Evaluating with Statistics

Which Outcome Measures Differentiate Among Matching Mechanisms?

Marek Pycia

Discussed by Andy Haupt

MIT Market Design Reading Group

July 2020

A Slide from 14.125: Market Design

Performance of Boston mechanism

Sample year, 2001-2002

	K2	6	9
1st choice	2,598	4,157	5,497
2nd choice	301	415	428
3rd choice	131	294	100
4th choice	61	61	42
5th choice	33	26	11
Unassigned	202	476	302

- Roughly 80% get their top choice, 8% get 2nd choice, ..., 5-9% unassigned
- Similar patterns across the years before 2005

Abdulkadiroglu, Che, Pathak, Roth, and Tercieux ’17

Table 2. Comparison of Mechanisms for Main Transition Grades (K1, K2, 6, and 9) in Boston

	TTC-Counters (1)	TTC-Clinch and Trade (2)	Serial Dictatorship (3)	Student-Proposing Deferred Acceptance (4)
A. Choice Assigned				
1	1240	1240	1236	1227
2	322	323	315	336
3	134	134	132	138
4	56	55	51	57
5+	39	39	34	40
Unassigned	102	101	124	96
Total	1893	1893	1893	1893
B. Statistics on Blocking				
Blocks defined by priority and lottery number				
students with justified envy (i)	389	368	280	0
blocking pairs (i,s)	538	506	369	0
instances of justified envy (i, (j,s))	1943	1752	3650	0
schools with justified envy (s)	30	29	44	0
Blocks defined by priority				
students with justified envy	129	126	280	0
blocking pairs (i,s)	160	156	369	0
instances of just envy (i, (j,s))	768	711	3650	0
schools with justified envy (s)	18	18	44	0

The Main Takeaway

In a large market, Pareto efficient matching mechanisms produce similar anonymous aggregate statistics.

Implications for MD

- If we care about efficient mechanisms, we should look beyond anonymous statistics
  - Simplicity
  - Blocking
  - Comparisons to Status Quo
- If we care about anonymous statistics, we cannot improve in efficiency with ordinal mechanisms

Related Literature

- Continuum Limits (Azevedo, Leshno '16, Leshno, Lo '19...)
- Outcome Equivalence (Abdulkadiroğlu, Sönmez '98, Pycia, Liu '11, Caroll '15...)
- Comparing Statistics of mechanisms (Abdulkadiroglu, Che, Pathak, Roth, and Tercieux '17...)

- N set of applicants, $\prec_i$, $i \in N$ (strict) preference orders
- Θ set of preference types, $\prec \in \Theta^n$ preference profile
- A set of schools, $|a|$, $a \in A$ capacity, $|A|$ maximal capacity
- Mechanisms $\phi, \psi \colon \Theta^n \to \{\text{matchings}\}$
- PE, SP, stability, constrained efficiency canonically defined
- outcome code: $f \colon N \times \Theta \times A \to K := \{1, 2, \ldots, k\}$
- Statistic $F(\prec, a) \in [0,1]^K$ is empirical distribution corresponding to tuple $(f(\prec, a, i))_{i \in N}$
- $|F(\prec, a) - F(\prec', a')| = \sum_{i=\ell}^{k} |F_\ell(\prec, a) - F_\ell(\prec', a')|$; (think $k = 2$)

Anonymity

- A code f is anonymous if $f(i, \prec, a) = f(j, \prec, a)$ for any $i, j \in N, \prec \in \Theta, a \in A$.
- A statistic f is anonymous if derived from an anonymous code.

Examples

- Did you get k^{th} choice?
- Are you assigned to a school in Queens?

Non-Examples

- Part of blocking pair?
- Better than status quo?
- In SEG and k^{th} choice?

Robustness

- Mechanism ϕ is c-robust at $\prec$ if a change of report by one agent changes outcome only for c other agents.
- ϕ is c-robust if it is c-robust at any $\prec$ for any $\prec$.

Examples

- TTCs
- SD
- Boston

Non-Examples

- DA

High-Probability Bounds ("Positive Results")

Theorem

Let F be anonymous, and fix $\varepsilon, c > 0$, and $|A|$. For large enough $|N|$, for ϕ, ψ that are Pareto efficient, strategy-proof and c-robust, for a $1-\varepsilon$-fraction of all strategy profiles,

$$|F(\prec, \phi(\prec)) - F(\prec, \psi(\prec))| \leq \varepsilon.$$

Partial Converse

Theorem

For the class of TTCs, we have for $\mathbb{P} \in \Delta(\Theta^n)$ any iid distribution,

$$\mathbb{P}[|F(\prec, \phi(\prec)) - F(\prec, \psi(\prec))| > t] \leq 8e^{\frac{t^2|N|}{16|A|^2}}.$$

Proof Strategy

- UC Matching: 0.8% of preference profiles deviate > 10%.

What about DA?

- Assume that rankings of schools only depend on group membership in priority groups P (renamed from paper, as ambiguous notation.)

Theorem

Let F be anonymous, and fix $\varepsilon, c > 0$, and $|A|$. If priority groups are uniformly large enough, then for any stable, constrained efficient mechanisms then for a $1 - \varepsilon$-fraction of all strategy profiles where ϕ and ψ are c-robust.

$$|F(\prec, \phi(\prec)) - F(\prec, \psi(\prec))| \leq \varepsilon.$$

Expectation Bounds ("Normative Results")

Theorem

Let F be anonymous, and fix $\varepsilon, c > 0$, and $|A|$. For large enough $|N|$, for ϕ, ψ that are Pareto efficient, strategy-proof and c-robust,

$$\mathbb{E}[|F(\prec, \phi(\prec)) - F(\prec, \psi(\prec))|] \leq \varepsilon.$$

where $\mathbb{E}$ is expectation wrt the uniform distribution.

- Exact iid might not be reasonable

Exchangeability

Type distribution exhangeable: $\prec_N \overset{\mathcal{D}}{=} \prec_{\sigma(N)}$

Examples

- iid distributions
- mixtures of iid distributions

Non-Examples

- The exact top-10 percent of students prefer a certain school

Main Result for Exchangeable Distributions

Theorem

Let F be anonymous, and fix $\varepsilon, c > 0$, and $|A|$. For large enough $|N|$, for ϕ, ψ that are Pareto efficient, strategy-proof and c-robust,

$$\mathbb{E}[|F(\prec, \phi(\prec)) - F(\prec, \psi(\prec))|] \leq \varepsilon.$$

where $\mathbb{E}$ is expectation wrt an exchangeable distribution for which with constant probability a constant fraction of students have identical preferences. Formal Definition

- Extension to replica economies (first draw random type, then replicate) without additional requirements.

Main Technical Tool

- Let ϕ be a mechanism.
- Define its symmetrization by
  $\phi^S(i, a)(\prec) = \sum_{\sigma \in S_n} \frac{1}{|N|!} \phi(\sigma(i), a)(\prec_\sigma)$.

Lemma

- *Two mechanisms whose symmetrizations have identical outcome distributions have identical anonymous statistics under any exchangeable preference distribution*
- *If two mechanisms have the same mean under any anonymous statistics and any exchangeable preference distribution. Then the symmetrizations of the mechanisms have the same distributions.*

Result Dependence Map

Necessity of Properties

Robustness Crucial for All finite-sample and non-iid

Strategyproofness Conceptually necessary, to be able to control the distribution of submitted distributions; no use beyond that

Pareto Efficiency Crucial to prove the main result

The paper focuses on Pareto efficient mechanisms but the equivalence insight is also valid for stable and constrained efficient mechanisms such as Gale and Shapley's Deferred Acceptance

- How reasonable is the assumption of many students per priority class?

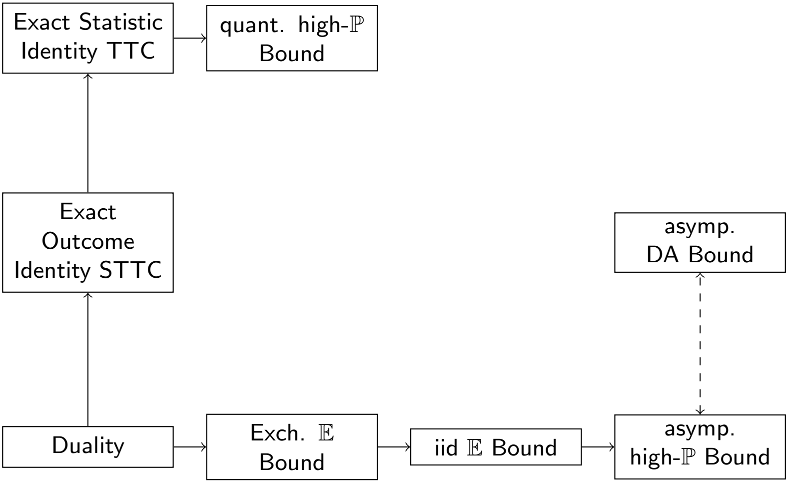
The Role of Strategy-Proofness

- The paper does not use strategyproofness beyond the fact that strategyproofness allows to control distributions of stated preferences.
- Therefore, we can read the statements for Boston as well.
- Generalization to Boston would be possible.

Thank you for your time! Here is one more backup:

An Exact Identity Result

Result Dependency Graph



Back

A Result of Exact Identity

Theorem

The population mean and median of anonymous statistics with respect to any exchangeable distribution are identical for TTC mechanisms.

Back

A more general distribution that gives stability

$$\mathcal{P}_\delta = \bigcup_{\prec \in \Theta} \{\prec_N : |\{i \in N| \prec_i = \prec\}| > \delta |N|\}$$

- It should not be too unlikely that a constant fraction of students have the same preferences.

Back

- Determine expectation using results below
- Use concentration inequalities for Lipschitz functions of several variables
- Talagrand's ('95) inequality gives an exponential bound

Back

- Partial converse using asymptotic anonymity: Sequence of statistics such that for any large enough N, and for all but a fraction of $\frac{1}{N}$ of preference profiles $\prec$

$$|F^N(\prec, \phi(\prec)) - F^N(\prec_{\sigma(N)}, \sigma(\phi(\prec)))| < \frac{1}{N}$$

Back