school of r' with that on $r(k)$. Since the continuation values are the same for both ROLs, this change should (weakly) increase the expected utility. If the increase is only weak, iterate this process going further up the ROL r' until the first time that the increase is strict (which must occur since r' is suboptimal).[26]

□

Remark 8. With a convex cost function, one can use the decreasing marginal return for portfolio size (Chade and Smith, 2006) to extend the proof to portfolios of general size: the optimal portfolio of size k' will be improved by dropping a school if k' is "too large," and by adding a school if it is too small (as shown by Chade and Smith, 2006).

Proposition 9. *(multiplicity of local optima) There exist a decision problem and a suboptimal ROL, r, such that any change to r that changes the composition of schools on r by at most one school only leads to a reduction in utility.*

Proof. Follows from Example 3. □

Example 3. There are four schools, $X = \{1, 2, 3, 4\}$, such that $c_1 = .9$, $c_2 = .55$, $c_3 = .45$, and $c_4 = .099$, and $u_1 = .9001$, $u_2 = .55$, $u_3 = .45$, and $u_4 = .1$.

The expected utility from the portfolio $\{1, 3\}$, $V_{13} = .1 \times .9001 + .45 \times .45 = .29251$, which is higher than that from any of the "single-swap" portfolios:

[26] While the argument here is reminiscent of the single deviation principle (e.g., Fudenberg and Tirole, 1991), note that in the first case the deviation may require that the swapped school appear in a different position on the ROL.

$$V_{12} = .1 \times .9001 + .35 \times .55 = .28251,$$
$$V_{14} = .1 \times .9001 + .801 \times .1 = .17011,$$
$$V_{23} = .45 \times .55 + .1 \times .45 = .2925,$$
$$V_{34} = .55 \times .45 + .351 \times .1 = .2826.$$

But another portfolio, $\{2, 4\}$, yields a higher payoff:

$$V_{24} = .45 \times .55 + .451 \times .1 = .2926.$$

5 Top Trading Cycles

Variants of the top trading cycles algorithm are strategy-proof when the message space is not constrained. However, whenever this mechanism was implemented in the field, the message space was constrained (Abdulkadiroğlu et al., 2017*c*).

In a recent study, Leshno and Lo (2017) present a large market approach to TTC. Using their framework, one can show that regardless of the structure of priorities (single lottery, multiple lotteries, or more general structures), applicants are facing a problem that is equivalent to the case of a single lottery, studied in Section 2. For this reason, all of the results that were proved about this setting (in Sections 2 and 4) apply. Specifically, when a constraint is in place, TTC is not strategically simple in the senses studied in Section 4, regardless of the way priorities are randomized.

Remark 9. In the case of multiple lotteries, an MLRP improvement to the distribution of scores in one school (holding the rest of the distribution fixed) is sufficient to ensure that the distribution in the equivalent single-lottery problem is MLRP higher.

6 The Immediate Acceptance (Boston) Mechanism

This section studies the Immediate Acceptance (Boston) mechanism. Variants of this mechanism are used by school systems around the world to determine the assignment of students. Mechanisms in this class take as input ROLs, and then proceed to match as many students as possible to their first choice. If some schools are overdemanded, they reject low-priority applicants and the algorithm matches as many of the rejected students as possible to their second choice, and so on. Variants of the mechanism include school priorities that are based on exam scores, a single lottery, multiple independent lotteries, or on a combination of one of these with the identity of the first-ranked school. Schools may also have strict priority classes based, for example, on neighborhood of residence, or on the applicant having a sibling currently enrolled in the school. Additionally, the maximal allowed length of the preference list may vary.

I restrict attention to large markets, where priorities do not depend on one's ROL. Such environments can be summarized by the set of options, X, and for each $i \in X$ a pair $t_i = (z_i, l_i)$, where if option i is ranked higher than z_i then admission to i is guaranteed if the applicant is not assigned to a higher ranked option, if i is ranked equal to z_i then the applicant's (i-specific) priority score must be greater than l_i to secure admission, and if i is ranked lower than z_i then the probability of admission is zero (see Abdulkadiroğlu et al., 2017*a*; Ergin and Sönmez, 2006). By Observation 2, the optimal portfolio in all of the cases mentioned above can be found quickly by using a dynamic program, as proved by Casalmiglia, Fu and Güell (2014).

Proposition 10. *For any agent, b, any constraint, k, and for both a single lottery and independent lotteries, there exists a b-optimal ROL that is monotonic in b's preferences and in z_i. With a single lottery, the ROL is monotonic in t_i according to the lexicographic order.*

Proof. See Appendix. □

6.1 Beliefs

Kapor, Neilson and Zimmerman (2018) find that in New Haven, where a variant of immediate acceptance is in place, students with optimistic (subjective) beliefs applied substantially more aggressively than students with pessimistic (subjective) beliefs (see also Pan, 2019). This section makes such a prediction.

Definition. Given two ROLs, r and r', let k be the rank on r of the first option to which r guarantees assignment if such an option exists and otherwise $|r|$, and let k' be similarly defined. We say that r *is at least as aggressive as* r' if for all $j \leq \min\{k, k'\}$, $r^j \leq r'^j$ (r^j is at least as desirable as r'^j).

Assume that admissions lotteries are independent, but some agents draw from different distributions (e.g., because of their place of residence). The following holds:

Proposition 11. *Let s and b be two agents who are identical, except that b's lottery draw for each option $(\beta_1, \beta_2, ..., \beta_N)$ dominates s's $(\alpha_1, \alpha_2, ..., \alpha_N)$ in the sense that* $\Pr\{\beta_i \geq l_i\} \geq \Pr\{\alpha_i \geq l_i\}$ *for all i, and* $\frac{\Pr\{\beta_i \geq l_i\}}{\Pr\{\alpha_i \geq l_i\}} \geq \frac{\Pr\{\beta_{i+1} \geq l_{i+1}\}}{\Pr\{\alpha_{i+1} \geq l_{i+1}\}}$.[27] *Then for any constraint on the length of the ROL, b has an optimal portfolio that is at least as aggressive as the optimal portfolio of s.*

Proof. See Appendix. □

When ties are broken using a single lottery (or a centralized exam) whose result is unknown at the time the portfolio is chosen, the following holds:

Proposition 12. *Let b and s be agents with identical preferences, but b has more optimistic beliefs about her score in the sense of MLRP. Then for any constraint on the length of the ROL, there exists a b-optimal ROL that is at least as aggressive as any s-optimal ROL.*

[27] To economize on notation, I do not consider cases where $z_i's$ differ between agents, but extending the condition to accommodate such cases is not difficult.

Proof. See Appendix. □

6.2 Risk Aversion

Proposition 13. *Let s and b be agents with identical beliefs and ordinal preferences, but b is more risk averse than s. Then for any length constraint there exists an s-optimal ROL that is at least as aggressive as any b-optimal ROL.*

Proof. See Appendix. □

6.3 Outside Options

Calsamiglia and Güell (2014) find that in Barcelona, where seats in public schools are assigned centrally using a variant of immediate acceptance, families with access to private schools were more likely to choose an aggressive portfolio. The following theorem make such a prediction.

Theorem 9. *Let s and b be agents who are identical, except that b's outside option MLRP-dominates s's outside option (and both outside options are independent of any lottery/score). Then for any length constraint k, there exists a b-optimal size-k portfolio that is at least as aggressive as any s-optimal size-k portfolio.*[28]

Proof. The proof is identical to the proof of Theorem 3, with the only difference being that Proposition 11 (or 12) is invoked instead of Proposition 5. □

Corollary 4. *Let s and b be agents who are identical, except that b has a more desirable (deterministic) outside option than s. Then for any length constraint k, there exists a b-optimal ROL that is at least as aggressive as any s-optimal ROL.*

7 Other Decision Problems

7.1 More General Cost Functions

For convenience, I assumed that $C(\cdot)$ is a function of the length of the ROL. But there are instances where the cost function is more complex. For example, in Hungary, the cost of an ROL depends on the number of programs ranked on the ROL, but applicants can rank multiple tracks in the same program (Biró, 2012), and, in Israel, the cost depends on the number of institutions, but each institution offers multiple tracks (Hassidim, Romm and Shorrer, 2016*a*,*b*). The approach presented in this paper allows the optimal portfolio to be calculated in such settings, as long as $C(\cdot)$ does not take "too many" different values. In what follows I discuss several cases of special interest.

[28] Until recently, seats in Magnet schools in Wake County were assigned centrally using a variant of immediate acceptance, but students had an outside option in the form of a default public school assignment, and were able to apply for seats in application schools, where admission was based on test scores (Dur, Hammond and Morrill, 2018).

7.1.1 Cadet-Branch Matching

There are many real-world allocation mechanisms that use "familiar" algorithms, but their bidding language does not allow applicants the liberty to express their preferences consistently with their preferences. Specifically, applicants are sometimes only able to rank a lottery over options. Famous examples include mechanisms used by the US army, where cadets can only choose between extreme degrees of aversion to additional service time (Sönmez, 2013; Sönmez and Switzer, 2013) and school choice in some cities in China, where families can only express extreme levels of willingness to pay additional tuition (Wang and Zhou, 2018). The approach presented in this paper allows the optimal portfolio to be calculated in such settings.

7.1.2 Job Market Signaling

Several entry-level labor markets have introduced technologies that allow applicants to signal their interest. Famous examples include the job market for new economists (Coles et al., 2010), *early action* in American college admissions (Avery, Fairbanks and Zeckhauser, 2009), and dating websites (Lee and Niederle, 2015). In these markets, there is an additional constraint: not only are applications (potentially) costly, but the number of signals is constrained (Coles, Kushnir and Niederle, 2013).

In the job market for new economists, for example, sending a signal to department i increases the applicant's chance of receiving an offer from i relative to the counterfactual of not sending a signal to i. This decision problem therefore satisfies Condition 1, with the corresponding ROL listing signals first (up to two, in the case of the job market for new economists), and then all applications in order of preference (including the institutions to which the applicant has sent his signal). When the number of signals is limited, Condition 2 is also satisfied when the market clears in any of the ways this paper has covered. Many comparative statics can also be derived. For example, one can show that more cautious students signal less aggressively.

7.1.3 Early Decision

In American college admissions, *early decision* is a scheme offered by some colleges (Avery, Fairbanks and Zeckhauser, 2009). It differs from early action in that the applicant is committed to accept the college's offer, if one is made. This situation can thus be represented by an ROL where ranking a college first increases admission chances, but the applicant is assigned to the first-ranked college in her ROL if accepted (in contrast to early action, where "signals" only affect admission chances but not "choices" and the applicant is assigned to the highest ranked option among "non-signal entries").

From the perspective of a single agent, early decision is a specific case of a Chinese parallel mechanism (which is a combination of immidiate acceptance and deferred acceptance Chen and Kesten, 2013). The approach presented in this paper could be used in such environments to find the optimal portfolio quickly, and comparative statics can also be derived. For example, one can show that when applications become costless, the optimal early-decision choice becomes more aggressive.

7.2 Deferred Acceptance with Field-Specific Priorities

In many college admissions markets students are ranked differently by different fields of study. Rankings are often based on weighted averages of exam scores with field-specific weighting schemes.[29] In these settings, admission successes are neither independent nor perfectly aligned. Yet, the optimal ROL of any size can still be calculated quickly as long as the number of fields is small, as this assures that the conditions of Observation 2 are met. In this kind of decision problems, the arguments used above show that a cost reduction leading to a larger portfolio weakly increases the desirability of the first ranked school.

7.3 Anti-application

I have shown that Condition 1 is satisfied by many centralized matching markets. It is tempting to think that this mild condition is satisfied by "any centralized market one can conceive of." This, however, is not the case. For example, the condition does not hold in the Israeli Medical Interns Match (Bronfman et al., 2015; Roth and Shorrer, 2015).

8 Discussion

There are many important simultaneous search problems that present a nontrivial correlation structure in the probabilities of success. I showed that many important cases can be solved using dynamic programming. This approach it sheds light on findings of previous studies and, more importantly, provides novel predictions in several important settings and a practical tool for theoretical and empirical research.

I have shown that when applicants are vertically differentiated but face uncertainty about their standing relative to others, an application-cost reduction leading to an increase in the number of schools one applies to makes the optimal portfolio wider. This stands in contrast to the prediction of the model where admission is based on independent lotteries, and is consistent with behavior in the (centralized) Ghanaian high-school admissions and in the (decentralized) U.S. college admissions, suggesting that applicants are facing uncertainty about the strength of their application.

The representation of the main application as a max-coverage problem is a special case of Segal (1989; 1993). It sheds light on the relation between the concavity of the utility function and of beliefs, through a notion of duality, not unrelated to that of Yaari (1987). This relationship, which is present in all of the applications I study, has been documented in other contexts in numerous studies of non-expected utility theory (e.g., Yaari, 1987; Hong, Karni and Safra, 1987).

[29] Examples include Hungary (Biró, 2012; Shorrer and Sóvágó, 2017) and Turkey (Krishna, Lychagin and Robles, 2015).

Centralized clearinghouses determine the school assignment of millions of students around the world. Pathak and Sönmez (2008) show that strategy-proof assignment mechanisms, ones that give no incentives for applicants to misrepresent their preferences, "level the playing field" by protecting strategically unsophisticated applicants. I provide an alternative argument in support of strategy-proof mechanisms: when popular manipulable mechanisms are in place, strong applicants, ones with good outside options (such as private schools, or access to a separate admissions process), will be over represented in the most-desirable schools, even when all applicants are strategically sophisticated. The same holds for applicants with differing levels of confidence or risk preferences, which may lead to disparities across gender lines (Dargnies, Hakimov and Kübler, forthcoming; Pan, 2019).

This paper also connects the literature on priority-tie-breaking in school choice to the discussion on protecting unsophisticated agents (e.g., Rees-Jones, 2017; Hassidim et al., 2017). Tie-breaking methods have been studied extensively, especially comparing single lottery with multiple independent lotteries (e.g., Abdulkadiroğlu, Pathak and Roth, 2009; Ashlagi, Nikzad and Romm, 2015; De Haan et al., 2015). While this literature focused on centralized mechanisms, admission decisions are also independent in decentralized markets where oversubscribed schools use independent (school-specific) lotteries to break priority ties between applicants (e.g., Dobbie and Fryer, 2011). Similarly, if all schools use a single lottery (or exam) to break priority ties, admission decisions are perfectly aligned.

I show that when applications are restricted or costly there may exist a tension between efficiency, which may be improved by the use of a single lottery, and the desire to protect unsophisticated agents. Specifically, I show that identifying the optimal portfolio under constrained deferred-acceptance mechanisms is simpler when priority ties are broken with multiple independent lotteries rather than with a single lottery. Furthermore, when the market is cleared using TTC, Bayesian sophistication is required regardless of the way priority ties are broken.

There are many interesting research directions that are beyond the scope of this paper. First, evaluating agents' performance under different correlation regimens in the field and and the experimental lab. Second, revisiting empirical studies that make the assumption of independence and relaxing this assumption. Third, analyze other simultaneous search environments using the dynamic programming approach. Finally, study equilibrium behavior in labor kets

References

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93(3): 729–747.

Abdulkadiroğlu, Atila, Joshua D Angrist, Yusuke Narita, and Parag A Pathak. 2017*a*. "Research design meets market design: Using centralized assignment for impact evaluation." *Econometrica*, 85(5): 1373–1432.

Abdulkadiroğlu, Atila, Joshua D Angrist, Yusuke Narita, Parag A Pathak, and Roman A Zarate. 2017*b*. "Regression Discontinuity in Serial Dictatorship: Achievement Effects at Chicago's Exam Schools." *American Economic Review*, 107(5): 240–45.

Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. 2009. "Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match." *American Economic Review*, 99(5): 1954–1978.

Abdulkadiroğlu, Atila, Yeon-Koo Che, Parag A Pathak, Alvin E Roth, and Olivier Tercieux. 2017*c*. "Minimizing Justified Envy in School Choice: The Design of New Orleans' OneApp." National Bureau of Economic Research.

Agarwal, Nikhil, and Paulo Somaini. 2018. "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism." *Econometrica*, 86(2): 391–444.

Ajayi, Kehinde, and Modibo Sidibe. 2015. "An empirical analysis of school choice under uncertainty." Boston University Working paper, http://people. bu. edu/kajayi/research. html.

Ajayi, Kehinde F. 2011. "School choice and educational mobility: Lessons from secondary school applications in Ghana." *University of California-Berkeley Working Paper*.

Akbarpour, Mohammad, and Winnie van Dijk. 2018. "School Choice with Unequal Outside Options." Working paper.

Albrecht, James, Pieter A Gautier, and Susan Vroman. 2006. "Equilibrium directed search with multiple applications." *The Review of Economic Studies*, 73(4): 869–891.

Arnosti, Nick. 2016. "Centralized clearinghouse design: A quantity-quality tradeoff."

Ashlagi, Itai, Afshin Nikzad, and Assaf Romm. 2015. "Assigning more students to their top choices: A tiebreaking rule comparison." *EC '15*, 755–756. New York, NY, USA:ACM.

Ashlagi, Itai, and Afshin Nikzad. 2016. "What matters in school choice tie-breakings? How competition guides design." 767–768, ACM.

Athey, Susan, and Glenn Ellison. 2011. "Position auctions with consumer search." *The Quarterly Journal of Economics*, 126(3): 1213–1270.

Avery, Christopher, and Jonathan Levin. 2010. "Early admissions at selective colleges." *American Economic Review*, 100(5): 2125–56.

Avery, Christopher, Andrew Fairbanks, and Richard J Zeckhauser. 2009. *The early admissions game.* Harvard University Press.

Avery, Christopher, Jessica S Howell, and Lindsay Page. 2014. "A Review of the Role of College Applications on Students' Postsecondary Outcomes. Research Brief." *College Board*.

Azevedo, Eduardo M., and Jacob D. Leshno. 2016. "A supply and demand framework for two-sided matching markets." *Journal of Political Economy*, 124(5): 1235–1268.

Biró, Péter. 2012. "University admission practices-hungary." *www. matching-in-practice. eu accessed July*, 21: 2014.

Bobba, Matteo, and Veronica Frisancho. 2014. "Learning about oneself: The effects of signaling academic ability on school choice." *Unpublished working paper.*

Bronfman, Slava, Avinatan Hassidim, Arnon Afek, Assaf Romm, Rony Shreberk, Ayal Hassidim, and Anda Massler. 2015. "Assigning Israeli medical graduates to internships." *Israel journal of health policy research*, 4(1): 6.

Calsamiglia, Caterina, and Maia Güell. 2014. "The illusion of school choice: empirical evidence from Barcelona."

Calsamiglia, Caterina, Francisco Martínez-Mora, and Antonio Miralles. 2015. "School choice mechanisms, peer effects and sorting." *Univ. of Leicester Dept. Ecoomics WP*, , (15/01).

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn. 2010. "Constrained school choice: An experimental study." *American Economic Review*, 100(4): 1860–1874.

Card, David, and Alan B Krueger. 2005. "Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas." *ILR Review*, 58(3): 416–434.

Casalmiglia, Caterina, Chao Fu, and Maia Güell. 2014. "Structural estimation of a model of school choices: The Boston mechanism vs. its alternatives." *Documentos de trabajo (FEDEA)*, 21: 1–63.

Chade, Hector, and Lones Smith. 2005. "Simultaneous search."

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

Chade, Hector, Gregory Lewis, and Lones Smith. 2014. "Student portfolios and the college admissions problem." *Review of Economic Studies*, 81(3): 971–1002.

Chen, Yan, and Onur Kesten. 2013. "From boston to chinese parallel to deferred acceptance: theory and experiments on a family of school choice mechanisms." WZB Discussion Paper.

Che, Yeon-Koo, and Olivier Tercieux. 2017. "Top Trading Cycles in Prioritized Matching: An Irrelevance of Priorities in Large Markets."

Coles, Peter, Alexey Kushnir, and Muriel Niederle. 2013. "Preference signaling in matching markets." *American Economic Journal: Microeconomics*, 5(2): 99–134.

Coles, Peter, and Ran I. Shorrer. 2014. "Optimal truncation in matching markets." *Games and Economic Behavior*, 87: 591–615.

Coles, Peter, John Cawley, Phillip B Levine, Muriel Niederle, Alvin E Roth, and John J Siegfried. 2010. "The job market for new economists: A market design perspective." *Journal of Economic Perspectives*, 24(4): 187–206.

Dargnies, Marie-Pierre, Rustamdjan Hakimov, and Dorothea Kübler. forthcoming. "Self-confidence and unraveling in matching markets." *Management Science*.

De Haan, Monique, Pieter A Gautier, Hessel Oosterbeek, and Bas Van der Klaauw. 2015. "The performance of school assignment mechanisms in practice."

Delacroix, Alain, and Shouyong Shi. 2006. "Directed search on the job and the wage ladder." *International Economic Review*, 47(2): 651–699.

De Los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest. 2012. "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *American Economic Review*, 102(6): 2955–80.

Dobbie, Will, and Jr. Fryer, Roland G. 2011. "Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics*, 3(3): 158–87.

Dur, Umut, Robert G. Hammond, and Thayer Morrill. 2018. "Identifying the Harm of Manipulable School-Choice Mechanisms." *American Economic Journal: Economic Policy*, 10(1): 187–213.

Ergin, Haluk, and Tayfun Sönmez. 2006. "Games of school choice under the Boston mechanism." *Journal of public Economics*, 90(1-2): 215–237.

Fu, Chao. 2014. "Equilibrium tuition, applications, admissions, and enrollment in the college market." *Journal of Political Economy*, 122(2): 225–281.

Fudenberg, Drew, and Jean Tirole. 1991. "Game theory, 1991." *Cambridge, Massachusetts*, 393(12): 80.

Gale, David, and Lloyd S. Shapley. 1962. "College admissions and the stability of marriage." *American Mathematical Monthly*, 69(1): 9–15.

Galenianos, Manolis, and Philipp Kircher. 2009. "Directed search with multiple job applications." *Journal of Economic Theory*, 144(2): 445–471.

Gross, Betheny, Michael DeArmond, and Patrick Denice. 2015. "Common Enrollment, Parents, and School Choice: Early Evidence from Denver and New Orleans. Making School Choice Work Series." *Center on Reinventing Public Education*.

Haeringer, Guillaume, and Flip Klijn. 2009. “Constrained school choice.” *Journal of Economic Theory*, 144(5): 1921–1947.

Hassidim, Avinatan, Assaf Romm, and Ran I. Shorrer. 2016*a*. “Redesigning the Israeli Psychology Master’s Match.” mimeo.

Hassidim, Avinatan, Assaf Romm, and Ran I. Shorrer. 2016*b*. “’Strategic’ Behavior in a Strategy-proof Environment.” mimeo.

Hassidim, Avinatan, Marciano Déborah, Assaf Romm, and Ran I. Shorrer. 2017. “The mechanism is truthful, why aren’t you?” *American Economic Review: Papers and Proceedings*, 107(5).

Hastings, Justine, Thomas J Kane, and Douglas O Staiger. 2009. “Heterogeneous preferences and the efficacy of public school choice.” *NBER Working Paper*, 2145: 1–46.

Hochbaum, Dorit S. 1996. *Approximation algorithms for NP-hard problems.* PWS Publishing Co.

Hong, Chew Soo, Edi Karni, and Zvi Safra. 1987. “Risk aversion in the theory of expected utility with rank dependent probabilities.” *Journal of Economic theory*, 42(2): 370–381.

Immorlica, Nicole S, Jacob D Leshno, Irene Y Lo, and Brendan J Lucier. 2018. “Costly Information Acquisition and Stable Matching Mechanisms.”

Istomin, Roman, Kala Krishna, and Sergey Lychagin. 2018. “Non-Parametrically Identifying Preferences withStrategic Reports: An Application to Turkish Undergraduate Admissions.” Penn State.

Kapor, Adam, Christopher A Neilson, and Seth D Zimmerman. 2018. “Heterogeneous Beliefs and School Choice Mechanisms.” National Bureau of Economic Research Working Paper 25096.

Kim, Matthew. 2010. “Early decision and financial aid competition among need-blind colleges and universities.” *Journal of Public Economics*, 94(5): 410 – 420.

Kircher, Philipp. 2009. “Efficiency of simultaneous search.” *Journal of Political Economy*, 117(5): 861–913.

Krishna, Kala, Sergey Lychagin, and Verónica Frisancho Robles. 2015. “Retaking in High Stakes Exams: Is Less More?” National Bureau of Economic Research.

Larroucau, Tomás, and Ignacio Rios. 2018. “Demand Estimation in the College Admissions Problem: Should We Assume Truth Telling?”

Lee, Soohyung, and Muriel Niederle. 2015. “Propose with a rose? signaling in internet dating markets.” *Experimental Economics*, 18(4): 731–755.

Leshno, Jacob, and Irene Lo. 2017. "The Simple Structure of Top Trading Cycles in School Choice."

Lucas, Adrienne M., and Isaac M. Mbiti. 2012. "The Determinants and Consequences of School Choice Errors in Kenya." *American Economic Review*, 102(3): 283–88.

Nagypal, Eva. 2004. "Optimal application behavior with incomplete information." *Manuscript, Econ. Dept., Northwestern Univ*, 1.

Olszewski, Wojciech, and Rakesh Vohra. 2016. "Simultaneous selection." *Discrete Applied Mathematics*, 200: 161–169.

Oosterbeek, Hessel, Sándor Sóvágó, and Bas van der Klaauw. 2019. "Why Are Schools Segregated? Evidence from the Secondary-School Match in Amsterdam."

Pallais, Amanda. 2015. "Small differences that matter: Mistakes in applying to college." *Journal of Labor Economics*, 33(2): 493–520.

Pan, Siqi. 2019. "The instability of matching with overconfident agents." *Games and Economic Behavior*, 113: 396 – 415.

Pathak, Parag A., and Tayfun Sönmez. 2008. "Leveling the playing field: Sincere and sophisticated players in the Boston mechanism." *American Economic Review*, 98(4): 1636–1652.

Rees-Jones, Alex. 2017. "Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match." *Games and Economic Behavior*.

Roth, Alvin E, and Ran I Shorrer. 2015. "The redesign of the medical intern assignment mechanism in Israel." *Israel journal of health policy research*, 4(1): 11.

Schmelzer, André. 2016. "Single versus multiple randomization in matching mechanisms."

Segal, Uzi. 1989. "Anticipated utility: a measure representation approach." *Annals of Operations Research*, 19(1): 359–373.

Segal, Uzi. 1993. "The measure representation: A correction." *Journal of Risk and Uncertainty*, 6(1): 99–107.

Shorrer, Ran I., and Sándor Sóvágó. 2017. "Obvious Mistakes in a Strategically Simple College Admissions Environment: Causes and Consequences." mimeo.

Sönmez, Tayfun. 2013. "Bidding for army career specialties: Improving the ROTC branching mechanism." *Journal of Political Economy*, 121(1): 186–219.

Sönmez, Tayfun, and Tobias B. Switzer. 2013. "Matching with (branch-of-choice) contracts at the United States Military Academy." *Econometrica*, 81(2): 451–488.

Wang, Tong, and Congyi Zhou. 2018. "Purchasing Seats for High School Admission in China."

Wong, Adam Chi Leung. 2014. "The choice of the number of varieties: Justifying simple mechanisms." *Journal of Mathematical Economics*, 54: 7–21.

Wu, Binzhen, and Xiaohan Zhong. 2014. "Matching mechanisms and matching quality: Evidence from a top university in China." *Games and Economic Behavior*, 84: 196–215.

Yaari, Menahem E. 1987. "The dual theory of choice under risk." *Econometrica: Journal of the Econometric Society*, 95–115.

A Forward-Looking PSA

Algorithm 2 Forward-Looking PSA

Step 1 For all possible values of r^2, find $r^1(r^2) = \arg\max_{c_{r^1} \geq c_{r^2}} (1 - c_{r^1})(u_{r^1} - u_{r^2})$. Set $v_1(r^2) := \max_{c_{r^1} \geq c_{r^2}} (1 - c_{r^1})(u_{r^1} - u_{r^2})$.

Step i<k For all possible values of r^{i+1}, find $r^i(r^{i+1}) = \arg\max_{c_{r^i} \geq c_{r^{i+1}}} (1 - c_{r^i})(u_{r^i} - u_{r^{i+1}}) + v_{i-1}(r^i)$. Set $v(r^{i+1}) = \max_{c_{r^i} \geq c_{r^{i+1}}} (1 - c_{r^i})(u_{r^i} - u_{r^{i+1}}) + v(r^i)$.

Step k Set $r^k = \arg\max_i (1 - c_{r^k})u_{r^k} + v_{k-1}(r^k)$, and for all $j < k$, $r^j = r^j(r^{j+1}(...(r^k)))$.

B Nondecreasing Marginal Benefit

Example 4. Consider the following six rectangles:[30]

$A = ((0,0),(3,0),(0,3),(3,3))$
$B_1 = (-0.4,0),(-0.4,1.5),(3,0),(3,1.5))$
$B_2 = ((-0.4,1.5),(-0.4,3),(3,1.5),(3,3))$
$C_1 = ((0,0),(0,4),(1,0),(1,4))$
$C_2 = ((1,0),(1,4),(2,0),(2,4))$
$C_3 = ((2,0),(2,4),(3,0),(3,4))$

Direct calculation shows that the maximal coverage by a single rectangle is by A with a covered area of 9, the maximal coverage by a pair of rectangles is by $B_1 \cup B_2$ with a covered area of 10.2, and the maximal coverage by three rectangles is by $C_1 \cup C_2 \cup C_3$ with a covered area of 12. Thus, the marginal benefit from relaxing the constraint from 1 to 2 is 1.2, but the marginal benefit from relaxing the constraint from 2 to 3 is $1.8 > 1.2$.

[30] I thank Avinatan Hassidim for providing this example.

C Additional Proofs

Proposition 3

Proof. Let b be an agent with full-support beliefs whose optimal portfolio does not consist of the top k choices. Fixing preferences, it is easy to identify such beliefs, by making high scores implausible.

Let r_b^k be the last-ranked option on b's ROL. Mix b's belief (with weight α) with her belief conditional on her score being above c_k (with weight $1-\alpha$). For low enough values of α, the optimal portfolio has a strictly more selective last choice (the k-th option). Also, for any value of α, the optimal portfolio is at least as aggressive as b's (by Proposition 2). Denote by $\hat{\alpha}$ the supremum of values of α for which the last choice is strictly more selective than r_b^k. Comparing the optimal ROLs for $\hat{\alpha}-\delta$ and $\hat{\alpha}+\delta$ for a small $\delta>0$ gives an example as required, since admission chances increase only marginally (thanks to improved beliefs) but decrease discontinuously (due to the shift to a more selective last choice). □

Proposition 4

Proof. Assume that b has a uniform belief on $[0,1]$, and the marginal cost is low enough so that b's optimal portfolio consists of more than one option. If s's belief is uniform on $[c_1,1]$, her optimal portfolio consists only of the most selective school (since she is certain she will be admitted).

Similarly, let s have a uniform belief on $[0,1]$, and let c be low enough so that s's optimal portfolio consists of more than one option. If $b's$ belief is uniform on $[0,c_{N-1}]$, her optimal portfolio consists of either the least selective school or no school at all (since she is certain she will be rejected by any other school).

The third case follows from continuity of the expected utility in beliefs. □

Proposition 6 (Completed)

Proof. In the main text I assumed that $r_b^{i^*}$ does not belong to r_s. Here, I lay out the argument for when this is not the case.

Since r_s is not at least as aggressive as r_b, there exists $i\le k$ such that $r_s^i > r_b^i$. Let i^* be the maximal such index. Let m be the largest index such that $m\le i^*$, $r_s^m > r_b^m$ and $r_s^{m-1}\le r_b^{m-1}$. I will consider the "block" of options between m and i^*, B_x^m for each agent x. For each agent, $x\in\{s,b\}$, normalize her utility function (using a different normalization) so that her expected utility from the portfolio r_x conditional on her being rejected from the first i^* options (i.e., the continuation value at i^*) is equal to 0, and her expected utility conditional on some option in B_b^m being realized to 1 (that is, both agents assign to b's block, B_x^m, a value of 1 conditional on being assigned to some option).[31]

[31] This normalization is possible since ranking consistently with one's true preferences is optimal, thus the expected value from b's continuation portfolio is the average of utilities of less preferred options than $r_b^{i^*}$, and s's continuation is such average over options less preferred than $r_s^{i^*}$ which, by assumption, is less preferred than $r_b^{i^*}$, wheres the utility from B_b^m conditional on success is the average of utilities from more preferred options.

Since i^* is the maximal index with the property $r_s^i > r_b^i$, and it is suboptimal to rank options in a way that is inconsistent with true preferences (thus, $r_b^j \geq r_s^j > r_s^{i^*}$ for all $j > i^*$), $r_s^{i^*}$ does not appear on r_b (otherwise, there would be a higher index with $r_s^i > r_b^i$). Hence, had B_b^m been changed to B_s^m, b's (original) optimal continuation would still be available.

Denote by p the probability of acceptance to some option in the block B_b^m ($p = 1 - \prod_{i=m}^{i^*}(1 - p_{r^i})$) and by p' the probability of acceptance to some option in B_s^m. Since b chose B_b^m over B_s^m we have that

$$p = p \cdot 1 + (1-p) \cdot 0 \geq p' \cdot u_b(B_s^m) + (1-p) \cdot 0 = p' \cdot u_b(B_s^m),$$

where the zeros stand for the expected utility from the optimal continuation, which is still available after the deviation. Similarly, since s chose B_s^m and her optimal continuation over B_b^m and the same continuation (which may be suboptimal), we have that

$$p' \cdot u_s(B_s^m) = p' \cdot u_s(B_s^m) + (1-p') \cdot 0 \geq p \cdot 1 + (1-p) \cdot 0 = p.$$

These two inequalities together imply that

$$u_s(B_s^m) \geq \frac{p}{p'} \geq u_b(B_s^m).$$

And this implies that for at least one index, k, between m and i^* we have that $u_s(r_s^k) \geq u_b(r_s^k)$. The rest of the argument is as in the body of the paper: showing that there is no loss of utility for s from changing the entire block B_s^m to B_b^m while keeping the rest of his ROL unchanged. □

Proposition 10

Proof. Consider an optimal ROL r. Assume that r is not consistent with true preferences. Then there is a highest i ($< N$) such that $u_b(r^i) < u_b(r^{i+1})$. If the probability that r^{i+1} will be realized is 0, removing it from the ROL results in a new, shorter, optimal ROL with a (weakly) smaller highest index with "reversal." Otherwise, there is a positive probability of b being assigned to r^{i+1}, which means that by ranking r^{i+1} on the i-th entry of the ROL, b would assure admission to this option. Since r^{i+1} is preferred to r^i and to r^{i+j} for all $j > 1$, replacing r^i with r^{i+1} and omitting all higher-ranked options would result in a new optimal ROL with a lower highest index with "reversal." Repeating the process described above must result in an optimal ROL with no reversals.

Let r be an optimal ROL that is consistent with b's preferences, constructed as above. If one of the options on the ROL, r^j, is ranked higher than z_{r^j}, truncate the ROL after the minimal index with this property. This is inconsequential since the agent is assured to be assigned to this option or a higher-ranked (i.e., lower-index) one, and thus the resulting ROL, $\bar{r}$, is still optimal. Denote by L the length of the resulting ROL, $L = |\bar{r}|$. Then $z_{\bar{r}^i} = i$ for all $i < L$ and $z_{\bar{r}^L} \geq L$. Furthermore, for the case of a single lottery, it is not possible that for some $i < L$ we have that $l_{\bar{r}^i} < l_{\bar{r}^{i+1}}$, since in this case the probability of assignment to the option $\bar{r}^{i+1}$ is 0. □

Proposition 11

Proof. Consider r_s and r_b, two optimal portfolios that are consistent with the preferences and the selectivity as described above.

When b considers r_b^1, her continuation value from any choice is higher; thus, her optimal continuation yields a higher payoff.[32] Furthermore, her admission chances are disproportionately higher in more-preferred options. Thus, b has an unambiguous incentive to rank more-preferred options first. To see this, note that, by fixing the suffix of the ROL, the continuation payoff, which is independent of the first-ranked choice, is unchanged, this problem is no different from the problem of choosing an optimal portfolio of size 1 in the environment studied in Section 3. Hence, $r_b^1 \leq r_s^1$ by Theorems 7 and 8.

Finally, note that for each agent, the optimal continuation is a solution to the same problem, removing options with $z_i = 1$ from the menu, and reducing all other z_i's by 1. The result follows by induction on the size of the menu (i.e., $|X|$). $\square$

Proposition 12

Proof. (sketch) The proof proceeds by induction on the length constraint, k. As usual, the case of $k = 1$ is easy (it coincides with previous environments). For larger values of k, there are two cases to consider.

First, given a b-optimal ROL, r_b, if for any s-optimal ROL, r_s, we have that $r_b^1 \leq r_s^1$, then the desired result follows by induction as the continuation ROLs are optimal ROLs under the length constraint of $k-1$ in a smaller problem (described above), where unless b is assigned with certainty (in which case the proof is simple), b's beliefs MLRP-dominate s's beliefs (as they dominate b's beliefs conditional on rejection from r_s^1, which dominate s's beliefs in this case).

Second, consider the case where $r_b^1 > r_s^1$ for some r_s. In this case it is possible to change r_b to $\bar{r}_b$ such that the resulting ROL, $\bar{r}_b$, is at least as aggressive as both r_b and r_s. The argument is completely analogous to the one used in the proof of Proposition 2. $\square$

Proposition 13

Proof. First, consider the case of independent lotteries. Given an s-optimal ROL, remove from it any payoff-irrelevant options (i.e., options that are realized with probability 0). Denote the resulting ROL by r_s.

Assume that r_s is not at least as aggressive as some optimal ROL for b, r_b. Let i^* be the highest index, i, in which $r_s^i > r_b^i$. Consider the normalizations under which $u_s(r_b^{i^*}) = u_b(r_b^{i^*}) = 2$ and $u_s(r_s^{i^*}) = u_b(r_s^{i^*}) = 1$. Observe that b's continuation payoff is clearly not higher than s's, as s can imitate b and get at least as high a payoff (by Proposition 10, since $u_s(z) \geq u_b(z)$ for all

[32] Allowing ROLs to include repetitions of the same option is inconsequential in this environment: if the applicant has a chance of being admitted in the later round, she is surely assigned in an earlier round.

$z \geq r_s^i$). But this means that s will do (weakly) better by changing $r_s^{i^*}$ to $r_b^{i^*}$.[33] Making this change repeatedly and repeating the process enumerating over all b-optimal portfolios yields the desired s-optimal ROL.

The proof for the case of a single lottery proceeds by induction. Given an optimal ROL for s, remove from it any payoff-irrelevant options (i.e., options that are realized with probability 0). As usual, the case of $k = 1$ is easy, as it coincides with previous environments. For larger values of k, there are two cases to consider. First, if for any b-optimal ROL, r_b, we have that $r_s^1 \leq r_b^1$, then the desired result follows by induction as the continuation ROLs are optimal ROLs under the length constraint of $k-1$ in a smaller problem (described above), where unless s is assigned with certainty (contradicting $r_b's$ optimality), s's beliefs MLRP-dominate b's beliefs since s's beliefs MLRP-dominate her beliefs conditional her being rejected from r_b^1 (b's beliefs), which means that s has an optimal continuation ROL that is at least as aggressive as any optimal continuation ROL under b's beliefs (by Proposition 12), and there is an optimal continuation ROL under s's preferences and b's beliefs that is at least as aggressive as any b-optimal continuation ROL by the inductive hypothesis.

Second, consider the case where $r_s^1 > r_b^1$. In this case, by the argument in the previous paragraph, there exists $\hat{\imath} \geq 1$ such that r_b ranks relevant options up to $\hat{\imath}$ (at least), $r_s^i > r_b^i$ for all $i \leq \hat{\imath}$, and $r_s^i \leq r_b^i$ for higher values of i (or b is assigned with certainty before i).

Note that for all $1 < i \leq \hat{\imath}$, the inequality $\max\{r_s^{i-1}, r_b^{i-1}\} \leq \min\{r_s^i, r_b^i\}$ holds. This is true since if b's probability of being assigned to r_b^i is not 0, then s would receive it with certainty by ranking it $i-1$. Therefore, $r_s^{\hat{\imath}} > r_b^{\hat{\imath}} \geq ... r_s^2 > r_b^2 \geq r_s^1 > r_b^1$. This also means that the associated l's are ordered in the same way for all $i < \hat{\imath}$ and also for $\hat{\imath}$ unless $z_{r_s^{\hat{\imath}}} > \hat{\imath}$ (since the choices are payoff relevant in these histories, each option is ranked in the corresponding z. See the proof of Proposition 10 for details).

Next, observe that s has the possibility to change any part of the prefix of her ROL of length $\hat{\imath}$ to equal b's ROL, but she does not exercise this option. Note that in the event that s gains from choosing $r_s^{\hat{\imath}}$ over $r_b^{\hat{\imath}}$ she gains $u_s(r_s^{\hat{\imath}}) - u_s(r_s^{\hat{\imath}+1})$ (where $u_s(r_s^{\hat{\imath}+1})$ is equal to 0 if the ROL is of length $\hat{\imath}$) and in the event that she loses, her loss is $u_s(r_b^{\hat{\imath}}) - u_s(r_a^{\hat{\imath}})$. Gains are more substantial for b for two reasons: $r_b^{\hat{\imath}+1}$ is less attractive and u_b is more concave. Formally

$$\frac{u_b(r_s^{\hat{\imath}}) - u_b(r_b^{\hat{\imath}+1})}{u_s(r_s^{\hat{\imath}}) - u_s(r_s^{\hat{\imath}+1})} \geq \frac{u_b(r_s^{\hat{\imath}}) - u_b(r_s^{\hat{\imath}+1})}{u_s(r_s^{\hat{\imath}}) - u_s(r_s^{\hat{\imath}+1})} \geq \frac{u_b(r_b^{\hat{\imath}}) - u_b(r_s^{\hat{\imath}})}{u_s(r_s^{\hat{\imath}}) - u_s(r_s^{\hat{\imath}})}.$$

Similarly, one can consider changing the $(\hat{\imath}-1)$-th entry of the agent's ROL in addition to the change in the $\hat{\imath}$-th entry. Again, when the change matters, b derives higher utility gains (proportionally). Continuing this argument up to the first entry, one gets that b has a stronger incentive than s to use s's prefix. Changing b's prefix to s's and repeating the process, enumerating over all b-optimal portfolios, yields the desired s-optimal ROL.

[33] Holding the rest of the optimal portfolio fixed, the problem is no different than choosing the optimal size-1 portfolio in the environment studied in Section 3.

$\square$