Correlation of rankings in matching markets*

Rémi Castera†, Patrick Loiseau‡, Bary S.R. Pradelski§

March 2025 (First version: May 2022)

We study the role of correlation in matching markets, where multiple decision-makers simultaneously face selection problems from the same pool of candidates. We propose a model in which a candidate's priority scores across different decision-makers exhibit varying levels of correlation dependent on the candidate's sociodemographic group. Such *differential correlation* can arise in school choice due to the varying prevalence of selection criteria, in college admissions due to test-optional policies, or due to algorithmic monoculture, that is, when decision-makers rely on the same algorithms and data sets to evaluate candidates. We show that higher correlation for one of the groups generally improves the outcome for all groups, leading to higher efficiency. However, students from a given group are more likely to remain unmatched as their own correlation level increases. This implies that it is advantageous to belong to a low-correlation group. Finally, we extend the tie-breaking literature to multiple priority classes and intermediate levels of correlation. Overall, our results point to differential correlation as a previously overlooked systemic source of group inequalities in school, university, and job admissions.

---

* This paper was previously circulated under the title "Statistical discrimination in stable matching" and an extended abstract was published in the proceedings of the ACM Conference on Economics and Computation (EC'22).

We would like to thank Nick Arnosti, Itai Ashlagi, Jean-Paul Carvalho, Julien Combe, Vitalii Emelianov, Simon Finster, Jacob Leshno, Ravi Jagadeesan, Simon Jantschgi, Negar Matoorian, Meg Meyer, Faidra Monachou, Thanh Nguyen, Marek Pycia, Bassel Tarbush, Jakob Weissteiner, Julien Fixary, and the participants of various workshops and conferences for their remarks and advice, all of which greatly helped us improving this work. All remaining errors are ours. This work was partially supported by MIAI @ Grenoble Alpes (ANR-19-P3IA-0003), by the French National Research Agency (ANR) through grants ANR-19-CE48-0018 and ANR-20-CE23-0007, by the graduate school MSTII that funded Rémi Castera's PhD, and by the National Science Foundation under Grant No. DMS-1928930 and by the Alfred P. Sloan Foundation under grant G-2021-16778 that funded Bary Pradelski's residency at the Simons Laufer Mathematical Sciences Institute (formerly MSRI) in Berkeley, California, during the Fall 2023 semester.

† Moroccan Center for Game Theory, Université Mohammed VI Polytechnique; remi.castera@protonmail.com

‡ Inria, Fairplay joint team; patrick.loiseau@inria.fr

§ CNRS, Maison Française d'Oxford; Department of Economics, University of Oxford; bary.pradelski@cnrs.fr

1. Introduction

Outcome inequalities for different demographic or social groups are ubiquitous, for example, in college admission, job assignment, or investment allocation. Arcidiacono, Kinsler, and Ransom (2022) find that Asian-American applicants have lower admission chances at Harvard than white applicants with similar academic records, Niessen-Ruenzi and Ruenzi (2019) find significantly lower inflows in female managed mutual funds than in male managed mutual funds, and Bertrand and Mullainathan (2004) find race-based discrimination in callback decisions by job advertisers. Consequently, the sources of observed outcome inequities remain the subject of frequent and ongoing controversy and political debate. A common concern is that the causes of outcome inequalities in matching markets are poorly understood, making them difficult to address (Longhofer 1995).

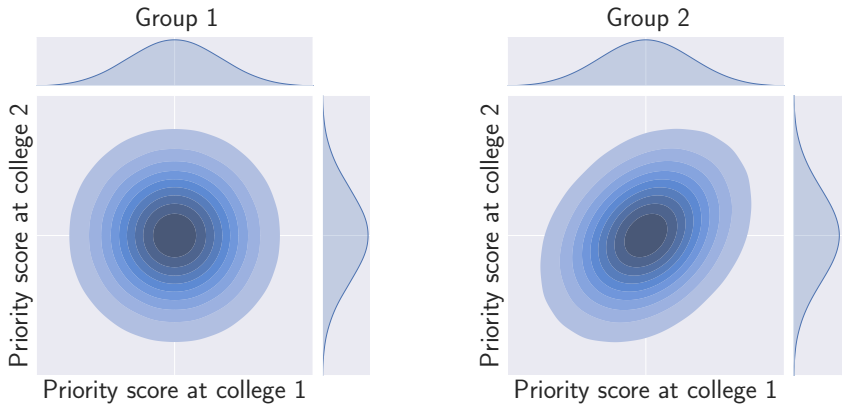
We study how different correlations of priority scores, and thus rankings, between different sociodemographic groups—*differential correlation*—affect outcome inequalities and efficiency in matching markets. Our findings point to a previously unacknowledged source of inequity between different groups that is specific to matching markets. This may be the case, as most policies, including those aiming to correct for biases or implement affirmative action, are designed with one decision-maker in mind. In particular, we find that differential correlation across groups leads to outcome inequities even when the rankings by each college are "fair", i.e., all groups are represented at all levels of each college's ranking in the same proportion as in the applicant population. Furthermore, we identify the impact of variations of the correlation of rankings on the efficiency of the resulting matching.

Figure 1 illustrates differential correlation for two groups at two colleges, showing distributions of priority scores, from which ordinal rankings can be derived. The marginal distributions are the same for all students (Gaussian), however, the joint distributions differ due to differential correlation. Looking ahead, we will show that, ceteris paribus, the group with lower correlation is better off.

Differential correlation arises when different decision-makers, such as colleges, use different information on candidates from different (sociodemographic) groups, such as students, when assigning priority scores to rank and admit them. Decision-makers hereby have to rely on observable, partial information, as candidates' latent qualities are unknown to the decision-maker.[1] Differences in information between groups may be due to the varying

[1] The idea to study a model of latent quality plus noise can be traced to Phelps (1972). More recently, this idea has been used by Chade, Lewis, and Smith (2014), Garg, Li, and Monachou (2021), Emelianov, Gast, Gummadi, and Loiseau (2022).

Figure 1 Differential correlation between two groups at two colleges.



Note. The left and right panels represent the priority score distributions for Groups 1 and 2, respectively. The color shading indicates density, with darker shades representing higher probability density. The marginal distributions of priority scores at each college are shown along the top and right edges of each plot. While both groups share the same marginal distributions, their joint distributions differ due to differential correlation, with Group 2 exhibiting higher correlation than Group 1.

cost of information acquisition, the availability of information, or varying prevalence of attributes between different groups.

For concreteness, consider school choice and selection criteria that are more or less prevalent between different groups, e.g., diversity with respect to the current student body, proximity, or sibling priority. The latter is, for example, used in the centralized school choice mechanism in Chile (Correa et al. 2022), where, in 2023, 21.3% students among the economically disadvantaged received sibling priority, while only 15.6% of the remaining students received sibling priority. As a result, a group for which the criterion is more prevalent will, ceteris paribus, exhibit lower correlation than a group for which the criterion is less prevalent (assuming that a student only receives sibling priority at one school).

Differential correlation may also result from test-optional policies for standardized tests, such as the SAT used in the United States for college admissions (cf., e.g., Leonhardt 2024). Standardized testing increases the correlation of assessment results between different colleges. Moreover, if different demographic groups exhibit different participation patterns in standardized tests, then differential correlation arises.

Finally, algorithms are increasingly used to support selection tasks, ranging from bail decisions (cf., e.g., Angwin, Larson, Mattu, and Kirchner 2016) to screening and selection

of candidates for university admission, employment, tenancy, or mortgages (cf., e.g., Citron and Pasquale 2014). The increased use of the same or similar machines, algorithms, or data sets has been termed "algorithmic monoculture" (Goth 2003).[2] Various risks have been associated with algorithmic monoculture, e.g., vulnerability to flaws, hacker attacks, or favoring persistent biases (cf., e.g., Goth 2003, Citron and Pasquale 2014, Kleinberg and Raghavan 2021, Peng and Garg 2024a). Algorithmic monoculture leads to increased correlation between decision-makers' assessments of candidates and can also lead to differential correlation.

Our results identify differential correlation as an overlooked source of inefficiency and inequity in matching markets. Policy-makers and managers who orchestrate matching markets should be aware that only considering and enforcing the *fairness* of each decision-maker separately is insufficient to ensure overall equity (and efficiency). In practice, school choice mechanisms have been frequently redesigned over the past several decades and researchers have actively contributed to these efforts (cf. e.g., Abdulkadiroğlu and Sönmez 2003, Abdulkadiroğlu 2005, Abdulkadiroğlu, Pathak, and Roth 2009, Correa et al. 2022, Kamada and Kojima 2024). According to our results, differential correlation should be accounted for when (re)designing these and other markets. The increasing use of algorithms to facilitate screening and selection tasks may lead to changes in (differential) correlation that should be taken into account; optimistically, new interventions and regulations may be available, for example, against monopolies by software and data providers.

1.1. Our contribution

We study the college admissions problem, where multiple decision-makers select a subset of applicants from an applicant pool with stability as the solution concept (Gale and Shapley 1962, Azevedo and Leshno 2016). Specifically, we suppose that an infinite population of students divided into groups $G_1, \ \ldots, \ G_d$ applies to colleges $C^1, \ \ldots, \ C^m$. The groups represent, for example, protected attributes such as gender or race. Each college assigns a priority score to each student and thus each student receives a priority score at each college. We propose an original model for the distributions of these priority scores to study the correlation between the rankings produced by different colleges. Our model allows for any number of groups, different marginals, and student preferences that depend on the priority scores.

[2] The terminology is borrowed from its use in agriculture.

To formalize correlation and thus capture the vague notion of "a connection between two things in which one thing changes as the other does",[3] we leverage prior work on copulas and their relation with classical notions of correlation via *coherence*. This allows us to model correlation without a specific functional form and, in particular, nest classical notions as special cases, e.g., Spearman's and Kendall's correlation indices. With this, we assume that the correlation between the priority scores at different colleges depends on a candidate's group; we call this feature *differential correlation*.

We investigate two main questions:

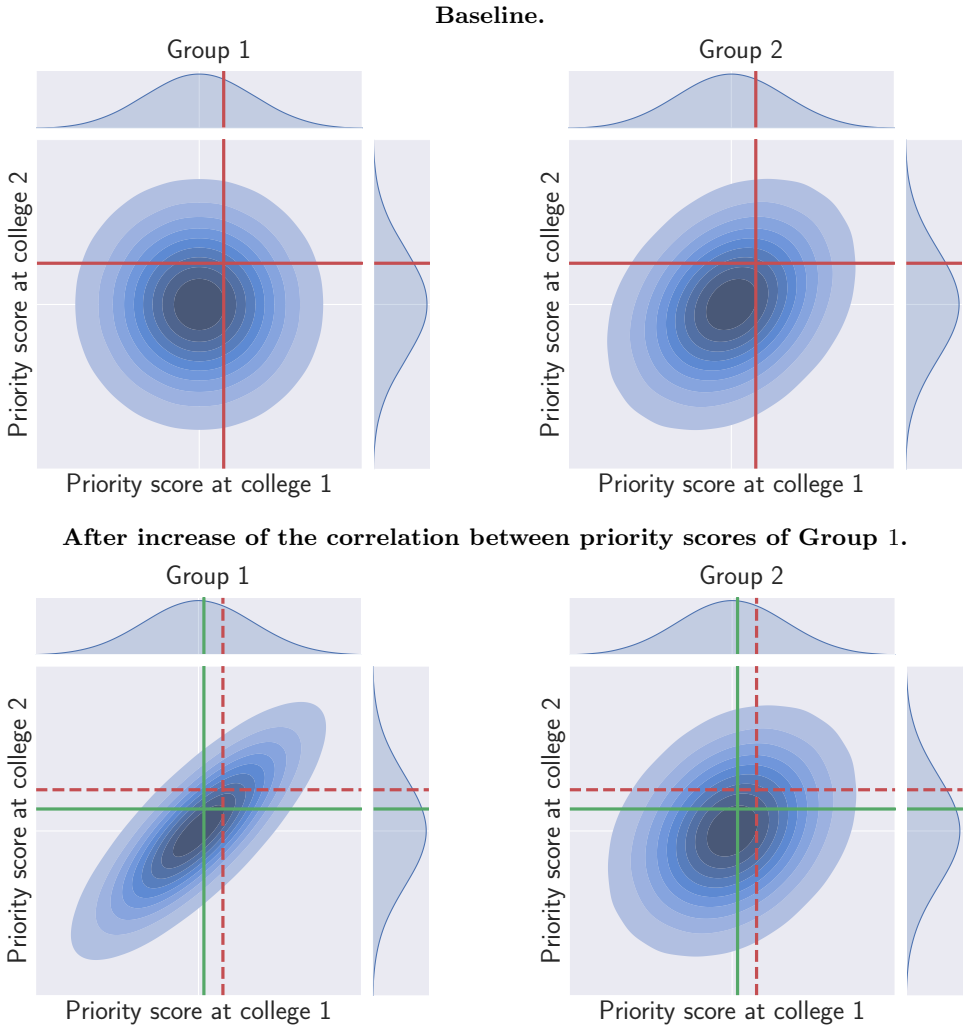
- *How does correlation impact efficiency, i.e., the number of students who obtain their first choice (or one of their top k choices)?*
- *How does differential correlation impact inequality, i.e., the difference between groups in the number of students remaining unmatched?*

First, we show that efficiency increases with each group's correlation level; i.e., increasing the correlation level of any group increases the number of students who obtain their first choice in all groups, and also increases the number of students who obtain one of their top k choices, for all k, in all groups except the one whose correlation increased (Theorem 1). Intuitively, this is the case because increasing correlation leads to a decrease in admission cutoffs, otherwise the number of unmatched students would increase. As a consequence, it becomes easier for students to be assigned to one of their preferred colleges. However, the benefit is countered for the group whose correlation increased, since their probability of being below the cutoff of several colleges also increases. Figure 2 illustrates the impact of increasing correlation on the cutoffs for two colleges and two groups.

Considering inequality, we show that the proportion of students in a given group who remain unmatched increases in its own correlation level and decreases in the correlation level of all other groups (Theorem 2). This implies that it is advantageous to belong to a low-correlation group. Intuitively, this is the case because an "independent second chance" is preferable to "carrying over the bad signal from previous rejection". However, as stated before, low correlation also leads to higher admission cutoffs. From an individual's viewpoint, to benefit from both effects, it is best to have a low correlation in their own priority scores, while everyone else's priority scores exhibit high correlation.

[3] Oxford Advanced Learner's Dictionary, 2023.

Figure 2 Illustration of the effect of correlation increase on cutoffs and efficiency.



Note. The left and right columns represent the priority score distributions for Groups 1 and 2, respectively. The top row shows the baseline scenario, while the bottom row illustrates the effect of increasing correlation for Group 1. The color shading indicates density, with darker shades representing higher probability density. The red lines indicate the initial cutoffs (shown dashed for reference in the bottom row), while the green lines show the updated cutoffs after increasing correlation. As correlation increases, cutoffs decrease, allowing more students to be matched to their first choice, as can be seen from the plots of the marginal distributions (top and right of each plot). The impact differs between groups since Group 1's priority score distribution changed while Group 2's remained the same; for the former less students are above either cutoff, while for the latter the overall change is entirely beneficial.

Next, we show that a given efficiency level can be reached by a continuum of different correlation vectors, generating different levels of inequality; this shows that differential

correlation has a distinct impact on efficiency on the one hand and on inequality on the other (Proposition 4).

Finally, our results imply extensions of known results on tie-breaking (cf. Ashlagi, Nikzad, and Romm 2019, Ashlagi and Nikzad 2020, and Arnosti 2023), particularly for multiple priority classes and intermediate levels of correlation (Proposition 5).

All the aforementioned results hold under the assumption that admission cutoffs decrease in the correlation parameters. We show that the assumption always holds for two colleges (Theorem 3). Moreover, under standard symmetry conditions from the literature, the assumption holds for any number of colleges (Proposition 6). Finally, we show through simulations that our results remain qualitatively valid across the parameter space.

1.2. Outline

The remainder of the article is organized as follows. Section 2 introduces the model and the concept of differential correlation. Section 3 introduces our welfare metrics and presents preliminary results. Our main results are presented in Section 4: we first derive comparative statics regarding the impact of correlation on our metrics, under the assumption that admission cutoffs are decreasing in the correlation parameters (Section 4.1), then apply them to the tie-breaking problem in Section 4.2, and, in Section 4.3, study the validity of the decreasing cutoffs assumption. Finally, in Section 5, we provide an extended overview of the related work and Section 6 concludes.

2. Setup

We introduce the college admissions problem with a *continuum* of students and *correlated priority scores*, formalize the notion of correlation, and introduce the supply and demand framework to identify stable matchings. A table of notation is provided in Appendix A.1.1 for convenience.

2.1. Model

A finite set of m colleges $\mathcal{C} = \{C^1, \dots, C^m\}$ and a continuous set of students $\mathcal{S}$ with unit mass, endowed with a measure η, need to be matched.[4] Colleges have respective capacities $(\alpha^1, \dots, \alpha^m) := \boldsymbol{\alpha} \in (0,1)^m$, which represent the maximal mass of students each can admit, and we assume that capacity is constrained, i.e., $\sum_{i \in [m]} \alpha^i < 1$, using the notation $[k] :=$

[4] The formal definition of the measure is deferred to Appendix A.1.2.

$\{1,2,\dots,k\}$ for $k \in \mathbb{N}$. The set of students $\mathcal{S}$ is divided into d groups $G_1,\dots,G_d$, with a mass $\gamma_j \in [0,1]$ of students belonging to G_j, such that $\sum_{j=1}^{d} \gamma_j = 1$ and each student belongs to exactly one group. Define the vector $\boldsymbol{\gamma} \coloneqq (\gamma_j)_{j\in[d]}$. Let $G(s)$ be the group student $s \in \mathcal{S}$ belongs to.

Each student has strict preferences over colleges: when student s prefers college C to college C', we write $C \succ_s C'$. Preferences can also be represented by a permutation $\sigma \in \Sigma([m])$, where $\Sigma([m])$ is the set of all permutations of m elements, and $\sigma(1)$ represents the favorite college and $\sigma(m)$ the least favorite college. In group G_j a share $\beta_j^\sigma \in (0,1)$ of students has preference list σ (with the share thus depending on the group).[5] We assume that each preference list is used by a non-zero mass of students, and that all students prefer attending some college to remaining unmatched. Let $\boldsymbol{\beta}_j = (\beta_j^\sigma)_{\sigma\in\Sigma([m])}$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_j)_{j\in[d]}$.

Each college assigns a priority score to each student, and the higher the priority score, the better the student's evaluation. Thus, each student s is assigned a vector of priority scores $(W_s^1,\dots,W_s^m)$. College C^i prefers $s \in \mathcal{S}$ to s' if and only if $W_s^i > W_{s'}^i$. The (marginal) distribution of priority scores W^i given by college C^i to students in G_j is described by a probability density function[6] (pdf) f_j^i defined on the support $I_j^i \subseteq \mathbb{R}$, assumed to be an interval $(\underline{I}_j^i, \bar{I}_j^i)$.[7] Let $I_j = \prod_{i=1}^{m} I_j^i$. Denote by F_j^i the cumulative distribution function (cdf) associated with f_j^i. Let $\mathbf{f} \coloneqq (f_j^i)_{j\in[d]}^{i\in[m]}$ be the set that contains all the information on the marginal priority score distributions.

Differential correlation. Consider the joint distribution of the vectors $(W_s^1,\dots,W_s^m)$. A joint distribution can be characterized by its marginals, i.e., the distribution of each component of the vector, and the shape of the joint distribution, captured by a coupling function, called *copula*. A copula is a cdf over $[0,1]^m$ with uniform marginals. Sklar (1959) shows that any joint distribution can be decomposed into (independent) marginals and a unique copula.

Sklar (1959, Theorems 1, 2, and 3) *Let F be an m-dimensional cdf with marginal cdfs $F^1,\dots,F^m$. Then there exists a unique m-dimensional copula $H : [0,1]^m \to [0,1]$ such that*

$$F(x^1,\dots,x^m) = H(F^1(x^1),\dots,F^m(x^m)).$$

[5] Notice that β_j^σ is a share that is conditional on the group, and not a mass: $\eta(\{s \in G_j : \sigma_s = \sigma\}) = \gamma_j \beta_j^\sigma$.

[6] This implies that ties happen with probability zero.

[7] The bounds of this interval can be finite or not.

Conversely, for any m-dimensional copula H and any set of 1-dimensional cdfs $F^1, \dots, F^m$, $F(x^1, \dots, x^m) \coloneqq H(F^1(x^1), \dots, F^m(x^m))$ is an m-dimensional cdf with marginals $F^1, \dots, F^m$.

The priority score vector of each group G_j can then be represented by its marginals and a unique copula H_j. We consider a family of m-dimensional copulas $(H_\theta)_{\theta \in \Theta}$ (Θ being an interval of $\mathbb{R}$) such that, for all $j \in [d]$, there exists a parameter $\theta_j \in \Theta$ such that H_{θ_j} is the copula associated with G_j's distribution; i.e., $H_{\theta_j} = H_j$ (such a family always exists by Sklar 1959). This feature of the model allows us to vary the shape of the joint distribution without changing the marginals; this is crucial to model correlation and isolate its effects. We see this as our principal modeling contribution.

Figure 3 provides an example, showing the Gaussian copula family with covariance used as parameter θ.

Denote by $\boldsymbol{\theta} \coloneqq (\theta_j)_{j \in [d]}$ the vector that contains the parameter for each group. Note that each group has a different θ_j and thus a different joint distribution, in addition to possibly different marginals. With some foresight to the explanations provided in Section 2.2, we call this feature of the model *differential correlation*.[8] Finally, we assume that all copulas in $(H_\theta)_{\theta \in \Theta}$ have full support over $[0,1]^m$. We write f_{j,θ_j} and F_{j,θ_j} for the joint multivariate pdf and cdf of group G_j.

Given a family of copulas $(H_\theta)_{\theta \in \Theta}$, we refer to the tuple $(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{f}, \boldsymbol{\theta})$ as a *college admissions problem*. Notice that so far, we only assume that distributions admit a density and have full support. This model allows each group to have different marginal distributions at each college. We implicitly assumed that the priority scores were independent of preferences within each group, which only causes a very limited loss of generality.[9]

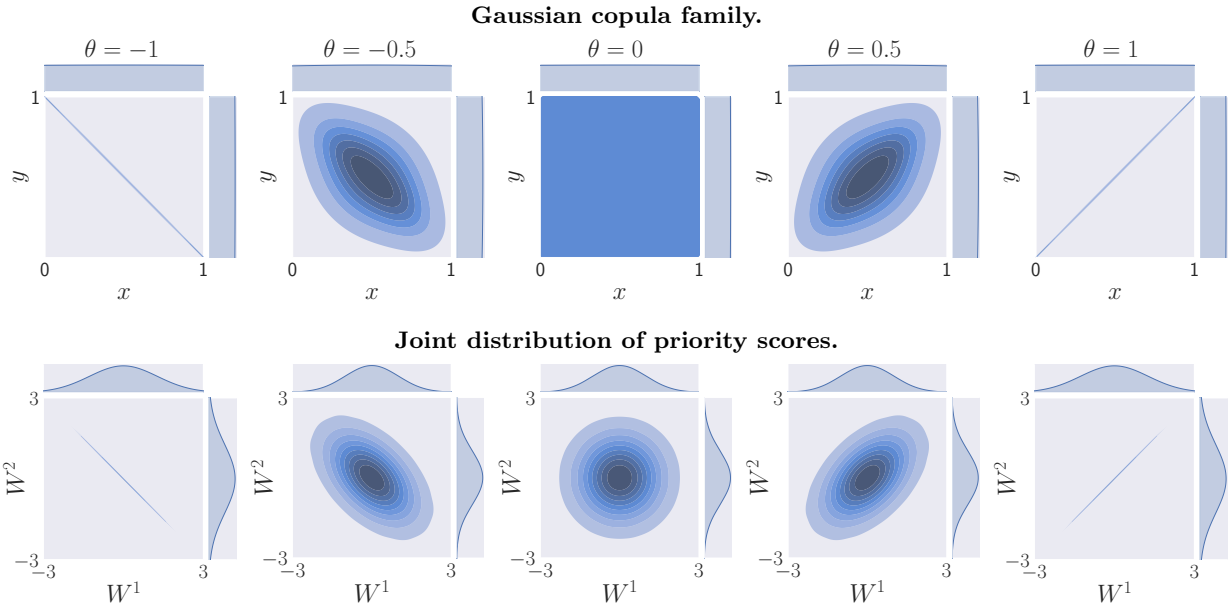
2.2. Correlation and coherence

The purpose of our model is to use θ as a proxy for correlation of rankings, rather than some specific functional form. However, all classical measures of correlation are defined for couples of variables, and there is no standard generalization of the notion of correlation to vectors of higher dimension. We employ the notion of *coherence*, that ensures that θ is

[8] This is in the spirit of the notion of differential variance studied in Emelianov, Gast, Gummadi, and Loiseau (2022) and Garg, Li, and Monachou (2021).

[9] If the independence assumption is not verified, we can recover it as long as there exist a finite subdivision of the groups such that preferences and priority scores are independent inside of each newly defined subgroup.

Figure 3 Gaussian copula and joint distributions for two colleges.



Note. The columns represent the priority score distributions for different correlation levels. The x-axis represents the priority scores assigned by college C^1, and the y-axis represents the priority scores assigned by college C^2. The top row illustrates the Gaussian copula for five different correlation levels θ, demonstrating how correlation influences the dependence structure between priority scores. The bottom row presents the corresponding joint distributions of priority scores, assuming standard Gaussian marginals for both colleges. The color shading indicates density, with darker shades representing higher probability density. As correlation increases, the distribution shifts toward alignment along the diagonal (higher correlation) or anti-diagonal (negative correlation), meaning that a student's ranking at one college becomes more predictive of their ranking at another. The marginal distributions of priority scores at each college are shown along the top and right edges of each plot.

aligned with classical measures of correlation for two colleges ($m = 2$), and still encompasses the intuitive idea of correlation for more than two colleges ($m > 2$).

ASSUMPTION 1 **(Coherence)**. *We say that the family of copulas* $(H_\theta)_{\theta\in\Theta}$ *is* coherent *if for all* $x \in (0,1)^m$, $H_\theta(x)$ *is increasing in* θ *on* Θ.

When $m = 2$, coherence induces an equivalence between θ and two classical ordinal correlation measures; using Scarsini (1984), we can show that Spearman's and Kendall's correlation coefficients are increasing functions of θ (cf. Appendix A.2 for a further discussion). The Gaussian copula depicted in Figure 3 is an example of a coherent family. When $\theta = 0$, the variables are independent; when θ is positive, the joint distribution gets

closer to the diagonal $x = y$; and when θ is negative, the joint distribution gets closer to the diagonal $y = -x$.

Finally, we introduce a technical assumption that will be required for some of our results, especially when considering comparative statics in θ. Denote by $\mathring{\Theta}$ the interior of Θ.

ASSUMPTION 2 **(Differentiability)**. *We say that the family of copulas $(H_\theta)_{\theta\in\Theta}$ is* differentiable *if for all $x \in (0,1)^m$ and for all $\theta \in \mathring{\Theta}$, $H_\theta(x)$ is differentiable in the variable θ.*

The coherence and differentiability assumptions are not particularly restrictive; for instance, the Gaussian copula (with covariance chosen as θ) verifies them, and so do other commonly occurring copulas (cf. discussion on distributional assumptions in Appendix A.3). Moreover, our model nests the setting where students are assumed to have an—unobserved—latent quality and decision-makers only observe a noisy measurement (cf. Appendix B.1).

2.3. Stable matching and the supply and demand framework

We now introduce the key elements of matching theory used throughout the paper, following Azevedo and Leshno (2016).[10]

DEFINITION 1 (MATCHINGS, CUTOFFS, AND STABILITY).

- A matching $\mu : \mathcal{S} \to \mathcal{C} \cup \emptyset$ is a mapping associating each student to a college, or to the empty set if they are unmatched, such that $\eta(\mu^{-1}(C^i)) \leq \alpha^i$ for all $i \in [m]$.
- For a matching μ, for all $i \in [m]$, let $P^i := \inf\{W_s^i : \mu(s) = C^i\}$, be the *cutoff* of college C^i.
- A matching μ is *stable* if each student s is matched to their favorite college among those for which their priority score is above the respective cutoff, i.e., $W_s^i \geq P^i$.

Given a cutoff vector $\mathbf{P} = (P^1, \dots, P^m)$, let $D_s(\mathbf{P}) \in \mathcal{C} \cup \emptyset$ be the *demand* of student s, i.e., their favorite college among those where their priority score is above the respective cutoff (or the empty set if there is none). The *aggregate demand* at college C^i is the mass of students demanding it: $D^i(\mathbf{P}) = \eta(\{s : D_s(\mathbf{P}) = C^i\})$. Finding a stable matching then amounts to finding cutoffs such that demand is equal to supply, the latter being the capacity of each college.

[10] For further details on matching theory and the deferred acceptance algorithm, see Appendix A.5.

DEFINITION 2 (MARKET-CLEARING). The cutoff vector $\mathbf{P}$ is *market-clearing* if for $i \in [m]$, $D^i(\mathbf{P}) \leq \alpha^i$, with equality if $P^i > \min_{j\in[d]} \underline{I}^i_j$.

A cutoff vector is therefore market-clearing if it induces a demand that is equal to colleges' capacities when they reach their capacity constraint and lower for colleges that are not full ($P^i = \min_{j\in[d]} \underline{I}^i_j$ means that C^i rejects no one, and therefore is not full). When the constraint is reached in all colleges, i.e., when $\sum_{i\in[m]} \alpha^i < 1$, the system

$$\mathbf{D}(\mathbf{P}) = \boldsymbol{\alpha} \tag{1}$$

is called the *market-clearing equation*, and the corresponding *market-clearing cutoffs* can be computed by solving the equation.

The following results from Azevedo and Leshno (2016) establish the link between market-clearing cutoffs and stable matchings (for an illustration, see Appendix A.4, Figure 7).

Azevedo and Leshno (2016, Lemma 1, Theorem 1).[11] *For any college admissions problem* $(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{f}, \boldsymbol{\theta})$*:*

1. *A matching* μ *is stable if and only if the associated cutoff vector* $\boldsymbol{P}$ *is market-clearing.*
2. *There exists a unique stable matching.*

This result allows us to study stable matchings in the college admissions problem through the unique cutoff vector $\mathbf{P}(\boldsymbol{\theta})$ associated with it.[12] We study $\mathbf{P}$ as a function of $\boldsymbol{\theta}$ since we assume all other parameters to be fixed.

3. Welfare metrics and preliminary results

In selection problems, inequalities between groups are measured by the proportion of admitted candidates in each group. In a matching setting, the situation is more complex; the proportion of students who get one of their top k choices is a relevant metric for all $k \in [m]$, and some group might be advantaged compared to another for some k, while being disadvantaged for another k'.

[11] Note that the original theorem specifies conditions on the distribution of students' types, such as being continuous and having full support, which hold in our definition of a college admissions problem. The result stated here is thus a special case of the theorem as stated in Azevedo and Leshno (2016).

[12] Azevedo and Leshno (2016) further show that the stable matching varies continuously in the parameters of the problem and that the set of stable matchings from a college admissions problem with a finite number of students converges to the unique stable matching of the continuum problem with the same parameters. The latter result justifies the approximation of large finite instances by their limit. For a better approximation of markets with a small number of students, Arnosti (2022) proposes a related framework.

DEFINITION 3 (RANK FUNCTIONS). Let $\boldsymbol{\theta}$ be a correlation vector. In the matching induced by $\boldsymbol{\theta}$, let $\text{Rank}(s)$ be the rank of the match of student s in their preference list, e.g., $\text{Rank}(s) = 1$ means that s got their first choice. For any $k \in [m]$, let

$$R_j^{k,\sigma}(\boldsymbol{\theta}) \coloneqq \mathbb{P}(\text{Rank}(s) \leq k | s \in G_j, \sigma_s = \sigma) = \frac{\eta(\{s \in \mathcal{S} : \text{Rank}(s) \leq k, s \in G_j, \sigma_s = \sigma\})}{\gamma_j \beta_j^\sigma} \tag{2}$$

be the proportion of students from group G_j with preferences σ who get one of their top k choices.

For example, $R_j^{1,\sigma}(\boldsymbol{\theta})$ is the relative mass of students who obtain their first choice among those in group G_j who have preferences σ, or equivalently, it is the probability of a randomly drawn student obtaining their first choice conditionally on belonging to G_j and having preferences σ. Note that we condition on preferences to isolate the effect of differential correlation from differences in preferences. The rank metrics can be easily computed from the cutoff vector, as shown in Appendix A.5.

PROPOSITION 1. *If two groups, G_j, G_ℓ, have the same marginal distribution at college C^i, then for students whose first choice is college C^i, the probability of obtaining this college is the same for both groups, i.e., $R_j^{1,\sigma} = R_\ell^{1,\sigma}$.*

The proof is provided in Appendix C.1.[13] Proposition 1, though simple, is an important property of the model. If two students prefer the same college, then their probabilities of obtaining it only depend on their respective groups' marginals and not on their correlations; thus, differential correlation has no effect on this metric.

Another metric of interest is the proportion of students who remain unmatched, which is the complement of the proportion of students getting one of their top m choices:

$$R_j^{\emptyset,\sigma}(\boldsymbol{\theta}) \coloneqq \mathbb{P}(\mu(s) = \emptyset | s \in G_j, \sigma_s = \sigma) = 1 - R_j^{m,\sigma}(\boldsymbol{\theta}) \tag{3}$$

We can derive a simple yet important result.

PROPOSITION 2. *The probability that a student remains unmatched depends only on their group and is independent of their preferences, i.e., for all j, σ, σ', $R_j^{\emptyset,\sigma}(\boldsymbol{\theta}) = R_j^{\emptyset,\sigma'}(\boldsymbol{\theta})$. Moreover, the total mass of unmatched students is constant in the correlation of any group and is equal to $1 - \sum_{i \in [m]} \alpha^i$.*

[13] Note that Proposition 1 remains valid when capacity is not constrained, i.e., $\sum_{i \in [m]} \alpha^i \geq 1$.

The proof is provided in Appendix C.2.[14] With Proposition 2 at hand, we use the notation $R_j^{\emptyset}$ since these quantities do not depend on the preferences of the students. For all the metrics we defined, when there is no ambiguity, we also omit the dependence on $\boldsymbol{\theta}$ and write $R_j^{k,\sigma}$ and $R_j^{\emptyset}$ instead.

We next define two global metrics, i.e., metrics that are not conditioned on the groups and preferences of students.

Definition 4 (Efficiency and Inequality). For $\boldsymbol{\theta} \in \Theta^d$ and $\mu_{\boldsymbol{\theta}}$ the induced stable matching, define the *efficiency* $E(\boldsymbol{\theta})$ of a matching as the proportion of students obtaining their first choice, and the *inequality* $L_{j,\ell}(\boldsymbol{\theta})$ between two groups, $j, \ell \in [d]$ as the difference in the probability of remaining unmatched between those two groups:

$$E(\boldsymbol{\theta}) = \eta\left(\{s \in \mathcal{S} : \mu_{\boldsymbol{\theta}}(s) = C^{\sigma_s(1)}\}\right) \tag{4}$$

$$L_{j,\ell}(\boldsymbol{\theta}) = |R_j^{\emptyset}(\boldsymbol{\theta}) - R_\ell^{\emptyset}(\boldsymbol{\theta})|. \tag{5}$$

We sometimes focus on these metrics as they represent the two extremes, getting one's first choice or not getting admitted at any college, and are thus arguably of particular importance. Also, note that when there are only two colleges, efficiency and inequality fully capture all rank metrics defined above.

Before we turn to our main results concerning the impact of correlation, we characterize the impact of marginal distributions on our metrics.

Proposition 3. *Suppose the priority score distributions of group G_j first-order stochastically dominate those of group G_ℓ at all colleges*[15] *and correlations are equal ($\theta_j = \theta_\ell$). Then, group G_j is favored, that is, $R_j^{k,\sigma} > R_\ell^{k,\sigma}$ for all $k \in [m]$ and all $\sigma \in \Sigma([m])$ and there is positive inequality, $L_{j,\ell} > 0$.*

The proof is provided in Appendix C.3. This proposition shows that differences in the marginal distributions create inequalities between students in the probability of being matched, but also in the probability of getting a top k choice, for all k. Marginal distributions impose a baseline level of inequality that will be increased or mitigated by differential correlation, as we shall see in the following section.

[14] Proposition 2 relies on the assumption that students prefer any college over staying unmatched, otherwise students with a short preference list would have a higher probability of staying unmatched than those with a long preference list.

[15] This means that $F_j^i < F_\ell^i$ for all $i \in [m]$.

4. Main results

We consider college admissions problems where the group sizes ($\boldsymbol{\gamma}$), the colleges' capacities ($\boldsymbol{\alpha}$), the students' preferences ($\boldsymbol{\beta}$), and the marginal priority score distributions ($\mathbf{f}$) are fixed and study the impact of the correlation, $\boldsymbol{\theta}$, on the stable matching. Section 4.1 contains general comparative statics, and Section 4.2 examines tie-breaking. We assume that Assumptions 1 (coherence) and 2 (differentiability) hold. Finally, the results presented here rely on an additional, crucial assumption.[16]

ASSUMPTION 3 **(Decreasing cutoffs)**. *For fixed parameters $\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{f}$, we say that* decreasing cutoffs *holds if for all $\boldsymbol{\theta} \in \Theta^d$, for all $i \in [m]$, for all $j \in [d]$, the cutoff $P^i(\boldsymbol{\theta})$ decreases when θ_j increases.*

We show in Section 4.3 that this assumption always holds for two colleges and holds for more than two colleges under common symmetry assumptions on preferences and capacities from the literature. Moreover, we show through simulations that our results remain qualitatively valid across the parameter space.

4.1. Comparative statics

We first consider how the efficiency of the matching varies when changing the correlation for one group.

THEOREM 1 **(Efficiency is increasing in correlation)**. *For all $k \in [m]$ and all preferences σ, for $j \in [d]$,*

1. *the proportion of students in group G_j with preferences σ who obtain one of their top k choices, $R_j^{k,\sigma}$, is increasing in all the other groups' correlations θ_ℓ for $\ell \neq j$, and*
2. *the proportion of students in group G_j with preferences σ who obtain their first choice, $R_j^{1,\sigma}$, is also increasing in G_j's own correlation θ_j.*

Consequently, the global efficiency $E(\boldsymbol{\theta})$ is increasing in all components of $\boldsymbol{\theta}$.

Proof. Lemma 4 in Appendix A.6 states that $\mathbf{P}(\boldsymbol{\theta})$ is of class $\mathcal{C}^1$.[17] Moreover, under Assumption 3 the cutoffs are decreasing in all θ_j. Since $R_j^{k,\sigma} = \mathbb{P}_{j,\theta_j}(W^{\sigma(1)} \geq P^{\sigma(1)} \cup \cdots \cup W^{\sigma(k)} \geq P^{\sigma(k)})$[18], if θ_ℓ increases for $\ell \neq j$, the distribution $\mathbb{P}_{j,\theta_j}$ does not change but all

[16] Recall also that we assume that capacity is constrained, which is needed for our results. See Appendix B.2 for a result on unconstrained capacity.

[17] This is not necessarily the case if $\boldsymbol{\theta}$ is not in the interior of Θ^d. However, since the $R_j^{k,\sigma}$ are continuous, they are still increasing over the whole interval Θ.

[18] The notation $\mathbb{P}_{j,\theta_j}$ denotes the probability distribution of priority scores of group G_j, with correlation θ_j.

cutoffs decrease, therefore $R_j^{k,\sigma}$ increases. If θ_j increases, since $R_j^{1,\sigma} = \mathbb{P}_j(W^{\sigma(1)} \geq P^{\sigma(1)})$, we can conclude with the same argument. Moreover, E is a convex combination of all the $R_j^{1,\sigma}$ and therefore is also increasing in all θ_ℓ. □

Theorem 1 implies that generally the outcome improves as correlation increases. In particular, students from a given group always benefit from an increase in correlation in other groups. However, if one's own group's correlation increases, this is only guaranteed to be beneficial in terms of admittance at one's favorite college, because the mass of its priority score distribution accumulates below all cutoffs. This can be observed by going back to Figure 2. Finally, note that Theorem 1 also provides insight into the impact of correlation on the efficiency of a matching without groups (i.e., only one group).

We next consider how the inequality of a matching varies when changing the correlation for one group. Recall from Proposition 2 that the probability that a student remains unmatched is independent of their preferences, and thus we consider $R_j^\emptyset$.

THEOREM 2 **(Low-correlation groups are advantaged)**. *For $j \in [d]$, the proportion of students in group G_j remaining unmatched, $R_j^\emptyset$,*

1. *is decreasing in all θ_ℓ, for $\ell \neq j$, and*
2. *is increasing in G_j's own correlation, θ_j.*

Consequently, the inequality between any two groups $\ell \neq j$, $L_{j,\ell}(\boldsymbol{\theta})$, decreases in the correlation of the group with the lower rate of unmatched students and increases in the correlation of the other group.

Proof sketch. Fix $\ell \in [d]$. We show that the $R_j^\emptyset$ are increasing in θ_ℓ for $j \neq \ell$ using Assumption 3 (decreasing cutoffs). We then use the fact that total capacity is constant to deduce that $R_\ell^\emptyset$ has to be decreasing in θ_ℓ to compensate for the others. The complete proof is provided in Appendix C.4.

Theorem 2 implies that, ceteris paribus, a group's students are more likely to be matched when having low correlation. This is the case because an "independent second chance" is preferable to "carrying over the bad signal from previous rejections". Since a given group's students are more likely to be matched as their correlation decreases, and capacity is limited, it follows that all other students become less likely to be matched. Consequently, inequality decreases when the correlation of the group with the lower proportion of unassigned students (that is, the better-off group) increases or when the correlation of the group with the higher proportion of unassigned students (that is, the worse-off group) decreases.

Note that the effect of one's own group's correlation on obtaining a match other than one's favorite college (but not remaining unmatched) is neither covered by Theorems 1 and 2; the effect is not monotone in correlation and depends on the other parameters of the college admissions problem.

Theorem 2 also allows us to formulate a corollary relating the inequality from different marginal distributions to the inequality from differential correlation.

COROLLARY 1. *If groups G_j and G_ℓ have the same marginals, $R_j^\emptyset < R_\ell^\emptyset$ if and only if $\theta_j < \theta_\ell$. If G_j first-order stochastically dominates G_ℓ at all colleges, for any fixed θ_ℓ, there exists $\hat{\theta} > \theta_\ell$ such that $R_j^\emptyset < R_\ell^\emptyset$ if and only if $\theta_j < \hat{\theta}$, i.e., the inequalities from different marginal distributions and from differential correlation are cumulative.*

Proof. If marginal distributions are equal, and $\theta_\ell = \theta_j$ there is no inequality. As θ_ℓ increases the proportion of unmatched students increases for G_ℓ and decreases for G_j by Theorem 2, proving the first part. If G_j first-order stochastically dominates G_ℓ at all colleges, and $\theta_j = \theta_\ell$, Proposition 3 states that the inequality is in favor of G_j. By Theorem 2, a decrease in θ_j further increases inequality, while an increase in θ_j decreases the inequality until it reaches zero (if it does), then the inequality is reverted. □

Corollary 1 shows that the proportion of unmatched students is higher in the higher-correlation group, even with identical marginals. This proves the existence of inequalities specific to matching problems. Indeed, even when the rankings of each college are "fair", i.e., all groups are represented at all levels of each college's ranking in the same proportion as in the applicant population (which is equivalent to saying that all groups have the same marginals), inequality exists due to differential correlation. This finding contrasts with Proposition 1, which states that with identical marginals, the proportion of students who get their first choice is the same in all groups (conditional on preferences). The second part of Corollary 1 considers the case when marginals are different; hence, there is some "baseline" inequality when both groups have the same correlation. When the group with the higher marginal distribution also exhibits lower correlation, inequality increases further. Conversely, if it exhibits higher correlation, inequality is reduced and might even be reversed when the correlation gap becomes large.

So far, we examined efficiency and inequality separately. The following proposition describes their interaction. Specifically, different correlation vectors $\boldsymbol{\theta}$ can lead to the same

efficiency, while inducing different levels of inequality between groups. This shows that differential correlation has a distinct impact on efficiency on the one hand and on inequality on the other.

PROPOSITION 4. *Let* $\boldsymbol{\theta} \in \Theta^d$.

1. *There exists a continuum*[19] *of correlation vectors that achieve a given efficiency. Formally, the set of vectors* $\boldsymbol{\theta}'$ *such that* $E(\boldsymbol{\theta}') = E(\boldsymbol{\theta})$ *is a connected hypersurface of dimension* $d-1$.[20]

2. *Fixing efficiency, correlations are substitutes, i.e., for any two groups* G_j, G_ℓ, *there exists an interval* $U := [\underline{\theta}, \bar{\theta}]$ *and a decreasing function* $\phi : U \to \Theta$ *such that* $\theta_\ell = \phi(\theta_j)$ *for all* $\theta_j \in U$.

3. *Over the graph of* ϕ, $R_j^{\emptyset}$ *is minimized at* $(\underline{\theta}, \phi(\underline{\theta}))$ *and maximized at* $(\bar{\theta}, \phi(\bar{\theta}))$. *Moreover, there is a unique* $\hat{\theta} \in U$ *such that* $(\hat{\theta}, \phi(\hat{\theta}))$ *minimizes inequality* $L_{j,\ell}(\boldsymbol{\theta})$.

Proof sketch. We combine Theorem 1 with the implicit function theorem. For point 1, we fix the efficiency, arbitrarily choose some θ_j and express $(\theta_\ell)_{\ell \neq j}$ as a function of θ_j, the graph of this function is a connected hypersurface. For point 2, we do the same, but this time we only express one specific θ_ℓ as a function of θ_j. Finally, point 3 is obtained by applying Theorem 2 to this function. The complete proof is provided in Appendix C.5.

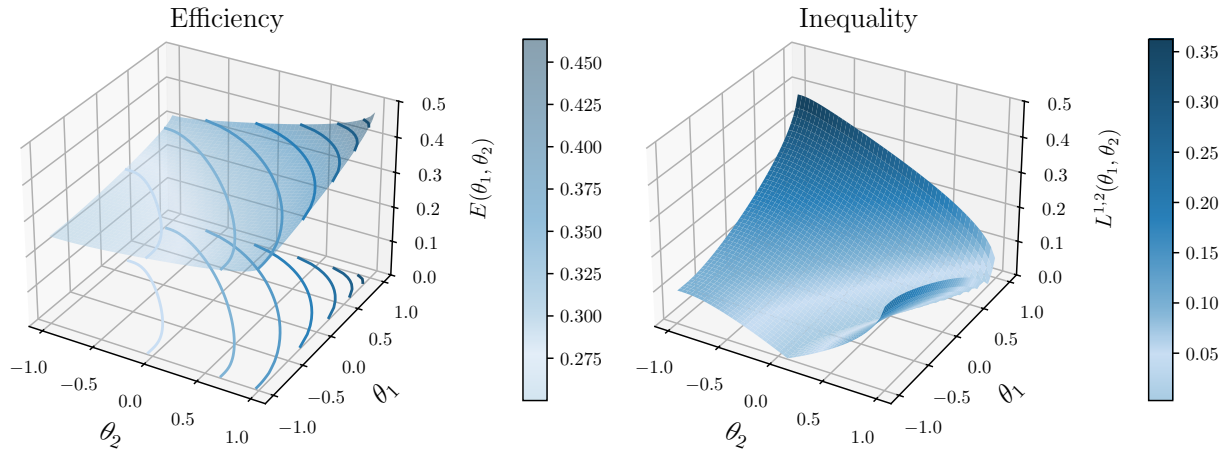
Beyond the intuition that correlation favors efficiency, Proposition 4 provides a precise insight into the relation between efficiency and inequality and the trade-off between the two. The first part states that, in general, there are infinitely many correlation vectors achieving the same level of efficiency. The second part considers the comparative statics between two groups. Fixing the efficiency, the correlation parameters behave as rival goods. As the correlation increases for one group, it necessarily decreases for the other group. Using Theorem 2, the third part of Proposition 4 states that this dynamic also implies that the proportion of unmatched students increases in the first group and decreases in the second; therefore there is a unique point minimizing inequality.

Figure 4 shows the levels of efficiency (left panel) and inequality (right panel) for two groups and two colleges with Gaussian distributions as functions of the correlation parameters θ_1, θ_2. On the left panel, the level lines show the decreasing relation between θ_1 and

[19] Unless $\boldsymbol{\theta} = (\inf\Theta, \dots, \inf\Theta)$ or $(\sup\Theta, \dots, \sup\Theta)$, in which case it is a singleton.

[20] In our context, a connected hypersurface is defined as the graph of a continuous function from Θ^{d-1} to Θ.

Figure 4 Variations of efficiency (E) and inequality (L).



Note. Illustration for two colleges and two groups, with bivariate Gaussian distributions (mean = (0, 0) for G_1, (0.2, 0.2) for G_2, variance = 1 for both) and θ_1, θ_2 chosen as each group's covariance. Other parameters: $\alpha^1 = \alpha^2 = 0.25$, $\gamma_1 = \gamma_2 = 0.5$, $\beta_{12} = \beta_{21} = 0.5$. Left: The surface shows the efficiency and the level lines indicate constant efficiency (also projected to the bottom of the figure). Right: The surface shows the inequality.

θ_2 when E is kept constant. On the right panel, following these level lines, observe that inequality is maximized when the gap between θ_1 and θ_2 is maximal, and minimized on a line located on the right of the diagonal $\theta_1 = \theta_2$ due to the fact that group G_2 first-order stochastically dominates group G_1, as predicted by Corollary 1.

4.2. Tie-breaking

A recent literature has studied the impact of tie-breaking rules on school choice problems, which is a special case of our model where colleges do not have preferences over students (or, equivalently, are indifferent). In this section, we extend some of the prior results and discuss the relation to the literature.

In a *tie-breaking problem*, colleges partition students into priority classes. Students belonging to the same priority class at some college are assumed to have the same priority at that college; however, due to limited capacity, the college might need to choose between them. To achieve this, colleges use a random ranking of students to which they refer each time they need to choose between students from the same priority class; this random ranking is called *tie-breaker*.

A natural question that has been actively studied in recent years is whether there is a difference in students' welfare if colleges all use the same tie-breaker (called a single

tie-breaker, or STB) instead of each producing an independent tie-breaker (called multiple tie-breakers, or MTB). Ashlagi, Nikzad, and Romm (2019), Ashlagi and Nikzad (2020), Arnosti (2023) show—with slightly different models and assumptions (and among other results)—that when the total capacity of colleges is lower than the number of students, then students are better off under STB than under MTB. To ease the comparison, we restate their results in a simplified form.

Ashlagi, Nikzad, and Romm 2019, Main Theorem: Suppose that students' preferences are drawn uniformly at random, there is only one priority class (the whole ranking is random), and the number of students is close to the number of seats; then, for any fixed k, the fraction of students matched to one of their top k choices approaches a positive constant under STB, but approaches 0 under MTB as the market gets large.

Ashlagi and Nikzad 2020, Theorem 3.1: Suppose there is one slot per college, students' preferences are drawn uniformly at random, there is only one priority class, and there is capacity shortage; then, with high probability, the fraction of students matched to one of their top k choices is higher under STB than MTB for all k.

Arnosti 2023, Theorem 2: Suppose there is only one priority class and students only list $l < m$ colleges picked uniformly at random in a uniform random order; then, for l small enough, the fraction of students matched to one of their top k choices is higher under STB than MTB for all k.

Our model, compared to prior work on tie-breaking, allows for any number of priority classes, intermediate levels of correlation or even negative correlation, and several groups of students with differently correlated tie-breaking rules. Intermediate correlations can arise in tie-breaking if, for example, student characteristics are introduced into rankings to break ties, e.g., sibling priority (Correa et al. 2022). This is commonly done to render algorithms more deterministic, and thus explainable. Consider sibling priority; ceteris paribus, a student who enjoys sibling priority at one school exhibits lower correlation between the priority scores at this and any other school than without sibling priority.

To this end, let $(H_\theta)_{\theta\in\Theta}$ be a family of copulas such that $\theta = 0$ gives independent random variables and $\theta = 1$ gives fully correlated variables. Define the θ-TB as the tie-breaker drawn according to H_θ. Thus, MTB corresponds to $\theta = 0$, and STB corresponds to $\theta = 1$. Moreover, we can assume the existence of several groups with different values of θ.

Let there be a continuum mass of students and assume that students prefer any college over being unmatched. Suppose that each college C^i has τ^i priority classes $\{Q^i_1, \dots, Q^i_{\tau^i}\}$, each student belongs to a product of priority classes (one for each college), and all products of priority classes contain a positive mass of students from each group. Furthermore, suppose that the students are divided into d groups, such that the θ_j-TB is used for group G_j. Define the group sizes ($\boldsymbol{\gamma}$), the colleges' capacities ($\boldsymbol{\alpha}$), and the students' preferences ($\boldsymbol{\beta}$) as before. Finally, assume that Assumption 3 (decreasing cutoffs) holds on the interior of each product of priority classes.

PROPOSITION 5. *For any $\ell \in [d]$, the metrics $R^{k,\sigma}_j$ (for all $j \neq \ell$ and $k \in [m]$), $R^{1,\sigma}_\ell$, and the efficiency, E, are almost surely increasing in θ_ℓ (otherwise constant),*[21] *and always increasing if there is only one priority class.*
Inequality between two groups, $L_{j,\ell}(\boldsymbol{\theta})$, is non-decreasing in the tie-breaker correlation of the group with the higher proportion of unassigned students and non-increasing in the other group's tie-breaker correlation.

Proof sketch. We build a distribution family that encompasses the priorities of students at each college, accounting for priority classes and tie-breakers, such that MTB and STB correspond to $\theta = 0$ and $\theta = 1$, respectively. The distribution obtained, while complex, still satisfies most of the assumptions required by our model; with some adjustments, we are able to apply Theorems 1 and 2 and conclude. The proof is provided in Appendix C.6.

This result shows that increasing the correlation of tie-breakers, for one or several groups, increases the proportion of students who obtain their first choice in all groups, as well as the proportion who obtain one of their top k choices in all groups except the one whose correlation increased. Moreover, it also shows that a policy-maker who is able to change the correlation of tie-breakers for some groups can use it to mitigate inequalities between groups.

Proposition 5 is in some regards more restrictive than the results from the literature presented above, because it requires decreasing cutoffs and that students prefer any college over being unmatched. It also does not prove that STB matches more students to one of their top k choices than MTB for $k > 1$, but it is worth noting that the aforementioned

[21] More precisely, the set of vectors $(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ such that those metrics are constant in some θ_j has Lebesgue measure 0.

results rely on stronger assumptions to do so. On the other hand, it is more general in other regards, as already discussed above: Proposition 5 allows for several priority classes, does not require students' preferences to be uniformly distributed, accommodates for several groups with different tie-breaking rules, and allows for intermediate tie-breaking rules that interpolate between MTB and STB, as well as for negatively correlated tie-breaking rules.

4.3. The decreasing cutoffs assumption

We explore the conditions under which Assumption 3 (decreasing cutoffs) holds, since it is required for all the results presented in Sections 4.1 and 4.2. We first provide theoretical guarantees for specific assumptions on the parameters, and then we show numerical experiments examining the results across the parameter space.

4.3.1. Theoretical guarantees

THEOREM 3. *Suppose there are two colleges and the family of copulas $(H_\theta)_{\theta\in\Theta}$ is coherent, then decreasing cutoffs always holds.*

Proof sketch. By coherence, at least one of the cutoffs has to be decreasing. By contradiction, we suppose that the other is not, and we show that the demand at some college is decreasing while it is supposed to be constant and equal to the capacity of this college. The key element that is only true for two colleges is the fact that $\mathbb{P}_{j,\theta_j}(W^1 < P^1, W^2 \geq P^2)$ is decreasing in θ_j due to the coherence assumption. The proof is provided in Appendix C.7.

From Theorem 3, it follows that all the results presented in Sections 4.1 and 4.2 are always true for two colleges. This observation allows to build intuition, since our results are easier to understand and to illustrate for two colleges. In addition, it is straightforward to show that in a model where all students prefer two colleges over all others, decreasing cutoffs holds for the two preferred colleges (see Appendix B.3 for a formal statement). Such preferences appear in practice; consider, for example, Oxford and Cambridge in the United Kingdom, Polytechnique and ENS in France, or, departing from colleges, McKinsey and BCG among consulting firms.

When there are more than two colleges, Assumption 3 holds under additional assumptions that are common in the literature (cf., e.g., Ashlagi, Nikzad, and Romm 2019, Ashlagi and Nikzad 2020, Peng and Garg 2024a), namely that colleges have identical capacities and students preferences are uniformly distributed (i.e., $\forall j \in [d], \forall \sigma \in \Sigma([m]),\ \beta_j^\sigma = 1/m!$).

PROPOSITION 6. *Assume that the family of copulas $(H_\theta)_{\theta\in\Theta}$ is coherent, all colleges have the same capacity, and student preferences are uniformly distributed. Then decreasing cutoffs holds for any $\boldsymbol{\gamma}$ and $\mathbf{f}$.*

Proof. Under the made assumptions all colleges are identical; then, by a symmetry argument, in the unique stable matching all cutoffs are equal. Since the coherence assumption implies that at least one cutoff has to decrease, they are in fact all decreasing. □

By a continuity argument decreasing cutoffs holds when appropriately relaxing the assumptions of Proposition 6.

COROLLARY 2. *Decreasing cutoffs holds on a subset of positive measure of the space of parameters $\boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\beta}$.*

Proof. By Lemma 4, cutoffs are $\mathcal{C}^1$ functions of $\boldsymbol{\theta}$, and by Azevedo and Leshno (2016, Theorem 2) they are continuous in all parameters $\boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\beta}$. Therefore, for all $i \in [m], j \in [d], dP^i/d\theta_j$ is continuous in $(\boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\beta})$, so it is negative in a neighborhood of the set of points that verify the assumptions of Proposition 6. □

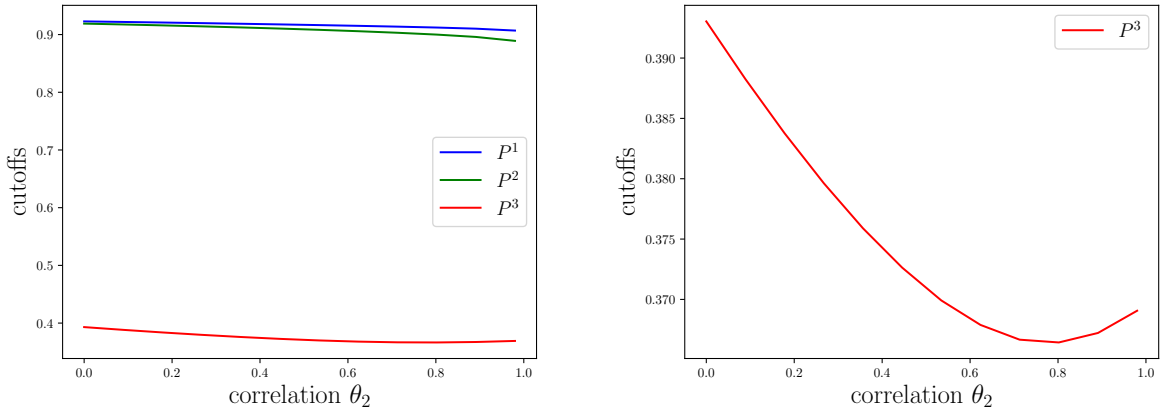
4.3.2. Numerical Experiments. We perform numerical experiments to examine whether Assumption 3 (decreasing cutoffs) and the results presented in Sections 4.1 and 4.2 hold beyond the settings covered by Theorem 3, Proposition 6, and Corollary 2. Our experiments show that, while the results sometimes fail, they remain qualitatively valid.

We analyze a broad range of parameter values for three colleges and two groups (and four colleges and two groups in Appendix D.2). We model priorities via Gaussian copulas, with covariance matrix $\Gamma_\theta = \begin{pmatrix} 1 & \theta & \theta \\ \theta & 1 & \theta \\ \theta & \theta & 1 \end{pmatrix}$ and marginals distributed uniformly over $[0, 1]$. We perform a grid search, varying the group sizes ($\boldsymbol{\gamma}$), the colleges' capacities ($\boldsymbol{\alpha}$), the students' preferences ($\boldsymbol{\beta}$), and the correlation of group G_1 (θ_1). Overall, we consider 324 different parameter combinations. For each parameter combination we compute all cutoffs, P^i, and ranks, $R_j^{k,\sigma}(\boldsymbol{\theta})$, as functions of θ_2. The list of parameters used and the detailed results of the experiments for three colleges are provided in Appendix D.1.

Cutoffs. We find that ca. 12% (38 of 324) of the scenarios considered do not exhibit decreasing cutoffs throughout the range of θ_2. This shows that Theorem 3 does not generalize beyond two colleges. Counterexamples occur when one college has significantly

higher capacity and lower demand than the others; then the cutoff of this college slightly increases when θ_2 moves towards 1. The increase for high levels of θ_2 is orders of magnitude smaller than the overall decrease across the whole range. This suggests that even when the assumption of decreasing cutoffs fails, the impact of differential correlation will qualitatively be mostly aligned with our theoretical results. Figure 5 displays one of those counterexamples, the shape of the cutoff that is not always decreasing is representative of all counterexamples.

Figure 5 Cutoffs as functions of groups 2's correlation θ_2.

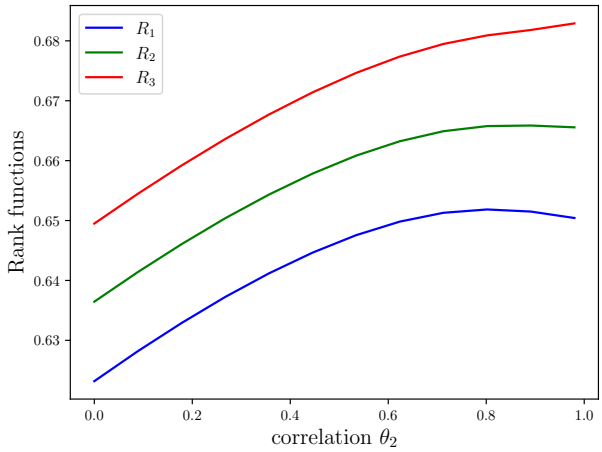


Note. On the left are the three cutoffs, and the right shows only cutoff P^3, zoomed in. Parameters: three colleges, two groups, each of size 1/2, preferences $(1/2, 1/16, 1/4, 1/32, 1/8, 1/32)$, $\theta_1 = 1/3$, and capacity $(1/15, 1/15, 8/15)$. The preference vector is to be read along the following ordering of permutations: $(123, 132, 213, 231, 312, 321)$.

Ranks. The counterexamples that violate Assumption 3 also imply that our results do not generally hold beyond two colleges. This is the case because when some cutoff P^i increases with θ_2, the same reasoning used to prove Theorem 1 shows that the proportion of students in group G_1 who prefer college C^i, and are assigned to it, decreases ($R_1^{1,\sigma}$, for preferences σ such that $\sigma(1) = i$). However, our results do not have implications for the variations of lower ranks. We consider the parameters for which we found non-decreasing cutoffs, i.e., where our theoretical results do not hold, and compute $R_1^{k,\sigma}$ for $k \in \{1, 2, 3\}$ and $\sigma = (312)$ (a preference profile for which we expect $R_1^{1,\sigma}$ to not be increasing). As expected, we find that $R_1^{1,\sigma}$ decreases exactly where P^3 increases. We also find that out of the 38 counterexamples, $R_1^{2,\sigma}$ decreases slightly in only six instances, whereas it increases

in the other scenarios. One of those instances is shown in Figure 6. The amount of students who receive their first choice decreases when θ_2 is high, as predicted, and the amount of students getting one of their top two choices also decreases for even higher values of θ_2.

Figure 6 **Rank functions for students of G_1 with preferences (312), as functions of groups 2's correlation θ_2.**



Note. Parameters: three colleges, two groups of size 1/2, capacities $(\frac{1}{15}, \frac{1}{15}, \frac{8}{15})$, $\theta_1 = 1/3$, and preference vector $(1/2, 1/16, 1/4, 1/32, 1/8, 1/32)$. The preferences vectors are to be read along the following ordering of permutations: $(123, 132, 213, 231, 312, 321)$.

Finally, note that, building on above counterexamples, it is straightforward to construct counterexamples where our results do not hold for any number of colleges. To this end, it suffices to take a counterexample with three colleges and add any number of additional colleges such that all students prefer the original three colleges over the added colleges.

5. Related Literature

Matching. The college admissions problem, i.e., how to assign prospective students to colleges given each student's preferences and colleges' priorities over students and capacities such that the outcome is stable, was introduced by Gale and Shapley (1962). The variant of this model where colleges do not have priorities over students is called the school choice problem (cf., e.g., Balinski and Sonmez 1999, Abdulkadiroğlu and Sönmez 2003, Abdulkadiroğlu 2005, Ergin and Sönmez 2006, Yenmez 2013). The idea of considering a continuum of students and a finite number of colleges has previously been exploited due to its analytical tractability (cf., Chade, Lewis, and Smith 2014, Abdulkadiroğlu, Che, and Yasuda 2015, Azevedo and Leshno 2016, Arnosti 2022).

Matching with correlated types. We study matching in the presence of correlation between the priority scores given by each college to a given student. A special case of this problem has been studied for school choice problems, where many students have the same priority and ties are broken at random. The implications of this feature for students' welfare have been studied by Erdil and Ergin (2008), Abdulkadiroğlu, Pathak, and Roth (2009) and Abdulkadiroğlu, Che, and Yasuda (2015). Most closely related and as discussed in Section 4.2, Ashlagi, Nikzad, and Romm (2019), Ashlagi and Nikzad (2020), and Arnosti (2023) compare the welfare of students in two settings, i.e., that in which either one common lottery is used by all colleges or all colleges draw independent lotteries. In our model, this corresponds to a correlation of either 1 or 0, and our results nest some findings of these prior papers. Another line of work has considered correlation between other features, e.g., between students' preferences and colleges' rankings. In this context, Brilliantova and Hosseini (2022) focus on one-to-one matching and identify a matching that does not favor one side over the other, while Che and Tercieux (2019) and Leshno and Lo (2020) study the stability-efficiency trade-off by comparing Deferred Acceptance and Top Trading Cycles (see Shapley and Scarf 1974), and how the magnitude of this trade-off depends on the correlation between agents' preferences and priorities. Considering a transferable utility model, Gola (2021) studies how workers sort into two competing sectors (such that their wages are maximized) and the impact of technological change. Although their model and analysis are quite distant from ours, they share the use of copulas to model correlation.

Fairness. The study of fairness in selection problems was initiated by Kleinberg and Raghavan (2018), who study the effect of bias and the efficiency of affirmative action policies.[22] Emelianov, Gast, Gummadi, and Loiseau (2020, 2022) and Garg, Li, and Monachou (2021) study statistical discrimination. Candidates have a latent quality, and the colleges or companies they apply to have access to only a biased or noisy estimator of this quality that varies depending on their group, which is called *differential variance*. We depart from those models by considering several decision-makers instead of one; that is, we consider the matching problem instead of the selection problem.

Studies on fairness in matching have considered various affirmative action policies, including upper and lower quotas, to reduce discrimination (Abdulkadiroğlu 2005, Kamada and

[22] Recently, reducing outcome inequalities in ranking rather than in final selection has been actively studied, cf. Celis, Mehrotra, and Vishnoi (2020), Yang, Loftus, and Stoyanovich (2021), Mehrotra, Pradelski, and Vishnoi (2022), and Zehlike, Yang, and Stoyanovich (2022).

Kojima 2015, 2024, Delacrétaz, Kominers, and Teytelboym 2023, Krishnaa, Limaye, Nasre, and Nimbhorkar 2022, Dur, Pathak, and Sönmez 2020). However, these works focus on finding stable matchings under some constraints, accounting for different fairness notions. In contrast, we aim to explain outcome inequalities that naturally occur in stable matchings without constraints.

Finally, the inequity we identify is a source of systemic discrimination, i.e., discrimination that arises only through the interaction of decision-makers—via the matching mechanism—and is not due to intentional or non-intentional discrimination by single decision-makers (cf. Pincus 1996, Feagin 2013, Bohren, Hull, and Imas 2025).

Individual fairness. All of the aforementioned studies consider what is generally termed *group fairness*, that is, the idea that all (relevant) groups should be treated similarly. In contrast, *individual fairness* posits that individuals with similar characteristics should be treated similarly. In this latter spirit Karni, Rothblum, and Yona (2022) broke new ground, showing that an individually fair ranking does not necessarily lead to an individually fair matching. This conclusion can also be drawn from our results in the context of group fairness. Devic, Kempe, Sharan, and Korolova (2023) also consider individual fairness and adapt the classical notion to incorporate agents' preferences; i.e., they require that similar agents be matched to a college in a similar position on their respective preference list. Our work is concerned with group fairness; however, we note that if the quality of applicants is similar in each group and the notion of individual fairness put forth by Devic, Kempe, Sharan, and Korolova (2023) is fulfilled, then group inequalities are automatically mitigated. In this sense, Devic, Kempe, Sharan, and Korolova (2023)'s notion is stricter.

Algorithmic monoculture. Our work also contributes to the recent literature on algorithmic monoculture, i.e., the fact that recommendations, choices, and preferences become homogeneous with the rise of algorithmic curation and analysis. Kleinberg and Raghavan (2021) study the utility of multiple decision-makers who use algorithms to evaluate candidates. They show that decision-makers are sometimes better off using different low-precision algorithms than when using the same high-precision algorithm. In contrast, we focus on the impact of correlation on applicant welfare and, moreover, introduce differential correlation between groups. Peng and Garg (2024a,b) recently study the impact of correlation in rankings in a model in which the number of decision-makers grows large. In contrast, our results hold in particular for a small number of decision-makers. Their findings on the

impact of correlation with many decision-makers show qualitative effects similar to ours. In further contrast to Peng and Garg (2024a,b), the main focus of our analysis is on the impact of differential correlation between groups.

Finally, in empirical work, Bommasani, Creel, Kumar, Jurafsky, and Liang (2022) find that outcomes are more homogeneous when models and training data sets are shared between decision-makers. Through our theoretical analysis, we thus elucidate the impact of algorithmic monoculture from the candidate's viewpoint: while increasing correlation increases the efficiency of the outcome, differential correlation may lead to increased inequality between different groups.

6. Conclusion

We have introduced a tractable model to study the impact of differential ranking correlation between different groups and studied its effect on outcome inequality and efficiency in matching markets. We have shown that efficiency generally improves as correlation increases; however, low-correlation groups are advantaged with regard to not remaining unmatched. Differential correlation leads to inequalities between groups, even when each college has a perfectly fair ranking. When there is only one group, our results quantify the role of correlation in student satisfaction and efficiency. Our framework can accommodate almost any priority score distribution, any number of groups with different distributions and different student preferences, and any number of colleges of any capacity. Although our results are most general when we limit ourselves to two colleges, they remain true with additional assumptions for any number of colleges, and simulations suggest that they quantitatively hold across the parameter space.

There is ample scope to study themes that have already been considered in single decision-maker settings in the matching context. Our analysis suggests that in matching new phenomena arise, which necessitate a deeper understanding. In models where decision-makers use noisy estimates of applicants' latent quality, interesting directions could include allowing applicants to invest in accurate assessment, e.g., by acquiring certifications or participating in in-person interviews, or considering the effects of risk aversion. Other possible research directions include making applications costly, allowing applicants to not list all colleges in their preferences, and an analysis of colleges' utility.

Appendix

Appendix A provides definitions and technical results, Appendix B additional results omitted in the main body, Appendix C omitted proofs, and Appendix D details on our numerical experiments.

Appendix A: Definitions and technical details

A.1. Definitions

A.1.1. Notation. Table 1 provides a summary of the notation used throughout the paper. In addition, we use the following convention: bold letters are used for vectors, superscripts represent colleges' indices or preference lists, and subscripts represent groups' or students' indices.

Table 1 Notation

Agents:	
$C^1, \dots, C^m$, $\mathcal{C}$	Colleges and set of all colleges
s, $\mathcal{S}$	An arbitrary student and the set of all students
$G_1, \dots G_d$	Groups of students, partition of $\mathcal{S}$
η	Measure over $\mathcal{S}$
Agents' features:	
$\alpha^i, \boldsymbol{\alpha}$	College C^i 's capacity $\in (0,1)$ and vector of all capacities
$\gamma_j, \boldsymbol{\gamma}$	Mass of students in group $G_j \in [0,1]$ and vector of all groups' masses
σ	Permutation of colleges (represents preferences), $\in \Sigma([m])$
$\beta_j^\sigma, \boldsymbol{\beta}$	Share of G_j students with preferences $\sigma \in (0,1)$ and vector of all the β_j^σ
Priority scores:	
W_s^i	Priority score at C^i of student s
$f_j^i, F_j^i, \mathbf{f}$	Marginal pdf and cdf of college C^i for group G_j and vector of all marginal pdfs
$(H_\theta)_{\theta\in\Theta}, (h_\theta)_{\theta\in\Theta}$	Copula family and associated pdfs, indexed by θ
Θ	Set of possible values for θ
θ_j, $\boldsymbol{\theta}$	Correlation level of group G_j and vector of all correlations
$f_{j,\theta_j}, F_{j,\theta_j}$	Group G_j's priority score vectors' joint pdf and cdf, $F_{j,\theta_j} = H_{\theta_j}(F_j^1, \dots, F_j^m)$
I_j^i, I_j	Support of f_j^i and f_{j,θ_j} respectively. $I_j = \prod_{m\in[m]} I_j^i$
$\underline{I}_j^i, \bar{I}_j^i$	Lower and upper bounds of I_j^i
Matching and metrics:	
μ	Matching
P^i, $\mathbf{P}$	Cutoff of college C^i and vector of all cutoffs
$R_j^{k,\sigma}$	Share of students of group G_j with preferences σ who get a top k choice
$R_j^\emptyset$	Share of students of group G_j who remain unassigned
E	Total mass of students getting their first choice
$L_{j,\ell}$	Inequality between G_j and G_ℓ, equal to $\lvert R_j^\emptyset - R_\ell^\emptyset \rvert$

A.1.2. Definition of the mass η. Here we formally define the notion of mass for a subset of students. This section is self-contained and is not necessary to understand the results of the paper; the notations introduced here are not used elsewhere. We identify $\mathcal{S}$ to $\mathtt{S} \coloneqq \mathbb{R}^m \times \{G_1, \dots, G_d\} \times \Sigma([m])$. We partition $\mathtt{S}$ into several subsets: $\mathtt{S}_{j,\sigma} \coloneqq \{(x, G_j, \sigma) \in \mathtt{S} : x \in \mathbb{R}^m\}$ is the subset of students belonging to group G_j with preferences σ. We say that a subset $J \subseteq \mathtt{S}$ is measurable if and only if for all j, σ, $J_{j,\sigma} \coloneqq \{\mathbf{W} : (\mathbf{W}, G_j, \sigma) \in J\}$ is Borel-measurable in $\mathbb{R}^m$. Let $\mathcal{B}(\mathtt{S})$ be the set of measurable subsets of $\mathtt{S}$. We assume that to each group G_j, there is an associated probability measure $\mathbb{P}_j : \mathbb{R}^m \to \mathbb{R}$. For all j, σ we define a measure $\eta_{j,\sigma}$ as follows: for $J \subseteq \mathtt{S}$ measurable,

$$\eta_{j,\sigma}(J) = \gamma_j \beta_j^\sigma \mathbb{P}_j(\mathbf{W} \in J_{j,\sigma}). \tag{6}$$

We define over $\mathcal{B}(\mathtt{S})$ the probability measure $\eta : \mathcal{B}(\mathtt{S}) \to [0,1]$ such that for any $J \in \mathcal{B}(\mathtt{S})$,

$$\eta(J) = \sum_{j \in [d]} \sum_{\sigma \in \Sigma([m])} \eta_{j,\sigma}(J). \tag{7}$$

This definition is consistent with the notations introduced in the model, as it verifies $\eta(G_j) = \gamma_j$, $\eta(\{s \in G_j : \sigma_s = \sigma\}) = \gamma_j \beta_j^\sigma$ and so on.

A.2. Elements of correlation theory

In this section, we present common measures of correlation used in the literature, and some of their properties.

Definition 5 (Common measures of correlation). Let (X, Y) be two random variables with respective cdfs F_X and F_Y. Define:

1. Pearson's correlation: assume X and Y have finite standard deviations σ_X and σ_Y. Then $r_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$.

2. Spearman's correlation: let $rk_X = F_X(X)$ and $rk_Y = F_Y(Y)$ be the quantile variables of X and Y. Then Spearman's correlation is $\rho_{X,Y} = r_{rk_X, rk_Y}$.

3. Kendall's correlation: let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables with the same joint distribution as (X, Y). Then Kendall's correlation is

$$\tau_{X,Y} = \mathbb{P}\left[(X_1 > X_2 \cap Y_1 > Y_2) \cup (X_1 < X_2 \cap Y_1 < Y_2)\right] -$$

$$\mathbb{P}\left[(X_1 > X_2 \cap Y_1 < Y_2) \cup (X_1 < X_2 \cap Y_1 > Y_2)\right].$$

The covariance of the standard bivariate Gaussian is equal to Pearson's correlation r. Moreover, for this distribution simple expressions exist for the two other correlation coefficients:

$$\rho = \frac{6}{\pi}\arcsin(r/2), \tau = \frac{2}{\pi}\arcsin(r).$$

A correlation measure should be zero when variables are independent, and reach its maximum when the variables are totally dependent on each other. The following lemma provides these properties for the measures we just introduced.

LEMMA 1 **(Scarsini 1984, Theorems 1, 4, and 5)**. *Let X, Y be two real random variables.*

1. $r_{X,Y}, \rho_{X,Y}, \tau_{X,Y} \in [-1, 1]$.
2. $\rho_{X,Y} = 1$ *if and only if* $Y = g(X)$ *with* $g : \mathbb{R} \to \mathbb{R}$ *increasing. The same holds for* $\tau_{X,Y}$. $r_{X,Y} = 1$ *if and only if the relation is affine.*
3. *If X and Y are independent, then* $r_{X,Y} = \rho_{X,Y} = \tau_{X,Y} = 0$.

LEMMA 2 **(Scarsini 1984, Theorems 4 and 5)**. Suppose $m = 2$, the family of copulas $(H_\theta)_{\theta\in\Theta}$ is coherent, and (X_θ, Y_θ) is a random vector drawn according to H_θ, then Spearman's and Kendall's correlation coefficients $\rho(X_\theta, Y_\theta)$ and $\tau(X_\theta, Y_\theta)$ are increasing functions of θ.

Lemma 2 originally applies to all functions that verify a set of assumptions and that the authors call *measures of concordance*. Assuming that $(H_\theta)_{\theta\in\Theta}$ is coherent makes θ a measure of concordance, which gives the lemma as stated above.

A.3. Discussion on distributional assumptions

We assume that priority score distributions admit a density and have full support over the product space of the marginal distributions' supports, and that they can be represented using a copula family that is coherent (Assumption 1) and differentiable (Assumption 2). We show here that these assumptions are not very restrictive by presenting canonical examples of classical copulas satisfying our assumptions (they are given for $m = 2$ for simplicity but extend naturally to any m).

1. *Gaussian copula*: The Gaussian copula is obtained by composing the cdf Φ_θ of a bivariate Gaussian with covariance matrix $\begin{pmatrix} 1 & \theta \\ \theta & 1 \end{pmatrix}$ and the univariate cdf ϕ of the standard Gaussian: $H_\theta(x, y) = \Phi_\theta(\phi(x), \phi(y))$. Here, the parameter θ controls the covariance.

Notice that when $m \geq 3$, many choices are possible. We could set the covariance of every pair of colleges as θ, or only use θ for one pair and keep a constant covariance for all other pairs, or in between those two extremes have the covariance of each pair of colleges equal to a (different) non decreasing function of θ, such that for any θ at least one of those functions is increasing. All those options give coherent and differentiable copula families.

2. *Archimedean copulas*: Archimedean copulas are a broad class of copulas, each member of this class being itself a parametric family of copulas with a real parameter θ. The general formula is

$$H_\theta(x,y) = \psi_\theta^{-1}\left(\psi_\theta(x) + \psi_\theta(y)\right)$$

where $\psi_\theta : [0,1] \to \mathbb{R}_+$ is a continuous decreasing and convex function such that $\psi_\theta(1) = 0$. Examples include:

- Clayton: $H_\theta(x,y) = \left(\max\{x^{-\theta} + y^{-\theta} - 1; 0\}\right)^{-1/\theta}$
- Frank: $H_\theta(x,y) = -\frac{1}{\theta}\log\left(1 + \frac{(\exp(-\theta x)-1)(\exp(-\theta y)-1)}{\exp(-\theta)-1}\right)$
- Gumbel: $H_\theta(x,y) = \exp\left(-((-\log(x))^\theta + (-\log(y))^\theta)^{1/\theta}\right)$

Clayton's, Frank's, Gumbel's and some other Archimedean copulas all satisfy our coherence and differentiability assumptions.

The only assumption our model makes on the marginals is that they are continuous. This is not particularly restrictive as long as there are no ties (see Section 4.2 for a treatment of ties).

A.4. Stable matching

We introduce elements of matching theory used throughout the paper. To define matching in a continuum context, we follow Azevedo and Leshno (2016).

Definition 6. A *matching* is an assignment of students to colleges, described by a mapping $\mu : \mathcal{S} \cup \mathcal{C} \to 2^{\mathcal{S}} \cup \mathcal{C}$, with the following properties:

1. for all $s \in \mathcal{S}$, $\mu(s) \in \mathcal{C} \cup \{\emptyset\}$;
2. for $i \in [m]$, $\mu(C^i) \subseteq \mathcal{S}$ is measurable and $\eta(\mu^{-1}(C^i)) \leq \alpha^i$;
3. $C^i = \mu(s)$ if and only if $s \in \mu(C^i)$;
4. for $i \in [m]$, the set $\{s \in \mathcal{S} : \mu(s) \preceq_s C^i\}$ is open.

The first three conditions are common to almost all definitions of matching in discrete or continuous models. Condition (1) ensures that a student is either matched to a college or to the empty set, which means that they remain unmatched. Condition (2) ensures

that colleges are assigned to a subset of students that respects the capacity constraints. Condition (3) ensures that the matching is consistent, i.e., if a student is matched to a college, then this college is also matched to the student. Condition (4) was introduced by Azevedo and Leshno (2016) and is necessary to ensure that there do not exist several stable matchings that only differ by a set of students of measure 0.

We next define the notions of blocking and stability.

DEFINITION 7 (STABILITY). The pair (s, C^i) *blocks* a matching μ if s would prefer C^i to her current match, and either C^i has remaining capacity or it admitted a student with a lower priority score than s; formally, if $\mu(s) \prec_s C^i$ and either $\eta(\mu(C^i)) < \alpha^i$ or $\exists s' \in \mu(C^i)$ such that $W^i_{s'} < W^i_s$. A matching is *stable* if it is not blocked by any student-college pair.

This definition is the classical definition of stability, and is equivalent to the one provided in Section 2.3 based on cutoffs. To find a stable matching, one can extend the classic deferred acceptance algorithm by Gale and Shapley (1962) to the continuum model. This algorithm is described in Algorithm 1.

Algorithm 1 Deferred acceptance algorithm (DA)

First step: All students apply to their favorite college, they are temporarily accepted. If the mass of students applying to college C is greater than its capacity α^C, then C only keeps the α^C best

while A positive mass of students are unmatched and have not yet been rejected from every college **do**

Each student who has been rejected at the previous step proposes to her preferred college among those that have not rejected them yet

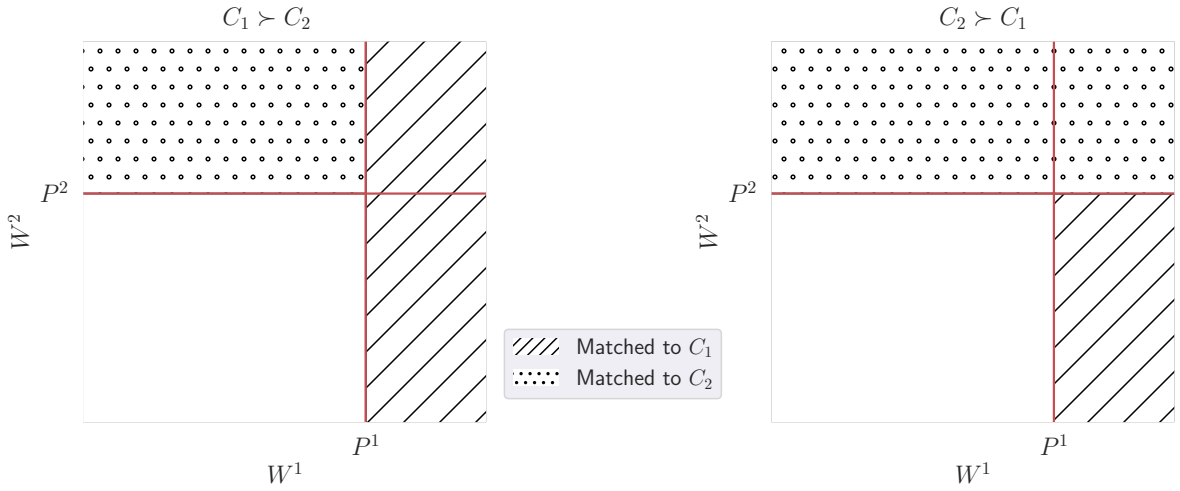
Each college C keeps the best α^C mass of students among those it had temporarily accepted and those who just applied, and rejects the others

end while

End: If the mass of students that are either matched or rejected from every college is 1, the algorithm stops. However it could happen that it takes an infinite number of steps to converge.

If the algorithm stops, the matching it outputs is stable; Abdulkadiroğlu, Che, and Yasuda (2015) show that even when the number of steps is infinite, the algorithm converges to a stable matching.

Figure 7 Cutoff description of stable matchings.



Note. Students in the hashed area are matched to college A, those in the dotted area to college B, and those in the white area remain unmatched.

Remark 1. Note that stable matchings do not only result from centralized algorithms but are often the result of a decentralized process (see, e.g., Roth and Vande Vate 1990).

Figure 7 illustrates the link between the cutoffs and the matching: students who prefer C^1 are admitted there if and only if their priority score W^1 is higher than the cutoff P^1. Otherwise, they are admitted to C^2 if their priority score W_2 is higher than P^2 and stay unmatched if it is not (left panel). The situation is symmetric for students who prefer college C^2 (right panel).

A.5. Computing metrics

We provide expressions for the metrics introduced in Section 3.

Lemma 3. *Let $j \in [d]$, $k \in [m]$, $\sigma \in \Sigma([m])$, $\boldsymbol{\theta} \in \Theta^d$, and $\boldsymbol{P}(\boldsymbol{\theta})$ be the cutoffs associated to the unique stable matching. We have:*

$$R_j^{k,\sigma}(\boldsymbol{\theta}) = \mathbb{P}_{j,\theta_j}\left(W^{\sigma(k)} \geq P^{\sigma(k)}(\boldsymbol{\theta}) \cap \bigcap_{l=1}^{k-1}(W^{\sigma(l)} < P^{\sigma(l)}(\boldsymbol{\theta}))\right). \tag{8}$$

The notation $\mathbb{P}_{j,\theta_j}$ is used as shorthand for $\mathbb{P}_{\mathbf{W}\sim f_{j,\theta_j}}$. Lemma 3 allows us to compare the satisfaction of different types of students and derive comparative statics with respect to differential correlation.

Proof. Consider student $s \in G_j$ who has preferences σ. By Azevedo and Leshno (2016, Lemma 1) (cf. Section 2.3), s is admitted to C^i which is their kth choice if and only if

$s \in D^i(\mathbf{P})$, i.e., if and only if they pass the cutoff at C^i and do not pass it at the colleges their prefer to C^i. Since the grades of students from group G_j are drawn according to F_{j,θ_j}, we obtain Equation (8). □

A.6. Regularity of cutoffs and metrics

The comparative statics results presented in Section 4 often require the cutoffs to be smooth, and all the metrics defined based on cutoffs as well. We prove here that they are.

LEMMA 4. *If Assumptions 1 and 2 hold (coherence and differentiability),* $\boldsymbol{P}(\boldsymbol{\theta})$ *is a* $\mathcal{C}^1$ *function over* $\mathring{\Theta}^d$*, and so are* $R_j^{k,\sigma}$*,* $R_j^{\emptyset}$*,* E *and* $L_{j,\ell}$*.*

Proof. The market-clearing equation (1) can be written as

$$\begin{cases} \sum\limits_{j\in[d]} \gamma_j \sum\limits_{k=1}^{m} \sum\limits_{\substack{\sigma\in\Sigma([m])\\ \sigma(k)=1}} \beta_j^{\sigma} \mathbb{P}_{\theta_j}\left(\bigcap\limits_{n<k} (W^{\sigma(n)} < P^{\sigma(n)}) \cap W^{\sigma(k)} \geq P^{\sigma(k)}\right) = \alpha^1 \\ \qquad\qquad\qquad\qquad \vdots \\ \sum\limits_{j\in[d]} \gamma_j \sum\limits_{k=1}^{m} \sum\limits_{\substack{\sigma\in\Sigma([m])\\ \sigma(k)=m}} \beta_j^{\sigma} \mathbb{P}_{\theta_j}\left(\bigcap\limits_{n<k} (W^{\sigma(n)} < P^{\sigma(n)}) \cap W^{\sigma(k)} \geq P^{\sigma(k)}\right) = \alpha^m. \end{cases} \tag{9}$$

We fix all θ_i except for θ_j for some j, and we want to study how the solution $\mathbf{P}(\boldsymbol{\theta})$ of the above equation varies as a function of θ_j. Define $T : \mathbb{R}^m \times \Theta \to \mathbb{R}^m$, $(\mathbf{P}, \theta_j) \mapsto (D^{C^1}(\mathbf{P}(\boldsymbol{\theta})) - \alpha^1, \dots, D^{C^m}(\mathbf{P}(\boldsymbol{\theta})) - \alpha^m)$. Denote by $T^1, \dots, T^m$ the components of T. Then for any $\theta_j \in \Theta$, the solution $\mathbf{P}$ of Equation (9) is also the solution of the equation $T(\mathbf{P}, \theta_j) = 0_{\mathbb{R}^m}$ (where $x_{\mathbb{R}^m}$ denotes the vector $(x, \dots, x)$ in $\mathbb{R}^m$). In order to show that $P^1, \dots, P^m$ are all decreasing in θ_j, we apply the implicit function theorem. Let $\mathbf{P} \in \mathbb{R}^m$ and $\theta_j \in \Theta$ such that $T(\mathbf{P}, \theta_j) = 0_{\mathbb{R}^m}$. Function T is of class $\mathcal{C}^1$ because $(H_\theta)_{\theta\in\Theta}$ is differentiable. We first verify that the partial Jacobian $J_{T,\mathbf{P}}(\mathbf{P}, \theta_j)$ is invertible, where

$$J_{T,\mathbf{P}} = \begin{pmatrix} \frac{\partial T^1}{\partial P^1} & \cdots & \frac{\partial T^1}{\partial P^m} \\ \vdots & & \vdots \\ \frac{\partial T^m}{\partial P^1} & \cdots & \frac{\partial T^m}{\partial P^m} \end{pmatrix}. \tag{10}$$

To prove this, we show that no (non-trivial) linear combination of the rows of $J_{T,\mathbf{P}}(\mathbf{P}, \theta_j)$ can be equal to zero. We start by proving the following inequality:

$$\forall i \in \{1, \dots, m\}, \sum_{n\in\{1,\dots,m\}} \frac{\partial T^n}{\partial P^i} < 0 \tag{11}$$

Note that T^n is a sum of probabilities of intersections containing terms of the type $W^i < P^i$ for all $i \neq n$, and $W^n \geq P^n$. Therefore, for every $i, n \in [m], i \neq n$, $\frac{\partial T^i}{\partial P^i} < 0$ and $\frac{\partial T^n}{\partial P^i} > 0$. So proving the relation (11) amounts to proving that for every $i \in [m]$, $-\frac{\partial T^i}{\partial P^i} > \sum_{n \neq i} \frac{\partial T^n}{\partial P^i}$. The first term can be written as:

$$-\frac{\partial T^i}{\partial P^i} = -\sum_{k=1}^{m} \sum_{\substack{\sigma \in \Sigma([m]) \\ \sigma(k)=i}} \sum_{\ell \in [d]} \gamma_\ell \beta_\ell^\sigma \frac{\partial \mathbb{P}_{\theta_\ell}\left(\bigcap_{p<k} (W^{\sigma(p)} < P^{\sigma(p)}) \cap W^{\sigma(k)} \geq P^{\sigma(k)}\right)}{\partial P^i}. \tag{12}$$

Notice that

$$\mathbb{P}_{\theta_\ell}\left(\bigcap_{p<k} (W^{\sigma(p)} < P^{\sigma(p)}) \cap W^{\sigma(k)} \geq P^{\sigma(k)}\right) = \mathbb{P}_{\theta_\ell}\left(\bigcap_{p<k} (W^{\sigma(p)} < P^{\sigma(p)})\right) - \mathbb{P}_{\theta_\ell}\left(\bigcap_{p\leq k} (W^{\sigma(p)} < P^{\sigma(p)})\right)$$

where the first term is constant in $P^i = P^{\sigma(k)}$. We then deduce that

$$\frac{\partial \mathbb{P}_{\theta_\ell}\left(\bigcap_{p<k} (W^{\sigma(p)} < P^{\sigma(p)}) \cap W^{\sigma(k)} \geq P^{\sigma(k)}\right)}{\partial P^i} = -\frac{\partial \mathbb{P}_{\theta_\ell}\left(\bigcap_{p\leq k} (W^{\sigma(p)} < P^{\sigma(p)})\right)}{\partial P^i}.$$

Inserting the latter in (12) gives:

$$-\frac{\partial T^i}{\partial P^i} = \sum_{k=1}^{m} \sum_{\substack{\sigma \in \Sigma([m]) \\ \sigma(k)=i}} \sum_{\ell \in [d]} \gamma_\ell \beta_\ell^\sigma \frac{\partial \mathbb{P}_{\theta_\ell}\left(\bigcap_{p\leq k} (W^{\sigma(p)} < P^{\sigma(p)})\right)}{\partial P^i}. \tag{13}$$

Now consider the terms in T^n for $n \neq i$:

$$\frac{\partial T^n}{\partial P^i} = \sum_{k=2}^{m} \sum_{\substack{\sigma \in \Sigma([m]) \\ \sigma(k)=n \\ \sigma^{-1}(i)<n}} \sum_{\ell \in [d]} \gamma_\ell \beta_\ell^\sigma \frac{\partial \mathbb{P}_{\theta_\ell}\left(\bigcap_{p<k} (W^{\sigma(p)} < P^{\sigma(p)}) \cap W^{\sigma(k)} \geq P^{\sigma(k)}\right)}{\partial P^i}.$$

The terms of T^n where P^i does not appear are constant in P^i and do not appear here, which is encompassed in the condition $\sigma^{-1}(i) < n$ in the second sum and explains why the first sum starts at $k = 2$. We want to prove that $\sum_{n \neq i} \frac{\partial T^n}{\partial P^i} < -\frac{\partial T^i}{\partial P^i}$. We can do so by comparing term by term what is inside the derivative on each side (by linearity). Let $\sigma \in \Sigma([m])$, $\ell \in [d]$, and consider the terms on each side of the inequality that have β_ℓ^σ as a factor. On the right side, we have

$$\mathbb{P}_{\theta_\ell}\left(\bigcap_{p\leq \sigma^{-1}(i)} (W^{\sigma(p)} < P^{\sigma(p)})\right).$$

On the left side, those terms exist only for $k > \sigma^{-1}(i)$:

$$\sum_{k>\sigma^{-1}(i)} \mathbb{P}_{\theta_\ell}\left(\bigcap_{n<k}(W^{\sigma(n)} < P^{\sigma(n)}) \cap W^{\sigma(k)} \geq P^{\sigma(k)}\right). \tag{14}$$

Consider the last term of the sum in (14), and upper-bound it by removing the last intersection:

$$\mathbb{P}_{\theta_\ell}\left(\bigcap_{n<m}(W^{\sigma(n)} < P^{\sigma(n)}) \cap W^{\sigma(m)} \geq P^{\sigma(m)}\right) < \mathbb{P}_{\theta_\ell}\left(\bigcap_{n<m}(W^{\sigma(n)} < P^{\sigma(n)})\right).$$

The penultimate term being

$$\mathbb{P}_{\theta_\ell}\left(\bigcap_{n<m-1}(W^{\sigma(n)} < P^{\sigma(n)}) \cap W^{\sigma(m-1)} \geq P^{\sigma(m-1)}\right),$$

we can add those two to obtain

$$\mathbb{P}_{\theta_\ell}\left(\bigcap_{n<m-1}(W^{\sigma(n)} < P^{\sigma(n)})\right).$$

We continue packing the terms together, and we finally obtain

$$\sum_{k>\sigma^{-1}(i)} \mathbb{P}_{\theta_\ell}\left(\bigcap_{n<k}(W^{\sigma(n)} < P^{\sigma(n)}) \cap W^{\sigma(k)} \geq P^{\sigma(k)}\right) < \mathbb{P}_{\theta_\ell}\left(\bigcap_{p\leq\sigma^{-1}(i)}(W^{\sigma(p)} < P^{\sigma(p)})\right),$$

which is the term associated to β_ℓ^σ in $-\frac{\partial T^i}{\partial P^i}$ (cf. Equation (13)). We finally conclude that $\forall i \in \{1,\dots,m\}$, $\sum_{n\in\{1,\dots,m\}} \frac{\partial T^n}{\partial P^i} < 0$.

We can now use this result to prove that $J_{T,\mathbf{P}}(\mathbf{P},\theta_j)$ is invertible. Let us call $R^i := (\frac{\partial T^i}{\partial P^k})_k$ the i-th row of $J_{T,\mathbf{P}}(\mathbf{P},\theta_j)$. Assume that there exist $\lambda^1,\dots,\lambda^m$, not all zero, such that $\sum \lambda^i R^i = 0_{\mathbb{R}^m}$. Let $i_0 \in \arg\max_i \lambda^i$. We assume without loss of generality that $\lambda^{i_0} > 0$. Then on the i_0-th column, using the inequality from Equation (11), we have

$$\sum_i \lambda^i \frac{\partial T^i}{\partial P^{i_0}} \leq \lambda^{i_0} \sum_i \frac{\partial T^i}{\partial P^{i_0}} < 0,$$

which contradicts $\sum \lambda^i R^i = 0_{\mathbb{R}^m}$. We conclude that no non-trivial linear combination of the rows of $J_{T,\mathbf{P}}(\mathbf{P},\theta_j)$ can be zero, therefore it is invertible.

The assumptions of the implicit function theorem are verified, therefore there exists a neighborhood $U \subseteq \mathbb{R}^m \times \Theta$ of $(\mathbf{P},\theta_j)$, a neighborhood $V \subseteq \Theta$ of θ_j, and a $\mathcal{C}^1$ function $\psi : V \to \mathbb{R}^m$ such that for all $(x,\theta) \in \mathbb{R}^m \times \Theta$,

$$(\,(x,\theta) \in U \text{ and } T(x,\theta) = 0\,) \Leftrightarrow (\,\theta \in V \text{ and } x = \psi(\theta)\,).$$

Applying this reasoning for every $\theta_j \in \mathring{\Theta}$, the function ψ is uniquely defined over each element of an open cover of $\mathring{\Theta}$, and therefore is well defined on $\mathring{\Theta}$. Since this is true for all $j \in [d]$, we finally conclude that $\mathbf{P}(\boldsymbol{\theta})$ is a $\mathcal{C}^1$ function. All the other metrics are defined as the composition of the cutoffs with $\mathcal{C}^1$ functions and are therefore also of class $\mathcal{C}^1$. $\square$

Appendix B: Additional results

B.1. The latent quality plus noise setting

We here discuss the classical model where an (unknown) latent quality is observed with an added noise term (cf. Phelps 1972, Emelianov, Gast, Gummadi, and Loiseau 2020, 2022, Garg, Li, and Monachou 2021). This provides an example of how differential correlation can arise.

Assume that all students have a latent quality W, and their priority score $\widehat{W}^i$ at college C^i is the sum of the latent quality and a noise term ε drawn independently at each college, that is,

$$\forall s \in \mathcal{S}, \text{ for } i \in [m], \widehat{W}^i_s = W_s + \varepsilon^i_s.$$

Further, assume that the latent qualities of all students (independently of their group) are drawn from the same Gaussian distribution, and that the noises are also normally distributed and depend on the group:

$$\forall s \in \mathcal{S}, \text{ for } i \in [m], W_s \sim \mathcal{N}(0, \chi^2),\ \varepsilon^i_s \sim \mathcal{N}(0, \sigma^2_{G(s)}).$$

The fact that the noise's variance is different for each group can be interpreted as colleges having different accuracies when evaluating students from different groups. Consider two examples: First, G_1 could consist of students from well-known high schools, which colleges can evaluate well since they have a lot of applicants from there each year; and G_2 could consist of students from unknown high schools, for which colleges do not have a lot of prior information. Second, each group G_j could consist of students from different demographic groups, defined by sensitive attributes such as gender, ethnicity, or social class.

Suppose that colleges know each student's group, and are aware of the difference that exists in noise variance across groups. Further, assume that colleges implement equal opportunity policies, that is, ceteris paribus the rank distribution of students must be the same for all groups.[23] To do so while maximizing the expected quality of admitted candidates, it is optimal for a college to not change the order of priority scores within groups, but to only fit each group's priority scores $\widehat{W}^i$ to the same, standardized distribution. The transformation that achieves this goal is the following

[23] Without the equal opportunity assumption, colleges would compute the expected true qualities based on the different variances for each group, see Emelianov, Gast, Gummadi, and Loiseau (2022), Garg, Li, and Monachou (2021).

$$\forall s \in \mathcal{S}, \text{ for } i \in [m],\ \widetilde{W}_s^i = \frac{\widehat{W}_s^i}{\sqrt{\chi^2 + \sigma^2_{G(s)}}}.$$

With these new standardized priority scores, the marginal priority distribution of each group is $\mathcal{N}(0,1)$ at each college. The priority vectors then follow a centered bivariate normal distribution with variance 1 and a correlation that is different between the two groups: formally,

$$\widetilde{\mathbf{W}}_s \sim \mathcal{N}\left(0_{\mathbb{R}^m}, \begin{pmatrix} 1 & r_{G(s)} & \cdots & r_{G(s)} \\ r_{G(s)} & \ddots & & \vdots \\ \vdots & & \ddots & r_{G(s)} \\ r_{G(s)} & \cdots & r_{G(s)} & 1 \end{pmatrix}\right),$$

with

$$r_{G(s)} = \frac{\chi^2}{\chi^2 + \sigma^2_{G(s)}}.$$

This setting satisfied the assumptions of our general model, and thus our analysis applies. Notice that this is also the case for the priority scores $\widehat{\mathbf{W}}$ before renormalization, the difference being that in this case the marginals are not identical. The parameter θ_j for each group G_j can then be chosen as $-\sigma^2_{G_j}$, or r_{G_j}, or any increasing function of one of these quantities.

Beyond this example, note that any priority score vector can be decomposed into an unknown latent quality and the remainder, interpreted as noise. As long as the observed priority score vectors' distributions depend on a parameter so that they form a coherent family, our results apply.

B.2. Excess capacity

We explore a case where correlation does not play any role in the matching.

PROPOSITION 7. *If for all pairs of colleges $C^i, C^{i'}$, $\alpha^i + \alpha^{i'} \geq 1$, then correlation has no effect on the matching, i.e., $\boldsymbol{P}$ is constant in $\boldsymbol{\theta}$, and every student gets either their first or second choice.*

Proof. If for all $i, i' \in [m]$, $\alpha^i + \alpha^{i'} \geq 1$, at most one college is full. Because if two colleges were full, the mass of students admitted to those two colleges would exceed the total mass of students. Assume college C^i is full, then students whose favorite college is not C^i get their first choice independently of their priority scores, and those who prefer C^i get it if

they pass the cutoff, otherwise they get their second choice. Therefore, the demand at each college depends only on preferences and marginal distributions, not the correlation, and the same holds for the cutoffs, and thus the matching. □

B.3. Two preferred colleges

Here we formalize the result mentioned in Section 4.3.1, where two colleges are preferred to all others, which technically violates the assumptions of the model since we assumed that all preference lists are used by a positive mass of students.

PROPOSITION 8. *Assume that two colleges, say, C^1 and C^2, are preferred to all other colleges by all students.*[24] *Further assume that the projection of $(H_\theta)_{\theta\in\Theta}$ on its two first variables is coherent. Then there exists a unique stable matching, P^1 and P^2 are decreasing, and all the results presented until now apply to the college admissions problem obtained by restricting $\mathcal{C}$ to $\{C^1, C^2\}$.*

Proof. Students only get matched to colleges $C^i, i \geq 3$ if they do not pass the cutoffs at C^1 and C^2. Therefore, we can compute the matching on the restricted set of colleges $\{C^1, C^2\}$, then remove the matched students and start again with the remaining colleges. As long as the projection of $(H_\theta)_{\theta\in\Theta}$ on the variables associated to C^1 and C^2 is coherent, we recover Theorem 3, so their respective cutoffs are decreasing, and all subsequent results apply to the reduced market $\{C^1, C^2\}$. □

[24] By that we mean that the mass of students using a preference list σ such that $\{\sigma(1), \sigma(2)\} \neq \{1, 2\}$ is zero.

Appendix C: Omitted proofs

C.1. Proof of Proposition 1

The result follows directly by applying Lemma 3 and noticing that for $k=1$ only the marginal distribution of the favorite college plays a role, not the joint distribution. Therefore, if groups G_j and G_ℓ have the same marginal at college C^i ($f_j^i = f_\ell^i$), then for all σ such that $\sigma(1)=i$, $R_j^{1,\sigma} = R_\ell^{1,\sigma}$. □

C.2. Proof of Proposition 2

It is sufficient to notice that $R_j^{\emptyset,\sigma} = \mathbb{P}_{j,\theta_j}(\bigcap_{i=1}^{m} W^i < P^i)$ does not depend on the preferences to obtain the first part of the lemma. The second part follows from the fact that either there is excess capacity and everyone is matched, or all colleges are full and the mass of matched students is the sum of the capacities. □

Remark 2. The first part of Proposition 2 could also be derived from the strategy-proofness for students of the student-proposing deferred acceptance algorithm (Roth 1985). Indeed, the fact that students cannot improve their outcome by modifying the order of their preferences implies that them being unmatched or not does not depend on their preferences.

C.3. Proof of Proposition 3

Let $\theta_j = \theta_\ell = \theta \in \Theta$, $k \in [m]$ and $\sigma \in \Sigma([m])$. The projection of $H_j = H_\ell$ on coordinates $\{\sigma(1),\dots,\sigma(k)\}$ induces a copula $\bar{H}$.

By Sklar's theorem,

$$\mathbb{P}_{j,\theta}\left(\bigcap_{i=1}^{k} W^{\sigma(i)} < P^{\sigma(i)}\right) = \bar{H}\left(F_j^{\sigma(1)}(P^{\sigma(1)}),\dots,F_j^{\sigma(k)}(P^{\sigma(k)})\right).$$

Moreover, $\bar{H}$ is increasing in all coordinates. Therefore, since $F_j^i < F_\ell^i$ for all $i \in [m]$,

$$\bar{H}\left(F_j^{\sigma(1)}(P^{\sigma(1)}),\dots,F_j^{\sigma(k)}(P^{\sigma(k)})\right) < \bar{H}\left(F_\ell^{\sigma(1)}(P^{\sigma(1)}),\dots,F_\ell^{\sigma(k)}(P^{\sigma(k)})\right).$$

Since

$$R_j^{k,\sigma} = 1 - \mathbb{P}_{j,\theta}\left(\bigcap_{i=1}^{k} W^{\sigma(i)} < P^{\sigma(i)}\right),$$

we conclude that $R_j^{k,\sigma} > R_\ell^{k,\sigma}$. Choosing $k=m$, we also get $R_j^{\emptyset} < R_\ell^{\emptyset}$ and therefore $L_{j,\ell} > 0$.

□

C.4. Proof of Theorem 2

By Assumption 3, all P^i are decreasing in all θ_j, thus for $\ell \neq j \in [d]$:

$$\frac{dR_j^\emptyset}{d\theta_\ell} = \frac{d\mathbb{P}_{j,\theta_j}\left(\bigcap_{i\in[m]} W^i < P^i\right)}{d\theta_\ell}$$

$$= \left(\frac{d\mathbb{P}_{j,\theta_j}\left(\bigcap_{i\in[m]} W^i < P^i\right)}{dP^1}, \ldots, \frac{d\mathbb{P}_{j,\theta_j}\left(\bigcap_{i\in[m]} W^i < P^i\right)}{dP^m} \right) \cdot \left(\frac{dP^1}{d\theta_\ell}, \ldots, \frac{dP^m}{d\theta_\ell} \right)^T$$

$$< 0.$$

Since the total capacity is constant, the mass of unmatched students must also be constant. Therefore, we have

$$\gamma_j R_j^\emptyset + \sum_{\ell \neq j} \gamma_\ell R_\ell^\emptyset = 1 - \sum_{i\in[m]} \alpha^i.$$

By differentiating this equation we get

$$\gamma_i \frac{dR_j^\emptyset}{d\theta_j} + \sum_{\ell\neq j} \gamma_\ell \frac{dR_\ell^\emptyset}{d\theta_j} = 0$$

$$\Leftrightarrow \frac{dR_j^\emptyset}{d\theta_j} = -\frac{1}{\gamma_j} \sum_{\ell\neq j} \gamma_\ell \frac{dR_\ell^\emptyset}{d\theta_j}$$

$$\Rightarrow \frac{dR_j^\emptyset}{d\theta_j} > 0$$

which proves the first part of the theorem. Moreover, since $L_{j,\ell} = |R_j^\emptyset - R_\ell^\emptyset|$, assume without loss of generality that $R_j^\emptyset > R_\ell^\emptyset$, then $L_{j,\ell} = R_j^\emptyset - R_\ell^\emptyset$, and we can use the first part of the theorem to deduce the variations of $L_{j,\ell}$. □

C.5. Proof of Proposition 4

$E(\boldsymbol{\theta})$ is a convex combination of the $R_j^{1,\sigma}(\boldsymbol{\theta})$ that are increasing in all coordinates θ_j. Moreover it is continuous, and we assumed Θ to be an interval, so the image set of $E(\boldsymbol{\theta})$ is an interval, say $[E^{\min}, E^{\max}]$. Fix $\hat{E} \in (E^{\min}, E^{\max})$, and consider the solutions of the

equation $E(\boldsymbol{\theta}) = \hat{E}$. By continuity, this equation has a solution. The implicit function theorem applied to express some θ_j (the choice of j does not matter) as a function ϕ of all the other coordinates θ_ℓ shows that the solutions of $E(\boldsymbol{\theta}) = \hat{E}$ is a connected subset of Θ^d, and also an hypersurface because the function ϕ is monotonous in all θ_ℓ (this comes from the fact that $E(\boldsymbol{\theta})$ is itself monotonous). This proves the first part of the proposition.

Let us choose two groups G_j, G_ℓ, and fix all θ_k for $k \neq j, \ell$. We apply the implicit function theorem to express θ_j as a function of θ_ℓ, which shows that there exists an interval $U := [\underline{\theta}, \bar{\theta}] \subseteq \Theta$ and a differentiable function $\phi : U \to \Theta$ such that $(\theta_\ell \in U \text{ and } \theta_j = \phi(\theta_\ell)) \Leftrightarrow (E(\boldsymbol{\theta}) = \hat{E} \text{ and } \boldsymbol{\theta} \in U \times \phi(U))$. Since $E(\boldsymbol{\theta})$ is increasing in all arguments, ϕ is necessarily decreasing. Along the line $\theta_j = \phi(\theta_\ell)$, then $\frac{dR^\emptyset_\ell}{d\theta_\ell} = \frac{\partial R^\emptyset_\ell}{\partial \theta_i} + \frac{\partial R^\emptyset_\ell}{\partial \theta_j}\phi'(\theta_\ell)$, which is positive by Corollary 1, so $(\theta_\ell, \theta_j) = (\underline{\theta}, \phi(\underline{\theta}))$ minimizes $R^\emptyset_\ell$, and $(\theta_\ell, \theta_j) = (\bar{\theta}, \phi(\bar{\theta}))$ maximizes it. The same reasoning shows that those two points respectively maximize and minimize $R^\emptyset_j$. Finally, since $R^\emptyset_\ell$ is increasing and $R^\emptyset_j$ decreasing, $L_{j,\ell} = |R^\emptyset_\ell - R^\emptyset_j|$ has a unique global minimum on the line $\theta_j = \phi(\theta_\ell)$. □

C.6. Proof of Proposition 5.

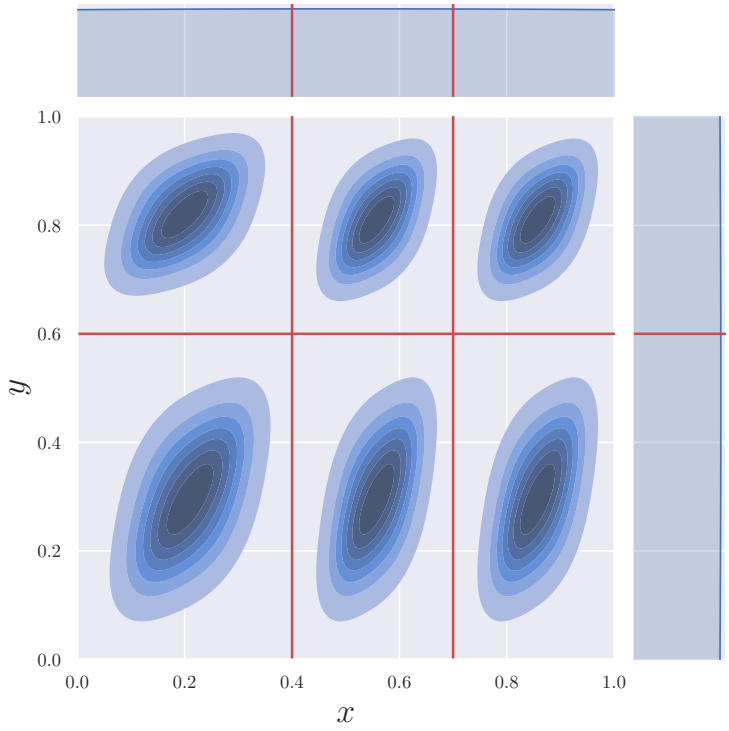
We start by building a distribution family that can represent both STB and MTB for two different values of the parameter. Let $(H_\theta)_{\theta \in \Theta}$ be a coherent copula family such that $\theta = 0$ gives independent variables and $\theta = 1$ gives equal variables, and let h_θ be the derivative of H_θ for all θ. We denote by $\kappa^i_p = \eta(Q^i_p)$ the mass of students inside class p of college C^i. For all $i \in [m]$, let $a^i_0 = 0, a^i_1 = \kappa^i_1, a^i_2 = \kappa^i_1 + \kappa^i_2, \ldots, a^i_{\tau^i} = 1$, such that they form a partition of $[0,1]$ with the p-th segment having length κ^i_p. Finally, for any $\mathbf{p} = (p^1, \ldots, p^m) \in \prod_{i \in [m]} [\tau^i]$, let $\kappa_{\mathbf{p}} = \eta(\prod_{i \in [m]} Q^i_{p^i})$ be the mass of students belonging to class p^1 at C^1, class p^2 at C^2, and so on.

Let $\tilde{h}_\theta : [0,1]^m \to \mathbb{R}$ be defined as:

$$\tilde{h}_\theta(x) = \kappa_{\mathbf{p}} h_\theta \left(\left(\frac{x_i - a^i_{p^i - 1}}{\kappa^i_{p^i}} \right)_{i \in [m]} \right), \text{ with } \mathbf{p} \text{ such that } \forall i \in [m], a^i_{p^i - 1} \leq x \leq a^i_{p^i}.$$

Let $\tilde{H}_\theta$ be the primitive of $\tilde{h}_\theta$ that cancels out on $0_{\mathbb{R}^m}$. Defined this way, $\tilde{H}_\theta$ is a copula since it is non-negative, increasing in all coordinates, marginals are uniform and do not depend on θ, and $\tilde{H}_\theta(1_{\mathbb{R}^m}) = 1$. Moreover, each hyper-rectangle $\prod_{i \in [m]} Q^i_{p^i}$ contains a "copy" of H_θ adjusted to its dimensions, that has mass $\kappa_{\mathbf{p}}$. $\tilde{H}_\theta$ is consistent with priority classes:

Figure 8 **Illustration of the distribution $\tilde{H}_\theta$.**



Note. Parameters: two colleges, three priority classes at C^1 (30% of applicants in the first class, 30% in the second, 40% in the third), two priority classes at C^2 (40% in the first class, 60% in the second). Base copula: Gaussian copula, with correlation chosen as the covariance (equal to 0.5 here).

if student s is in a higher priority class at college C^i than student s', then s will have a higher priority score with probability 1. This distribution is depicted in Figure 8.

We can verify that this definition recovers MTB and STB: if $\theta = 0$, if two students are in the same priority class for a college, they have the same ex-ante probability of getting a seat there, and if they also are in the same priority class for another college, the result of this second tie-breaker is independent from the first one. When $\theta = 1$, if two students are in the same priority class for a college, they have the same ex-ante probability of getting a seat there, but if they also are in the same priority class for another college, the winner of the tie-breaking is the same as in the first college since priority scores inside the rectangle are perfectly correlated. Therefore MTB corresponds to $\theta = 0$ and STB to $\theta = 1$.

Let us now prove the two parts of the proposition:

1. We have:

- The family $(\tilde{H}_\theta)_{\theta\in\Theta}$ is differentiable because $(H_\theta)_{\theta\in\Theta}$ is. It is also coherent (except for the x such that $x_i = a^i_p$ for some i and some p, i.e., the sides of the hyper-rectangles,

in which case the cdf is constant and not increasing). Therefore by applying Theorem 1, $R_j^{k,\sigma}, R_\ell^{1,\sigma}$ and E are either increasing or constant.

• Moreover, the case where it could be constant can only happen if there are several priority classes, so if there is only one they are increasing.

• Consider the case of multiple priority classes. Suppose that $\exists \boldsymbol{\theta} \in \Theta$ such that $P^i(\boldsymbol{\theta}) \neq a_p^i$ for all i and all p. We can apply Theorem 1, and deduce that the $R_j^{k,\sigma}$, $R_j^{1,\sigma}$ and E are increasing in θ_ℓ on the whole interval Θ. If there exists no such $\boldsymbol{\theta}$, it implies that $\mathbf{P}$ is constant in $\boldsymbol{\theta}$ and so are all the metrics. However, as any perturbation of either $\boldsymbol{\gamma}, \boldsymbol{\beta}$, or $\boldsymbol{\alpha}$ would change the cutoffs and resolve the issue, the set of problematic values of $(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ has Lebesgue measure 0.

2. Finally, Theorem 2 can be applied with the same adjustments, which proves the second part.

□

C.7. Proof of Theorem 3

Let $j \in [d]$ and $\boldsymbol{\theta} \in \Theta^2$. By the coherence assumption, the mass of students below fixed cutoffs increases in θ_j, however, since the mass of unmatched students is constant at least one of the cutoffs has to be decreasing. Without loss of generality, assume that P^1 is decreasing, and suppose that P^2 is non-decreasing. College C^2's component of the market-clearing equation 1 gives

$$\left(\sum_{\ell \in [d]} \beta_\ell^{(21)}\right) \mathbb{P}(W^2 \geq P^2) + \left(\sum_{\ell \in [d]} \beta_\ell^{(12)} \mathbb{P}_{\ell,\theta_\ell}(W^1 < P^1, W^2 \geq P^2)\right) = \alpha^2 \tag{15}$$

Let us analyze the variations of each term. $\mathbb{P}(W^2 \geq P^2)$ is decreasing in P^2, its partial derivative w.r.t. θ_j is zero, and P^2 is assumed to be non-decreasing in θ_j so $\mathbb{P}(W^2 \geq P^2)$ is overall decreasing in θ_j. For $\ell \neq j$, $\mathbb{P}_{\ell,\theta_\ell}(W^1 < P^1, W^2 \geq P^2)$ is increasing in P^1, decreasing in P^2, and its partial derivative w.r.t. θ_j is zero, so it is overall decreasing in θ_j. Finally, $\mathbb{P}_{j,\theta_j}(W^1 < P^1, W^2 \geq P^2)$ is increasing in P^1, decreasing in P^2, and its partial derivative w.r.t. θ_j is negative by the coherence assumption (because $\mathbb{P}_{j,\theta_j}(W^1 < P^1, W^2 \geq P^2) = \mathbb{P}(W^1 < P^1) - \mathbb{P}_{j,\theta_j}(W^1 < P^1, W^2 < P^2)$). Therefore, all terms are decreasing in θ_j, but the right-hand side of Equation (15) is a constant. We deduce by contradiction that P^2 is decreasing, which concludes the proof.

□

Appendix D: Numerical experiments

In this section, we show in more detail the results of the numerical experiments from Section 4.3.2. We compute cutoffs and rank metrics for various values of the parameters, for three and four colleges. For three colleges we consider three different preference vectors (i.e., distributions over the possible preference lists of students). For four colleges, we only consider uniform preferences, because accounting for 24 different preferences lists each used by a different proportion of students is computationally too costly. For the same reason, we stop our exploration at four colleges, since the number of preference lists is equal to $m!$.

D.1. Three colleges

We compute the cutoffs as functions of group G_2's correlation θ_2, for 324 different sets of values of the other parameters:

- **Group Proportions ($\gamma_1 : \gamma_2$):** 10:90, 30:70, 50:50
- **Total Capacity ($\sum_{i\in[3]} \alpha^i$):** 1/3, 1/2, 2/3
- **College Capacity Allocation ($\frac{100}{\sum_{i\in[3]} \alpha^i}(\alpha^1, \alpha^2, \alpha^3)$):** $(10, 10, 80)$, $(33.3, 33.3, 33.3)$, $(10, 45, 45)$
- **Correlation of Group 1 (θ_1):** 0, 1/3, 2/3, 0.99
- **Preference Vector ($\boldsymbol{\beta}$):** $\boldsymbol{\beta}^{\mathrm{I}}$ =Uniform, $\boldsymbol{\beta}^{\mathrm{II}}$ =(1/32; 1/8; 1/32; 1/4; 1/16; 1/2), $\boldsymbol{\beta}^{\mathrm{III}}$ =(1/2; 1/16; 1/4; 1/32; 1/8; 1/32).

The preferences vectors are to be read along the following order of permutations of colleges: $(123, 132, 213, 231, 312, 321)$.

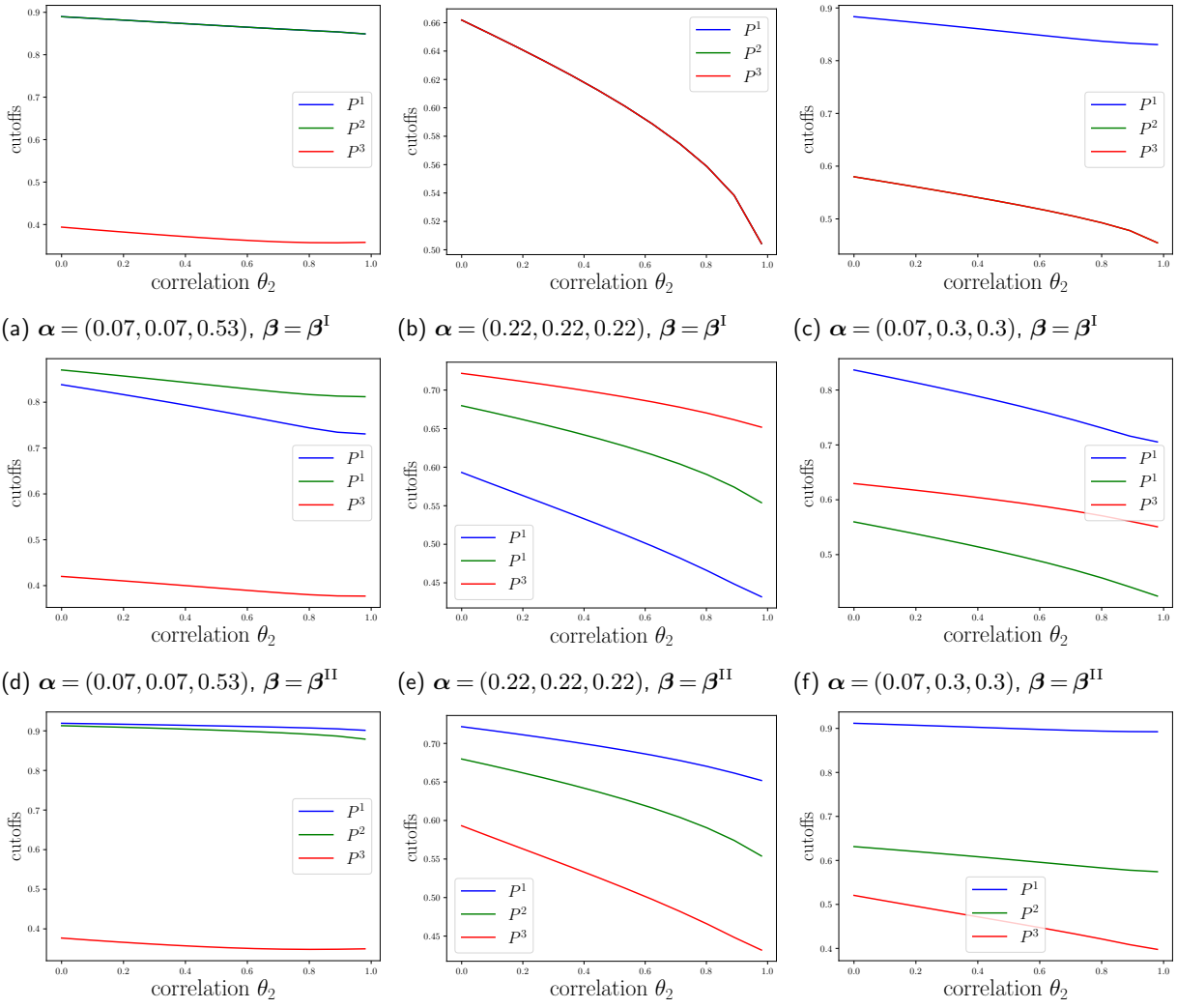
D.1.1. Cutoffs. Out of the 324 points in the grid, some give cutoff functions with very similar shapes. Specifically, group size, total capacity and correlation of group G_1 change the value of cutoffs but have very little impact on the shape of the functions; only students preferences and the allocation of the total capacity among colleges makes a noticeable difference.

Figure 9 therefore only displays nine points, corresponding to the three different preference vectors and three different capacity allocations; the choice of the other parameters has little importance and is as follows: group proportions: 50:50, total capacity: 2/3, correlation of group G_1: 1/3.

Note that all cutoffs decrease over the whole interval $[0, 1]$ for the middle and right columns, corresponding to capacity vectors $(0.22, 0.22, 0.22)$ and $(0.07, 0.3, 0.3)$. When the

capacity is $(0.07, 0.07, 0.53)$ (in the left column), however, we find counterexamples. In Figures 9a and 9g, that correspond respectively to uniform preferences and preference vector $(1/2; 1/16; 1/4; 1/32; 1/8; 1/32)$, the cutoff P^3 increases slightly as the correlation approaches 1. We can conjecture from this observation that counterexamples are found when preferences and capacity are greatly misaligned, i.e., when some college has low demand and very high capacity.

Figure 9 Variations of the cutoffs for three colleges.



(a) $\boldsymbol{\alpha} = (0.07, 0.07, 0.53)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{I}}$ (b) $\boldsymbol{\alpha} = (0.22, 0.22, 0.22)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{I}}$ (c) $\boldsymbol{\alpha} = (0.07, 0.3, 0.3)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{I}}$

(d) $\boldsymbol{\alpha} = (0.07, 0.07, 0.53)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{II}}$ (e) $\boldsymbol{\alpha} = (0.22, 0.22, 0.22)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{II}}$ (f) $\boldsymbol{\alpha} = (0.07, 0.3, 0.3)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{II}}$

(g) $\boldsymbol{\alpha} = (0.07, 0.07, 0.53)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{III}}$ (h) $\boldsymbol{\alpha} = (0.22, 0.22, 0.22)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{III}}$ (i) $\boldsymbol{\alpha} = (0.07, 0.3, 0.3)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{III}}$

Note. Note that in some plots the cutoffs overlap. Preference vectors: $\boldsymbol{\beta}^{\mathrm{I}} = (1/6, 1/6, 1/6, 1/6, 1/6, 1/6)$, $\boldsymbol{\beta}^{\mathrm{II}} = (1/32; 1/8; 1/32; 1/4; 1/16; 1/2)$, $\boldsymbol{\beta}^{\mathrm{III}} = (1/2; 1/16; 1/4; 1/32; 1/8; 1/32)$.

D.1.2. Ranks. We compute the rank metrics $R_1^{k,\sigma}$ with $\sigma = (312)$ for all points of the grid to study the validity of Theorem 1. We choose the preferences σ as it allows to highlight

potential counterexamples of Theorem 1, as it has college C^3 as the first choice and in our examples it is C^3's cutoff that is sometimes increasing. The results are displayed in Figure 10. For the values of the parameters where cutoffs are decreasing (middle and right columns, as well as Figure 10d), Theorem 1 applies and we find that all $R_1^{k,\sigma}$ are increasing as predicted. For the cases where the cutoffs do not always decrease, i.e., in Figures 10a and 10g, the proportion of students who get their first choice $R_1^{1,\sigma}$ decreases exactly where cutoffs increase. This is predicted because the proof of Theorem 1 implies that $R_1^{1,\sigma}$'s variations are the opposite of those of the first choice's cutoff, here P^3. Interestingly, for preference vector $(1/2; 1/16; 1/4; 1/32; 1/8; 1/32)$ (Figure 10g), $R_1^{2,\sigma}$ also decreases slightly for values of θ_2 very close to 1, which is not the case for uniform preferences (Figure 10a). This tends to further support the conjecture that what drives counterexamples is the misalignment between capacities and preferences.

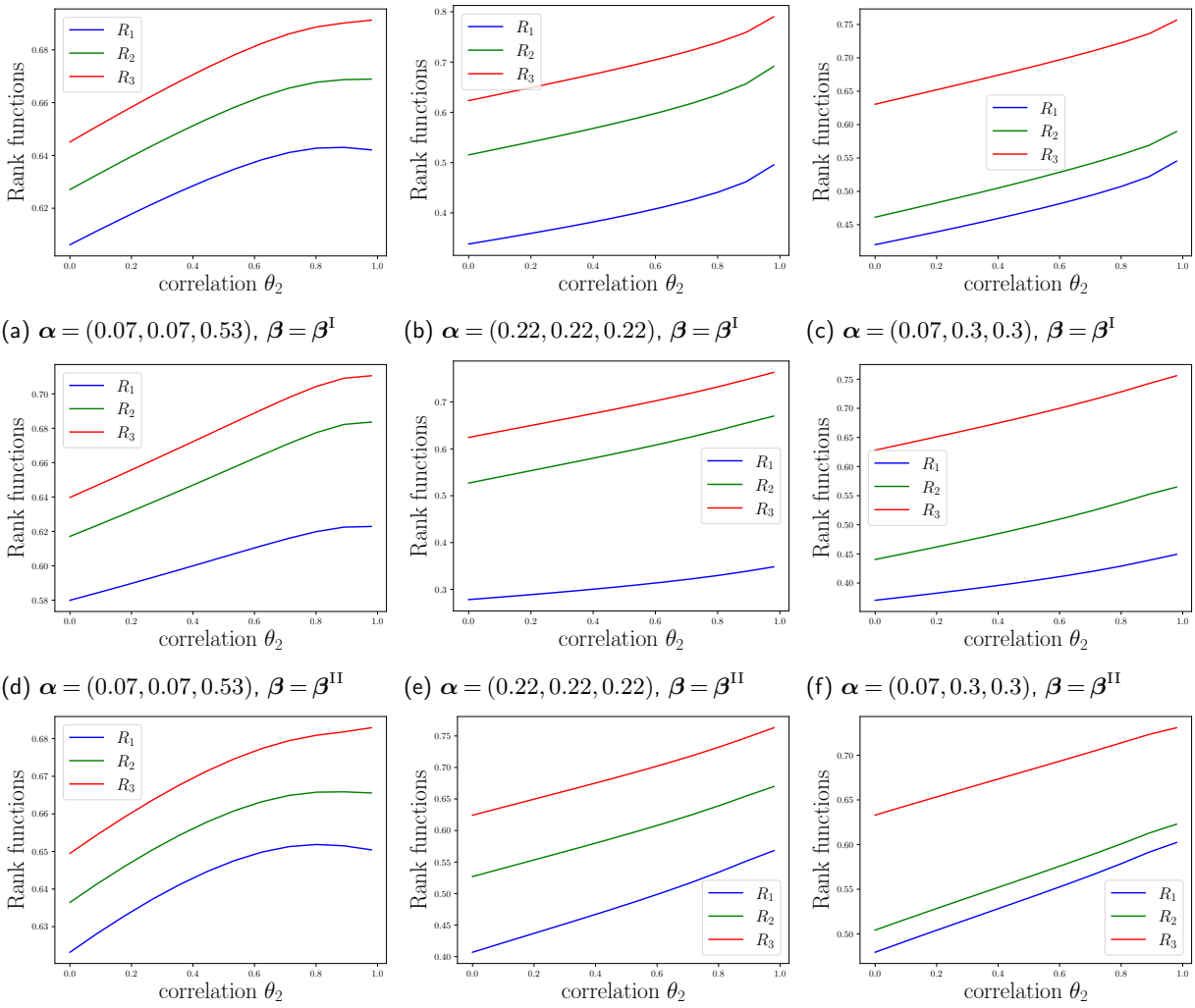
D.2. Four colleges

We compute the cutoffs as functions of group G_2's correlation θ_2, for 36 different sets of values of the other parameters:

- **Group Proportions ($\gamma_1 : \gamma_2$):** 10:90, 50:50
- **Total Capacity ($\sum_{i\in[4]} \alpha^i$):** 1/3, 1/2, 9/10
- **College Capacity Allocation ($\frac{100}{\sum_{i\in[4]} \alpha^i}(\alpha^1, \alpha^2, \alpha^3, \alpha^4)$):** $(10, 20, 30, 40)$, $(6.25, 12.5, 18.75, 62.5)$
- **Correlation of Group 1 (θ_1):** 0, 1/2, 0.99
- **Preference Vector ($\boldsymbol{\beta}$):** Uniform.

D.2.1. Cutoffs. Out of the 36 points in the grid, we find four (around 11%) where one cutoff (namely, P^4) is not always decreasing. Some sets of parameters give cutoff functions with very similar shapes. Specifically, as for the three colleges case, group size and correlation of group G_1 change the value of cutoffs but have very little impact on the shape of the functions; as before, the allocation of the total capacity among colleges makes a noticeable difference. Total capacity, while not having an equally sizeable impact, still sometimes plays a role in making cutoffs non-decreasing, which is why we also display the results for different values of total capacity. Figure 11 displays six points, corresponding to the three different total capacities and two different capacity allocations; the choice of the other parameters is of little importance and is as follows: group proportions 50:50, correlation of group G_1: 1/2.

Figure 10 Variations of rank metrics for three colleges.



(a) $\boldsymbol{\alpha} = (0.07, 0.07, 0.53)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{I}}$ (b) $\boldsymbol{\alpha} = (0.22, 0.22, 0.22)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{I}}$ (c) $\boldsymbol{\alpha} = (0.07, 0.3, 0.3)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{I}}$

(d) $\boldsymbol{\alpha} = (0.07, 0.07, 0.53)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{II}}$ (e) $\boldsymbol{\alpha} = (0.22, 0.22, 0.22)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{II}}$ (f) $\boldsymbol{\alpha} = (0.07, 0.3, 0.3)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{II}}$

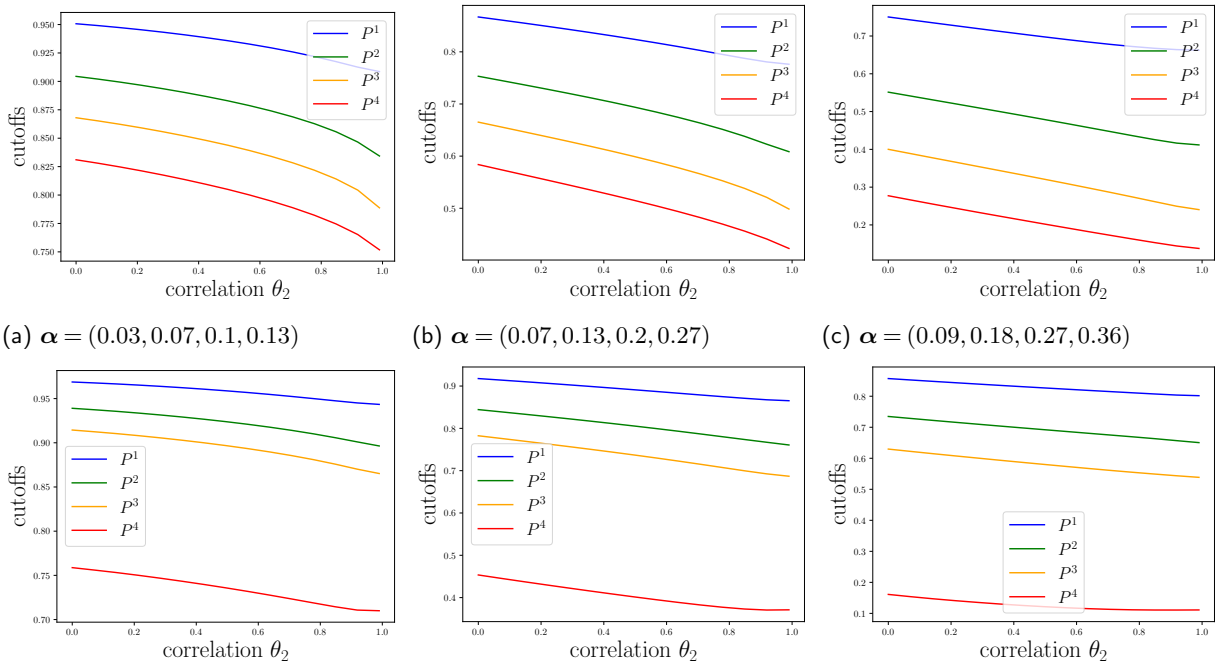
(g) $\boldsymbol{\alpha} = (0.07, 0.07, 0.53)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{III}}$ (h) $\boldsymbol{\alpha} = (0.22, 0.22, 0.22)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{III}}$ (i) $\boldsymbol{\alpha} = (0.07, 0.3, 0.3)$, $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{III}}$

Note. Preference vectors: $\boldsymbol{\beta}^{\mathrm{I}} = (1/6, 1/6, 1/6, 1/6, 1/6, 1/6)$, $\boldsymbol{\beta}^{\mathrm{II}} = (1/32; 1/8; 1/32; 1/4; 1/16; 1/2)$, $\boldsymbol{\beta}^{\mathrm{III}} = (1/2; 1/16; 1/4; 1/32; 1/8; 1/32)$.

Note that all cutoffs decrease over the whole interval $[0, 1]$ on the top row, corresponding to capacity allocation 10:20:30:40. Cutoffs are also decreasing on the left figure of the bottom row, Figure 11d, corresponding to capacity allocation 6.25:12.5:18.75:62.5 and total capacity 1/3. For total capacities 2/3 and 0.9, however, we find that cutoffs are not always decreasing. More specifically, in Figures 11e and 11f, the cutoff P^4 increases slightly when correlation approaches 1, similarly to what we observed in the three colleges case. With those parameters, college C^4 has a high capacity (respectively 0.42 and 0.56) compared to its demand (the proportion of students ranking it first is 0.25), which further confirms the

intuition we had in the three colleges case that counterexamples arise in colleges that have a capacity significantly larger than their demand.

Figure 11 Variations of the cutoffs for four colleges.



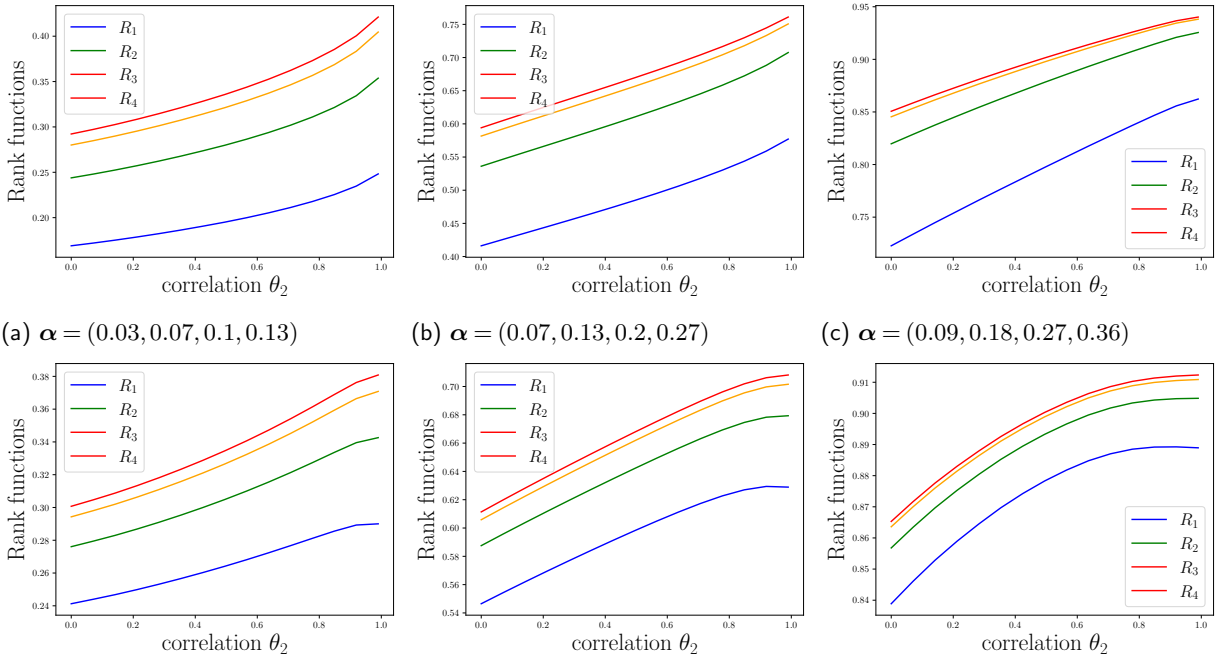
(a) $\boldsymbol{\alpha} = (0.03, 0.07, 0.1, 0.13)$ (b) $\boldsymbol{\alpha} = (0.07, 0.13, 0.2, 0.27)$ (c) $\boldsymbol{\alpha} = (0.09, 0.18, 0.27, 0.36)$

(d) $\boldsymbol{\alpha} = (0.02, 0.04, 0.06, 0.21)$ (e) $\boldsymbol{\alpha} = (0.04, 0.08, 0.12, 0.42)$ (f) $\boldsymbol{\alpha} = (0.06, 0.11, 0.17, 0.56)$

Note. First row: capacity allocation 10:20:30:40, second row: 6.25:12.5:18.75:62.5. Left column: total capacity 1/3, middle: 2/3, right:0.9.

D.2.2. Ranks. We perform the same analysis that we performed for three colleges. We compute the rank metrics $R_1^{k,\sigma}$ with $\sigma = (4321)$ for all points of the grid, to look for counterexamples to Theorem 1. The results are displayed in Figure 12. For the values of the parameters where cutoffs are decreasing (top row, as well as Figure 12d), Theorem 1 applies and we find that all $R_1^{k,\sigma}$ are increasing as predicted. For the cases where cutoffs do not always decrease, i.e., in Figures 12e and 12f, the proportion of students who get their first choice $R_1^{1,\sigma}$ decreases exactly where cutoffs increase as predicted. Contrary to the three colleges case, here the proportion of students getting one of their two top choices $R_1^{2,\sigma}$ is never decreasing. This may be explained by the fact that P^3, in the three colleges case, was increasing at a higher rate than P^4 is in the four colleges case. Making $R_1^{2,\sigma}$ decrease, thus probably requires a sufficiently high rate of increase.

Figure 12 Variations of the rank metrics for four colleges.



(a) $\boldsymbol{\alpha} = (0.03, 0.07, 0.1, 0.13)$ (b) $\boldsymbol{\alpha} = (0.07, 0.13, 0.2, 0.27)$ (c) $\boldsymbol{\alpha} = (0.09, 0.18, 0.27, 0.36)$

(d) $\boldsymbol{\alpha} = (0.02, 0.04, 0.06, 0.21)$ (e) $\boldsymbol{\alpha} = (0.04, 0.08, 0.12, 0.42)$ (f) $\boldsymbol{\alpha} = (0.06, 0.11, 0.17, 0.56)$

Note. First row: capacity allocation 10:20:30:40, second row: 6.25:12.5:18.75:62.5. Left column: total capacity 1/3, middle: 2/3, right: 0.9.

References

Abdulkadiroğlu A (2005) College admissions with affirmative action. *International Journal of Game Theory* 33(4):535–549.

Abdulkadiroğlu A, Che YK, Yasuda Y (2015) Expanding "Choice" in School Choice. *American Economic Journal: Microeconomics* 7(1):1–42.

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

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

Angwin J, Larson J, Mattu S, Kirchner L (2016) Machine bias. *ProPublica* URL `https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing`, May 23, 2024.

Arcidiacono P, Kinsler J, Ransom T (2022) Asian American discrimination in Harvard admissions. *European Economic Review* 144:104079.

Arnosti N (2022) A continuum model of stable matching with finite capacities. *Proceedings of the 23rd ACM Conference on Economics and Computation*.

Arnosti N (2023) Lottery design for school choice. *Management Science* 69(1):244–259.

Ashlagi I, Nikzad A (2020) What matters in school choice tie-breaking? how competition guides design. *Journal of Economic Theory* 190:105120.

Ashlagi I, Nikzad A, Romm A (2019) Assigning more students to their top choices: A comparison of tie-breaking rules. *Games and Economic Behavior* 115:167–187.

Azevedo EM, Leshno JD (2016) A Supply and Demand Framework for Two-Sided Matching Markets. *Journal of Political Economy* 124(5):1235–1268.

Balinski M, Sonmez T (1999) A Tale of Two Mechanisms: Student Placement. *Journal of Economic Theory* 84(1):73–94.

Bertrand M, Mullainathan S (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94(4):991–1013.

Bohren JA, Hull P, Imas A (2025) Systemic discrimination: Theory and measurement. *Quarterly Journal of Economics* Accepted.

Bommasani R, Creel KA, Kumar A, Jurafsky D, Liang PS (2022) Picking on the same person: Does algorithmic monoculture lead to outcome homogenization? Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K, Oh A, eds., *Advances in Neural Information Processing Systems*, volume 35, 3663–3678 (Curran Associates, Inc.).

Brilliantova A, Hosseini H (2022) Fair stable matching meets correlated preferences. *CoRR* abs/2201.12484.

Celis LE, Mehrotra A, Vishnoi NK (2020) Interventions for ranking in the presence of implicit bias. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*.

Chade H, Lewis G, Smith L (2014) Student Portfolios and the College Admissions Problem. *The Review of Economic Studies* 81(3):971–1002.

Che YK, Tercieux O (2019) Efficiency and stability in large matching markets. *Journal of Political Economy* 127(5):2301–2342.

Citron DK, Pasquale F (2014) The scored society: Due process for automated predictions. *Wash. L. Rev.* 89.

Correa J, Epstein N, Epstein R, Escobar J, Rios I, Aramayo N, Bahamondes B, Bonet C, Castillo M, Cristi A, Epstein B, Subiabre F (2022) School choice in Chile. *Operations Research* 70(2):1066–1087.

Delacrétaz D, Kominers SD, Teytelboym A (2023) Matching mechanisms for refugee resettlement. *American Economic Review* 113(10):2689–2717.

Devic S, Kempe D, Sharan V, Korolova A (2023) Fairness in matching under uncertainty. *Proceedings of the 40th International Conference on Machine Learning*, ICML'23.

Dur U, Pathak PA, Sönmez T (2020) Explicit vs. statistical targeting in affirmative action: Theory and evidence from chicago's exam schools. *Journal of Economic Theory* 187:104996.

Emelianov V, Gast N, Gummadi KP, Loiseau P (2020) On fair selection in the presence of implicit variance. *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*.

Emelianov V, Gast N, Gummadi KP, Loiseau P (2022) On fair selection in the presence of implicit and differential variance. *Artificial Intelligence* 302:103609.

Erdil A, Ergin H (2008) What's the matter with tie-breaking? improving efficiency in school choice. *American Economic Review* 98(3):669–89.

Ergin H, Sönmez T (2006) Games of school choice under the Boston mechanism. *Journal of Public Economics* 90(1-2):215–237.

Feagin J (2013) *Systemic racism: A theory of oppression* (Routledge).

Gale D, Shapley LS (1962) College Admissions and the Stability of Marriage. *The American Mathematical Monthly* 69(1):9–15.

Garg N, Li H, Monachou F (2021) Standardized tests and affirmative action: The role of bias and variance. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*.

Gola P (2021) Supply and demand in a two-sector matching model. *Journal of Political Economy* 129(3):940–978.

Goth G (2003) Addressing the monoculture. *IEEE Security & Privacy* 1(06):8–10.

Kamada Y, Kojima F (2015) Efficient Matching under Distributional Constraints: Theory and Applications. *American Economic Review* 105(1):67–99.

Kamada Y, Kojima F (2024) Fair Matching under Constraints: Theory and Applications. *The Review of Economic Studies* 91(2):1162–1199.

Karni G, Rothblum GN, Yona G (2022) On Fairness and Stability in Two-Sided Matchings. *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, 92:1–92:17, Leibniz International Proceedings in Informatics (LIPIcs).

Kleinberg J, Raghavan M (2018) Selection Problems in the Presence of Implicit Bias. *Proceedings of the 9th Innovations in Theoretical Computer Science Conference (ITCS '18)*.

Kleinberg J, Raghavan M (2021) Algorithmic monoculture and social welfare. *Proceedings of the National Academy of Sciences* 118(22).

Krishnaa P, Limaye G, Nasre M, Nimbhorkar P (2022) Envy-freeness and relaxed stability: hardness and approximation algorithms. *Journal of Combinatorial Optimization* 45.

Leonhardt D (2024) The misguided war on the sat. *The New York Times* URL `https://www.nytimes.com/2024/01/07/briefing/the-misguided-war-on-the-sat.html`, Jan. 7, 2024.

Leshno JD, Lo I (2020) The Cutoff Structure of Top Trading Cycles in School Choice. *The Review of Economic Studies* 88(4):1582–1623.

Longhofer SD (1995) Rooting out discrimination in home mortgage lending. *Economic Commentary, Federal Reserve Bank of Cleveland* (Nov).

Mehrotra A, Pradelski BS, Vishnoi NK (2022) Selection in the presence of implicit bias: The advantage of intersectional constraints. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 599–609.

Niessen-Ruenzi A, Ruenzi S (2019) Sex matters: Gender bias in the mutual fund industry. *Management Science* 65(7):3001–3025.

Peng K, Garg N (2024a) Monoculture in matching markets. *Advances in Neural Information Processing Systems*.

Peng K, Garg N (2024b) Wisdom and foolishness of noisy matching markets. *Proceedings of the 25th ACM Conference on Economics and Computation*.

Phelps E (1972) The statistical theory of racism and sexism. *American Economic Review* 62(4):659–661.

Pincus FL (1996) Discrimination comes in many forms: Individual, institutional, and structural. *American Behavioral Scientist* 40(2):186–194.

Roth AE (1985) The college admissions problem is not equivalent to the marriage problem. *Journal of Economic Theory* 36(2):277–288.

Roth AE, Vande Vate H (1990) Random paths to stability in two-sided matching. *Econometrica* 58:1475–1480.

Scarsini M (1984) On measures of concordance. *Stochastica* 8(3):201–218.

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

Sklar M (1959) Fonctions de répartition à N dimensions et leurs marges. *Annales de l'ISUP* VIII(3):229–231.

Yang K, Loftus JR, Stoyanovich J (2021) Causal intersectionality for fair ranking. *Proceedings of the Symposium on Foundations of Responsible Computing (FORC 2021)*.

Yenmez MB (2013) Incentive-Compatible Matching Mechanisms: Consistency with Various Stability Notions. *American Economic Journal: Microeconomics* 5(4):120–41.

Zehlike M, Yang K, Stoyanovich J (2022) Fairness in ranking, part i: Score-based ranking. *ACM Comput. Surv.* 55(6).