Evaluating with Statistics:

Which Outcome Measures Differentiate Among Matching Mechanisms?

Marek Pycia (U Zurich)

Allocation of school seats to students

Canonical case of allocation of resources without transfers (cf. also refugee resettlement, teacher assignment, etc).

Australia, Chile, China, Finland, France, Ghana, Hungary, Ireland, the Netherlands, Norway, Poland, Romania, Spain, Taiwan, Turkey, US, UK, etc.

Allocation mechanisms rely on rankings of schools provided by applicants.

The Contentious Choice of Mechanism

- Street protests in France and Taiwan.
- Frequent changes or discussed changes of mechanisms (Chicago, Boston, NYC, New Orleans).

The Contentious Choice of Mechanism

- Street protests in France and Taiwan.
- Frequent changes or discussed changes of mechanisms (Chicago, Boston, NYC, New Orleans).

Literature proposing school choice mechanisms:

- Balinski and Sonmez (1999)
- Abdulkadriglu and Sonmez (2003)
- Abdulkadiroglu, Che, and Yasuda (2015)
- Kesten (2010), Morrill (2014), Hakimov and Kesten (2015), Delacretaz, Kloosterman, and Troyan (2019)
- Pycia and Unver (2011)
- Budish, Che, Kojima, Milgrom (2013)
- Ashlagi and Shi (2014)
- He, Miralles, Pycia, and Yan (2018)
- Nguyen, Peivandi, and Vohra (2017)

Puzzle: Many Mechanisms Nearly Identical in the Data

New Orleans statistics from Abdulkadiroglu, Che, Pathak, Roth and Tercieux (2017):

	TTC	SD
1	772	777
2	126	121
3	46	44
4	18	17
5+	11	8
Unassigned	222	228

Puzzle: Many Mechanisms Nearly Identical in the Data

New Orleans statistics from Abdulkadiroglu, Che, Pathak, Roth and Tercieux (2017):

	TTC	SD
1	772	777
2	126	121
3	46	44
4	18	17
5+	11	8
Unassigned	222	228

Different mechanisms also give very similar standard statistics in:

- Amsterdam (de Haan, Gautier, Oosterbeek, and van der Klaauv 2015).
- Boston (Abdulkadiroglu, Che, Pathak, Roth and Tercieux 2017)
- New York (Abdulkadiroglu, Pathak, and Roth 2009, Abdulkadiroglu, Agarwal, and Pathak 2015).
- Teacher Assignment in France (Combe, Tercieux, Terrier 2017).

Not All Statistics Are Nearly the Same

- Improvements over reference mechanism
  - Calsamiglia and Miralles 2012, He 2012, Agarwal and Somaini 2016.
- Violations of stability
  - Kesten 2010, Abdulkadiroglu, Che, Pathak, and Roth 2017.

Theory?

Symmetric Mechanisms

- Abdulkadiroglu and Sonmez (1998), Knuth (1996), Pathak and Sethuraman (2011), Carroll (2014), Lee and Sethuraman (2011)
- Che and Kojima (2010), Miralles (2008)
- Liu and Pycia (2016), Pycia (2011)
- Ashlagi, Kanoria, Leshno (2017)
- Pycia and Troyan (2019)

Theory?

Symmetric Mechanisms

- Abdulkadiroglu and Sonmez (1998), Knuth (1996), Pathak and Sethuraman (2011), Carroll (2014), Lee and Sethuraman (2011)
- Che and Kojima (2010), Miralles (2008)
- Liu and Pycia (2016), Pycia (2011)
- Ashlagi, Kanoria, Leshno (2017)
- Pycia and Troyan (2019)

Asymptotic Population Mean Distribution of Expected Payoffs

- Che and Tercieux (2018)

Preview of Findings

- Any **anonymous** statistics is asymptotically the same for all standard Pareto-efficient mechanisms.
  - An analogue holds true for stable constrained-efficient mechanisms.
  - **Converse**: the asymptotic equivalence requires at least asymptotic anonymity of the statistics.

Preview of Findings

- Any **anonymous** statistics is asymptotically the same for all standard Pareto-efficient mechanisms.
  - An analogue holds true for stable constrained-efficient mechanisms.
  - **Converse**: the asymptotic equivalence requires at least asymptotic anonymity of the statistics.
- Equivalence bounds in realistic size markets.
- The means and medians of anonymous statistics are **exactly identical** when averaged over exchangeable distributions of preferences (in **any size market**).

Model

A – finite set of schools; each school $a \in A$ has $|a| > 0$ seats.

N – finite set of agents; each agent i demands a single seat and has a strict preference ranking $\succ_i$ over schools.

Θ – the set of preference rankings.

Model

A – finite set of schools; each school $a \in A$ has $|a| > 0$ seats.

N – finite set of agents; each agent i demands a single seat and has a strict preference ranking $\succ_i$ over schools.

Θ – the set of preference rankings.

An *allocation* μ specifies the school $\mu(i)$ assigned to agents $i \in N$.
μ is *Pareto efficient* = no other allocation is weakly better for all and strictly better for at least one agent.

Model

A – finite set of schools; each school $a \in A$ has $|a| > 0$ seats.

N – finite set of agents; each agent i demands a single seat and has a strict preference ranking $\succ_i$ over schools.

Θ – the set of preference rankings.

An *allocation* μ specifies the school $\mu(i)$ assigned to agents $i \in N$.
μ is *Pareto efficient* = no other allocation is weakly better for all and strictly better for at least one agent.

A *mechanism* ϕ maps profiles of rankings to allocations (or lotteries over allocations);

strategy-proof = reporting true ranking is dominant.

Example

Schools a, b, c; each school has $n \in \{1, 2, ...\}$ seats.

$3n$ agents $1, 2, ..., 3n$; each ranks schools $a \succ b \succ c$.

Serial Dictatorship with an ordering of agents:

- assigns to the first agent his/her most preferred school,
- assigns to the second agent his/her most preferred school with available seats, etc.

Example

Schools a, b, c; each school has $n \in \{1, 2, ...\}$ seats.

$3n$ agents $1, 2, ..., 3n$; each ranks schools $a \succ b \succ c$.

Serial Dictatorship with an ordering of agents:

- assigns to the first agent his/her most preferred school,
- assigns to the second agent his/her most preferred school with available seats, etc.

Anonymous statistics: What fraction of agents obtain their most preferred outcome?

- $\frac{1}{3}$ under any ordering.

Example

Schools a, b, c; each school has $n \in \{1, 2, ...\}$ seats.

$3n$ agents $1, 2, ..., 3n$; each ranks schools $a \succ b \succ c$.

Serial Dictatorship with an ordering of agents:

- assigns to the first agent his/her most preferred school,
- assigns to the second agent his/her most preferred school with available seats, etc.

Anonymous statistics: What fraction of agents obtain their most preferred outcome?

- $\frac{1}{3}$ under any ordering.

Non-anonymous statistics: What fraction of agents strictly improve over the outcome under the ordering $1, 2, ..., 3n$?

- 0 if we stick to this ordering;
- $\frac{1}{3}$ if we use the ordering $3n, 3n-1, ..., 1$.

Outcome Statistics

$K = \{1, ..., k\}$ set of codes; $k \geq 2$.

$f : N \times \Theta \times A \to K$ coding function.

Aggregate statistics $F : (\Theta \times A)_{i \in N} \to [0, 1]^K$ is an empirical distribution of f.

F is *anonymous* iff $f(i, \succ, a) = f(j, \succ, a)$ for all $i, j, \succ, a$.

Outcome Statistics

$K = \{1, ..., k\}$ set of codes; $k \geq 2$.

$f : N \times \Theta \times A \rightarrow K$ coding function.

Aggregate statistics $F : (\Theta \times A)_{i \in N} \rightarrow [0,1]^K$ is an empirical distribution of f.

F is *anonymous* iff $f(i, \succ, a) = f(j, \succ, a)$ for all $i, j, \succ, a$.

Examples:

- How many students obtain their top outcome, their two top outcomes, etc.
- The empirical distribution of ranks.
- How many students are assigned to school A, school B, etc.

Positive Equivalence: Bounds

$\mathcal{M}$ = Fixed-Endowment Hierarchical Exchange for School Choice (Papai 2000, Abdulkadiroglu and Sonmez 2003, Pycia and Unver 2011).

$\mathbb{P}$: a distribution on preference profiles that is iid across agents.

Theorem. If $\phi, \psi \in \mathcal{M}$ and statistics F is anonymous then:

$$\mathbb{P}\left(\left|F_\ell\left(\succ, \phi\left(\succ\right)\right) - F_\ell\left(\succ, \psi\left(\succ\right)\right)\right| \leq \epsilon\right) \geq 1 - 8\,\text{exp}\left(-\frac{\epsilon^2 N}{4|A|^2}\right),\ \forall \ell \in K.$$

$$\mathbb{P}\left(\sum_{\ell=1}^{|K|}\left|F_\ell\left(\succ, \phi\left(\succ\right)\right) - F_\ell\left(\succ, \psi\left(\succ\right)\right)\right| \leq \epsilon\right) \geq 1 - 8\,\text{exp}\left(-\frac{\epsilon^2 N}{16|A|^2}\right).$$

Positive Equivalence: Bounds

$\mathcal{M}$ = Fixed-Endowment Hierarchical Exchange for School Choice (Papai 2000, Abdulkadiroglu and Sonmez 2003, Pycia and Unver 2011).

$\mathbb{P}$: a distribution on preference profiles that is iid across agents.

Theorem. If $\phi, \psi \in \mathcal{M}$ and statistics F is anonymous then:

$$\mathbb{P}\left(\left|F_\ell\left(\succ, \phi\left(\succ\right)\right) - F_\ell\left(\succ, \psi\left(\succ\right)\right)\right| \leq \epsilon\right) \geq 1 - 8\,\text{exp}\left(-\frac{\epsilon^2 N}{4|A|^2}\right), \ \forall \ell \in K.$$

$$\mathbb{P}\left(\sum_{\ell=1}^{|K|}\left|F_\ell\left(\succ, \phi\left(\succ\right)\right) - F_\ell\left(\succ, \psi\left(\succ\right)\right)\right| \leq \epsilon\right) \geq 1 - 8\,\text{exp}\left(-\frac{\epsilon^2 N}{16|A|^2}\right).$$

Similar bounds for:

- Hierarchical Exchange for School Choice (not only fixed endowment),
- Pareto efficient and Li's 2017 OSP mechanisms.

Example: University of California

- 9 campuses admit undergraduates.
- 221,788 applicants in 2017.
- For any coding category of any anonymous statistics, mechanisms from $\mathcal{M}$ differ by less than 10% for at least

$$1 - 8\,\text{exp}\left(-\frac{.1^2 * 221788}{4 * 9^2}\right) \approx .991$$

of possible preference profiles.

Example: Refugees

$749'487$ refugees and asylum-seekers officially registered in Germany in 2015.

Suppose we elicit their preferences over Germany's 16 lands and run the matching by a mechanism from $\mathcal{M}$.

The choice of the mechanism would impact the coding categories of any anonymous statistics by less than 10% for at least

$$1 - 8\,\text{exp}\left(-\frac{.1^2 * 749487}{4 * 16^2}\right) \approx .995$$

of possible preference profiles.

Robustness

A mechanism is *robust* with ratio $c > 0$ if changing the report of one agent affects the allocations of at most c agents.

Robustness

A mechanism is *robust* with ratio $c > 0$ if changing the report of one agent affects the allocations of at most c agents.

The following mechanisms are robust with $c = |A|$:

- Serial Dictatorships
- Abdulkadiroglu and Sonmez (2003) Top Trading Cycles for School Choice
- Pycia and Unver's (2011) Fixed-Endowment Trading Cycles for School Choice.

Positive Equivalence: Asymptotics

Theorem. $\forall \epsilon, c > 0 \; \exists n^* \; \forall |N| \geq n^*$ and F anonymous:

If mechanisms ϕ and ψ are Pareto, strategy-proof, and c-robust, then:

$$\sum_{\ell=1}^{|K|} \left| F_\ell \left(\succ, \phi \left(\succ^\phi \right) \right) - F_\ell \left(\succ, \psi \left(\succ^\psi \right) \right) \right| < \epsilon.$$

for at least fraction $1 - \epsilon$ of all preference profiles.

Converse

- Asymptotic Positive Equivalence: suffices that a sequence of aggregate statistics is asymptotically anonymous on relevant Pareto-efficient allocations.

Converse

- Asymptotic Positive Equivalence: suffices that a sequence of aggregate statistics is asymptotically anonymous on relevant Pareto-efficient allocations.

- If a sequence of statistics F^N fails this weaker anonymity assumption, then the analogue of the Asymptotic Positive Equivalence fails.

Converse

- Asymptotic Positive Equivalence: suffices that a sequence of aggregate statistics is asymptotically anonymous on relevant Pareto-efficient allocations.
- If a sequence of statistics F^N fails this weaker anonymity assumption, then the analogue of the Asymptotic Positive Equivalence fails:
  $\forall c > 0 \; \exists \; \epsilon > 0 \; \forall \; n^* \; \exists \, |N| > n^*$ and $\exists$ Pareto and strategy-proof mechanisms ϕ and ψ such that

$$\sum_{\ell=1}^{|K|} \left| F_\ell^N \left(\succ_N, \phi \left(\succ_N \right) \right) - F_\ell^N \left(\succ_N, \psi \left(\succ_N \right) \right) \right| > \epsilon$$

  for least $1 - \epsilon$ fraction of all preference profiles $\succ$.

Proof Strategy

1. Prove a normative equivalence of ϕ and ψ.
2. Use concentration theory.

Exchangeable Distributions

A distribution on Θ^N is *exchangeable* if the probability of θ_N is the same as the probability of $\theta_{\sigma(N)}$ for any permutation $\sigma : N \to N$.

Examples

- IID distributions.
- Distributions that are IID conditional on an aggregate shock.

Exchangeable Distributions

A distribution on Θ^N is *exchangeable* if the probability of θ_N is the same as the probability of $\theta_{\sigma(N)}$ for any permutation $\sigma : N \to N$.

Examples

- IID distributions.
- Distributions that are IID conditional on an aggregate shock.

Exchangeability of preference profile distributions assumed throughout.

Normative Equivalence for Standard Mechanisms

Theorem. The population mean and median of any anonymous statistic do not vary on $\mathcal{M}$.

Proof: builds on a new equivalence result for symmetric mechanisms (next slide).

Symmetric Mechanisms

A random mechanism L is symmetric if

$$\tau^{-1} \circ L \circ \tau = L$$

for any transposition τ.

Symmetric Mechanisms

A random mechanism L is symmetric if

$$\tau^{-1} \circ L \circ \tau = L$$

for any transposition τ.

Theorem. The distribution over matchings is the same for any symmetric lottery over mechanisms from $\mathcal{M}$.

Symmetric Mechanisms

A random mechanism L is symmetric if

$$\tau^{-1} \circ L \circ \tau = L$$

for any transposition τ.

Theorem. The distribution over matchings is the same for any symmetric lottery over mechanisms from $\mathcal{M}$.

For marginal distributions the analogue of this theorem obtains for all school choice hierarchical exchange mechanisms.

Symmetric Mechanisms

A random mechanism L is symmetric if

$$\tau^{-1} \circ L \circ \tau = L$$

for any transposition τ.

Theorem. The distribution over matchings is the same for any symmetric lottery over mechanisms from $\mathcal{M}$.

For marginal distributions the analogue of this theorem obtains for all school choice hierarchical exchange mechanisms.

Analogous results for $|a| = 1$ at all schools: Abdulkadirglu and Sonmez 1998, Knuth 1996, Pathak and Sethuraman 2011, **Carroll 2014**, Lee and Sethuraman 2011, Pycia and Troyan 2019.

Normative Equivalence in Large Markets

Theorem. $\forall \epsilon, c > 0 \; \exists n^* \; \forall |N| \geq n^*$ and F anonymous:

If mechanisms ϕ and ψ are Pareto, strategy-proof, and c-robust then:

$$\mathbb{E} \sum_{\ell=1}^{|K|} \left| F_\ell \left(\succ_N, \phi \left(\succ_N \right) \right) - F_\ell \left(\succ_N, \psi \left(\succ_N \right) \right) \right| < \epsilon,$$

where the expectation is over the iid uniform distribution on all preference profiles.

Normative Equivalence in Large Markets

Theorem. $\forall \epsilon, c > 0 \; \exists n^* \; \forall |N| \geq n^*$ and F anonymous:

If mechanisms ϕ and ψ are Pareto, strategy-proof, and c-robust then:

$$\mathbb{E} \sum_{\ell=1}^{|K|} \left| F_\ell \left(\succ_N, \phi \left(\succ_N \right) \right) - F_\ell \left(\succ_N, \psi \left(\succ_N \right) \right) \right| < \epsilon,$$

where the expectation is over the iid uniform distribution on all preference profiles.

For sufficiently large N and for any $\delta > 0$, the same obtains for any exchangeable distribution $\mathbb{P}$ such that

$$\mathbb{P} \left(\{ \succ_N \colon \; (\forall \succ) \; \left| \{ i \in N : \succ_i = \succ \} \right| > \delta \left| N \right| \} \right) \geq 1 - \frac{\epsilon}{3}.$$

Normative Equivalence in Large Markets

Theorem. $\forall \epsilon, c > 0 \, \exists n^* \, \forall |N| \geq n^*$ and F anonymous:

If mechanisms ϕ and ψ are Pareto, strategy-proof, and c-robust then:

$$\mathbb{E} \sum_{\ell=1}^{|K|} \left| F_\ell \left(\succ_N, \phi \left(\succ_N \right) \right) - F_\ell \left(\succ_N, \psi \left(\succ_N \right) \right) \right| < \epsilon,$$

where the expectation is over the iid uniform distribution on all preference profiles.

For sufficiently large N and for any $\delta > 0$, the same obtains for any exchangeable distribution $\mathbb{P}$ such that

$$\mathbb{P} \left(\{ \succ_N : \, (\forall \succ) \; |\{ i \in N : \succ_i = \succ \}| > \delta |N| \} \right) \geq 1 - \frac{\epsilon}{3}.$$

Proof: builds on the asymptotic equivalence of symmetric mechanisms proven by Liu and Pycia 2011.

Population-Symmetry Duality

$\phi(i, \succ)(a)$=probability that i obtains a in random mechanism ϕ.

The *symmetrization* ϕ^S of mechanism ϕ is given by

$$\phi^S(i, \succ)(a) = \sum_{\sigma: N \overset{1-1}{\to} N} \frac{1}{|N|!} \phi(\sigma(i), \succ_\sigma)(a).$$

Population-Symmetry Duality

$\phi(i, \succ)(a)$=probability that i obtains a in random mechanism ϕ.

The *symmetrization* ϕ^S of mechanism ϕ is given by

$$\phi^S(i, \succ)(a) = \sum_{\sigma : N \overset{1-1}{\to} N} \frac{1}{|N|!} \phi(\sigma(i), \succ_\sigma)(a).$$

Duality Lemma. For any exchangeable distribution Λ over preference profiles $\succ$, the following are equivalent:

- $\forall$ anonymous statistics F: $\mathbb{E}F(\succ, \phi(\succ)) = \mathbb{E}F(\succ, \psi(\succ))$;
- $\forall \succ$ in support of Λ, $\forall i, a$: $\phi^S(i, \succ)(a) = \psi^S(i, \succ)(a)$.

Population-Symmetry Duality

$\phi(i, \succ)(a)$=probability that i obtains a in random mechanism ϕ.

The *symmetrization* ϕ^S of mechanism ϕ is given by

$$\phi^S(i, \succ)(a) = \sum_{\sigma : N \overset{1-1}{\to} N} \frac{1}{|N|!} \phi(\sigma(i), \succ_\sigma)(a).$$

Duality Lemma. For any exchangeable distribution Λ over preference profiles $\succ$, the following are equivalent:

- $\forall$ anonymous statistics F: $\mathbb{E}F(\succ, \phi(\succ)) = \mathbb{E}F(\succ, \psi(\succ))$;
- $\forall \succ$ in support of Λ, $\forall i, a$: $\phi^S(i, \succ)(a) = \psi^S(i, \succ)(a)$.

Analogous dualities for approximate and asymptotic statements.

Related result for medians of statistics.

Random Mechanisms

The results extend to:

- random mechanisms;
- asymptotic strategy-proofness.

Priorities: Model

- A finite set T of global priority types
- Priority ranking of two agents is strict iff they have different priority type.
- An allocation is *stable* if no pair of agents i and j such that i has higher priority at the school j is assigned and i prefers this school over his or her assignment. (For brevity, no unacceptable schools).

Priorities: Result

Theorem. $\forall \epsilon, c > 0 \ \exists n^* \ \forall F$ anonymous:

If there are at least n^* agents in each priority group and mechanisms ϕ, ψ are stable, constrained-Pareto, & strategy-proof then

$$\sum_{\ell=1}^{|K|} \left| F_\ell \left(\succ, \phi\left(\succ\right)\right) - F_\ell \left(\succ, \psi\left(\succ\right)\right) \right| < \epsilon$$

for at least $1 - \epsilon$ fraction of all preference profiles at which ϕ and ψ are c-robust.

Priorities: Result

Theorem. $\forall \epsilon, c > 0\ \exists n^*\ \forall F$ anonymous:

If there are at least n^* agents in each priority group and mechanisms ϕ, ψ are stable, constrained-Pareto, & strategy-proof then

$$\sum_{\ell=1}^{|K|} \left| F_\ell \left(\succ, \phi\left(\succ\right)\right) - F_\ell \left(\succ, \psi\left(\succ\right)\right) \right| < \epsilon$$

for at least $1 - \epsilon$ fraction of all preference profiles at which ϕ and ψ are c-robust.

Notice: local robustness assumption.

Implications for Market Design

- Non-Anonymous Statistics
  - Improving non-anonymous statistics is possible.
  - No (or little) adverse effect on anonymous statistics when we design for non-anonymous ones.

Implications for Market Design

- Non-Anonymous Statistics
  - Improving non-anonymous statistics is possible.
  - No (or little) adverse effect on anonymous statistics when we design for non-anonymous ones.

- Anonymous Statistics
  - Improving anonymous statistics calls for relaxing the equivalence assumptions and e.g. eliciting preference intensity (Hylland and Zeckhauser 1979) or using observable correlates of agents' utilility (Leshno and Lo 2018).

Summary

- Anonymous vs non-anonymous statistics have qualitatively different properties.
- Any two Pareto-efficient, strategy-proof, and robust mechanisms generate asymptotically the same anonymous statistics for asymptotically almost all preference profiles.
- Normative exact equivalence and meaningful bounds in positive results.
- Converse: the equivalence requires at least asymptotic anonymity of the statistics.