Ranking Correlation in Matching Markets

Remi Castera, Université Grenoble Alpes

Joint with with Patrick Loiseau, Inria & Bary Pradelski, CNRS

CIRM, Marseille, December 2023

Outcome inequality in matching

Student calls college application process unfair, calls for inclusion of marginalized communities

Published: Mar. 16, 2022, 5:59 a.m.

"Brise vocation", "Koh Lanta de l'orientation" : faut-il supprimer Parcoursup ?

par **Sonia Princet** publié le 18 janvier 2022 à 13h24

The New York Times — 29 June 2023

Supreme Court Affirmative Action Ruling | LIVE Updates 5m ago | Read the Decision | Highlights From the Ruling | How Admissions Could Change

Supreme Court Rejects Affirmative Action Programs at Harvard and U.N.C.

In earlier decisions, the court had endorsed taking account of race as one factor among many to promote educational diversity.

Causes of outcome inequality in the literature

Taste-based discrimination [Becker, 1957]: intentional discrimination from the decision-marker (i.e., colleges' evaluations of students)

Causes of outcome inequality in the literature

Taste-based discrimination [Becker, 1957]: intentional discrimination from the decision-marker (i.e., colleges' evaluations of students)

Statistical discrimination [Phelps, 1972, Arrow, 1973]: unintentional discrimination due to imperfect information of the decision-maker on candidates' qualities

- Implicit bias [Kleinberg and Raghavan, 2018]
- Differential variance [Garg et al., 2021, Emelianov et al., 2022]

Causes of outcome inequality in the literature

Taste-based discrimination [Becker, 1957]: intentional discrimination from the decision-marker (i.e., colleges' evaluations of students)

Statistical discrimination [Phelps, 1972, Arrow, 1973]: unintentional discrimination due to imperfect information of the decision-maker on candidates' qualities

- Implicit bias [Kleinberg and Raghavan, 2018]
- Differential variance [Garg et al., 2021, Emelianov et al., 2022]

Is there another source of outcome inequality in **matching problems**?

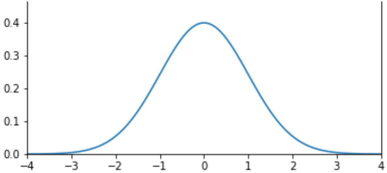
Motivating example

High School 1

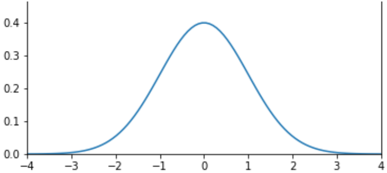
High School 2

Motivating example

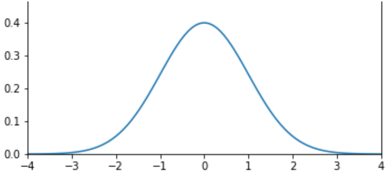
High School 1
Math grades distribution



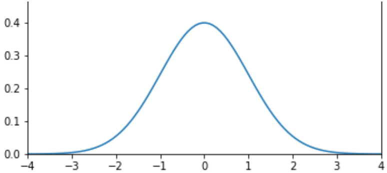
Physics grades distribution



High School 2
Math grades distribution



Physics grades distribution



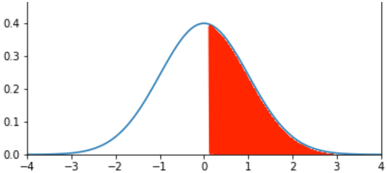
Motivating example

College A: selects based on math grades

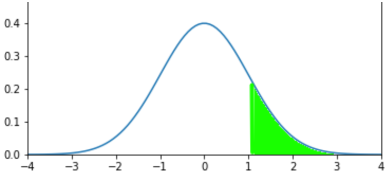
College B: selects based on physics grades

Motivating example

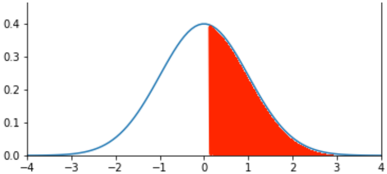
College A: selects based on math grades
Math grades distribution (HS1)



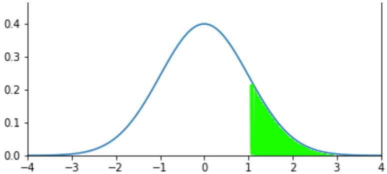
College B: selects based on physics grades
Physics grades distribution (HS1)



Math grades distribution (HS2)



Physics grades distribution (HS2)



Motivating Example

Intuitively, students from High Schools 1 and 2 should have similar admission rates, right?
Well...

Motivating Example

Intuitively, students from High Schools 1 and 2 should have similar admission rates, right?
Well...

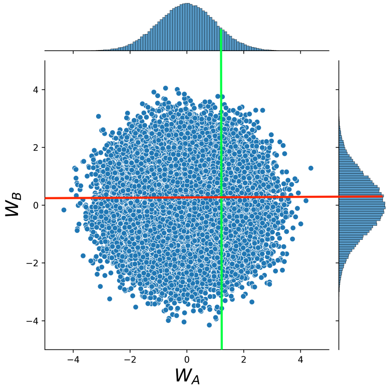
70% of students from High School 1 get a seat

55% of students from High School 2 get a seat

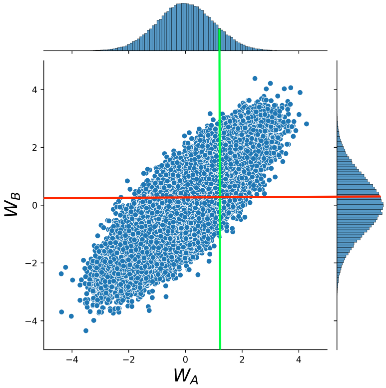
Motivating Example

Intuitively, students from High Schools 1 and 2 should have similar admission rates, right? Well...

70% of students from High School 1 get a seat



55% of students from High School 2 get a seat



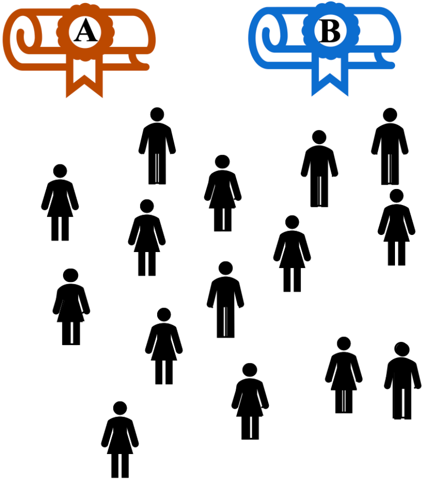
Differential correlation as a source of outcome inequality in matching

We study **matching** problems (i.e., multiple decision-makers) with multiple groups.

- ***Differential correlation***: Different ranking correlation between different groups.
- How does differential correlation affect outcome inequality and efficiency in matching?
- Key finding: differential correlation across groups leads to outcome inequalities even when the rankings by each college are 'fair'

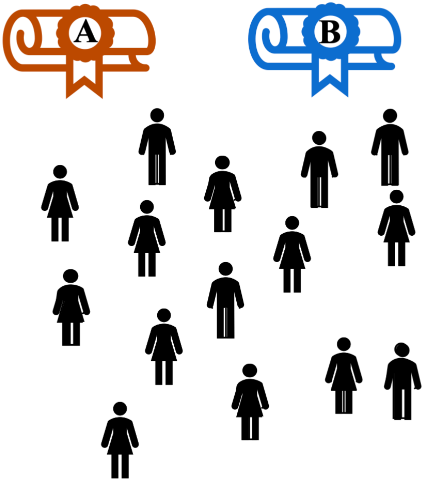
⇒ Identifies a new source of outcome inequalities that is specific to matching markets and should be included in assessments of, for example, school and university admissions

The model: basic setup



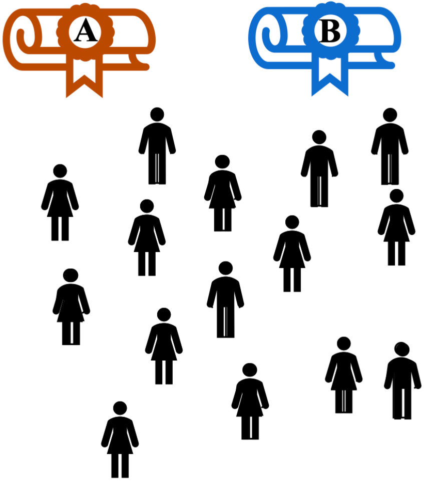
- We consider a **continuum of students** S.

The model: basic setup



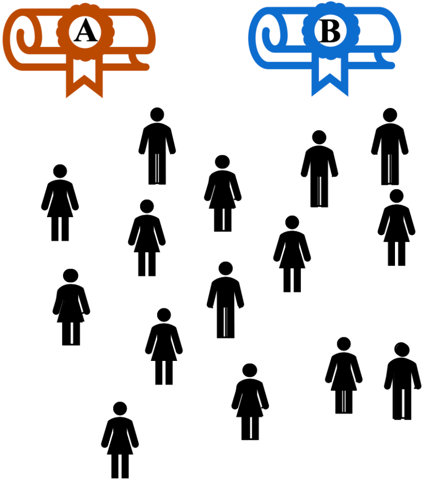
- We consider a **continuum of students** S.
- S is divided into **2 groups** G_1 and G_2.
  [Note: extends to k groups.]

The model: basic setup



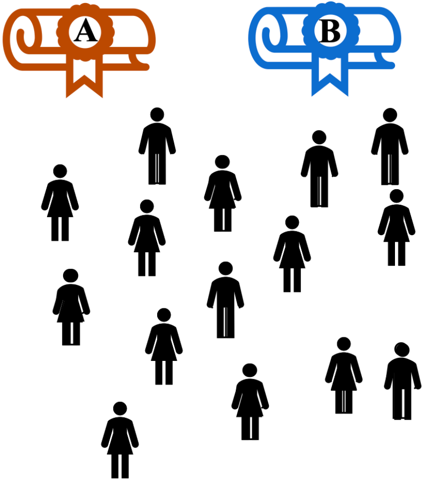
- We consider a **continuum of students** S.
- S is divided into **2 groups** G_1 and G_2.
  [Note: extends to k groups.]
- We consider **two colleges**, A and B.
  Capacities α_A and $\alpha_B \in [0, 1]$; $\alpha_A + \alpha_B < 1$.

The model: basic setup



- We consider a **continuum of students** S.
- S is divided into **2 groups** G_1 and G_2.
  [Note: extends to k groups.]
- We consider **two colleges**, A and B.
  Capacities α_A and $\alpha_B \in [0, 1]$; $\alpha_A + \alpha_B < 1$.
- **Student preference** : Some prefer A, some prefer B

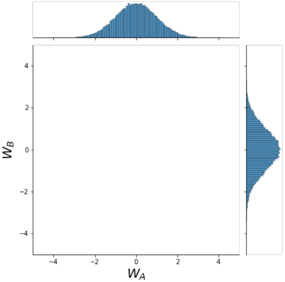
The model: basic setup

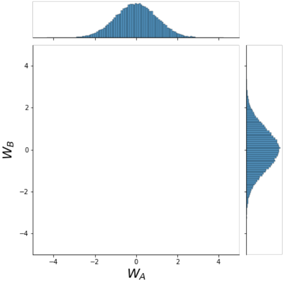


- We consider a **continuum of students** S.
- S is divided into **2 groups** G_1 and G_2.
  [Note: extends to k groups.]
- We consider **two colleges**, A and B.
  Capacities α_A and $\alpha_B \in [0, 1]$; $\alpha_A + \alpha_B < 1$.
- **Student preference** : Some prefer A, some prefer B
- **College ranking** by giving **grades**.
  Student s gets grades W_A^s and W_B^s.

Grades distribution: what we want to model

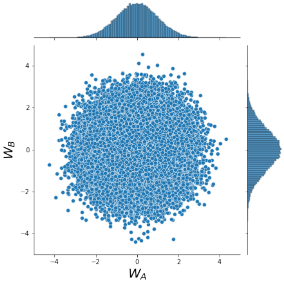
- Consider the bivariate (joint) distribution of the vector of grades (W_A^s, W_B^s)
- Example:



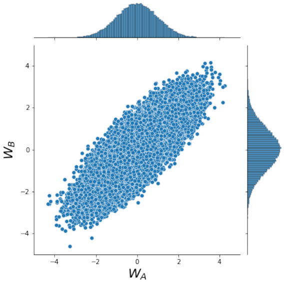


Grades distribution: what we want to model

- Consider the bivariate (joint) distribution of the vector of grades (W_A^s, W_B^s)
- Example:



Group “low correlation”



Group “High correlation”

- Differential correlation: the correlation depends on the group

Copulas

Multivariate distribution = marginals (independent) + coupling function (the copula)

Definition (Copula H)

A copula is a multivariate CDF (cumulative distribution function) on the unit cube with uniform marginals. (E.g., in 2 dimensions: $H : [0,1]^2 \to [0,1]$.)

Theorem (Sklar, 1959)

Let F be a bivariate CDF with marginals F_A and F_B. There exists a copula H s.t.

$$F(x_A, x_B) = H(F_A(x_A), F_B(x_B)), \qquad \textit{for all } (x_A, x_B) \in [-\infty, +\infty]^2.$$

(The converse is also true: given a copula H and univariate CDFs F_A and F_B, $H(F_A(x_A), F_B(x_B))$ is the CDF of a bivariate random variable with marginals F_A and F_B.)

Note: If X_A is a random variable with (univariate) CDF F_A, then $F_A(X_A)$ is uniform on $[0,1]$.

Leveraging coherence to measure correlation

- "A connection between two things in which one thing changes as the other does"[1]
- Formalizations include cardinal versus ordinal notions (e.g., quantiles)
- Coherence gives a functional-form free one-parameter proxy for correlation

Definition (Coherence)

The family $(H_\theta)_{\theta\in\Theta}$, $\Theta \subset \mathbb{R}$, is coherent iif $H_\theta(x_A, x_B)$ is increasing in θ for all x_A, x_B.

Lemma 1

Le (X_A, X_B) be a random vector with CDF $F_\theta = H_\theta(F_A, F_B)$. Assume that $(H_\theta)_{\theta\in\Theta}$ is coherent (and that the marginals F_A, F_B are independent of θ). Then:

- $\mathbb{P}_\theta(X_A < x_A, X_B < x_B)$ *and* $\mathbb{P}_\theta(X_A < x_A, X_B > x_B)$ *increasing in* θ
- $\mathbb{P}_\theta(X_A < x_A, X_B \geq x_B)$ *and* $\mathbb{P}_\theta(X_A \geq x_A, X_B < x_B)$ *decreasing in* θ

[1] Oxford Advanced Learner's Dictionary

Connection to common measures of correlation

Let (X, Y) be a couple of random variables with marginals F_X and F_Y and copula H_θ.

1. Pearson's correlation coefficient: Assume that (X, Y) have finite standard deviation σ_X, σ_Y. Then $r_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$
2. Spearman's rank correlation coefficient: Let $rank_X = F_X(X)$ and $rank_Y = F_Y(Y)$ be the rank of X and Y inside a sample. Then $\rho_{X,Y} = r_{rank_X, rank_Y}$
3. Kendall tau coefficient Let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables with the same distribution as (X, Y). Then $$\tau_{X,Y} = \mathbb{P}\left[(X_1 > X_2 \cap Y_1 > Y_2) \cup (X_1 < X_2 \cap Y_1 < Y_2)\right] - \mathbb{P}\left[(X_1 > X_2 \cap Y_1 < Y_2) \cup (X_1 < X_2 \cap Y_1 > Y_2)\right]$$

Lemma ([Scarsini, 1984])

If H_θ is coherent, then Spearman's and Kendall's correlation coefficients ρ and τ are strictly increasing functions of θ.

Important features

Recall:

- Each student receives a vector of grades (W_A, W_B) at colleges A and B
- Students belong to group G_1 or G_2

Important features

Recall:

- Each student receives a vector of grades (W_A, W_B) at colleges A and B
- Students belong to group G_1 or G_2

Denote:

- Each group G_i has a correlation level θ_i

Important features

Recall:

- Each student receives a vector of grades (W_A, W_B) at colleges A and B
- Students belong to group G_1 or G_2

Denote:

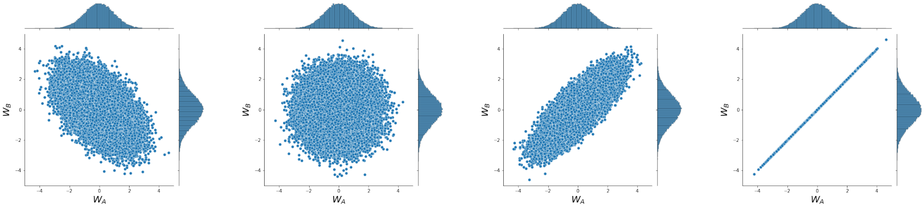
- Each group G_i has a correlation level θ_i

Important property:

- We can change the correlation of a group without changing the marginals (technical tool: copulas, not detailed here)

Examples of distributions satisfying our assumptions

- Bivariate Gaussian
  - Here θ can be taken as the covariance (or correlation) parameter ρ



$\theta = -0.5$ $\theta = 0$ $\theta = 0.8$ $\theta = 1$

Solution concept: stable matching

Definition (Stable Matching)

For each student s, for each college C such that s prefers C to the college they are matched with, all students matched to C were ranked better than s at C.

Solution concept: stable matching

Definition (Stable Matching)

For each student s, for each college C such that s prefers C to the college they are matched with, all students matched to C were ranked better than s at C.

Algorithm (Continuous Deferred Acceptance):

Initialize: 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 which 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

Theorem ([Abdulkadiroğlu et al., 2015])

The Continuous Deferred Acceptance Algorithm converges to a stable matching (possibly in infinitely many steps).

Cutoffs

A stable matching can be represented by a pair of cutoffs on the grade to get in each college.

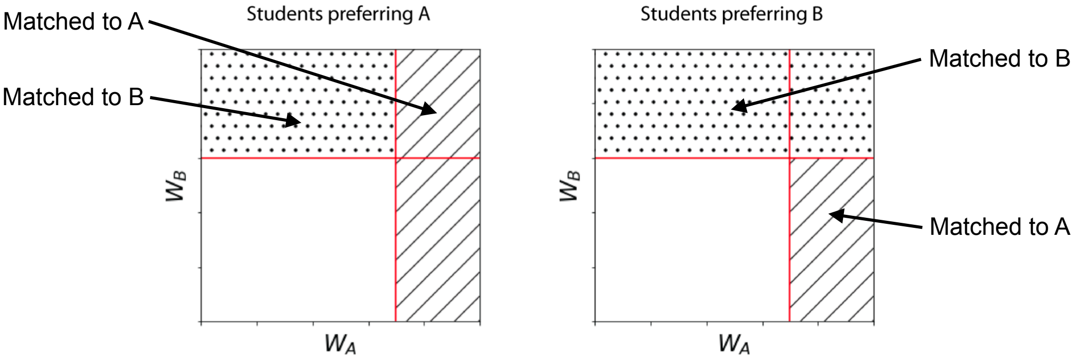


Figure: Dashed: matched to A, Dotted: matched to B, white: unmatched

The supply/demand framework for matching markets

Let $P_A, P_B \in \mathbb{R}$ be **cutoffs**, i.e., the grade of the 'worst' admitted student in resp. A and B.

The supply/demand framework for matching markets

Let $P_A, P_B \in \mathbb{R}$ be **cutoffs**, i.e., the grade of the 'worst' admitted student in resp. A and B.

Consider the **demand** at each college for these cutoffs:

$$\begin{aligned} D_A(P_A, P_B) = \quad & \#\text{Students who prefer A and with } W_A \geq P_A \\ & + \#\text{Students who prefer B and with } W_A \geq P_A,\ W_B < P_B \end{aligned}$$

$$\begin{aligned} D_A(P_A, P_B) = \quad & \#\text{Students who prefer B and with } W_B \geq P_B \\ & + \#\text{Students who prefer A and with } W_B \geq P_B,\ W_A < P_A \end{aligned}$$

The supply/demand framework for matching markets

Let $P_A, P_B \in \mathbb{R}$ be **cutoffs**, i.e., the grade of the 'worst' admitted student in resp. A and B.

Consider the **demand** at each college for these cutoffs:

$$\begin{aligned} D_A(P_A, P_B) = \quad & \#\text{Students who prefer A and with } W_A \geq P_A \\ & + \#\text{Students who prefer B and with } W_A \geq P_A,\ W_B < P_B \end{aligned}$$

$$\begin{aligned} D_A(P_A, P_B) = \quad & \#\text{Students who prefer B and with } W_B \geq P_B \\ & + \#\text{Students who prefer A and with } W_B \geq P_B,\ W_A < P_A \end{aligned}$$

We say that the cutoffs P_A and P_B are **market clearing** if

$$D_A(P_A, P_B) = \alpha_A \text{ and } D_B(P_A, P_B) = \alpha_B$$

The supply/demand framework for matching markets

Let $P_A, P_B \in \mathbb{R}$ be **cutoffs**, i.e., the grade of the 'worst' admitted student in resp. A and B.

Consider the **demand** at each college for these cutoffs:

$$D_A(P_A, P_B) = \#\text{Students who prefer A and with } W_A \geq P_A$$
$$+\#\text{Students who prefer B and with } W_A \geq P_A,\ W_B < P_B$$

$$D_A(P_A, P_B) = \#\text{Students who prefer B and with } W_B \geq P_B$$
$$+\#\text{Students who prefer A and with } W_B \geq P_B,\ W_A < P_A$$

We say that the cutoffs P_A and P_B are **market clearing** if

$$D_A(P_A, P_B) = \alpha_A \text{ and } D_B(P_A, P_B) = \alpha_B$$

Theorem ([Azevedo and Leshno, 2016])

There is a unique stable matching, and it is given by the unique pair of market clearing cutoffs.

- Hard to compute (in closed form), but we can still state qualitative properties

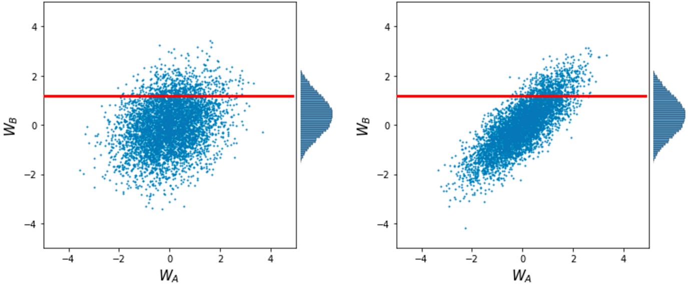
Is one group advantaged compared to the other? – *1st choice*

Proposition 1

If the marginals are the same for both groups at some college, the probability for a student to get it as their first choice is independent of the group they belong to.

(If $F_C^1 = F_C^2$ for some college C, then $V_1^{G_1,C} = V_1^{G_2,C}$.)

Consequence: If two groups have different proportions of students getting their first choice, it is only due to their marginals, not their correlation levels.

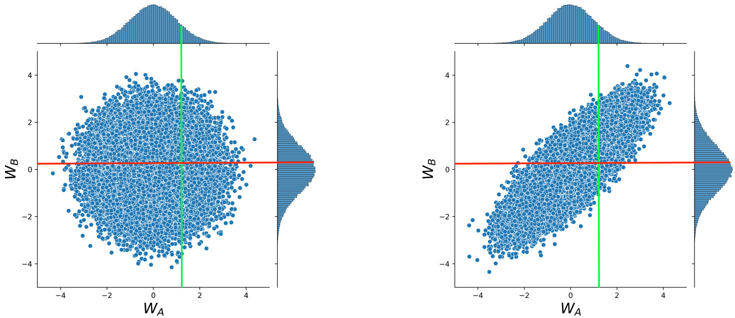


Is one group advantaged compared to the other? – *unassigned*

Proposition 2

If the marginals are the same for both groups at some college, the group with the highest correlation has the highest rate of unassigned students.

Consequence: Even when marginals are the same, i.e., each college ranks both groups identically, one group has more unassigned students than the other



Does capacity matter?

Proposition 3

If capacity is not constrained, i.e., $\alpha_A + \alpha_B \geq 1$, then differential correlation has no impact on the stable matching.

Intuition (for same marginals):

- By Proposition 1, first-choice admittance is the same for both groups
- But here there is enough capacity for every student to be admitted at (at least) their second choice

From now on, we assume $\alpha_A + \alpha_B < 1$.

What happens when correlation levels change? – *Efficiency*

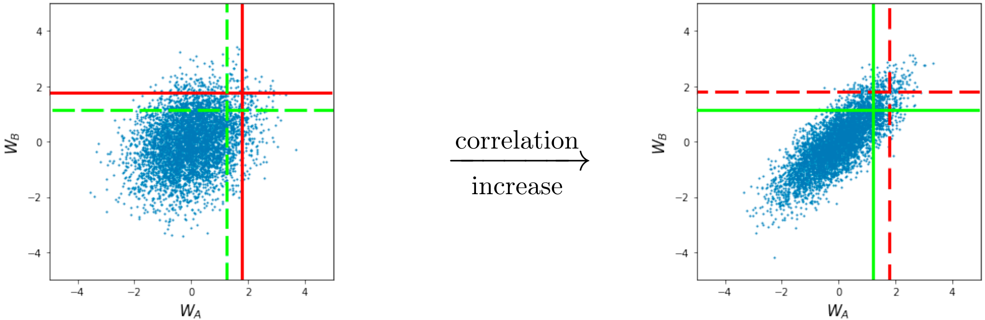
Theorem 1 (Efficiency increases in all correlation levels)

The total amount of students getting their first choice increases in all groups' correlation levels.

(In more detail: for any $G \in \{G_1, G_2\}$ and $C \in \{A, B\}$, $\frac{\partial V_1^{G,C}}{\partial \theta_G} > 0$.)

Intuition: increasing correlation decreases cutoffs

Formal proof: implicit function theorem on solution of market-clearing equations



What differs between the groups? – *Inequality*

Proposition 4

The probability of a student remaining unmatched is decreasing in the other group's correlation level and increasing in her own. (For $G \in \{G_1, G_2\}$, $\frac{\partial V_\emptyset^G}{\partial \theta_G} > 0$ and $\frac{\partial V_\emptyset^G}{\partial \theta_{\bar{G}}} < 0$.)

- Consequence: Increasing the correlation of the group with the lowest chance of staying unmatched (the "advantaged group") decreases inequality.
- If the groups have the same marginals
  - We recover Proposition 2.

Does efficiency predict inequality?

Proposition 5

For any reachable efficiency, there exists a continuum of correlation pairs achieving it, each extremity being optimal for one group and pessimal for the other.

If marginals are identical across groups, any given efficiency level is reachable with zero inequality.

- Any efficiency level can hide inequality
- Thus, efficiency loss alone does not capture the impact of differential correlation

Does efficiency predict inequality?

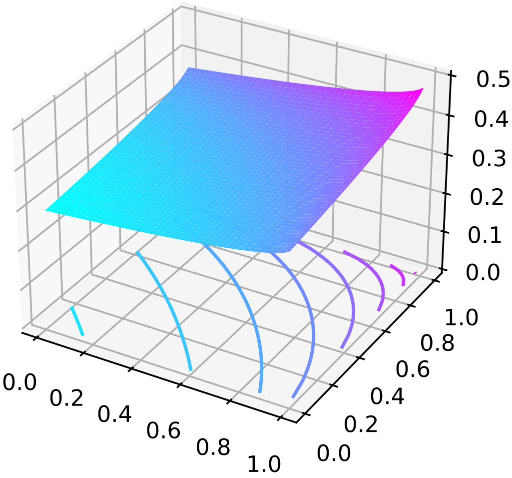


Figure: Achievable efficiency levels for all values of (θ_1, θ_2)

Link to the tie-breaking literature

Tie-breaking in school choice [Ashlagi et al., 2019, Ashlagi and Nikzad, 2020, Arnosti, 2022]:

- One demographic group, one "priority classes"
- Single tie-breaking vs multiple tie-breaking $\Rightarrow$ single tie-breaking rule is more efficient

Link to the tie-breaking literature

Tie-breaking in school choice [Ashlagi et al., 2019, Ashlagi and Nikzad, 2020, Arnosti, 2022]:

- One demographic group, one "priority classes"
- Single tie-breaking vs multiple tie-breaking $\Rightarrow$ single tie-breaking rule is more efficient

With our framework, we consider multiple priority classes:

- We recover and partially extend some of these results
  - Construct a bivariate distribution consistent with the priority classes, but with full support
  - Such that single tie-breaking corresponds to $\theta = 1$ and multiple tie-breaking to $\theta = 0$

$-$: limited to two colleges

$+$: arbitrary number of priority classes

$+$: arbitrary correlation level (not limited to 0 or 1)

- corresponds to an intermediate tie-breaker, e.g., colleges use "unimportant" criteria to break ties (e.g., proximity) and only then revert to fully random tie-breaking

Conclusion

- We identify a new source of outcome inequality in matching: differential correlation
  - Fairness in rankings does not guarantee fairness in matching
- We provide a novel framework to study correlation in matching problems (via coherence)
  - Differential correlation creates distinct effects on inequality and efficiency
    - No effect on good students, intermediate students are better off in the low correlation group.
    - Increasing both group's correlation levels as high as possible while keeping them equal provides efficiency and fairness simultaneously.
  - Allows to extend existing results on tie-breaking
- Some open questions
  - More complex models (marginals that change with the correlation, e.g. latent quality + noise)
  - Mechanism design (share estimates, fairness constraints)

Thank you!

Full paper: `https://hal.archives-ouvertes.fr/hal-03672270/`

APPENDIX & REFERENCES

Example: different criteria

Two colleges, A and B, with **different criteria**. Suppose college A is interested in the level of applicants in maths, and college B in physics. Applicants come from two high schools:

Example: different criteria

Two colleges, A and B, with **different criteria**. Suppose college A is interested in the level of applicants in maths, and college B in physics. Applicants come from two high schools:

High school 1: Physics is taught in a very theoretical way
$\rightarrow$ grades in maths and physics are **highly correlated**.

Example: different criteria

Two colleges, A and B, with **different criteria**. Suppose college A is interested in the level of applicants in maths, and college B in physics. Applicants come from two high schools:

High school 1: Physics is taught in a very theoretical way
$\rightarrow$ grades in maths and physics are **highly correlated**.

High school 2: Physics is taught through experiments
$\rightarrow$ grades in maths and physics are more **independent**.

Example: different criteria

Two colleges, A and B, with **different criteria**. Suppose college A is interested in the level of applicants in maths, and college B in physics. Applicants come from two high schools:

High school 1: Physics is taught in a very theoretical way
$\rightarrow$ grades in maths and physics are **highly correlated**.

High school 2: Physics is taught through experiments
$\rightarrow$ grades in maths and physics are more **independent**.

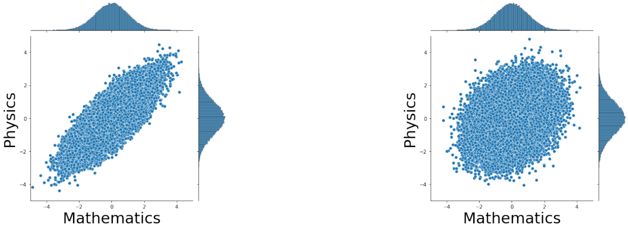


Figure: Example of distributions. Left: correlation 0.8, right: correlation 0.3

Example: Latent quality + noise with standardization

Colleges A and B have noisy estimates of applicants' qualities. Each applicant s has a **latent quality** $W^s \sim \mathcal{N}(0, \eta^2)$; and her grade at each college is:

$$\widehat{W}_A^s = W^s + \varepsilon_A^s, \ \widehat{W}_B^s = W^s + \varepsilon_B^s$$

Example: Latent quality + noise with standardization

Colleges A and B have noisy estimates of applicants' qualities. Each applicant s has a **latent quality** $W^s \sim \mathcal{N}(0, \eta^2)$; and her grade at each college is:

$$\widehat{W}_A^s = W^s + \varepsilon_A^s,\ \widehat{W}_B^s = W^s + \varepsilon_B^s$$

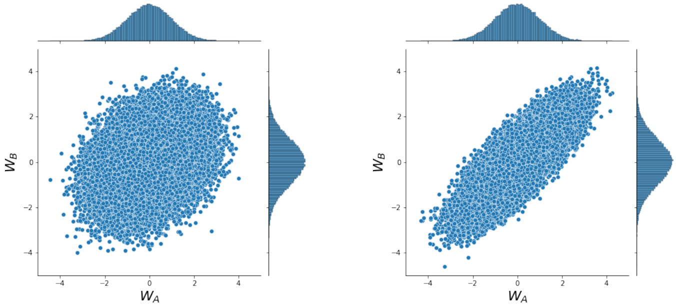
Two groups of applicants: **local and foreign**. Evaluation of local applicants is more precise than for foreign applicants. For a local applicant s, $\varepsilon^s \sim \mathcal{N}(0, \sigma_{loc}^2)$ and for a foreign applicant $\varepsilon^s \sim \mathcal{N}(0, \sigma_{for}^2)$, with $\sigma_{loc} < \sigma_{for}$.

Example: Latent quality + noise with standardization

For fairness purposes, colleges decide to standardize the grade distributions: grades of local students are fitted into $\mathcal{N}(0,1)$, and so are grades of foreign students:

$$\text{for any local student } s,\ \widetilde{W}_A^s = \widehat{W}_A^s / \sqrt{\eta^2 + \sigma_{loc}^2},\ \widetilde{W}_B^s = \widehat{W}_B^s / \sqrt{\eta^2 + \sigma_{loc}^2}$$

$$\text{for any foreign student } s,\ \widetilde{W}_A^s = \widehat{W}_A^s / \sqrt{\eta^2 + \sigma_{for}^2},\ \widetilde{W}_B^s = \widehat{W}_B^s / \sqrt{\eta^2 + \sigma_{for}^2}$$



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

Arnosti, N. (2022).
Lottery design for school choice.
Management Science, forthcoming.

Arrow, K. J. (1973).
The theory of discrimination.
In Ashenfelter, O. and Rees, A., editors, *Discrimination in Labor Markets*, pages 3–33. Princeton University Press.

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

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

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

Becker, G. S. (1957).
The economics of discrimination.
University of Chicago press.

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

Garg, N., Li, H., and Monachou, F. (2021).
Standardized tests and affirmative action: The role of bias and variance.