Desirable Rankings: A New Method for Ranking Outcomes of a Competitive Process[†]

Thayer Morrill[1] and Peter Troyan[2]

[1]Department of Economics, North Carolina State University

[2]Department of Economics, University of Virginia

October 4, 2023

Abstract

We consider the problem of aggregating individual preferences over alternatives into a social ranking. A key feature of the problems that we consider—and the one that allows us to obtain positive results, in contrast to negative results such as Arrow's Impossibililty Theorem—is that the alternatives to be ranked are outcomes of a competitive process. Examples include rankings of colleges or academic journals. The foundation of our ranking method is that alternatives that agents rank higher than the one they receive (and thus have been rejected by) should also be ranked higher in the aggregate ranking. We introduce axioms to formalize this idea, and call any ranking that satisfies our axioms a *desirable ranking*. We show that as the market grows large, any desirable ranking coincides with the true underlying ranking of colleges by quality. Last, we provide an algorithm for constructing desirable rankings, and show that the outcome of this algorithm is the unique ranking of the colleges that satisfy our axioms.

[†]Emails: tsmorril@ncsu.edu and troyan@virginia.edu. We are grateful to Kyle Woodward who helped us start this project and contributed many thoughtful ideas. Thayer would like to thank the Stanford market design coffee group, and especially the world's best host Al Roth, for many helpful suggestions during this projects early stages. We would like to thank Itai Ashlagi, Marcelo Fernandez, Ravi Jagadeesan, Maciej Kotowski, Andrew Mackenzie, and Alexander Teytelboym, as well as audiences at Stanford, Johns Hopkins, UC-Davis, EC22 and the 12th Conference on Economic Design.

1 Introduction

What college should a student choose? Should an assistant professor be granted tenure? For a wide range of important decisions, we rely on external rankings of quality to help us make a good decision.

One standard approach to constructing a ranking is to use a formula based on indicators of quality. An important example is the *US News and World Report* rankings of colleges, which receive a great deal of media attention and have been shown to influence the decisions of college applicants (Bowman and Bastedo, 2009; Griffith and Rask, 2007). While the precise formula behind this ranking is opaque, inputs include quality indicators such as the acceptance rate (percentage of applicants who are accepted) and the yield rate (percentage of those admitted who enroll).

Similarly, deciding whether an assistant professor's body of research is sufficient to justify granting tenure or determining what share of state funding a university should receive are strongly influenced by rankings of journals.[1] In turn, journal rankings are typically based on measures such as acceptance rates or citation counts.

As a third example, consider the National Resident Matching Program (NRMP), which every year uses a centralized algorithm to match graduating medical students to hospital residency programs. An informal metric often used by programs to judge quality is called the "rank-to-fill" ratio: in other words, how far down the program's submitted rank list did they have to go in order to fill all of their residency slots.[2] This creates incentives for programs to pressure candidates to reveal their rank lists to programs—in violation of NRMP rules—with an implicit "threat" that if the applicant does not commit to ranking the program highly, the program will demote the applicant in their own rankings (Sbicca et al., 2012). This causes much undue stress for the medical students, and, from an economic perspective, may lead to inefficient matching.

The examples above are susceptible to Goodhart's Law: when a measure becomes a target, it ceases to be a good measure. Indeed, there is evidence that universities purposefully solicit applications from students they know will be rejected in order to lower their admissions rates,

[1] For example, research funding is allocated to British universities in part using the "Research Excellence Framework" which aims to take a metrics based approach to measuring research impact.

[2] While this is not an officially published ranking, many programs do care about their rank-to-fill ratio. For instance, Jena et al. (2012) write "In our experience, programs that must move farther down their rank list of applicants to fill their openings may be viewed as having an unsuccessful Match; aggressive lobbying of applicants by programs may be a by-product of this motivation." Wu et al. (2015) echo a similar sentiment: "in a metric-driven world where program quality is gauged by rank-to-fill performance (ie, "how low did you go?") the pressures are manifold."

as well as reject highly qualified applicants that they fear may choose another school so as not to harm their matriculation rates (Toor, 2000; Golden, 2001; Belkin, 2019). Similarly, a journal interested in improving its journal impact factor (JIF) has a strong incentive to manipulate its acceptance rate and citation count, and there are well-known strategies for doing so.[3] Many medical students report that post-interview pressure from residency programs altered their own rankings (Jena et al., 2012).

Rather than relying on seemingly "objective" measures such as acceptance rates or citation counts, an alternative approach is to aggregate individual rankings to produce a social ranking. For instance, which colleges students have chosen in previous years can be informative for the decisions of the current cohort. Information about the best academic journals can be obtained from the order in which researchers submit to (and are rejected from) them. The NRMP collects rank-order lists of medical student rankings of residency programs.

Of course, not all agents will agree on the answers to these questions, and so the question remains of how to aggregate diverse individual preferences into an aggregate ranking. Many results in social choice theory (e.g., Arrow's Impossibility Theorem (Arrow, 1950), the Gibbard-Satterthwaite Theorem (Gibbard, 1973; Satterthwaite, 1975)), speak to the difficulty of such preference aggregation.

In this paper, we propose a new method of ranking alternatives. For concreteness, and with the above application in mind, we refer throughout to the alternatives to be ranked as *colleges*. However, our method is general, and can be applied to many other settings, such as ranking public elementary schools, medical residency programs, or academic journals. A key feature that will allow us to obtain positive results in contrast with the impossibility results discussed above is that the alternatives we are ranking are outcomes of a competitive process: a student matriculates at a college after being accepted by it (and rejected by others); a research article appears in a journal after it has been rejected from other journals that the authors prefer.

Consider a student i who attends a college B but prefers and was rejected from college A (alternatively, a researcher who submits an article to journal B after having been rejected at A, or a doctor who is matched to residency program B but prefers another program A). The fact that college admissions are the outcome of a process in which students make application decisions and colleges make acceptance/rejection decisions allows us to infer two pieces of information. First, student i *desires* A (over B); but second, it also reveals that college A prefers its student,

[3]See, for instance, "A user's guide to inflated and manipulated impact factors" (Ioannidis and Thombs, 2019); "Authorship and citation manipulation in academic research" (Fong and Wilhite, 2017); "Editors JIF-boosting stratagems–Which are appropriate and which not?" (Martin, 2016); and "Games academics play and their consequences: how authorship, h-index and journal impact factors are shaping the future of academia" (Chapman et al., 2019).

say j, to student i, because A rejected i in favor of j. In a sense, A agrees with i's assessment that it is the better college: A would not trade, even if offered. This is the basis of our aggregate rankings: an object that an agent desires over her own outcome should be ranked higher.

A second feature that will allow us to attain positive results is that, even though we assume the students have strict preferences, we will not seek an aggregate ranking that is strict. Indeed, it is probably difficult (and artificial) to discern an exact strict ranking among colleges that are "close". Instead, we define a ranking as a partition of the colleges into *tiers*, with the interpretation that any two colleges in the same tier are "tied" in the ranking.

We begin by introducing three axioms that formalize the idea that any college a student prefers should be ranked higher. First, given a ranking, we formalize our concept of desire. There may be two reasons a student may prefer a college: either because it is a higher quality school overall, or for idiosyncratic reasons (e.g., geographical preferences). An aggregate ranking should reflect the former component, but not the latter. With the intention of filtering idiosyncratic preferences, we seek a proxy for when a student "strongly" prefers a school. A ranking of schools provides this opportunity. We define a student to *desire* a school if she prefers the school to not only her college but also to any college in the same tier as her college. We say a ranking satisfies the *Axiom of Desire (AoD)* if any school a student desires is ranked higher than the college she attends. Our second axiom, Justification, is the converse. In order for a college's ranking to be *justified*, that college should be desired by a student in the tier below (and if it is not, then it's ranking should be lowered). AoD and justification are conditions across tiers. Our final axiom is a within-tier condition we call Balance. A ranking is *balanced* if, within a tier, no school is either overdemanded or undermanded. We call any ranking that satisfies AoD, Justfication, and Balance a *desirable ranking.*

Our main contribution is to show that as a market gets large, any ranking that satisfies our axioms is correct, in the sense that it coincides with the true underlying quality of the colleges. Specifically, we consider a model in which a student i's utility for a college c is a convex combination of a vertical component that we call c's quality and that is common across all students, and an idiosyncratic component that is specific to i and c. We show that as the market grows large, a college with a quality greater than p% of the colleges is ranked (by a desirable ranking) arbitrarily close to the p^{th} percentile among colleges (Theorem 1).[4]

Our second main contribution is to show that for any problem there is a unique desirable ranking. We show this constructively by introducing an algorithm to produce desirable rankings

[4]As part of proving this result, we also show that in the limit, the size of any individual ranking tier becomes small, as a percentage of the total market size (Theorem 2). Thus, while we do allow for ties in rankings, desirable rankings are still able to meaningfully distinguish between colleges in different tiers.

and then showing that any desirable ranking must coincide with the outcome of the algorithm. Intuitively, the algorithm starts by finding a set of colleges such that no student outside of the set desires any college in the set. These colleges are placed in the bottom tier, removed, and the process is repeated for the remaining colleges. Formally, the algorithm uses the Top Trading Cycles algorithm of Shapley and Scarf (1974) applied to the competitive outcome to identify these last-tier colleges. While in general it holds that the TTC outcome is independent of the particular cycle selection rule used, the ordering of cycles is key for constructing a ranking. We define a cycle to be a *last cycle* if there exists an ordering of cycles in TTC where this cycle is chosen last. A desirable ranking must rank these colleges last, as no student outside of the cycle desires a college in the cycle. We thus identify all last cycles, rank these colleges last, remove them, and repeat. We call this algorithm the *delayed trading cycles* (DTC) algorithm. We show that the ranking produced by the DTC algorithm is a desirable ranking, and thus, by Theorem 1, the DTC algorithm ranks colleges correctly in the limit. Further, the DTC ranking is the unique ranking that is desirable (Theorem 4).

There is a relationship between desirable rankings and competitive equilibria. A competitive equilibrium is a vector of prices and an allocation such that each agent's allotment is her favorite object that she can afford. A classic result is that any implementation of TTC corresponds to a competitive equilibrium (Roth and Postlewaite, 1977). Since a desirable ranking is constructed through a particular cycle selection rule of TTC, a desirable ranking corresponds to a competitive equilibrium. Indeed, consider a desirable ranking, allocate each student her favorite college in her tier, and make prices monotonic in a college's ranking (i.e., higher-ranked colleges have a higher price than lower-ranked colleges). By AoD, any college a student strictly prefers to her allocation must be ranked higher, and therefore, be unaffordable. Thus, a corollary to our result on rankings is a novel result on competitive equilibria: as the market grows large, the price for an object in a competitive equilibrium converges to the average utility for that object.

There are other methods that can be used to produce rankings of alternatives. One typical approach used by economists is to make inferences using revealed preference: if a student is admitted to colleges A and B and chooses A, then we infer that she prefers A to B. This is the basis of the rankings introduced by Avery et al. (2013). Each student i is viewed as a tournament among the colleges they were admitted to, and the college i selects among these is the "winner" of the tournament for student i. Colleges accumulate points for each student tournament that they "win", and rankings are determined in a manner analogous to how chess or tennis players are ranked.

There is an important conceptual distinction between revealed preference rankings and desirable rankings. Revealed preference makes inferences by looking "down" an agent's preference list: it uses information about colleges that are *worse* than the one an agent is assigned to form an aggregate ranking. Desirability on the other hand, looks “up” an agent’s preference list: it uses information about the objects an agent *desires* to construct an aggregate ranking.

We think there are several advantages to desirability. First, desirability is less likely to make incorrect inferences in the presence of idiosyncratic individual preferences. Consider a professor who leaves university A to accept a position at university B. Revealed preference infers that B is better than A. However, it is common for a professor to leave a relatively high-ranked university for a lower-ranked one. For instance, if the professor is returning to her home country or she has a two-body problem, she might make this choice, and yet, we would not want to conclude that this choice reveals the relative quality of the institutions; rather, it is likely the result one person’s idiosyncratic preferences. Desirability internalizes this; in particular, desirability only ranks college B over A when a professor at A wants a position at B, and B is unwilling to hire her. A similar argument can be applied to students choosing relatively lower-ranked colleges for idiosyncratic reasons: desirability only ranks college B over college A when a student attending A would prefer B, and B was not willing to admit her.

A second advantage of desirability over revealed preference is that it is more general. For revealed preference to be applicable, (at least some) students must receive multiple offers. In settings where assignments are done through a centralized process, such as in medical residency matching, public school choice, or university admissions in many countries (e.g., China, India, Turkey), the students/doctors do not receive more than one offer. Even for some decentralized markets, revealed preference is not always applicable. For example, to use revealed preference to rank academic journals, a researcher would need to submit an article to multiple journals simultaneously, which is considered unethical.[5] On the other hand, it is easy to discern which journals an author desires: it is those that she submitted the article to (and was rejected from) prior to the one at which it was accepted. Therefore, even though the assignment of articles to journals is decentralized, it is not possible to use revealed preference to rank journals. Desirability may be used to rank journals, as well as any of the other above applications.

[5] An exception is law review journals, which often permit simultaneous submissions.

2 Model

2.1 Agents, Preferences, and Outcomes

There is a set of agents $I = \{i_1, \ldots, i_N\}$ and a set of institutions to which they can be assigned, $C = \{c_1, \ldots, c_N\}$. Throughout the paper, we refer to the agents I as *students* and the institutions C as *colleges* for concreteness, though there are many other potential applications (researchers and journals, doctors and residencies, elementary students and public schools, etc.). Each college c has a capacity of $q_c = 1$, and a strict ranking of the students $\succ_c$. We use $\succsim_c$ for the corresponding weak ranking, i.e., $i \succsim_c j$ if either $i \succ_c j$ or $i = j$. Each student i has a strict ordinal preference relation P_i over the colleges, where we write $c \; P_i \; c'$ to denote that c is strictly preferred to c'. We write R_i for the corresponding weak preference relation.

An **outcome** μ is an assignment of students to colleges; formally, $\mu : I \cup C \to I \cup C$ is a function such that $\mu(i) \in C$ for all $i \in I$, $\mu(c) \in I$ for all $c \in C$, and $\mu(i) = c$ if and only if $\mu(c) = i$. An outcome μ is **Pareto efficient** if there is no alternative outcome ν such that $\nu(i) \; R_i \; \mu(i)$ for all i and $\nu(i) \; P_i \; \mu(i)$ for some i. Throughout the paper, we fix an outcome μ. Our construction of rankings takes any outcome μ as an input, but we do not need to make any particular assumptions on how μ is determined (though it will sometimes be helpful to do so below, for illustrative purposes).[6] For instance, μ may arise from a fully decentralized process as in the U.S. college admissions market, a more structured process such as in academic journal submissions, in which an article can be sent to only one journal at a time, or fully a centralized process similar to college admissions in China, the National Resident Matching Program, or public school choice.

2.2 Rankings

A **ranking** of the colleges is a weak ordering on the set C, denoted $\trianglerighteq$. If $a \trianglerighteq b$ but $b \ntrianglerighteq a$, then we write $a \triangleright b$ and say that a is ranked **higher** than b. We allow for ties in rankings, and write $a \simeq b$ when $a \trianglerighteq b$ and $b \trianglerighteq a$. Any ranking $\trianglerighteq$ induces an ordered partition of the colleges

[6] In particular, we do not make a any assumption that μ is Pareto efficient in general, and our results will hold for any μ. However, Pareto efficiency will make certain concepts easier to state below, which is why we introduce the definition here.

$\Pi^{\trianglerighteq} = \{\Pi_1^{\trianglerighteq}, \Pi_2^{\trianglerighteq}, \ldots, \Pi_K^{\trianglerighteq}\}$ that can be defined recursively as:

$$\Pi_1^{\trianglerighteq} = \{c \in C : c \trianglerighteq c' \text{ for all } c' \in C\}$$
$$\Pi_k^{\trianglerighteq} = \left\{c \in C \setminus \cup_{k'=1}^{k-1} \Pi_{k'}^{\trianglerighteq} : c \trianglerighteq c' \text{ for all } c' \in C \setminus \cup_{k'=1}^{k-1} \Pi_{k'}^{\trianglerighteq}\right\}$$

Note that we use the convention that smaller numbers correspond to "better" rankings: $\Pi_1^{\trianglerighteq}$ is the set of colleges that are ranked the highest, $\Pi_2^{\trianglerighteq}$ is the set of colleges that are ranked higher than all others except those in $\Pi_1^{\trianglerighteq}$, etc. We refer to each $\Pi_k^{\trianglerighteq}$ as a **tier**, and the colleges in $\Pi_k^{\trianglerighteq}$ as the **tier-k colleges**. Similarly, we refer to the students assigned to these colleges, $\mu(\Pi_k^{\trianglerighteq}) = \{i : \mu(i) \in \Pi_k^{\trianglerighteq}\}$, as the **tier-$k$ students**. Define the function $\tau : I \cup C \to \mathbb{N}$ such that $\tau(x) = k$, where k is the tier to which agent x (which may be a student or a college) belongs. We use K to denote the lowest tier, and write $\Pi^{\trianglerighteq} = \{\Pi_1^{\trianglerighteq}, \Pi_2^{\trianglerighteq}, \ldots, \Pi_K^{\trianglerighteq}\}$ for the partition of colleges into tiers induced by $\trianglerighteq$. When no confusion arises, we will also refer to the induced tier partition $\Pi^{\trianglerighteq}$ as a "ranking". Phrases such as "college c is ranked k^{th}" or "c is a k^{th} ranked college" are interpreted as $\tau(c) = k$.

2.3 Desire

When considering the outcome of a competitive process, looking "up" a student's preference list gives two pieces of information: the student prefers the college and the college rejected the student in favor of another applicant. Ideally, we would like a ranking $\trianglerighteq$ that satisfies the following property: for every student i and every college c that i prefers to her outcome ($c \ P_i \ \mu(i)$), we have $c \triangleright \mu(i)$. A moment's reflection reveals that it is not possible to satisfy this condition in general. If student i prefers j's college, while student j prefers i's college—which may happen because of the admissions criteria of the colleges—then this condition would require both $\mu(i) \triangleright \mu(j)$ and $\mu(j) \triangleright \mu(i)$, which is a contradiction. (Recall that we do not assume the outcome is Pareto efficient in general, nor is there any reason to believe that in practice that outcomes will be Pareto efficient.)

Due to the impossibility outlined in the above paragraph, we must use a stronger condition than simple preference. A ranking itself provides a natural way of strengthening preference. Fix a ranking. Consider a tier-k student i; that is, i is assigned to a tier-k college, $\mu(i) \in \Pi_k^{\trianglerighteq}$. Let $\mu^*(i)$ denote i's favorite college in her tier: $\mu^*(i) = \max_{P_i}\{c \mid \tau(c) = \tau(i)\}$. Note that μ^* need not be a well-defined outcome (multiple students could have the same favorite college), but clearly $\mu^*(i) \ R_i \ \mu(i)$. Therefore, i preferring a college to $\mu^*(i)$ is a stronger condition on preferences than i preferring a college to $\mu(i)$. We call this "strong" preference for an object desire.

Definition 1. *Given a ranking* $\trianglerighteq$*, let* $\mu^*(i) = \max_{P_i}\{c \mid \tau(c) = \tau(i)\}$*. Student* i ***desires*** *college* c *if* $c\ P_i\ \mu^*(i)$*. Student* i ***weakly desires*** *college* c *if* $c\ R_i\ \mu^*(i)$*.*

In words, a student desires a college c if it is preferred not only to her own outcome, $\mu(i)$, but also is preferred to every college that is in her tier (note that $\mu(i)$ is in student i's tier by construction). It should be clear that if i desires college c, then she also weakly desires c. Further, if i weakly desires c, then $c\ R_i\ \mu(i)$. In other words, strong desire implies weak desire, which in turn implies weak preference. Therefore, desire is a stronger condition than preference.

Example 1. Consider a student i with the following preferences, and whose outcome is $\mu(i) = c_5$. Further, say that under a ranking $\trianglerighteq$, $\tau(i) = k$, and $\Pi_k^{\trianglerighteq} = \{c_3, c_5\}$. This tier is indicated by the boxes in the table below.

$$P_i : c_1, c_2, \boxed{c_3}, c_4, \boxed{c_5}, c_6, \ldots$$

Then, relative to her outcome c_5, student i (weakly) prefers colleges c_1, c_2, c_3, c_4, and c_5; she weakly desires colleges c_1, c_2, and c_3; and she desires colleges c_1 and c_2.

3 Axioms

In this section, we introduce three axioms that we want our ranking to satisfy. Informally, we look for rankings that satisfy the following properties:

1. Any college c that a student desires over her own assignment $\mu(i)$ should be ranked higher than $\mu(i)$.
2. Colleges in higher tiers should be desired by lower tier students (and if not, they should be moved down in the rankings).
3. Within each tier, no college should be more desired than any other.

In the remainder of this section, we formally define and discuss each of these in turn.

3.1 The Axiom of Desire

Consider a student who receives her favorite school at an outcome μ. It may be that this student is a very good student attending a very good school; alternatively, it may be that she prefers a less good school simply because she has idiosyncratic preferences (e.g., locational preferences),

$P_{a^{high}}$	$P_{a^{low}}$	$P_{b^{high}}$	$P_{b^{low}}$
B^{good}	A^{good}	A^{good}	B^{good}
B^{bad}	$\boxed{A^{bad}}$	A^{bad}	$\boxed{B^{bad}}$
$\boxed{A^{good}}$	B^{good}	$\boxed{B^{good}}$	A^{good}
A^{bad}	B^{bad}	B^{bad}	A^{bad}

$\succ_{A^{good}}$	$\succ_{A^{bad}}$	$\succ_{B^{good}}$	$\succ_{B^{bad}}$
a^{high}	a^{high}	b^{high}	b^{high}
a^{low}	a^{low}	b^{low}	b^{low}
b^{high}	b^{high}	a^{high}	a^{high}
b^{low}	b^{low}	a^{low}	a^{low}

Table 1: The student and college preferences for Example 2. The boxes indicate the unique stable assignment of students to colleges.

and no other students are interested in it. In the latter case, she is likely to be admitted to this school as her first choice, yet we would not want the aggregate ranking to rank this school highly. Therefore, instead of basing our ranking on which school a student attends, we base our ranking on which schools a student desires.

The following example will help motivate our formal axiom.

Example 2. Suppose there are two states, A and B, and each state has two state universities, a "good" university and a "bad" university. Thus, the set of colleges is $C = \left\{A^{good}, A^{bad}, B^{good}, B^{bad}\right\}$. There are four students: a^{high}, a^{low}, b^{high} and b^{low} (representing a high and low performing resident student in each state). All students agree that within a state, the good school is better than the bad school. However, high performing students wish to get away from home, and so prefer any out of state school to any in state school. Low performing students prefer in state colleges to out of state. Colleges prefer in-state students to out of state students, and, within each category, prefer high-performing students to low-performing students. Specifically, the preferences of the students and colleges are given in Table 2. In this example, it is easy to calculate that the unique stable assignment is for the high-performing students to attend the good school in their state, and the low-performing students to attend the bad school in their state. This is shown by the boxes in Table 2.[7]

Given the heterogeneity among the students, the difficulty of mapping arbitrary preferences into a strict social ranking is well-understood. At the same time, the structure of our problem provides additional information that we can make use of: the resulting assignment that is the outcome of some competitive admissions process, which allows us to base our rankings on the outcomes the agents desire relative to their own assignment.

Example 2 again makes it clear why we must strengthen preference to desire: requiring $c \rhd \mu(i)$

[7]As discussed above, we do not in general model how μ is determined, and we will provide a method for determining a ranking given any input μ. However, a common equilibrium criterion for college admissions markets such as this is stability, where an assignment μ is *stable* if there is no student i and college c such that $c\ P_i\ \mu(i)$ and $i \succ_c \mu(c)$. The outcome μ defined is indeed the stable assignment for this example.

whenever $c \; P_i \; \mu(i)$ would imply both $A^{good} \rhd B^{good}$ and $B^{good} \rhd A^{good}$, a contradiction. Notice that this issue arises because the outcome depends not only on the student preferences, but also on the admissions criteria of the colleges: students a^{high} and b^{high} would prefer to trade their assignments, but the colleges do not.

At the same time, in this example, there is a clear natural ranking of the colleges:

$$\Pi_1^{\trianglerighteq} = \{A^{good}, B^{good}\}$$
$$\Pi_2^{\trianglerighteq} = \left\{A^{bad}, B^{bad}\right\}$$

Under this ranking, the good students are tier-1 students (each attends a tier-1 college), and the bad students are tier-2 students. Since a student *desires* a college c if it is better than all of the schools in its tier, students a^{high} and b^{high} do not desire any college, while student a^{low} desires A^{good} and student b^{low} desires B^{good}.[8] Thus, this ranking satisfies the property that all of the colleges that students desire are ranked in a higher tier. We formalize this property as our first axiom, the axiom of desire.

Definition 2. *A ranking $\trianglerighteq$ satisfies the* ***axiom of desire (AoD)*** *if for every student and every college $a \in C$ that i desires, we have $a \rhd \mu(i)$.*

In our analysis, it will be convenient to use alternative formulations of AoD. While trivial to prove, we state them as a lemma, for ease of reference. Given a student i and a subset of colleges $C' \subseteq C$, define i's **favorite** school in C' as:

$$\text{fav}_i(C') = \{c \in C' : c \; R_i \; c' \text{ for all } c' \in C'\}.$$

Lemma 1. *The following are equivalent to the Axiom of Desire:*

1. *For all k, all tier-k students i, and all colleges c in tiers lower than k ($\tau(c) > k$): $\text{fav}_i(\Pi_k^{\trianglerighteq}) \; P_i \; c$.*

2. *For all k and all colleges $c \in \Pi_k^{\trianglerighteq}$: $\text{fav}_{\mu(c)}(\Pi_k^{\trianglerighteq}) = \text{fav}_{\mu(c)}(\cup_{\ell \geq k} \Pi_\ell^{\trianglerighteq})$.*

In words, this says that any tier-k student must prefer at least one of the tier-k schools to any of the lower tier schools. This school may be the student's actual assignment, $\mu(i)$, or it may be some other tier-k school $c \neq \mu(i)$. In the latter case, the reason that i is not assigned to c

[8] Student a^{high} prefers college B^{bad} to her assignment, due to her locational preferences to go out of state. But, since there is another college in her tier, B^{good}, that she prefers even more than B^{bad}, while she prefers B^{bad} (to $\mu(a^{high}) = A^{good}$), she does not *desire* it.

is because school c prefers its own student. This is further explored in the following example, which may help to better understand the axiom of desire.

Example 3. There are three students $I = \{i, j, k\}$ and three schools $C = \{A, B, C\}$. The student preferences and the school rankings of the students are shown in the table below.

P_i	P_j	P_k
$A*$	$C*$	A
B	$\boxed{A}$	C
$\boxed{C}$	B	$\boxed{B*}$

$\succ_A$	$\succ_B$	$\succ_C$
j	j	i
i	k	k
k	i	j

Consider the outcome $\mu^\square$ indicated by the boxes.[9] Imagine a "shadow economy" in which students were allowed to trade admissions to a college. In this shadow economy, agents i and j would trade, resulting in each of them now getting their first choice. This *shadow assignment* μ^* is denoted by stars in the table.

In reality, students i and j are not getting their first choice because of the admissions criteria of the colleges, and students are not allowed to simply trade admissions to a college. While the shadow assignment μ^* is thus not actually implementable, it contains information on how the students themselves view the colleges, and, as this is the basis for our rankings, also suggests a natural way to construct a ranking of the colleges that is in line with AoD:

$$\{A, C\} \triangleright \{B\}.$$

Indeed, this ranking does satisfy AoD. The tier-1 students are students i and j, and both of these students' favorite tier-1 colleges are preferred to any college in a lower tier (in this case, B):

$$\text{fav}_i(\Pi_1) = A \; P_i \; B \quad \text{fav}_j(\Pi_1) = C \; P_j \; B.$$

As we will discuss in more detail in the next section, using shadow assignments like μ^* will be a useful tool in constructing rankings that satisfy AoD.

3.2 Justification

Our approach is based on ranking colleges based on desire instead of revealed preference. Returning to the motivating example from the previous section, consider a student i who receives her favorite school c at an outcome μ. It may be that this student is a very good student

[9]Note that $\mu^\square$ is stable; cf. footnote 7.

attending a very good school; alternatively, it may be that she prefers a less good school simply because she has idiosyncratic preferences, and no other students are interested in it. We can use desire to determine which of the two cases applies: if other students desire c, then it is likely that c was a competitive school, and should be ranked highly. If no other students desire c, then i likely got her top choice due to her idiosyncratic preferences, and c should be ranked low.

The axiom of desire introduced in the previous section is not quite sufficient for this purpose. What is missing is a converse of AoD. AoD says that *if* a school is desired by a student, it should be ranked in a higher tier; it should also be the case that a school is ranked in a higher tier *only if* it is desired by a student from a lower tier. When this is the case, we say that the school's ranking has been *justified*. Obviously, this does not apply to the schools in the bottom tier: the rankings of these schools are "justified" vacuously. Consider a ranking $\trianglerighteq$, and suppose we have justified college c's ranking. If c's student i desires college c', then this is justification to rank c' *above* c. If i weakly desires c' but does not desire c', then it is justified to rank c' in the *same tier* as c. We formalize this as our next axiom.

Definition 3 (Justification). *Consider a ranking $\trianglerighteq$ with K tiers, $\Pi^{\trianglerighteq} = \{\Pi^{\trianglerighteq}_1, \ldots, \Pi^{\trianglerighteq}_K\}$ and a tier-k college $c \in \Pi^{\trianglerighteq}_k$. College c's ranking is **justified** if either:*

1. *k is the lowest ranking: $k = K$;*
2. *a tier-k or $k+1$ college $\tilde{c}$'s ranking has been justified, and $\mu(\tilde{c})$ weakly desires c.*

*A ranking $\trianglerighteq$ is **justified** if every college's ranking is justified.*

Returning to Example 3, recall that the ranking we considered there was $\{A, C\} \triangleright \{B\}$. As B is in the last tier, it's ranking is justified vacuously. For college A, the tier 2 student $k(= \mu(B))$ desires A, and so A's ranking of $\tau(A) = 1$ is justified. A similar argument applies to B. Thus, the rankings of all colleges are justified, and $\trianglerighteq$ satisfies both AoD and justification.

When the outcome μ is Pareto efficient, justification reduces to a simpler criterion, stated in the lemma below.

Lemma 2. *Suppose the outcome μ is Pareto efficient and let $\trianglerighteq$ be a ranking with K tiers, $\Pi^{\trianglerighteq} = \{\Pi^{\trianglerighteq}_1, \ldots, \Pi^{\trianglerighteq}_K\}$. A tier-$k$ college c's ranking is justified if and only if c is ranked last ($k = K$) or c is desired by a tier-$(k+1)$ student whose outcome's ranking is justified.*

3.3 Balance

Many rankings will satisfy AoD and justification. Indeed, one way to trivially satisfy both is to rank all colleges in the same tier. In this case, no student can desire any college; thus, AoD will hold trivially. Similarly, as all colleges are ranked last, no college's "high" ranking needs to be justified. However, this is not a useful ranking; indeed, it is unlikely that all colleges are equally good. The last thing we are missing is a criterion for when it is (or is not) acceptable to rank colleges the same.

Consider a tier of colleges $\Pi_k^{\trianglerighteq}$ and the corresponding tier-k students $\mu(\Pi_k^{\trianglerighteq})$. Imagine asking each student $i \in \mu(\Pi_k^{\trianglerighteq})$ what is her favorite college in $\Pi_k^{\trianglerighteq}$. If more students answer college c than college c', then there is more demand for c than c', and so we posit that colleges c and c' should not be ranked in the same tier. Formally, we define the **aggregate demand** for a college $c \in \Pi_k^{\trianglerighteq}$, $D(c)$ as

$$D(c) = \sum_{i \in \mu(\Pi_k^{\trianglerighteq})} 1\{\text{fav}_i(\Pi_k^{\trianglerighteq}) = c\}.$$

A tier $\Pi_k^{\trianglerighteq}$ is **balanced** if $D(c) = D(c')$ for all $c, c' \in \Pi_k^{\trianglerighteq}$. Our next axiom says that all tiers should be balanced; if not, then some college is more demanded than another, and these two colleges should not be ranked the same.

Definition 4 (Balance)**.** *A ranking $\trianglerighteq$ is **balanced** if each tier k is balanced.*

Because in our model each college has one seat (and thus, the number of tier-k students is the same as the number of tier-k colleges), another way to restate the balanced axiom is that for any two tier-k students i, i', $\text{fav}_i(\Pi_k^{\trianglerighteq}) \neq \text{fav}_{i'}(\Pi_k^{\trianglerighteq})$. Note that on its own, balance is a weak condition on a ranking: any strict ranking of the colleges is trivially balanced.

3.4 Desirable Rankings

We have now introduced three axioms for rankings: the axiom of desire, justification, and balance. While we argue that each axiom is a useful desiderata for rankings to satisfy, taken individually, each of the three axioms introduced are weak conditions. Indeed, the trivial ranking that places all colleges in the same tier satisfies both AoD and justification, and any strict ranking over the objects is trivially balanced.

However, as we show in the remainder of the paper, combining these three axioms pins down a unique and accurate ranking of the colleges. We call any ranking that satisfies all three axioms a *desirable ranking.*

Definition 5 (Desirable rankings). *A ranking $\trianglerighteq$ is **desirable** if it justified, balanced, and satisfies the axiom of desire.*

4 Results

In this section, we present our main results on desirable rankings. Section 4.1, shows that any desirable ranking is "correct", in the sense that, as the market grows large, any desirable ranking coincides with the true underlying ranking of college qualities. Section 4.2 shows that for any outcome μ (and any market size), there is a unique desirable ranking, and provides a constructive algorithm for finding it.

4.1 Desirable Rankings Are Correct

In this section, we show that desirable rankings are correct, in the sense that as the market grows large, any desirable ranking coincides with the true underlying quality rankings of colleges. To make this point formally, we must first expand our model to define what is meant by a college's underlying 'quality'.

Specifically, let each student i's preferences over colleges be determined according to the following utility function:

$$U_{i,c} = \alpha\theta_c + (1-\alpha)\eta_{i,c}.$$

$U_{i,c}$ is student i's utility for attending college c. The random variable θ_c represents college c's intrinsic quality. This component of the utility function is the same for all students, and thus θ_c induces vertical preferences over colleges. The random variable $\eta_{i,c}$ is the idiosyncratic utility specific for student i if she attends college c, and thus corresponds to horizontal preferences (e.g., geographic preferences). We assume that all of the random variables are drawn independently from the uniform distribution on $[0,1]$. The uniform distribution is not important to our results: all of our results continue to hold in a more general framework.[10] The number $\alpha \in (0,1)$ represents the weight students place on college quality relative to the idiosyncratic component. For a student with utility function $U_{i,c}$ we define ordinal preferences in the standard way: $c \; P_i \; c'$ if $U_{i,c} > U_{i,c'}$ and $c \; R_i \; c'$ if $U_{i,c} \geq U_{i,c'}$.

A correct ranking of colleges should reflect the common quality component, rather than id-

[10]This is because we can define a general utility function $U_{i,c} = U(\theta_c, \eta_{i,c})$, and then transform any probability distribution into a uniform distribution, while at the same time monotonically transforming the utility function; cf. Lee (2016).

iosyncratic preferences. That is, we say a ranking is **correct** if $c \trianglerighteq c'$ if and only if $\theta_c \geq \theta_{c'}$. Since the random variables θ_c are uniform[0,1], as the market grows large, a correct ranking will rank a college with quality θ_c above approximately $(100 \times \theta_c)\%$ of the other college. Formally, given a ranking of the colleges $\trianglerighteq$, we can define the induced **percentile ranking**,

$$\rho(c) := \frac{|\{c' \in C | c \triangleright c'\}|}{n},$$

where $\rho(c)$ is the percentage of colleges that college c is ranked strictly ahead of. Then, a college with quality θ_c should have a ranking $\rho(c) \approx \theta_c$.

Of course, it is easy to construct examples of realized random variables where any desirable ranking will not reflect the underlying college qualities. This will occur when students draw idiosyncratic utilities that "reverse" their individual total utilities relative to the college qualities. Indeed, this can happen with any ranking system, not just desirable rankings. However, if the market is large, then these events should be relatively rare, and a good ranking method should be able to uncover the underlying college qualities with high probability. The next result shows that this is indeed the case for desirable rankings.

Given a ranking of the colleges $\trianglerighteq$ and a number $\epsilon > 0$, let

$$C^{\rho}(\epsilon) = \{c \in C : \theta_c - \epsilon < \rho(c) < \theta_c + \epsilon\}.$$

We can think of θ_c as college c's true or "target" ranking, and the set $C^{\rho}(\epsilon)$ then is the set of colleges whose calculated rankings are within ϵ of their target. If $c \in C^{\rho}(\epsilon)$, then we say that college c's ranking is **ϵ-correct**. Our first main result says that if $\trianglerighteq$ is a desirable ranking, then for any $\epsilon > 0$, as the market grows large, all college rankings are ϵ-correct.

Theorem 1 (Correct Rankings)**.** *Let $\trianglerighteq$ be a desirable ranking, and $\rho(\cdot)$ the induced percentage ranking. Then, for any $\epsilon > 0$, as n grows large, the expected proportion of colleges whose rankings are ϵ-correct approaches 1. That is,*

$$\lim_{n\to\infty} E\left(\frac{|C^{\rho}(\epsilon)|}{n}\right) = 1.$$

The full proof of this theorem is in the appendix; in the next section, we provide an outline and discuss the intuition.

A key first step in proving Theorem 1 is to show that for any desirable ranking, the ranking tiers become small in the limit. We state this result as a theorem, because it is also important in its own right, as it allays what might be a concern with desirable rankings. The concern is that because we allow for ties in rankings, a desirable ranking could in principle have too many colleges ranked the same, and a ranking that does not distinguish among colleges is not useful. Theorem 2 says that as the market grows large, the size of all ranking tiers, as a percentage of the market size, goes to zero. In other words, desirable rankings do meaningfully distinguish colleges into different tiers. Formally:

Theorem 2 (Small tiers)**.** *Let $\trianglerighteq$ be a desirable ranking. Then,*

$$\max_k \frac{1}{n}|\Pi_k^{\trianglerighteq}| \xrightarrow{p} 0 \text{ as } n \to \infty.$$

An important corollary of Theorem 2 that we will use in the proof of Theorem 1 is that, if all rankings tiers are small, then the percentage of colleges ranked above (below) any given $\theta \in [0,1]$ converges to $1-\theta$ (θ).[11]

Corollary 1. *Fix $\tilde{\theta} \in [0,1]$, and define $D(\tilde{\theta}) = \{c \in C | \rho(c) \geq \tilde{\theta}\}$ and $D'(\tilde{\theta}) = \{c \in C | \rho(c) < \tilde{\theta}\}$. Then,*

$$\frac{|D(\tilde{\theta})|}{n} \xrightarrow{p} 1-\tilde{\theta} \quad \text{and} \quad \frac{|D'(\tilde{\theta})|}{n} \xrightarrow{p} \tilde{\theta}.$$

Theorems 1 and 2 give support for desirable rankings, as they will meaningfully differentiate among the colleges and with high probability will uncover the true college qualities in the limit.

Intuition of the proofs

The proofs of Theorem 1 and 2 are inspired by an innovative technique introduced by Lee (2016). He uses a result from Dawande et al. (2001) on the size of bicliques in random bipartite graphs to analyze incentives in stable matching mechanisms. We use a similar technique to analyze desirable rankings.[12] We first discuss some basic concepts from random graph theory that we will need.

A **graph** $G = (V, E)$ is a pair that consists of a set V of **nodes** and a set E of **edges**, where each edge $e \in E$ is an unordered pair $e = (i, j)$ or (j, i) for $i, j \in V$. A graph G is **bipartite** if V can be partitioned as $V = V_1 \cup V_2$, where each edge has one node in V_1 and one node in V_2;

[11]That $|D(\tilde{\theta})|/n \leq 1-\tilde{\theta}$ follows from the definition of $\rho(\cdot)$. The converse is not immediate, i.e., it may be that $|D(\tilde{\theta})|/n > 1-\tilde{\theta}$. The corollary uses Theorem 2 to show that if rankings tiers are small, then $|D(\tilde{\theta})|/n$ becomes arbitrarily close to $1-\tilde{\theta}$ as n grows large.

[12]For another application of Lee's technique to interviews in the NRMP medical residency market, see Echenique et al. (2022).

for our purposes, it will generally be that $V_1 \subseteq I$ is a subset of the students and $V_2 \subseteq C$ is a subset of the colleges.

Definition 6. *Given a set of nodes $V = I' \cup C'$ where $I' \subseteq I$ and $C' \subseteq C$ and $p \in (0,1)$, a* ***random, bipartite graph*** *is a graph that is constructed as follows: each edge $(i,c) \in I' \times C'$ is included in the graph independently with probability p. We use G^p to denote a random bipartite graph.*

A biclique of a random graph is a complete connected subgraph. That is, given a graph G^p with nodes $V = I' \cup C'$ and edges E, a subset of nodes $I'' \cup C''$, where $I'' \subseteq I'$ and $C'' \subseteq C'$, is a **biclique** of G^p if $(i'', c'') \in E$ for all $i'' \in I''$ and all $c'' \in C''$. A biclique is **balanced** if $|I''| = |C''|$, and we call $B =: |I''| = |C''|$ the **size** of a balanced biclique. We make use of the following result from random graph theory.

Theorem 3 (Dawande et al. (2001))**.** *Consider a random bipartite graph G^p, where $V = V_1 \cup V_2$ is a partitioned set of nodes, $p \in (0,1)$ is a constant, $|V_1| = |V_2| = n$, and $\beta_n = 2\log(n)/\log(1/p)$. If a maximal balanced biclique of this graph has size $B \times B$, then*

$$Pr(\beta_n/2 \leq B \leq \beta_n) \to 1 \text{ as } n \to \infty.$$

We first discuss how we prove Theorem 2, since it is a stepping stone to proving Theorem 1. The overall strategy is to show that if there is a "large" tier, then there is also a "large" balanced biclique in an associated random graph. By Theorem 3, the probability of a large balanced biclique vanishes as n grows large, and thus so also must the probability of a large tier.

For simplicity, consider the case that $U_{i,c} = \theta_c + \eta_{i,c}$.[13] To construct this random graph, we take the colleges in a tier $\Pi_k^{\trianglerighteq}$ and divide them into three intervals (Low, Medium, and High) based on their common values (see Figure 4.1 for an example):

$$L = \{c \in \Pi_k^{\trianglerighteq} : \theta_c < a\} \quad M = \{c \in \Pi_k^{\trianglerighteq} : a < \theta_c \leq b\} \quad H = \{c \in \Pi_k^{\trianglerighteq} : \theta_c > b\},$$

where $0 \leq a < b \leq 1$. We then consider a set of students $U := \{i \in I | \theta_{\mu^*(i)} \in L\}$, where $\mu^*(i) = \text{fav}_i(\Pi_k^{\trianglerighteq})$. These students are "unusual", in the sense that their favorite tier-k college is one with a relatively low common value. Letting $\Delta = b - a$, draw an edge between a student i and a college c if i's idiosyncratic utility for c is low:

$$\eta_{i,c} < 1 - \Delta.$$

[13] This can be achieved by setting $\alpha = 1/2$ and then rescaling the utility function; in the proof in the appendix, we allow for arbitrary α.

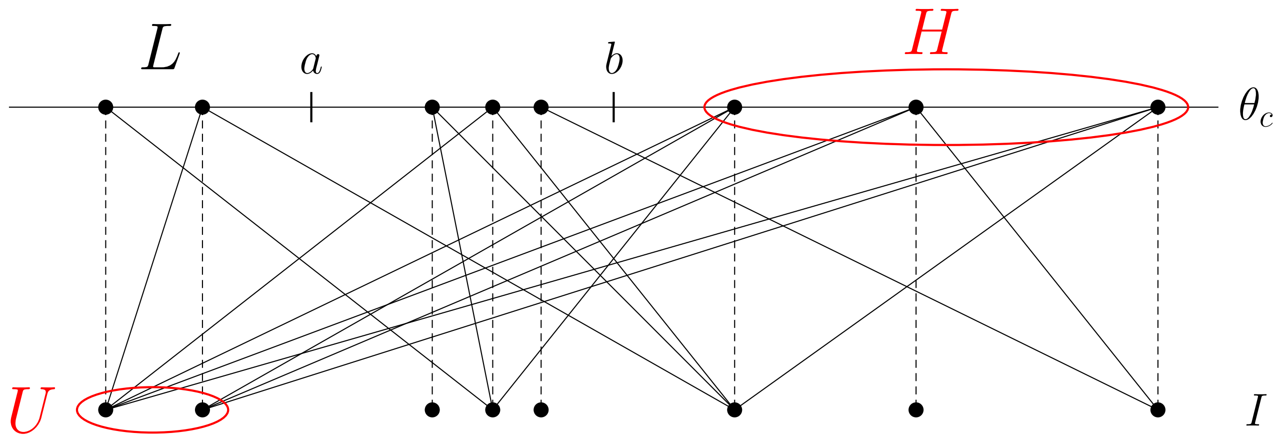


Figure 1: The dashed lines represent the matching μ^*, while the solid lines are the bipartite graph G^p, where student i and college c are connected if $\eta_{i,c} < p = 1 - \Delta$. All nodes in U are connected to all nodes in H, and thus $U \cup H$ forms a biclique. In this example, the size of the maximum balanced biclique is $\min\{|U|, |H|\} = \min\{2, 3\} = 2$.

That is, we draw an edge between each student i and college c independently with probability $p = 1 - \Delta$, and so this fits into the random bipartite graph model. Next, notice that $U \cup H$ is a biclique in this random graph. To see why, take some $i \in U$ and $c \in H$, and note that

$$\theta_{\mu^*(i)} + 1 \geq \theta_{\mu^*(i)} + \eta_{i,\mu^*(i)} > \theta_c + \eta_{i,c},$$

where the first inequality follows because $1 \geq \eta_{i,\mu^*(i)}$ and the second follows because $\mu^*(i)$ is i's favorite tier-k college. The first and last inequalities can be rearranged to

$$1 - (\theta_c - \theta_{\mu^*(i)}) > \eta_{i,\mu^*(i)}.$$

As $\theta_c > b$ (because $c \in H$) and $\theta_{\mu^*(i)} \leq a$ (because $\mu^*(i) \in L$), we have $\Delta < \theta_c - \theta_{\mu^*(i)}$, and so

$$1 - \Delta > \eta_{i,\mu^*(i)},$$

which is precisely our condition for an edge above. Thus, there is an edge between every student in U and every college in H, i.e., $U \cup H$ form a biclique. This means there is a biclique of size at least $B = \min\{|U|, |H|\} = \min\{|L|, |H|\}$, where $|U| = |L|$ follows by balancedness of $\trianglerighteq$. The remainder of the proof shows that if the tier is large, with high probability, we can always find some interval $[a, b]$ such that at least $1/4$ of the colleges have common value less than a, and at least $1/4$ of the colleges have common values above b. This implies that B is large (on the order of n), in violation of Theorem 3, which says that the size of the maximum balanced biclique grows only on the order of $\log(n)$.

After showing Theorem 2 and its corollary, Corollary 1, we use them together with Theorem 3 again to prove Theorem 1 as follows. For any $\theta \in [0,1]$ and $\epsilon > 0$, define two sets of colleges:

$$L = \{c \in C | \theta_c \geq \theta \text{ and } \rho(c) \leq \theta - \epsilon\} \quad \text{and} \quad W = \{c \in C | \theta_c < \theta' \text{ and } \rho(c) > \theta - \epsilon\},$$

where $\theta - \epsilon < \theta' < \theta$. In words, L is a set of colleges who "lose" in the ranking in the sense that they are at least ϵ below their target, and W is a set of colleges that are "wrongly ranked" in the sense that their quality is (relatively) low, but they are ranked above any college in L.

The goal is to show that $|L|/n \xrightarrow{p} 0$. We once again define a set of students,

$$U = \{i \in I : \mu^*(i) \in W\}$$

who are "unusual", in the sense that their favorite colleges are relatively overranked. We construct another bipartite graph, connecting a student i and a college c if:

$$\eta_{i,c} < 1 - (\theta - \theta')$$

By AoD, for any $i \in U$ and $c \in L$, we have $\mu^*(i)\ P_i\ c$, which can be used to show that $\eta_{i,c} < 1 - (\theta - \theta')$. This is precisely the condition for an edge in our graph, and so the set $L \cup U$ forms a biclique. Because $|U| = |W|$, there is a balanced biclique of size

$$\min\left\{\frac{|L|}{n}, \frac{|W|}{n}\right\} \xrightarrow{p} 0, \tag{1}$$

where the "$\xrightarrow{p}$" follows from Theorem 3. The last step is to show that $|W|/n \xrightarrow{p} \omega$ for some $\omega > 0$. Define $A = \{c \in C | \rho(c) > \theta - \epsilon\}$. By Corollary 1 to Theorem 2, $|A|/n \xrightarrow{p} 1 - (\theta - \epsilon)$. Finally, because $\theta - \epsilon < \theta'$, a non-vanishing proportion of the colleges in A must have qualities θ_c less than θ', which is the set W.[14] Thus, $|W|/n \xrightarrow{p} \omega > 0$, which, combined with (1) implies that $|L|/n \xrightarrow{p} 0$

The above is an outline of how we show that the proportion of colleges with ranking $\rho(c)$ more than ϵ below their target becomes vanishingly small as n grows large. We also must show that the proportion of colleges whose ranking $\rho(c)$ is more than ϵ above their target becomes vanishingly small as well. The argument proceeds similarly, and the details can be found in the full proof of Theorem 1 in the appendix.

[14]Because the θ_c's are uniformly distributed, as n grows large, the proportion of colleges with $\theta_c \geq \theta'$ approaches $1 - \theta'$. Because $1 - \theta' < 1 - (\theta - \epsilon)$ and $|A|/n \xrightarrow{p} 1 - (\theta - \epsilon)$, intuitively, in order to "fill" the set A, it must include some non-trivial proportion of colleges with qualities less than θ'.

4.2 Computing Desirable Rankings

In this section, we will show that a desirable ranking both exists and is unique (for any market size), and provide a constructive algorithm for finding this unique desirable ranking. We start by developing further properties of a desirable ranking. We make no restriction on the actual outcome μ, and in particular, μ need not be Pareto efficient. However, in developing our ranking, it is useful to consider a closely related alternative assignment that is Pareto efficient. We call this assignment a "shadow assignment" because it is the assignment that would arise in a shadow economy if, after being admitted to a college, students within a tier were able to trade these admissions amongst themselves. We first introduced this idea in Example 3 above; the next definition formalizes this concept.

Definition 7. *Fix an outcome μ, and let $\trianglerighteq$ be a desirable ranking. Define a new matching, μ^*, as follows: for each tier k and each tier-k student i, $\mu^*(i) = \mathrm{fav}_i(\Pi_k^{\trianglerighteq})$. We call μ^* the **shadow assignment** (relative to $\trianglerighteq$).*

It is not obvious that μ^* is even a valid assignment (i.e., it could be that more than one student has the same favorite school). However, in the proof of Lemma 3 below, we show that if $\trianglerighteq$ is desirable, then μ^* is a valid assignment, and further, it is Pareto efficient.

The shadow assignment μ^* has the property that, within each tier k, the tier k schools are redistributed amongst the tier k students in a manner that is the best from the student's perspective, ignoring any preferences or admissions criteria of the colleges.

While μ^* is not actually implementable due to the admissions criteria of the colleges (students are not allowed to "trade" their admissions to college), it has important properties that are useful in computing desirable rankings. We next establish that for any matching μ, there is a unique shadow assignment that is independent of the desirable ranking.

Lemma 3. *Fix an outcome μ, and let $\trianglerighteq_1$ and $\trianglerighteq_2$ be two desirable rankings, with corresponding shadow assignments μ_1^* and μ_2^*. Then, $\mu_1^* = \mu_2^*$.*

Proof. We prove this result by showing that for any desirable ranking $\trianglerighteq$, the corresponding shadow assignment μ^* is equivalent to the assignment that results from applying the top trading cycles algorithm to the original outcome μ, which we denote by μ^{TTC}.[15] To see this, notice

[15]The top trading cycles algorithm is due to Shapley and Scarf (1974). Briefly, it works by constructing a graph in which each college c points at college $c' = \mathrm{fav}_{\mu(c)}(C)$, i.e., each college points to the favorite college of the student currently assigned to it at μ. For any cycles that form (of which there must be at least one), the indicated trades are implemented, and the relevant students and colleges are removed. The process is repeated until no students or colleges remain. A formal definition can be found in the appendix.

that for any tier k, we can partition the tier k schools into cycles. Specifically, have each tier-k school c point to $\mu(c)$ and have each tier-k student i point to her favorite tier k college (i.e. $\mu^*(i)$). By balancedness of $\trianglerighteq$, each tier k student has a unique favorite school, and so each tier k student and tier k college belongs to exactly one cycle. We call these the tier-k cycles.

We claim that every tier-1 cycle is a top trading cycle. Consider a particular tier-1 cycle $(i_1, \mu^*(i_1), i_2, \ldots, \mu^*(i_L))$. By construction of the cycles, $\mu^*(i_\ell) = \mu(i_{\ell+1})$ for all ℓ. Each tier 1 student i is pointing at $\mu^*(i)$, her favorite tier-1 college. By AoD, i cannot prefer any lower-ranked college to her favorite tier-1 college. Therefore, $\mu^*(i_\ell)$ is i_ℓ's favorite college overall, and indeed $(i_1, \mu^*(i_1), i_2, \ldots, \mu^*(i_L))$ is a top-trading cycle.

For the inductive step, fix $m > 1$ and suppose that for each $\ell < m$, that every tier-ℓ cycle is a top trading cycle of $C \setminus \cup_{k<\ell}\Pi_k^{\trianglerighteq}$. In words, if we remove the colleges ranked higher than k, then each tier-ℓ cycle is a top trading cycle. We will show that each tier-m cycle is a top trading cycle. The argument is the same as for the base step. Each student i is pointing at her favorite tier-m college. By AoD, she must prefer this college to any college ranked m or lower. Therefore, this is her favorite college if we remove the colleges ranked higher than m.

Thus, we have shown that $\mu^* = \mu^{TTC}$, where μ^{TTC} is the outcome of applying the TTC algorithm to the original μ. Since this argument applies to the shadow assignment corresponding any desirable ranking $\trianglerighteq$, we conclude that $\mu_1^* = \mu_2^*$. $\square$

In the proof of Lemma 3, we actually showed something slightly stronger, which is that, not only is there a unique shadow assignment that is independent of the desirable ranking, this shadow assignment is equivalent to μ^{TTC}, the assignment obtained by running the top trading cycles algorithm on the original outcome μ.[16] Thus, a further implication of the above lemma is the following corollary.

Corollary 2. *Run the top trading cycles algorithm on an outcome μ, and consider two schools c, c' that belong to the same cycle. Then, in any desirable ranking, schools c and c' must be in the same tier: $\tau(c) = \tau(c')$.*

The above results suggest the following two-step procedure for finding a desirable ranking: (1) compute the TTC cycles and (2) determine how to rank the *cycles* (which will also determine the rankings of the colleges within the cycles). There will in general be multiple ways to rank the cycles. However, as we will show, there is only one way to do so that is consistent with justification. We first show how to do this in a small example, and then introduce the general algorithm.

[16]Cf. footnote 15; a formal definition of TTC can be found in Appendix A.

Example 4. There are 5 students $I = \{i_1, i_2, i_3, i_4, i_5\}$ and 5 colleges $C = \{A, B, C, D, E\}$. The table below shows the student preferences and the boxes indicate the outcome $\mu^{\square}$ with respect to which we calculate the desirable ranking. The bolded letters indicate μ^{TTC}, the outcome from running TTC on the original outcome.

P_1	P_2	P_3	P_4	P_5
A	C	B	$\mathbf{A}$	$\mathbf{B}$
B	A	$\boxed{\mathbf{C}}$	$\boxed{B}$	$\boxed{A}$
$\mathbf{D}$	$\mathbf{E}$	A	C	C
$\boxed{E}$	$\boxed{D}$	D	D	E
C	B	E	E	D

The trading cycles are (A, B), (C), and (D, E). Note that i_1 strictly prefers A and B to her TTC assignment, D. Therefore, A and B must be ranked higher than D which must be ranked the same as E. Similarly, C must be ranked higher than D or E as i_2 prefers it to E, her TTC assignment. Note that no student prefers D or E to her TTC assignment. Therefore, there is no justification for ranking these colleges above any other college. We set $C^1 = \{D, E\}$. We rank these colleges last, and remove them. This leaves the following submarket:

P_3	P_4	P_5
B	$\mathbf{A}$	$\mathbf{B}$
$\boxed{\mathbf{C}}$	$\boxed{B}$	$\boxed{A}$
A	C	C

As A and B are a trading cycle, they must be ranked the same. As B is strictly preferred by i_3 to her TTC assignment, C, B must be ranked higher than C. As no student prefers C to her TTC assignment, it must be ranked last. We set $C^2 = \{C\}$. Thus, college C is ranked one step above the colleges in C^1, and it is removed, along with its student i_3, leaving:

P_4	P_5
$\mathbf{A}$	$\mathbf{B}$
$\boxed{B}$	$\boxed{A}$

The remaining colleges form a cycle. Therefore, we set: $C^3 = \{A, B\}$. These colleges are thus placed one step ahead of the colleges in C^2, and all colleges have now been ranked. The final ranking is:

$$\{A, B\} \triangleright \{C\} \triangleright \{D, E\}$$

As the above example suggests, we find a desirable ranking by identifying the colleges that must be ranked last. We remove these schools and repeat.

Definition 8. *Given an outcome μ, a set of colleges C' is a* ***last cycle*** *if C' is a trading cycle under TTC and for every student i such that $\mu(i) \notin C'$, $\mu^{TTC}(i)\ P_i\ c'$ for every $c' \in C'$.*

There are two related reasons why we call such a cycle a "last" cycle. First, it is well known that there is not a unique ordering of cycles for TTC; however, the outcome of TTC is independent of the order in which cycles are removed. The last cycles, as defined above, are exactly the cycles for which there exists an ordering of cycles such that it is the last cycle removed under TTC. Second, as the next lemma shows, colleges in a last cycle must be ranked last by any desirable ranking.

Lemma 4. *Under a desirable ranking, a college c is ranked last if and only if c is in a last cycle.*

Proof. Let $\trianglerighteq$ be a desirable ranking. Let C' be a last cycle. As C' is a trading cycle, all colleges must be in the same tier. Their ranking cannot be justified by a lower-tier student: no student desires a college in C'. The ranking of the colleges in C' cannot be justified by a college $c \notin C'$ that is in the same tier. By the definition of a last cycle, the student $\mu^{TTC}(c)$ strictly prefers c to any college in C'. Therefore, it is only possible to justify the ranking of a college in C' if the college is ranked last.

Now, suppose c is not in a last cycle. Let C' be c's trading cycle. As C' is not a last cycle, there exists a student i and a $c' \in C'$ such that $c'\ P_i\ \mu^{TTC}(i)$. By AoD, $c' \triangleright \mu^{TTC}(i)$. Therefore, c' is not ranked last. Colleges c and c' are in the same cycle and therefore must be ranked the same. Therefore, c is not ranked last. □

It is straightforward to iterate this process. We call a cycle a **second-to-last cycle** if it is a last cycle when all of the last cycles are removed. An analogous argument shows that a desirable ranking ranks a college second to last if and only if it is part of a second-to-last cycle.

We summarize this in the following algorithm, which we call **delayed trading cycles (DTC)**.

Definition 9 (Delayed Trading Cycles)**.** *Given an outcome μ, let μ^{TTC} be the outcome obtained by running TTC on μ. Recursively define the sets C^ℓ as follows:*

- *Step $\ell = 1$: C^1 is the set of last-cycles.*

- *Step ℓ: If $C \setminus \cup_{\ell'=1}^{\ell-1} C^{\ell'} \neq \emptyset$, then C^ℓ is the set of last-cycles of $C \setminus \cup_{\ell'=1}^{\ell-1} C^{\ell'}$. Otherwise, stop.*

Let $C^1, \ldots, C^L$ be the resulting partition of the colleges. The ***DTC ranking*** *of the colleges, $\trianglerighteq$, is given as follows: for any two colleges $a, b \in C$, where $a \in C^\ell$ and $b \in C^{\ell'}$, $a \trianglerighteq b$ if and only if $\ell \geq \ell'$. The tier-k colleges are $\Pi_k^{\trianglerighteq} = C^{L-k+1}$.*

Our final result shows that the DTC ranking produces a desirable ranking of the colleges, and further, this is the unique desirable ranking.

Theorem 4. *The DTC ranking is the unique desirable ranking of the colleges.*

Proof. Let $\trianglerighteq$ be the DTC ranking. We first show that DTC is desirable, and then show uniqueness.

To see that $\trianglerighteq$ satisfies AoD, consider a tier k student i. Let $c_i = \mu^{TTC}(i)$ be i's TTC college, and let C_i be the cycle that includes college c_i in the running of TTC. Note that since c_i is included in the same cycle as $\mu(i)$, it is removed at the same step as $\mu(i)$, and so c_i is also ranked k^{th}. By definition of the TTC algorithm, $c_i \; P_i \; c'$ for any other c' in C_i (because each student always points to their most preferred remaining school at each step of the algorithm). Similarly, $c_i \; P_i \; c'$ for any other c' that was removed at a step of the algorithm prior to or including the step in which c_i was removed.[17] In other words, i strictly prefers school c_i, which is ranked k^{th}, to all lower ranked colleges. Thus, i does not desire any such college, and so AoD holds. For balancedness, note that the argument above shows that $\mu^{TTC}(i)$ is i's favorite tier-k college. A similar argument shows that for all other tier-k students j, j's favorite tier k college is $\mu^{TTC}(j)$. Since each student j in the tier k cycles points to a unique college—namely, their TTC college, $\mu^{TTC}(j)$—each student has a distinct favorite college, and the ranking is balanced.

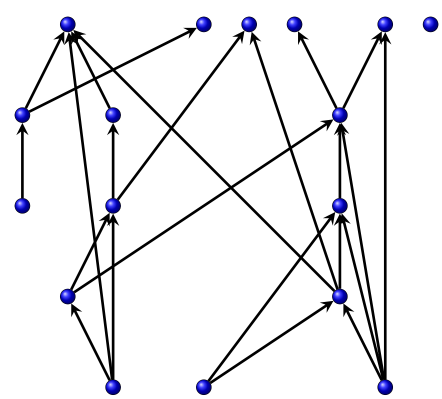
To see that $\trianglerighteq$ is justified, first notice that every college in C^1 is ranked K^{th} (i.e., ranked last), and so are justified by definition. Consider a tier $K-1$ college c', which means that $c' \in C^2$, and let C' be the top trading cycle of this college. Since C' was not a last cycle in step 1, there must be some i such that $\mu(i) \notin C'$ and $c'' \in C'$ such that $c'' \; P_i \; \mu^{TTC}(i)$. As C' is a last cycle in step 2, for all i such that $\mu(i) \in \cup_{\ell=2}^{L} C^\ell$, we have $\mu^{TTC}(i) \; P_i \; c'$ for every $c' \in C'$, which means that the student i from the previous sentence must have been assigned to a college that was removed in step 1. In other words, there is some i such that $\mu(i) \in C^1$ and $c'' \; P_i \; \mu^{TTC}(i)$ for

[17] This follows because colleges are only removed when they are part of a *last cycle* (see Definition 8). For all colleges removed prior to or including the step in which c_i was removed, c_i is still present in the market, and so by Definition 8, $c_i \; P_i \; c'$ for any such c'.

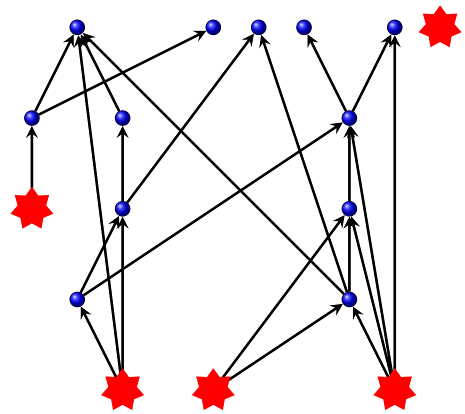
some $c'' \in C''$. Since $\mu(i) \in C^1$, it is a last-ranked college, and so $\mu(i)$'s ranking is justified. Further, $\mu^{TTC}(i)$ is student i's favorite tier K college, and since $c'' \; P_i \; \mu^{TTC}(i)$, student i desires c'', and so college c'''s ranking is justified. College c'' can be used to justify the ranking of the college it points to in the cycle C', that college justifies the ranking of the next college in the cycle, and so forth. Thus, all colleges in cycle C' have rankings that are justified. An analogous argument holds for every cycle that was removed in step 2, and so, all of the tier $K-1$ colleges have rankings that are justified. Repeating this same argument for $C^3, C^4, \ldots, C^L$ shows that all college rankings are justified. Therefore, $\trianglerighteq$ is balanced, justified, and satisfies AoD, i.e., it is a desirable ranking.

Uniqueness follows from Corollary 2, and Lemma 4: By Corollary 2, any two colleges in the same TTC cycle must be in the same tier in any desirable ranking. Iterated applications of Lemma 4 imply that there is a unique way to rank these cycles (and the colleges within them) that coincides with the DTC ranking. $\square$

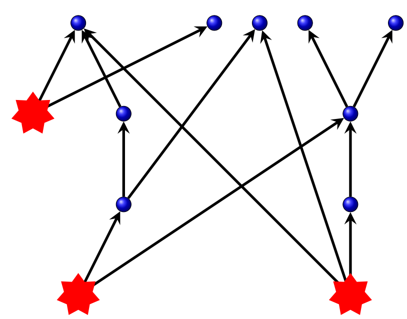
The DTC Algorithm is perhaps easiest to understand pictorially. In the graph below, each node represents a particular trading cycle that arises when we run the TTC algorithm on μ (so, each node represents a set of colleges). There is a directed arrow from node χ_1 to node χ_2 if there is a student in cycle χ_1 that desires one or more of the schools in χ_2. Note that by construction, the cycles from TTC produce a directed tree. No student in a top trading cycle desires any schools outside the cycle, so there cannot be an edge from an earlier cycle to a later cycle. Suppose TTC produces the following graph:



The last cycles are simply the nodes at the "bottom" of the tree. Namely, if a cycle is not pointed at by another cycle, it is a last-cycle. We indicate the last-cycles with large red stars in the graph below.



The DTC algorithm ranks these schools last and then removes them. This produces the following graph (with the new last cycles again indicated by large red stars).



These new "last" cycles are ranked second-to-last, removed, and the process is repeated.

5 Conclusion

We consider the problem of how to rank alternatives that are the outcome of a competitive process, such as rankings of colleges, medical residency programs, or academic journals. We introduce a new paradigm for constructing rankings that is based on the notion of desirability: alternatives that an agent desires (relative to what she receives) should be ranked higher.

We introduce several axioms that formalize the notion of desirability, and build an algorithm, the Delayed Trading Cycles algorithm, that can be used to calculate desirable rankings. Further, we characterize the output of Delayed Trading Cycles as the unique desirable ranking, and show that in the limit, it coincides with the true underlying ranking of college qualities.

We think there are several appealing features of desirability-based rankings methods compared to others, such as those based on revealed preference. First, desirable rankings are more robust to idiosyncratic preferences that are unique to a particular agent, rather than the underlying college quality. By basing the ranking on the colleges an agent desires (looking "up" an agent's preference list, rather than down), we are less likely to make incorrect inferences when a student prefers a lower quality college for purely idiosyncratic reasons.[18] Second, as discussed in the introduction, standard methods of calculating rankings (acceptance rates, citation counts) are problematic because they are susceptible to gaming. Notice that under a desirable ranking, for a college to be ranked higher, it must be desired by a student at a lower-ranked school. Thus, if a college wants to raise its rank, it must increase its quality to become a more attractive college. Thus, desirable rankings should provide better incentives for colleges to invest in quality improvement as opposed to strategies that artificially infltate their rankings. Formalizing this and showing that desirable rankings are less susceptible to gaming on the part of the colleges is an interesting question for future work.

References

Arrow, K. J. (1950): "A Difficulty in the Concept of Social Welfare," *Journal of Political Economy*, 58, 328–346.

Avery, C. N., M. E. Glickman, C. M. Hoxby, and A. Metrick (2013): "A revealed preference ranking of US colleges and universities," *The Quarterly Journal of Economics*, 128, 425–467.

Belkin, D. (2019): "For Sale: SAT-Takers' Names. Colleges Buy Student Data and Boost Exclusivity," *Wall Street Journal*.

Bowman, N. A. and M. N. Bastedo (2009): "Getting on the front page: Organizational reputation, status signals, and the impact of US News and World Report on student decisions," *Research in Higher Education*, 50, 415–436.

Chapman, C. A., J. C. Bicca-Marques, S. Calvignac-Spencer, P. Fan, P. J. Fashing, J. Gogarten, S. Guo, C. A. Hemingway, F. Leendertz, B. Li, et al. (2019):

[18] Cf. the introduction, where we discussed this issue. Note also that it is possible for desirability to make incorrect inferences. For instance, say the college qualities are $\theta_A > \theta_B > \theta_C$, but, due to the idiosyncratic component student i ranks $B\ P_i\ A\ P_i\ C$. If the student is accepted by A but rejected by B, then desirability will incorrectly infer that B is better than A. However, this requires both the student to prefer the lower quality college B, and the student be admitted to the (better) college A, but not the (worse) college B. In other words, both the student and the college must have idiosyncratic and coordinate preferences for lower-ranked alternatives. This will be a rarer occurrence than this happening for one side only, and so desirability is less likely to make incorrect inferences.

"Games academics play and their consequences: how authorship, h-index and journal impact factors are shaping the future of academia," *Proceedings of the Royal Society B*, 286, 20192047.

Dawande, M., P. Keskinocak, J. M. Swaminathan, and S. Tayur (2001): "On bipartite and multipartite clique problems," *Journal of Algorithms*, 41, 388–403.

Echenique, F., R. González, A. J. Wilson, and L. Yariv (2022): "Top of the Batch: Interviews and the Match," *American Economic Review: Insights*, 4, 223–38.

Fong, E. A. and A. W. Wilhite (2017): "Authorship and citation manipulation in academic research," *PloS one*, 12, e0187394.

Gibbard, A. (1973): "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41, 587–601.

Golden, D. (2001): "Glass floor colleges reject top applicants, accepting only the students likely to enroll," *Wall Street Journal*, A1.

Griffith, A. and K. Rask (2007): "The influence of the US News and World Report collegiate rankings on the matriculation decision of high-ability students: 1995–2004," *Economics of Education Review*, 26, 244–255.

Ioannidis, J. P. and B. D. Thombs (2019): "A user's guide to inflated and manipulated impact factors," *European Journal of Clinical Investigation*, 49.

Jena, A. B., V. M. Arora, K. E. Hauer, S. Durning, N. Borges, N. Oriol, D. M. Elnicki, M. J. Fagan, H. E. Harrell, D. Torre, et al. (2012): "The prevalence and nature of postinterview communications between residency programs and applicants during the match," *Academic Medicine*, 87, 1434–1442.

Lee, S. (2016): "Incentive compatibility of large centralized matching markets," *The Review of Economic Studies*, 84, 444–463.

Martin, B. R. (2016): "Editors JIF-boosting stratagems–Which are appropriate and which not?" *Research Policy*, 45, 1–7.

Roth, A. E. and A. Postlewaite (1977): "Weak Versus Strong Domination in a Market with Indivisible Goods," *Journal of Mathematical Economics*, 4, 131–137.

Satterthwaite, M. (1975): "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," 10, 187–216.

Sbicca, J. A., E. S. Gorell, D. H. Peng, and A. T. Lane (2012): "A follow-up survey of the integrity of the dermatology National Resident Matching Program," *Journal of the American Academy of Dermatology*, 67, 429–435.

Shapley, L. and H. Scarf (1974): "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, 23–37.

Toor, R. (2000): "Pushy parents and other tales of the admissions game," *Chronicle of Higher Education*, 47, B18.

Wu, A. J., N. Vapiwala, S. J. Chmura, P. Das, R. H. Decker, S. A. Terezakis, and A. L. Zietman (2015): "Taking the game out of the match: A simple proposal," *International Journal of Radiation Oncology, Biology, Physics*, 93, 945–948.

A Top Trading Cycles

In this appendix, we provide a formal definition of the top trading cycles algorithm.

Fix an outcome μ and a set of college $C' \subseteq C$. A **top trading cycle of** C' is a list of distinct colleges $\chi = (c^1, c^2, \ldots, c^n)$ such that, for all k, $c^k \in C'$ and $c^{k+1} = \text{fav}_{\mu(c^k)}(C')$, where the superscripts k are taken modulo n, i.e., $c^{n+1} = c^1$. Given a set of colleges C' and a subset $C'' \subseteq C'$, we call C'' a **set of top trading cycles of** C' if C'' can be decomposed into top trading cycles of C'.

The Top Trading Cycles Algorithm

Let μ be an assignment of students to colleges.

Step 1: Let $C^1 = C$. Draw a directed graph as follows: for each $c \in C^1$, college c points to college $\text{fav}_{\mu(c)}(C^1)$. Let $\chi^1 = \{\chi^1_1, \ldots, \chi^1_M\}$ denote the set of M top trading cycles that form.[19] Select some subset $\chi' \subseteq \chi^1$ of the top trading cycles. For each college c included in a top trading cycle in χ', set $\mu^{TTC}(c) = \mu_{c'}$, where c' is the school that points to c' (i.e., school c's TTC assignment is the student at the school that points to c). Remove all colleges included in a top trading cycle in χ'.

[19] As there are a finite number of colleges, at least one top trading cycle exists, and each college can be part of at most one top trading cycle.

Step k: Let C^k be all of the schools that have not been removed in steps prior to k, and repeat the procedure from step 1: each $c \in C^k$ points to $\text{fav}_{\mu(c)}(C^k)$, the top trading cycles $\chi^k = \{\chi_1^k, \ldots, \chi_M^k\}$ are identified, and a selection of them are removed. For the removed colleges, their TTC assignment $\mu^{TTC}(c)$ is the student who, under μ, is assigned to the college c' that points to c.

The procedure above produces a new matching μ^{TTC}. Notice that in the description of the algorithm above, we do not require that all cycles be removed in any given step. However, any cycle that is not removed simply remains a top trading cycle in all later steps, and so can be removed at any point. In other words, no matter the order in which cycles are removed, the algorithm always results in the same final matching μ^{TTC} at the end.

B Proofs

The proofs in this appendix are presented in a different order than the theorems are presented in the main text. In particular, Theorem 2 and Corollary 1 are key tools used to prove Theorem 1. Thus, we prove these results first, and then prove Theorem 1.

B.1 Proof of Theorem 2

For each market realization, let X denote the largest tier of schools under desirable ranking $\trianglerighteq$, i.e., $X = \Pi_k^{\trianglerighteq}$ for the tier k such that $|\Pi_k^{\trianglerighteq}| \geq |\Pi_{k'}^{\trianglerighteq}|$ for all k' (if there are multiple largest tiers, we pick one of them). An equivalent way to state Theorem 2 is as follows: For any $\tau, \lambda > 0$, there exists N such that

$$\Pr\left(\frac{1}{n}|X| > \tau\right) < \lambda \text{ for every } n > N.$$

This is the statement that we will show.

Let $I_X = \{i \in I | \mu(i) \in X\}$ be the students assigned to the colleges in X. Take some number $\Delta \in (0,1)$ such that $\frac{\alpha}{1-\alpha} \times \Delta < 1$, and let $p_\Delta = 1 - \frac{\alpha}{1-\alpha}\Delta$ (note that $p_\Delta \in (0,1)$). Define a bipartite graph G^{p_Δ} as follows: $X \cup I_X$ is a bipartitioned set of nodes. Two vertices $c \in X$ and $i \in I_X$ are joined by an edge if and only if

$$\eta_{i,c} \leq p_\Delta. \tag{2}$$

Since the $\eta_{i,c}$'s are iid uniform$[0,1]$, this says that every student and college is joined by an edge independently with probability p_Δ. The next lemma relates to the size of balanced bicliques in the random graph G^{p_Δ}. To state it, given two numbers $0 \leq a < b \leq 1$, define

$$\begin{aligned} L(a) &= \{c \in X | \theta_c \leq a\} \\ H(b) &= \{c \in X | \theta_c \geq b\} \end{aligned}$$

to be the subsets of colleges with low (below a) and high (above b) common values, respectively.

Lemma 5. *Let $\Delta \in (0, \min\{1, \frac{1-\alpha}{\alpha}\})$, and let G^{p_Δ} be the random bipartite graph as defined in equation (2). For any $a, b \in (0,1)$ such that $\Delta = b - a$, the graph G^{p_Δ} has a balanced biclique of size min$\{|L(a)|, |H(b)|\}$.*

Proof of Lemma 5. For a student $i \in I_X$, let $\mu^*(i)$ denote i's favorite assignment in X. Since $\trianglerighteq$ is a balanced ranking, and since the colleges in X are ranked the same, μ^* is a proper matching between I_X and X. Let

$$U = \{i | \mu^*(i) \in L(a)\}$$

be the set of "unusual" students, in the sense that their favorite college in X has a low common value.

We claim that $H(b)$ and U form a biclique in the graph G^{p_Δ}. Intuitively, a student whose favorite college has a low common-value must have low idiosyncratic draws with each of the high common-value colleges. Consider any $c \in H(b)$ and student $i \in U$. Let $d = \mu^*(i)$, i.e. d is i's favorite college in X. By assumption, $d \in L(a)$. As d is i's favorite college in X, $d \; P_i \; c$. Therefore,

$$\begin{aligned} \alpha\theta_d + (1-\alpha)\eta_{i,d} &> \alpha\theta_c + (1-\alpha)\eta_{i,c} \\ \alpha\theta_d + (1-\alpha) &> \alpha\theta_c + (1-\alpha)\eta_{i,c} \\ \alpha(\theta_d - \theta_c) + (1-\alpha) &> (1-\alpha)\eta_{i,c} \\ 1 - \frac{\alpha}{1-\alpha}(\theta_c - \theta_d) &> \eta_{i,c} \\ 1 - \frac{\alpha}{1-\alpha}\Delta &> \eta_{i,c} \\ p_\Delta &> \eta_{i,c} \end{aligned}$$

where the second to last line follows because $\Delta = b - a$, $\theta_c > a$, and $\theta_d < b$. So, indeed, given the construction of G^{p_Δ} (see the edge condition given by Equation 2), there is an edge between every college in $H(b)$ and every student in U. Therefore, the vertices form a biclique, and thus there exists a balanced biclique of size $\min\{|U|, |H(b)|\}$. Since $|L(a)| = |U|$ (the assignment μ^* is a bijection), there is a balanced biclique of size at least $\min\{|L(a)|, |H(b)|\}$.

$\square$

Lemma 6. *For every $\tau, \delta \in (0,1)$, there exists $\epsilon, \Delta \in (0,1)$ and $N > 0$ such that for all $n > N$,*

$$Pr\left(\frac{1}{n}|X| > \tau\right) \leq Pr\left(\frac{B}{n} > \epsilon\right) + \delta,$$

where B is the size of the maximal balanced biclique in the associated graph G^{p_Δ}.

Proof of Lemma 6. Fix any $\tau \in (0,1)$, and let M be an integer such that $\frac{1}{M} < \min\{\frac{\tau}{4}, \frac{1-\alpha}{\alpha}\}$. Further, choose $\epsilon > 0$ such that $\epsilon < \frac{\tau}{4M}$, and set $\Delta = 1/M$. Divide the unit interval into M subintervals, each with length $1/M$, where we refer to $\left[\frac{m-1}{M}, \frac{m}{M}\right)$ as the m^{th} **subinterval**. Let

$$C^m = \left\{c \in C : \frac{m}{M} \leq \theta_c < \frac{m+1}{M}\right\}$$

be the colleges with common values in the m^{th} subinterval. Let E_n be the event that $\frac{1}{n}|X| > \tau$, and F_n the event that $\frac{|C^m|}{n} < \frac{\tau}{4}$ for all subintervals $m = 1, \ldots, M$.

Claim 1. *Assume that the events E_n and F_n are true. Then, in the graph G^{p_Δ}, there is a balanced biclique of size $B = \lceil \epsilon n \rceil$.*

Proof of Claim 1. Since F_n is true, we have $|C^m| < (\tau/4) \times n$ for all m, which also implies that $|C^m \cap X| < (\tau/4) \times n$ for all m. The event E_n is $\frac{1}{n}|X| > \tau$, and a necessary condition for this to be true is that there is some m^{th} subinterval that contains the common values of at least $\epsilon \times n$ of the colleges in X.[20] Let m be the index of the first such subinterval that contains the common values of at least $\epsilon \times n$ of the colleges in X. Combining these observations, we have:

$$\begin{aligned}
\left(\sum_{m'=1}^{m-1} |C^{m'} \cap X|\right) + |C^m \cap X| + |C^{m+1} \cap X| &\leq (m-1)\epsilon n + \frac{\tau}{4}n + \frac{\tau}{4}n \\
&< (m-1)\frac{\tau}{4M}n + \frac{\tau}{4}n + \frac{\tau}{4}n \\
&= \frac{m-1}{M}\frac{\tau}{4}n + \frac{\tau}{4}n + \frac{\tau}{4}n \\
&< \frac{\tau}{4}n + \frac{\tau}{4}n + \frac{\tau}{4}n \\
&= \frac{3}{4}\tau n. \qquad (3)
\end{aligned}$$

[20] Indeed, if not, then $|X| < M \times \epsilon \times n < M \times \tau/(4M) \times n < \tau \times n$, which implies $|X| < \tau n$.

The first inequality follows from the fact that each of the first $m-1$ intervals has less than ϵn of the colleges in X (by definition of m) and intervals m and $m+1$ have at most $(\tau/4)n$ total colleges (and therefore, must have less than that number of X colleges). The second inequality comes from the fact that we chose $\epsilon < \frac{\tau}{4M}$. The remaining inequalities are simple algebra.

As $\sum_{m'=1}^{M} |C^{m'} \cap X| = |X| > \tau n$, combining this with equation (3) implies that

$$\begin{aligned} \sum_{m'=m+2}^{M} |C^{m'} \cap X| &> \frac{\tau n}{4} \\ &> \epsilon n, \end{aligned}$$

which follows because we chose $\epsilon < \tau/4$. In sum, we have shown that there are at least ϵn colleges in X with common values in the interval $\left[0, \frac{m+1}{M}\right)$, and at least ϵn colleges with common values in the interval $\left[\frac{m+2}{M}, 1\right]$. Therefore, by Lemma 5, setting $a = \frac{m+1}{M}$, $b = \frac{m+2}{M}$, and $\Delta = \frac{1}{M}$,[21] we conclude that in the random graph G^{p_Δ}, there is a balanced biclique of size at least ϵn.

$\square$

Recall that E_n is the event that $\frac{1}{n}|X| > \tau$, and F_n is the event that $\frac{|C^m|}{n} < \frac{\tau}{4}$ for all subintervals $m = 1, \ldots, M$. Our goal is to show that $\Pr(E_n) \leq \Pr(B/n > \epsilon) + \delta$. Note that we can write:

$$\begin{aligned} \Pr(E_n) &= \Pr(E_n \cap F_n) + \Pr(E_n \cap F_n^c) \\ &\leq \Pr(E_n \cap F_n) + \Pr(F_n^c), \end{aligned} \tag{4}$$

where F_n^c is the complement of F_n, i.e., F_n^c is the event that there is some subinterval m' such that $|C^{m'}|/n > \tau/4$. Because all θ_c are iid uniform, $\frac{|C^m|}{n} \xrightarrow{p} \frac{1}{M}$ for each m, and thus for any $\epsilon', \delta' > 0$, for all n large enough,

$$\Pr\left(\left\{\left|\frac{|C^m|}{n} - \frac{1}{M}\right| < \epsilon' \text{ for all } m\right\}\right) > 1 - \delta'. \tag{5}$$

Further, because $1/M < \tau/4$, this also implies that for all n large enough,

$$\Pr\left(F_n\right) > 1 - \delta'. \tag{6}$$

for all $\delta' > 0$. In other words, with arbitrarily high probability, for all n large enough, every

[21] Recall that we chose M large enough such that $1/M < \min\{\tau/4, (1-\alpha)/\alpha\}$, and so the conditions of Lemma 5 are satisfied.

interval contains at most $\frac{\tau}{4} \times n$ of the colleges. This can be restated as

$$\Pr\left(F_n^c\right) < \delta' \text{ for all sufficiently large } n. \tag{7}$$

Combining equations (4) and (7), and choosing $\delta' = \delta$, we have

$$\Pr(E_n) \leq \Pr(E_n \cap F_n) + \delta \text{ for all } n \text{ sufficiently large.} \tag{8}$$

By Claim 1, $\Pr(E_n \cap F_n) \leq \Pr(B/n > \epsilon)$, and so

$$\Pr(E_n) \leq \Pr(B/n > \epsilon) + \delta \text{ for all } n \text{ sufficiently large,} \tag{9}$$

which is what we wanted to show. □

To finish the proof of Theorem 2, let $\beta_n = 2\log(n)/\log(1/(p_\Delta))$, where Δ is chosen as in Lemma 6, and let B be the size of the maximal balanced biclique in the graph G^{p_Δ}. Theorem 3 implies that

$$\Pr(B \leq \beta_n) \to 1 \text{ as } n \to \infty$$

and, since $\beta_n/n \to 0$, we have

$$\frac{B}{n} \xrightarrow{p} 0 \text{ as } n \to \infty. \tag{10}$$

Now, by Lemma 6, there exists N such that for all $n > N$, we have

$$\Pr\left(\frac{1}{n}|X| > \tau\right) \leq \Pr\left(\frac{B}{n} > \epsilon\right) + \delta \tag{11}$$

By (10), the first term on the RHS of (11) converges to 0, and thus by choosing δ sufficiently small and n sufficiently large, we can make the RHS of (11) smaller than any $\lambda > 0$, which completes the proof of Theorem 2.

□

B.2 Proof of Corollary 1

Fix $\tilde{\theta} \in [0,1]$. We first show that $|D(\tilde{\theta})|/n \xrightarrow{p} 1 - \tilde{\theta}$. First notice that $|D(\tilde{\theta})|/n \leq 1 - \tilde{\theta}$, which follows simply from the definition of $\rho(\cdot)$. However, the converse does not immediately follow, i.e., it may be that $|D(\tilde{\theta})|/n > 1 - \tilde{\theta}$. We show that as n grows large, $|D(\tilde{\theta})|/n$ becomes arbitrarily close to $1-\tilde{\theta}$ with high probability, i.e., we show that, for any $\epsilon > 0$,

$$\Pr\left(\frac{|D(\tilde{\theta})|}{n} > (1-\tilde{\theta}) - \epsilon\right) \to 1 \text{ as } n \to \infty. \tag{12}$$

Define the following quantities:

$$\underline{\rho} := \max_c \{\rho(c) | \rho(c) < \tilde{\theta}\}$$
$$\bar{\rho} := \min_c \{\rho(c) | \rho(c) \geq \tilde{\theta}\}.$$

By construction, $\bar{\rho} \geq \tilde{\theta} \geq \underline{\rho}$ (notice that it is possible that either or both of these inequalities are strict). Also, by the definition of $\rho(\cdot)$, we have

$$\bar{\rho} - \underline{\rho} = \frac{|c \in C | \rho(c) = \underline{\rho}|}{n}.$$

By Theorem 2, the RHS of the above equation converges in probability to 0, and thus $\bar{\rho} \xrightarrow{p} \underline{\rho}$. Combined with $\bar{\rho} \geq \tilde{\theta} \geq \underline{\rho}$, we have

$$\bar{\rho} \xrightarrow{p} \tilde{\theta}. \tag{13}$$

Define

$$E = \{c \in C | \rho(c) \geq \bar{\rho}\}.$$

Since $\bar{\rho} \geq \tilde{\theta}$, we have $E \subseteq D(\tilde{\theta})$, which implies that

$$\Pr\left(\frac{|D(\tilde{\theta})|}{n} > (1-\tilde{\theta}) - \epsilon\right) \geq \Pr\left(\frac{|E|}{n} > (1-\tilde{\theta}) - \epsilon\right) \tag{14}$$

for any $\epsilon > 0$. Next, notice that by definition of $\bar{\rho}$, $|E|/n = 1 - \bar{\rho}$, so the RHS of (14) becomes

$$\Pr(1 - \bar{\rho} > (1-\tilde{\theta}) - \epsilon) = \Pr(\bar{\rho} < \tilde{\theta} + \epsilon).$$

By (13), $\Pr(\bar{\rho} < \tilde{\theta} + \epsilon) \to 1$. So, the RHS of (14) converges 1, and thus

$$\Pr\left(\frac{|D(\tilde{\theta})|}{n} > (1-\tilde{\theta}) - \epsilon\right) \geq \Pr\left(\frac{|E|}{n} > (1-\tilde{\theta}) - \epsilon\right) \to 1,$$

and thus,

$$\Pr\left(\frac{|D(\tilde{\theta})|}{n} > (1-\tilde{\theta}) - \epsilon\right) \to 1,$$

which is what we wanted to show. The second statement that $|D'(\tilde{\theta})|/n \xrightarrow{p} \tilde{\theta}$ is an immediate consequence of the first.

$\square$

B.3 Proof of Theorem 1

To prove Theorem 1, we first show the following two propositions.

Proposition 1. *Let $\trianglerighteq$ be a desirable ranking and let $\rho(\cdot)$ be its induced percentage ranking. For every $\epsilon > 0$ and any $\theta \in (0,1]$*

$$\frac{1}{n}|\{c \in C|\theta_c \geq \theta \textit{ and } \rho(c) \leq \theta - \epsilon\}| \xrightarrow{p} 0 \textit{ as } n \to \infty.$$

Proposition 2. *Let $\trianglerighteq$ be a desirable ranking and let $\rho(\cdot)$ be its induced percentage ranking. For every $\epsilon > 0$ and any $\theta \in (0,1]$*

$$\frac{1}{n}|\{c \in C|\theta_c \leq \theta \textit{ and } \rho(c) \geq \theta + \epsilon\}| \xrightarrow{p} 0 \textit{ as } n \to \infty.$$

Proof of Proposition 1. Fix any $\theta \in (0,1]$ and any $\epsilon > 0$. Choose θ' so that $\theta > \theta' > \theta - \min\{\epsilon, \frac{1-\alpha}{\alpha}\}$.

For any market realization and induced ranking ρ, define the set of colleges that "lose" in ranking ρ by

$$L := \{c \in C|\theta_c \geq \theta \text{ and } \rho(c) \leq \theta - \epsilon\}.$$

These colleges "lose" because their quality θ_c is above θ, but their resulting ranking is more than ϵ below. Next, define the following set of colleges:

$$W := \{c \in C|\theta_c < \theta' \text{ and } \rho(c) > \theta - \epsilon\}.$$

We label this set W because these colleges are "wrongly ranked", in the sense that their true value θ_c is less than the colleges in L, and yet they are ranked higher than any of the colleges in L. As a reminder, $\mu^*(i)$ denotes i's favorite college that is in the same tier as $\mu(i)$. We define a set of "unusual" students as follows:

$$U := \{i|\mu^*(i) \in W\}.$$

These students are unusual in the sense that the colleges in W are overranked, and thus have relatively lower common values, and yet each of these students chose a college in W as her favorite college in her tier.

Claim 1. *For every college $c \in L$ and every student $i \in U$, $\mu^*(i)\ P_i\ c$.*

This follows from definitions of W, L, and AoD. Since $i \in U$, $\mu^*(i) \in W$. Therefore, $\rho(\mu^*(i)) > \rho(c)$ (from the definitions of W and L), and by AoD, $\mu^*(i)\ P_i\ c$.

In fact, we can place an upper bound on the idiosyncratic draw between an unusual student and a loser college.

Claim 2. *For every college $c \in L$ and every student $i \in U$, $\eta_{i,c} < 1 - \frac{\alpha}{1-\alpha}(\theta - \theta')$.*[22]

To see Claim 2, recall that $\theta_c > \theta$ (by the definition of L). By definition of U, $\mu^*(i) \in W$ and therefore, $\theta' > \theta_{\mu^*(i)}$, by definition of W. Since $\theta' > \theta_{\mu^*(i)}$ and $1 > \eta_{i,\mu^*(i)}$

$$\begin{aligned}
\alpha\theta' + (1-\alpha) &> \alpha\theta_{\mu^*(i)} + (1-\alpha)\eta_{i,\mu^*(i)} \\
&> \alpha\theta_c + (1-\alpha)\eta_{i,c} \\
&> \alpha\theta + (1-\alpha)\eta_{i,c}
\end{aligned}$$

The second line follows since $\mu^*(i)\ P_i\ c$ (Claim 1). The last line follows by construction: All colleges in L have a common value greater than θ. Thus,

$$\alpha\theta' + (1-\alpha) > \alpha\theta + (1-\alpha)\eta_{i,c},$$

and rearranging terms we find that

$$1 - \frac{\alpha}{1-\alpha}(\theta - \theta') > \eta_{i,c},$$

which is Claim 2.

We once again rely on random bipartite graphs, and show that our "unusual" occurrence corresponds to the formation of a balanced biclique in a random graph. Draw an edge between student i and college c if

$$\eta_{i,c} < 1 - \frac{\alpha}{1-\alpha}(\theta - \theta'). \tag{15}$$

[22] Recall that θ' is chosen such that $\theta' > \theta - \min\{\epsilon, \frac{1-\alpha}{\alpha}\}$, which implies that $1 - \frac{\alpha}{1-\alpha}(\theta - \theta') > 0$.

That is, every pair is joined by an edge independently with probability $p' = 1 - \frac{\alpha}{1-\alpha}(\theta - \theta')$,[23] and we denote the resulting random graph $G^{p'}$.

Claim 3. *The vertices L and U form a biclique in $G^{1-p'}$.*

This follows immediately from Claim 2 and equation (15). Thus, in the random graph, there exists a balanced biclique of size $\min\{|L|, |U|\}$. Because $\tilde{\mu}$ is a bijection between U and W, we have $|U| = |W|$, and so there exists a balanced bicliique of size $\min\{|L|, |W|\}$. By Theorem 3,[24]

$$\min\left\{\frac{|L|}{n}, \frac{|W|}{n}\right\} \xrightarrow{p} 0. \tag{16}$$

Finally, we have the following claim.

Claim 4. $|W|/n \xrightarrow{p} b > 0$.

Claim 4 is sufficient to prove Proposition 1, because if $|W|/n$ converges to something strictly positive, then by (16), we must have $|L|/n \xrightarrow{p} 0$. To show Claim 4, let

$$A = \{c \in C | \rho(c) \geq \theta - \epsilon\}.$$

Notice that by definition, $W \subseteq A$. Let $W^c = A \setminus W$, i.e.,

$$W^c = \{c \in C : \theta_c \geq \theta' \text{ and } \rho(c) \geq \theta - \epsilon\}$$

Since $A = W \cup W^c$ and $W \cap W^c = \emptyset$, we have

$$\frac{|A|}{n} = \frac{|W|}{n} + \frac{|W^c|}{n} \tag{17}$$

By Corollary 1, $|A|/n \xrightarrow{p} 1 - (\theta - \epsilon)$. Define ω by $|W^c|/n \xrightarrow{p} \omega$, and we have

$$\frac{|W|}{n} \xrightarrow{p} 1 - (\theta - \epsilon) - \omega. \tag{18}$$

It follows from the definition of W^c that $\omega \leq 1 - \theta'$, and as $\theta' > \theta - \epsilon$, the RHS of (18) is strictly positive.

$\square$

[23] Given how θ' is chosen, $p' \in (0, 1)$; cf. footnote 22.
[24] The full argument for (16) is analogous to that used to show (10) in the proof of Theorem 2.

Proof of Proposition 2. The proof of Proposition 2 proceeds in an analogous manner to the proof of Proposition 1. Fix $\theta \in (0, 1]$ and $\epsilon > 0$. Choose θ' such that $\theta < \theta' < \theta + \min\{\epsilon, \frac{1-\alpha}{\alpha}\}$. Define the following sets:

$$\begin{aligned} W &:= \{c \in C | \theta_c \leq \theta \text{ and } \rho(c) \geq \theta + \epsilon\} \\ L &:= \{c \in C | \theta_c > \theta' \text{ and } \rho(c) < \theta + \epsilon\} \\ U &:= \{i \in I : \mu^*(i) \in W\} \end{aligned}$$

Arguments analogous to Claims 1-3 in the proof of Proposition 1 deliver the same conclusion as equation (16):

$$\min\left\{\frac{|L|}{n}, \frac{|W|}{n}\right\} \xrightarrow{p} 0.$$

We now show that in this case, it is $|L|/n$ that converges to something strictly positive, which will imply that $|W|/n \xrightarrow{p} 0$ and complete the proof.

Claim 5. $|L|/n \xrightarrow{p} b > 0$.

To see this claim, define $A = \{c \in C | \rho(c) < \theta + \epsilon\}$ and $L^c = \{c \in C | \theta_c \leq \theta' \text{ and } \rho(c) < \theta + \epsilon\}$. Notice that $A = L \cup L^c$ and $L \cap L^c = \emptyset$, and therefore,

$$\frac{|A|}{n} = \frac{|L|}{n} + \frac{|L^c|}{n}.$$

By Corollary 1, $|A|/n \xrightarrow{p} \theta + \epsilon$, and thus,

$$\frac{|L|}{n} \xrightarrow{p} \theta + \epsilon - \beta,$$

where $\beta \xleftarrow{p} |L^c|/n$. It is clear from the definition of L^c that $\beta \leq \theta'$, and, as $\theta' < \theta + \epsilon$, we have $\theta + \epsilon - \beta > 0$, which completes the proof. $\square$

Finally, we use Propositions 1 and 2 to complete the proof of Theorem 1. Recall that $C^\rho(\epsilon)$ is the set of schools whose percentile ranking, $\rho(c)$, is within ϵ of their common values, θ_c, i.e., $\theta_c - \epsilon < \rho(c) < \theta_c + \epsilon$. We show that, for any $\epsilon, \delta > 0$,

$$\Pr\left(\frac{|C \setminus C^\rho(\epsilon)|}{n} > \delta\right) \to 0 \text{ as } n \to \infty \tag{19}$$

Choose a positive integer M such that $1/M < \min\{\delta, \epsilon/2\}$. Now, we have the following:

$$\begin{aligned}\frac{|C \setminus C^{\rho}(\epsilon)|}{n} \quad \leq \quad & \sum_{m=1}^{M} \frac{1}{n}|\{c \in C : \theta_c \geq m/M \text{ and } \rho(c) \leq m/M - \epsilon/2\}| + \\ & \sum_{m=1}^{M} \frac{1}{n}|\{c \in C : \theta_c \leq m/M \text{ and } \rho(c) \geq m/M + \epsilon/2\}| + \\ & \frac{1}{n}|\{c \in C : \theta_c \leq 1/M\}|. \end{aligned} \tag{20}$$

To see this, take a college $c \in C \setminus C^{\rho}(\epsilon)$. We claim that c belongs to one of the three terms on the right-hand side of (20). If $\theta_c < 1/M$, this is obvious. If not, then there is some $m_c \in \{1, 2, \dots, M-1\}$ such that $\theta_c \in [m_c/M, (m_c+1)/M]$. Since $c \notin C^{\rho}(\epsilon)$, we have that either: (i) $\rho(c) \geq \theta_c + \epsilon$ or (ii) $\rho(c) \leq \theta_c - \epsilon$. If (i) holds, then

$$\rho(c) \geq \theta_c + \epsilon \geq m_c/M + 1/M + \epsilon/2 = (m_c+1)/M + \epsilon/2,$$

where the middle inequality holds because $\theta_c \geq m_c/M$ and we chose M such that $\epsilon/2 > 1/M$. In this case, college c belongs in one of the sets on the second line of (20), namely the summation term for $m = m_c + 1$. If (ii) holds, then

$$\rho(c) \leq \theta_c - \epsilon < m_c/M - \epsilon/2,$$

in which case, school c belongs to one of the sets on the first line of (20), namely the summation term for $m = m_c$.

Therefore, each school c belongs to one of the sets on the right-hand side of (20), and so inequality (20) holds. Now, by Propositions 1 and 2 and the weak law of large numbers, the right-hand side of (20) converges to $1/M$, which is strictly less than δ, which shows equation (19) and completes the proof of Theorem 1.