Stable matching as transport

A welfarist perspective on market design

Federico Echenique
Berkeley

Joseph Root
Chicago

Fedor Sandomirskiy
Princeton

Fedor and Joe

Motivation

Match agents in two populations.

Agents have *types* $x_i \in X$ and $y_j \in Y$.

When i is matched with j, the utility to i is

$$u(i,j) = w(x_i, y_j) + g(y_j) + \varepsilon(i,j),$$

Motivation

Match agents in two populations.

Agents have *types* $x_i \in X$ and $y_j \in Y$.

When i is matched with j, the utility to i is

$$u(i,j) = \underbrace{w(x_i, y_j)}_{\text{match quality}} + \underbrace{g(y_j)}_{\text{vertical/quality}} + \underbrace{\varepsilon(i,j)}_{\text{idiosync.}},$$

Motivation

Match agents in two populations.

Agents have *types* $x_i \in X$ and $y_j \in Y$.

When i is matched with j, the utility to i is

$$u(i,j) = \underbrace{w(x_i, y_j)}_{\text{match quality}} + \underbrace{g(y_j)}_{\text{vertical/quality}} + \underbrace{\varepsilon(i,j)}_{\text{idiosync.}},$$

and the utility to j is

$$v(i,j) = w(x_i, y_j) + h(x_i) + \eta(i,j),$$

Motivation

"potential function:"

$$w(x_i, y_j) + g(y_j) + h(x_i).$$

The utility to i is

$$u(i,j) = \underbrace{w(x_i, y_j) + g(y_j)}_{+h(x_i)=\text{potential}} + \varepsilon(i,j)$$

and the utility to j is

$$v(i,j) = \underbrace{w(x_i, y_j) + h(x_i)}_{+h(x_i)=\text{potential}} + \eta(i,j),$$

We show: in a large market, when idiosync. drawn iid, can match agents near perfectly on the idiosyncratic component.

So all the action is in the potential.

Motivation

"potential function:"

$$w(x_i, y_j) + g(y_j) + h(x_i).$$

The utility to i is

$$u(i,j) = \underbrace{w(x_i, y_j) + g(y_j)}_{+h(x_i) = \text{potential}}$$

and the utility to j is

$$v(i,j) = \underbrace{w(x_i, y_j) + h(x_i)}_{+h(x_i) = \text{potential}}$$

We show: in a large market, when idiosync. drawn iid, can match agents near perfectly on the idiosyncratic component.

So all the action is in the potential.

Aligned markets

Let $u(x, y)$ be the utility that a type x and a type y agent obtain when they match.

- Same utility from matching
- $\neq$ same preferences.
- Ex: a potential.
- Ex: distance (as in school choice):

Aligned markets

Ex: school choice

$$\underbrace{w(x_i, y_j)}_{\text{distance}} + \underbrace{g(y_j)}_{\text{school/student ``quality''}} + \underbrace{\varepsilon(i, j)}_{\text{idiosync.}}$$

Distance is key driver of preferences in school choice (Walters, 2018).

Main result: optimal transport

Measure inequality by means of a social welfare function.

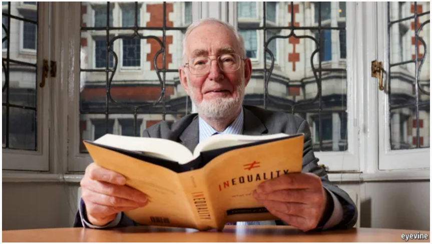
$A_\alpha(u)$

measure of inequality aversion: α.

If $\alpha < 0$, then inequality averse.

If $\alpha > 0$, then inequality loving.

(A_α is CRRA)



Anthony Atkinson

Main result: optimal transport

Optimization problem
w/parameter α.

$$\begin{array}{ll} \max & \int A_\alpha[u(x,y)]\,\mathrm{d}\pi \\ \text{s.t.} & \pi \text{ is a matching} \end{array}$$

Where A_α is Atkinson's social welfare function.
Atkinson's measure of inequality: $\alpha < 0$.

stability ($\alpha = +\infty$)
utilitarian welfare ($\alpha = 0$)
fairness ($\alpha = -\infty$)

Main result: take-aways

- Mkt. design focuses on stability: implicitly a planner who likes inequality.
- Uncover a new kind of unfairness: two arbitrarily similar agent on the same side of the market receive disparate outcomes.
- Result comes from standard techniques in OT. Handle finite and infinite markets.
- Argue that aligned preferences are interesting and useful for practical questions in mkt. design.

Related literature

- **Aligned Preferences:** decentralized dynamics (**Ferdowsian, Niederle, Yariv** 2020), random preferences (**Lee, Yariv** 2018), greedy algorithms and uniqueness for finite markets (**Eeckhout** 2000, **Clark** 2006, **Galichon, Ghelfi, Henry** 2023). Single-peaked preferences and uniqueness (**Flanders** 2018).
- **Large Markets:** **Azevedo, Leshno** (2016), **Ashlagi, Shi** (2016); **Lee** (2017); **Leshno, Lo** (2021), **Arnosti** (2022), **Greinecker, Kah** (2021).
- **Optimal Transport in Econ:** markets with transfers (**Galichon, Salanié** 2022, **Boerma, Tsyvinski, Wang, Zhang** 2023), mechanism design (**Daskalakis, Deckelbaum, Tzamos** 2015, **Kolesnikov, Sandomirskiy, Tsyvinski, Zimin** 2022, **Perez-Richet, Skreta** 2023), information design (**Malamud, Cieslak, Schrimpf** 2021, **Arieli, Babichenko, Sandomirskiy** 2023)

Warmup: Matching on the line

Matching on the line

Warning

This is just an example. It is not the paper.

Matching on the line

Model

- Agents are described by their "types" in $\mathbb{R}$
- Two sets of types $X \subset \mathbb{R}$ and $Y \subset \mathbb{R}$
- Two populations $\mu \in \Delta(X)$ and $\nu \in \Delta(Y)$
- Distance-based preferences: if $x \in X$ and $y \in Y$ match, each get utility

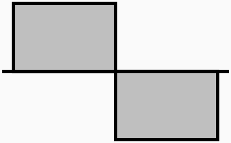
$$u(x, y) = -|x - y|$$

Matchings

Definition

$\pi \in \Delta(X \times Y)$ is a **matching** if it has marginal μ on X and ν on Y
Denote by $\Pi(\mu, \nu)$ the set of all matchings

Example: μ uniform on $[-1, 0]$, and ν uniform on $[0, 1]$



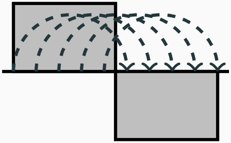
Matchings

Definition

$\pi \in \Delta(X \times Y)$ is a **matching** if it has marginal μ on X and ν on Y
Denote by $\Pi(\mu, \nu)$ the set of all matchings

Example: μ uniform on $[-1, 0]$, and ν uniform on $[0, 1]$

- π_f: $x \to y = x + 1$



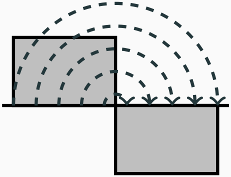
Matchings

Definition

$\pi \in \Delta(X \times Y)$ is a **matching** if it has marginal μ on X and ν on Y
Denote by $\Pi(\mu, \nu)$ the set of all matchings

Example: μ uniform on $[-1, 0]$, and ν uniform on $[0, 1]$

- π_f: $x \to y = x + 1$
- π_s: $x \to y = -x$



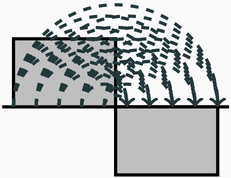
Matchings

Definition

$\pi \in \Delta(X \times Y)$ is a **matching** if it has marginal μ on X and ν on Y
Denote by $\Pi(\mu, \nu)$ the set of all matchings

Example: μ uniform on $[-1, 0]$, and ν uniform on $[0, 1]$

- π_f: $x \to y = x + 1$
- π_s: $x \to y = -x$
- random: $\pi = \mu \times \nu$



Stability

Definition

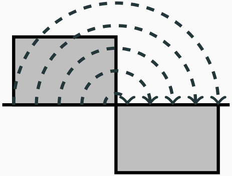
A matching π is **stable** if for any for any $(x, y),\ (x', y') \in \mathrm{supp}(\pi)$,

$$u(x, y') \le \max\left\{u(x, y),\ u(x', y')\right\}$$

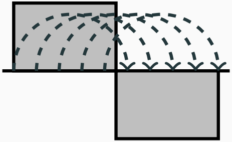
At least one member in the mismatched pair (x, y') prefers their current partner.
((x, y') is not a blocking pair.)

Stability

Example: μ uniform on $[-1, 0]$, and ν uniform on $[0, 1]$



stable



unstable

Note 1: π_s is stable and π_f is fair!

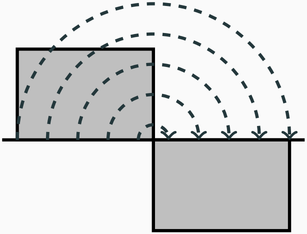
Note 2: For $u(x, y) = -|x - y|$, stability is related to **no-crossing**

No-crossing

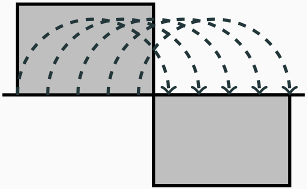
For interval $(z_1, z_2) \subset \mathbb{R}$, denote the circle in $\mathbb{R}^2$ w/interval as the diameter by $O(z_1, z_2)$

Definition

A matching π satisfies **no-crossing** if, for any $(x, y), (x', y') \in \mathrm{supp}(\pi)$, the circles $O(x, y)$ and $O(x', y')$ do not cross



satisfies no-crossing



violates no-crossing

Stability and no-crossing

Lemma

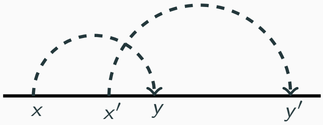
Any stable matching satisfies no-crossing

Stability and no-crossing

Lemma

Any stable matching satisfies no-crossing

Proof. We need to rule out the following two patterns in stable matching



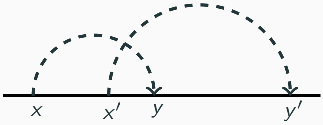
blocked by (x', y)

Stability and no-crossing

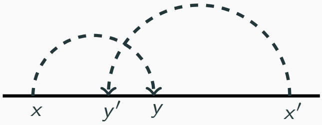
Lemma

Any stable matching satisfies no-crossing

Proof. We need to rule out the following two patterns in stable matching



blocked by (x', y)



blocked by (x, y') and (x', y)

$\square$

No-crossing II

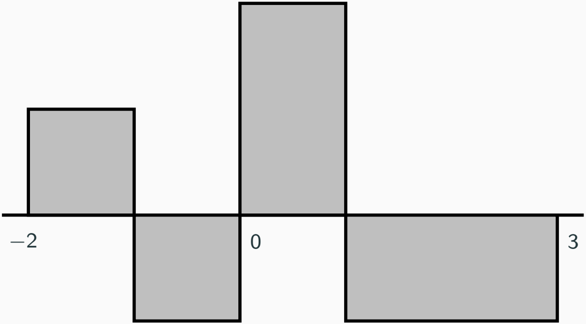
Structure of no-crossing matching (McCann 1999)

Consider μ, ν w/densities f and g.
Any no-crossing matching is a cvx. comb. of 2 deterministic matchings:

- Match $x = y$ as much as possible.
    - All common mass $h = \min\{f, g\}$ is eliminated
- No-crossing matchings of residual populations $(f - h)$ and $(g - h)$ form a finite number of parametric families

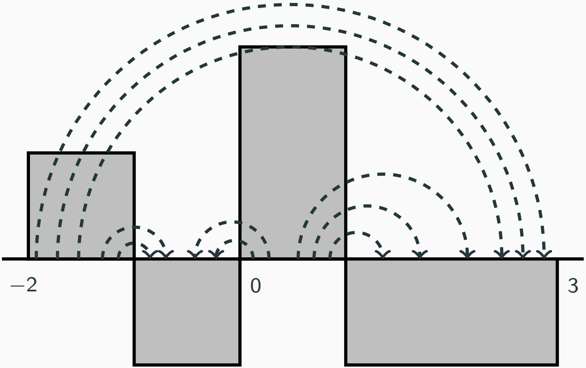
- The no-crossing condition makes the problem parametric!

An example



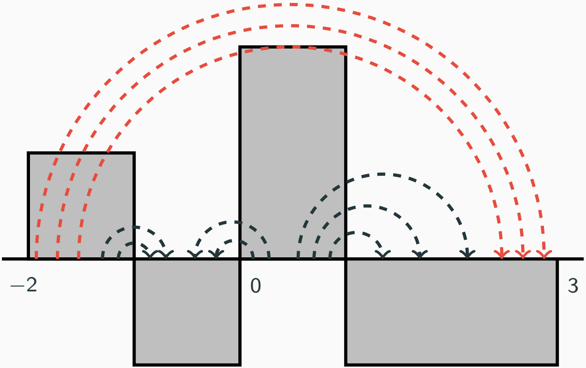
- No crossing matchings form a one-parametric family

An example



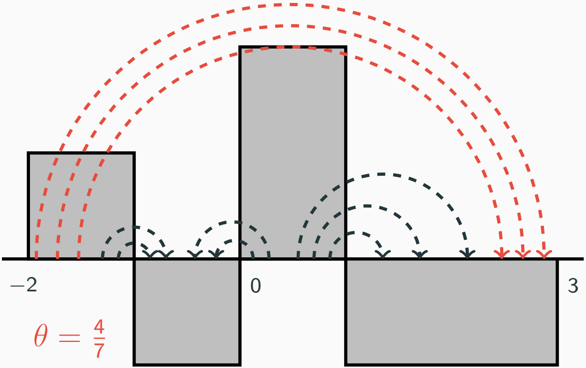
- No crossing matchings form a one-parametric family

An example



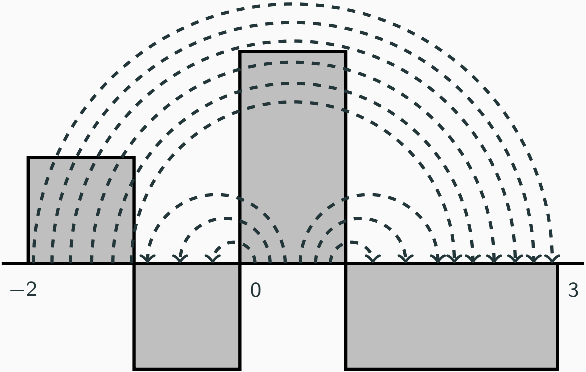
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally

An example



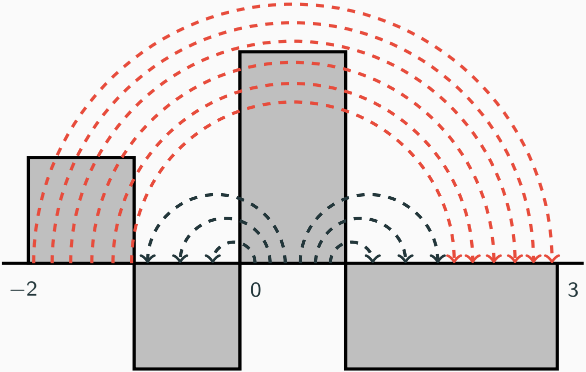
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally

An example



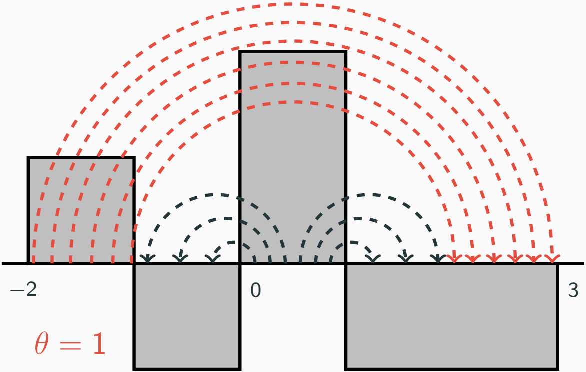
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally

An example



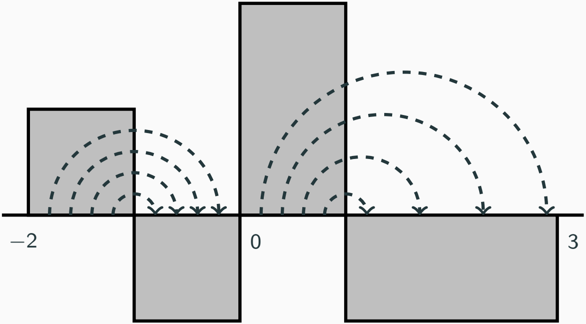
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally

An example



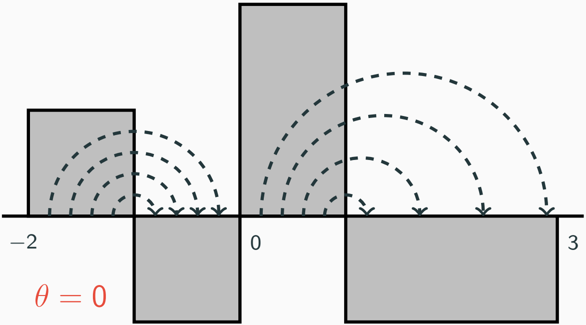
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally

An example



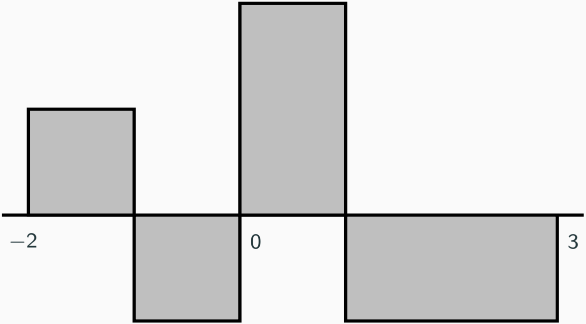
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally

An example



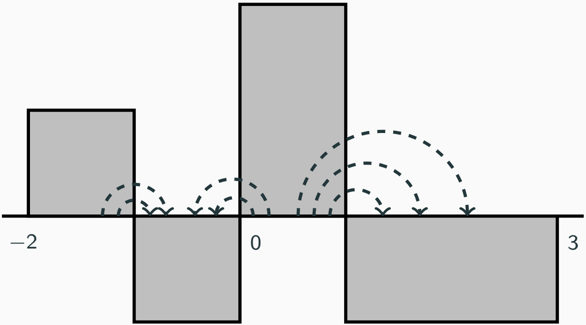
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally

An example



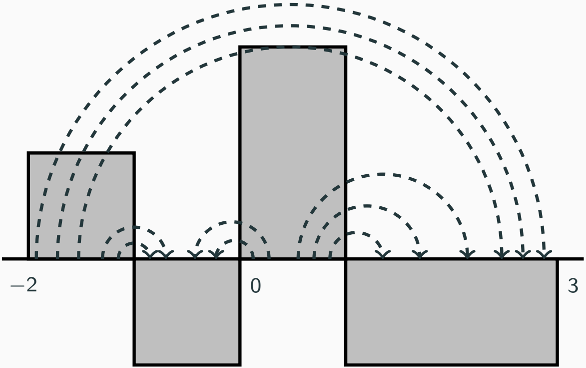
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally

An example



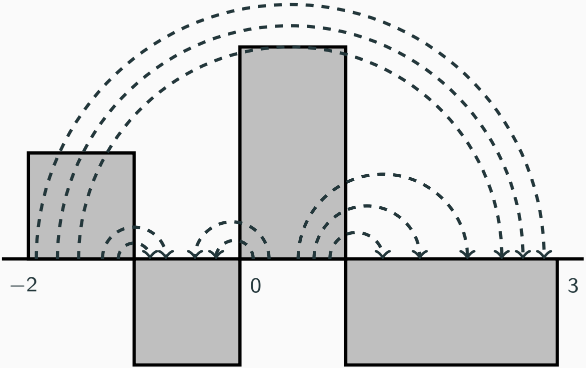
- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally
- Stable matching corresponds to $\theta = 4/7$

An example



- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally
- Stable matching corresponds to $\theta = 4/7$
- Non-local matches $\Rightarrow$ inequality & welfare loss. Quantify later

An example



- No crossing matchings form a one-parametric family
- $\theta \in [0, 1]$ is the fraction of the interval $[-2, -1]$ matched non-locally
- Stable matching corresponds to $\theta = 4/7$
- Non-local matches $\Rightarrow$ inequality & welfare loss. Quantify later
- **Angrist, Gray-Lobe, Idoux, Pathak (2022):** Deferred Acceptance in NYC and Boston $\Rightarrow$ 50% increase in travel expenditure

Stable matching on the line

Corollaries of no-crossing:

- A stable matching=a convex combination of two deterministic ones:
  x is matched with the ideal partner $y = x$ or at most one other y'
- It can be searched for within a finite number of parametric families

Bad news: The number of families blows up exponentially with the number of times $\mu - \nu$ changes sign

Stable matching on the line

Proposition

For non-atomic $\mu, \nu \in \Delta(\mathbb{R})$, a stable matching exists and is unique, and can be constructed via a simple algorithm. For piecewise-constant densities with m intervals of constancy, it requires $O(m^2)$ operations

Proof idea. Find a "simple independent submarket"

- is to be matched independently of the rest of the population
- a no-crossing matching is unique and thus is stable
- after eliminating, the number of sign changes decreases by 1

Repeat $\square$

Optimal transport and general markets with aligned preferences

Optimal transport

Given:

- measurable spaces X and Y;
- distributions $\mu \in \Delta(X),\ \nu \in \Delta(Y)$;
- payoff $p\colon X \times Y \to \mathbb{R}$.

General optimal transport problem

$$\max_{\pi \in \Pi(\mu,\nu)} \int_{X \times Y} p(x,y)\ \mathrm{d}\pi(x,y)$$

Optimal transport

- Often formulated for cost minimization ($c = -p$)
- **Standard interpretation:** μ and ν are spatial distributions of production and demand; π is cheapest way to transport supplied quantities to satisfy demands.

—**McCann (1990):** $X, Y \subset \mathbb{R}$, convex $p \Rightarrow$ no-crossing π
—Stability for $\mathbb{R}$ and $u(x, y) = -|x - y| \ \Rightarrow$ no crossing π

Question: Any direct connection between stability and transport?

Yes, and it is not limited to $\mathbb{R}$ and distance-based utility

General markets with aligned preferences

Model

- X and Y are Polish spaces with Borel σ-algebra.
- Two populations $\mu \in \Delta(X)$ and $\nu \in \Delta(Y)$.
- If x and y match, both obtain utility/payoff $u(x, y)$.
- Assume $u\colon X \times Y \to \mathbb{R}$ is cont. and bounded.
    - often, measurability is enough (in the paper)
    - “acyclicity” of ordinal preferences $\Rightarrow$ existence of u (in the paper)

Designer's objectives

Criteria for matchings

- Approximate stability.
- Approximate egalitarianism.
- Utilitarian welfare.

Approximate stability

Definition

A matching π is ε-**stable** with $\varepsilon \geq 0$ if for any $(x, y), (x', y') \in \text{supp}(\pi)$,

$$u(x, y') \leq \max\left\{u(x, y),\ u(x', y')\right\} + \varepsilon$$

- At least one partner in any mismatched pair can't benefit by $> \varepsilon$ from leaving current partner.
- for $\varepsilon = 0$, get the usual notion of stability.
- ε-stability $\simeq$ stability in the presence of ε-friction

Approximate egalitarianism

- For each matching $\pi \in \Pi(\mu, \nu)$ define

$$U_{\min}(\pi) = \min_{(x,y) \in \mathrm{supp}(\pi)} u(x, y)$$

  - Well-defined for compact X and Y
  - For non-compact, replace minimum with infimum
- Egalitarian lower bound

$$U^*_{\min}(\mu, \nu) = \max_{\pi \in \Pi(\mu, \nu)} U_{\min}(\pi)$$

Approximate egalitarianism

Definition

A matching $\pi \in \Pi(\mu, \nu)$ is ε-**egalitarian** if there is a subset $S \subset X \times Y$ with $\pi(S) \geq 1 - \varepsilon$ such that

$$u(x, y) \geq U^*_{\min}(\mu, \nu) - \varepsilon \qquad \text{for all} \quad (x, y) \in S$$

- All agents except ε-fraction have utilities above the ε-relaxed egalitarian bound

Utilitarianism

- The utilitarian welfare of a matching π by

$$W(\pi) = \int_{X \times Y} u(x,y) \, \mathrm{d}\pi(x,y)$$

- Optimal welfare

$$W^*(\mu,\nu) = \max_{\pi \in \Pi(\mu,\nu)} W(\pi)$$

- Welfare-max. $\simeq$ opt. transport w/payoff $p = u$
- The other objectives correspond to p equal to a **transformation** of u

Utility transformation

For a matching market with utility u, define the transformation

$$p_\alpha(x, y) = \frac{\exp(\alpha \cdot u(x, y)) - 1}{\alpha}$$

- p_α is convex in u for $\alpha > 0$ and concave for $\alpha < 0$
- for $\alpha \to 0$, the limit is $p_0(x, y) = u(x, y)$

Main result

Consider the transportation problem with payoff p_α

$$\max_{\pi \in \Pi(\mu,\nu)} \int_{X \times Y} p_\alpha(x, y) \, \mathrm{d}\pi(x, y)$$

- For $\alpha = 0$, this is utilitarian welfare-maximization
- What do we get for $\alpha \neq 0$?

Theorem proof

Let π be a solution to the optimal transport problem with payoff p_α.

- If $\alpha > 0$ then π is ε-stable, with $\varepsilon = (\ln 2)/\alpha$.
- If $\alpha < 0$ then π is ε-egalitarian, with $\varepsilon = \max\{1, \ln|\alpha|\}/|\alpha|$

Main result

Implications:

- Changing α, we interpolate between the three objectives:

  fairness ($\alpha = -\infty$), welfare ($\alpha = 0$), stability ($\alpha = +\infty$)
- Fairness and stability are on the opposite sides of the spectrum
- Provides stability with a (perhaps unintentional) social welfare objective: a convex Atkinson inequality index.
- Stability with aligned preferences $\simeq$ an **inequality-loving designer** prioritizing high-utility agents & ignoring externalities on low-utility agents.

The result extends to k-sided markets: replace "$\ln 2$" with "$\ln k$"

Holds for $h \circ u$ with $\frac{h'}{h} \geq \alpha$.

Main result

Let $\Pi^{u}_{+\infty}(\mu, \nu)$ be the set of matchings π that can be obtained as the weak limit $\pi = \lim_{n \to +\infty} \pi_{\alpha_n}$ of sequences of solutions π_{α_n} to the transportation problem for some seq. $\alpha_n \to +\infty$.

Define $\Pi^{u}_{-\infty}(\mu, \nu)$ to be the weak limits for some seq. $\alpha_n \to -\infty$.

Corollary

For continuous and bounded utility u, the sets $\Pi^{u}_{+\infty}(\mu, \nu)$ and $\Pi^{u}_{-\infty}(\mu, \nu)$ are non-empty, convex, and weakly closed. All matchings in $\Pi^{u}_{+\infty}(\mu, \nu)$ are stable, and all matchings in $\Pi^{u}_{-\infty}(\mu, \nu)$ are egalitarian.

Welfare and fairness of stable matchings

Theorem proof

If utility $u \geq 0$ and a matching π is ε-stable, then

- π guarantees approximately half of optimal welfare:
$$W(\pi) \geq \frac{1}{2}\left(W^*(\mu,\nu) - \varepsilon\right)$$
- π is ε'-egalitarian with $\varepsilon' = \max\{1/2,\ \varepsilon\}$

- Any stable matching guarantees $1/2$ of the optimal welfare and is $1/2$-egalitarian
- These conservative bounds are concerned with ε-stable matchings with lowest welfare or that are least egalitarian
- "2" is the number of sides of the market

What if the market is not aligned?

Non-aligned markets

First, a very simple point.

If each x and y's utility from matching is within $\varepsilon > 0$ of an aligned utility $u(x, y)$,

then any matching that is ε-stable for the aligned market is 3ε stable in the non-aligned market.

So an an approximately stable matching remains approximately stable for nearby non-aligned markets.

Non-aligned preferences

Let $X = Y = \mathbb{R}$. Assume non-atomic distributions $\mu \in \Delta(X)$ and $\nu \in \Delta(Y)$.

Two finite populations: $X_n = \{x_1, \ldots, x_n\} \subset X$ and $Y_n \in \{y_1, \ldots, y_n\} \subset Y$.

Assume X_n and Y_n are i.i.d. samples from μ and ν.

If a pair $(x_i, y_j) \in X_n \times Y_n$ is formed, agents i and j enjoy utilities

$$u_{i,j} = w(x_i, y_j) + \xi_{i,j} \qquad \text{and} \qquad v_{i,j} = w(x_i, y_j) + \eta_{i,j}.$$

$w\colon X \times Y \to \mathbb{R}$ is a continuous function capturing the aligned component of agents' preferences.

Idiosyncratic components $\xi_{i,j}$ and $\eta_{i,j}$ are independent shocks with cont. dist. F_i and G_j.

Non-aligned preferences

Notation:

Let π_n be a deterministic matching of X_n and Y_n.

Then,

$$\pi_n\big([a,b]\times[c,d]\big) = \frac{\Big\{(x_i, y_j) \in [a,b] \times [c,d] \,:\, x_i \text{ and } y_j \text{ are matched in } \pi_n\Big\}}{n}$$

Non-aligned preferences

Theorem

For $\pi \in \Pi(\mu, \nu)$, $\exists$ sequence $\delta_n \to 0$ s.t., with prob. $\geq 1 - \delta_n$,
$\exists$ a deterministic π_n with

$$\Big|\pi_n\big([a,b] \times [c,d]\big) - \pi\big([a,b] \times [c,d]\big)\Big| \leq \delta_n$$

for all $[a,b] \subseteq X$, $[c,d] \subseteq Y$.
Moreover, for all x_i and y_j matched under π_n,

$$F_i\big((-\infty, \xi_{i,j}]\big) \;\geq\; 1 - \delta_n \quad \text{and} \quad G_j\big((-\infty, \eta_{i,j}]\big) \;\geq\; 1 - \delta_n.$$

In a large market, any matching is, with high probability, close to a matching in which all agents' idiosyncratic match utilities are high (at quantile close to 1).
Result: related to SM Lee's work.

Application: school choice

Application: school choice

School choice: matching students to schools.

- Distance is a key component of student preferences (Walters, 2018).
- Distance is a key component of school preferences (priorities).
- Aligned distance-based preferences is a good approximation.

Application: school choice

Suppose that:

- Preferences have a distance and a "vertical" component.
- Students care about distance to school, and school quality q_s.
- Schools care about distance, and student achievements q_i.
- Additively.

Let $u(i,s) = -d(i,s) + f(q_s) + g(q_i)$.

Then,

$$u(i,s) - u(i,s') = d(i,s') - d(i,s) + f(q_s) - f(q_{s'}) \text{ and}$$
$$u(i,s) - u(j,s) = d(j,s) - d(i,s) + g(q_i) - g(q_j),$$

Hence, aligned preferences.

Application: school choice

Implications:

- We replicate some stylized facts.
- Increase in travel times after district switch to deferred acceptance.
- Unfairness in travel times.
- (Angrist et al 2022)

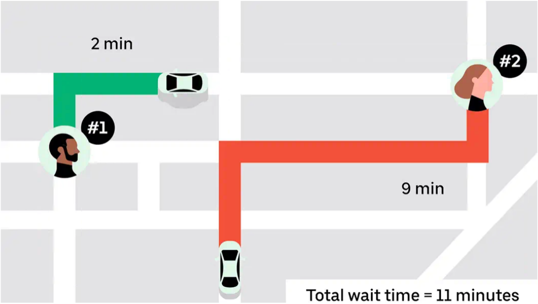
And is the objective really what we want to maximize? (Note this is a question we couldn't even ask without our results.)

Application: Ride-sharing

Matching

Uber Marketplace

In the seconds after a rider requests a ride, we evaluate nearby drivers and riders in one batch. We then pair riders and drivers in the distribution, aiming to reduce the average wait time for everyone, not just the closest pair. This helps keep things moving and rides reliable across the network.



First to request

In the early days, a rider was immediately matched with the closest available driver. It worked well for most riders but sometimes led to long wait times for others. Across a whole city, those longer wait times really added up.

Application: Bargaining with transfers and no comittment

Bargaining

Markets with transfers but lack of commitment power

- Becker's (1973) marriage market model:
  - A couple (x, y) generates surplus $s(x, y)$ and can share it as
  $$s(x, y) = \hat{u}(x, y) + \hat{v}(x, y)$$
  - Shares $\hat{u}(x, y)$ and $\hat{v}(x, y)$ are determined at the time of the match
- Transfers are negotiated and committed to, as part of the bargaining over the match
- **Question:** What if no commitment power?
- Partners use Nash bargaining with weights (1/2, 1/2) to split surplus after the match is formed
- Aligned preferences with $u(x, y) = s(x, y)/2$

Distance-based matching in $\mathbb{R}^d$

Distance-based matching in $\mathbb{R}^d$: fairness-welfare tension

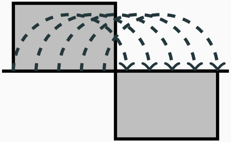
- $X = Y = \mathbb{R}^d$, utility $u(x, y) = -\|x - y\|$
- The payoff

$$p_\alpha(x, y) = \frac{\exp(\alpha \cdot \|x - y\|) - 1}{\alpha}$$

  is convex in the distance for $\alpha > 0$ and concave for $\alpha < 0$
- Optimal transport with $p(x, y) = f(\|x - y\|)$ is well-understood for convex/concave f

Let's focus on $d = 1$:

- Concave f $\Rightarrow$ assortative matching
- Thus $\alpha < 0 \Rightarrow$ assortative matching



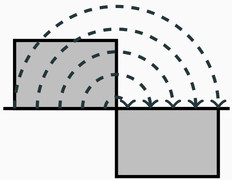
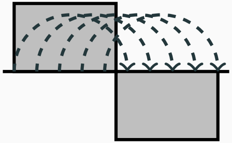
Corollary

For $d = 1$, there is no fairness-welfare tension. Both objectives are attained by the assortative matching.

- For $d > 1$, fairness-welfare tension emerges

Distance-based matching in $\mathbb{R}^d$: stability-fairness tension

Is there stability-fairness tension for $d = 1$? **Yes**

Both have the same welfare. Maybe there is no stability-welfare tension?

Distance-based matching in $\mathbb{R}^d$: stability-welfare tension

McCann (1999):

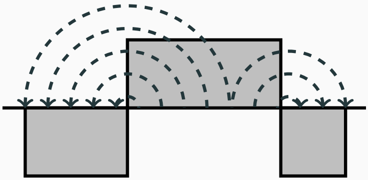
- For $d = 1$ and $p(x, y) = f(|x - y|)$ with strictly convex f, the optimal matching satisfies no-crossing
- If $\mu - \nu$ changes sign at most twice, a no-crossing matching is unique

For $\alpha > 0$ and ≤ 2 sign changes, the optimum does not depend on α

Corollary

If $\mu - \nu$ changes sign at most twice, there is no stability-welfare tension

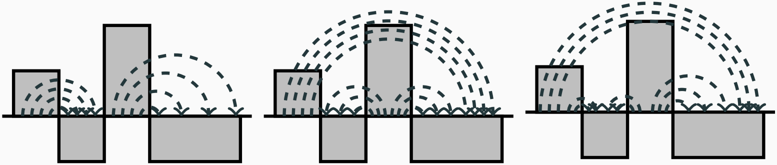
Example:



The conclusion extends to a round city in $\mathbb{R}^2$

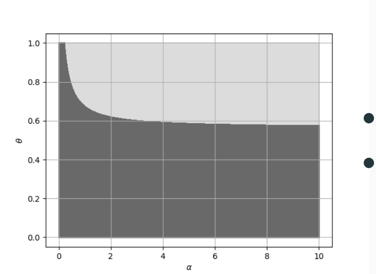
Distance-based matching in $\mathbb{R}^d$: stability-welfare tension II

If there are ≥ 3 sign changes, stability-welfare tension emerges



$\theta = 0$ $\theta = 1$ $\theta = 4/7$

The optimal θ depends on α in the optimal transport problem



- stability $\Rightarrow \theta = 4/7 \approx 0.57$
- welfare-maximization $\Rightarrow \theta = 1$

Conclusion

- Aligned preferences emerge when
  - match quality is common to both sides (distance in school choice)
  - there are transfers but no commitment power
- Connection to transport: a parametric family of objectives captures

  stability ($\alpha = +\infty$), welfare ($\alpha = 0$), fairness ($\alpha = -\infty$)
- Stability $\simeq$ prioritizing high-utility matches over low-utility ones
- Welfare and fairness losses, at most $1/2$ of each
- For particular spatial distributions no loss in welfare
  - Stability is OK if low-utility agents are compensated

Thank you!

Proof of Theorem 1: ε-stability as transport back to theorem

Definition: Given $p\colon X \times Y \to \mathbb{R}$, a set $\Gamma \subset X \times Y$ is p-**cyclic monotone** if

$$\sum_{i=1}^{n} p(x_i, y_i) \geq \sum_{i=1}^{n} p(x_i, y_{i+1})$$

for all $(x_1, y_1), \ldots, (x_n, y_n) \in \Gamma$ with $y_{n+1} = y_1$

Proof of Theorem 1: ε-stability as transport

Theorem (Beiglbock, Goldstern, Maresch, Schachermayer 2009)

If π solves an optimal transport problem

$$\pi \in \Pi(\mu, \nu) \quad : \quad \int_{X \times Y} p(x, y) \, \mathrm{d}\pi(x, y) \quad \to \quad \max,$$

then $\mathrm{supp}(\pi)$ is p-cyclic monotone

Proof of Theorem 1: ε-stability as transport

Use cyclic monotonicity for

$$p_\alpha(x, y) = \exp\left(\alpha \cdot u(x, y)\right)$$

On the support of the optimal matching π,

$$p_\alpha(x_1, y_2) + p_\alpha(x_2, y_1) \leq p_\alpha(x_1, y_1) + p_\alpha(x_2, y_2)$$

Equivalently,

$$\exp\left(\alpha \cdot u(x_1, y_2)\right) + \exp\left(\alpha \cdot u(x_2, y_1)\right) \leq \exp\left(\alpha \cdot u(x_1, y_1)\right) + \exp\left(\alpha \cdot u(x_2, y_2)\right)$$

Proof

Drop the second term on the LHS

$$\exp\left(\alpha \cdot u(x_1, y_2)\right) \leq \exp\left(\alpha \cdot u(x_1, y_1)\right) + \exp\left(\alpha \cdot u(x_2, y_2)\right)$$

Proof

Drop the second term on the LHS

$$\begin{aligned}\exp\left(\alpha \cdot u(x_1, y_2)\right) &\le \exp\left(\alpha \cdot u(x_1, y_1)\right) + \exp\left(\alpha \cdot u(x_2, y_2)\right) \\ &\le 2 \cdot \max\left\{ \exp\left(\alpha \cdot u(x_1, y_1)\right),\ \ \exp\left(\alpha \cdot u(x_2, y_2)\right)\right\}\end{aligned}$$

Proof

Drop the second term on the LHS

$$\begin{aligned} \exp\left(\alpha \cdot u(x_1, y_2)\right) &\leq \exp\left(\alpha \cdot u(x_1, y_1)\right) + \exp\left(\alpha \cdot u(x_2, y_2)\right) \\ &\leq 2 \cdot \max\left\{ \exp\left(\alpha \cdot u(x_1, y_1)\right), \quad \exp\left(\alpha \cdot u(x_2, y_2)\right)\right\} \\ &= 2 \cdot \exp\Big(\alpha \cdot \max\{u(x_1, y_1), \quad u(x_2, y_2)\}\Big) \end{aligned}$$

Proof

Drop the second term on the LHS

$$\begin{aligned}
\exp\left(\alpha \cdot u(x_1, y_2)\right) &\leq \exp\left(\alpha \cdot u(x_1, y_1)\right) + \exp\left(\alpha \cdot u(x_2, y_2)\right) \\
&\leq 2 \cdot \max\left\{ \exp\left(\alpha \cdot u(x_1, y_1)\right), \quad \exp\left(\alpha \cdot u(x_2, y_2)\right) \right\} \\
&= 2 \cdot \exp\Big(\alpha \cdot \max\{u(x_1, y_1), \quad u(x_2, y_2)\}\Big)
\end{aligned}$$

Take logarithm and divide by α

$$u(x_1, y_2) \leq \max\left\{u(x_1, y_1), \quad u(x_2, y_2)\right\} + \frac{\ln(2)}{\alpha} \qquad \square$$

Proof of Theorem 2 back to theorem

Let π be an ε-stable matching.

For any $(x_1, y_1), (x_2, y_2) \in \text{supp}(\pi)$,

$$u(x_1, y_2) \leq \max\{u(x_1, y_1),\ u(x_2, y_2)\} + \varepsilon.$$

By non-negativity of u, we get

$$u(x_1, y_2) \leq u(x_1, y_1) + u(x_2, y_2) + \varepsilon.$$

Proof of Theorem 2

Let π' be any other matching with marginals μ and ν.

Consider $\lambda \in \mathcal{M}_+\big((X \times Y) \times (X \times Y)\big)$ s.t. the marginals of λ on (x_1, y_1) and on (x_2, y_2) are equal to π and the marginal on (x_1, y_2) is π'.

We get

$$\begin{aligned} W(\pi') &= \int_{X \times Y} u(x_1, y_2)\ \mathrm{d}\pi'(x_1, y_2) = \int_{(X\times Y)\times(X\times Y)} u(x_1, y_2)\ \mathrm{d}\lambda(x_1, y_1, x_2, y_2) \\ &\le \int_{(X\times Y)\times(X\times Y)} \left(u(x_1, y_1) + u(x_2, y_2) + \varepsilon\right)\ \mathrm{d}\lambda(x_1, y_1, x_2, y_2) = \\ &= \int_{X \times Y} u(x_1, y_1)\ \mathrm{d}\pi(x_1, y_1) + \int_{X \times Y} u(x_2, y_2)\ \mathrm{d}\pi(x_2, y_2) + \varepsilon = \\ &= 2W(\pi) + \varepsilon. \end{aligned}$$

So:

$$W(\pi) \ge \frac{1}{2}\left(W(\pi') - \varepsilon\right)$$

for any matching π'. In particular, this inequality holds for π' maximizing welfare. Thus $W(\pi) \ge \frac{1}{2}\left(W^*(\mu, \nu) - \varepsilon\right)$.

Existence of a potential

Definition

A weak order (a complete and transitive binary relation) is termed a **preference**. If the weak order $\succeq$ is over a topological space Z, then we say that it is **continuous** if the upper contour sets $U_{\succeq}(z) = \{z' \in Z : z' \succ z\}$ and lower contour sets $L_{\succeq}(z) = \{z' \in Z : z' \prec z\}$ are open.

Existence of a potential

Primitives are a tuple $(X, Y, \succeq_X, \succeq_Y)$ in which:

- X and Y are topological spaces.
- $\succeq_X = \{\succeq_x : x \in X\}$, where for each $x \in X$, $\succeq_x$ is a continuous preference on Y;
- $\succeq_Y = \{\succeq_y : y \in Y\}$ for each $y \in Y$, $\succeq_y$ is a continuous preference on X.

A function $u \colon X \times Y \to \mathbb{R}$ is a **potential** for $(X, Y, \succeq_X, \succeq_Y)$ if

- $u(x, y) \geq u(x, y')$ iff $y \succeq_x y'$ for all x, y, y'
- and $u(x, y) \geq u(x', y)$ iff $x \succeq_y x'$ for all x, y, x'

Existence of a potential

The environment $(X, Y, \succeq_X, \succeq_Y)$ is **acyclic** if, for any sequence of couples,

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n),$$

with $n > 2$ and $(x_n, y_n) = (x_1, y_1)$, so that each couple (x_{i+1}, y_{i+1}) has exactly one agent in common with the preceding couple (x_i, y_i), whenever all the common agents prefer their partner in (x_{i+1}, y_{i+1}) to their partner in (x_i, y_i), all common agents are, in fact, indifferent between the two partners

Existence of a potential

1. **Continuity with respect to the agent:** If $b \succ_a b'$ then there is a neighborhood N_a of a for which $b \succ_c b'$ for any $c \in N_a$
2. **Local strictness:** If $b' \succeq_a b$ and $b \succeq_{a'} b''$ with $a \neq a'$ and $b \neq b', b''$, then, in any neighborhood of b, there exists $\hat{b}$ with $b' \succ_a \hat{b}$ and $\hat{b} \succ_{a'} b''$

Theorem

Let $(X, Y, \succeq_X, \succeq_Y)$ be such that X and Y are complete, separable and connected topological spaces. Suppose that acyclicity and properties (1) and (2) are satisfied. Then there is a potential for $(X, Y, \succeq_X, \succeq_Y)$.