Evaluating with Statistics: Which Outcome Measures Differentiate Among Matching Mechanisms?

Marek Pycia*

October 2019

Abstract

The selection of mechanisms to allocate school seats in public school districts can be highly contentious. At the same time the standard statistics of student outcomes calculated from districts' data are very similar for many mechanisms. This paper contributes to the debate on mechanism selection by explaining the similarity puzzle as being driven by the invariance properties of the standard outcome statistics: outcome measures are approximately similar if and only if they are approximately anonymous.

*U. Zürich. The paper was previously distributed under the title "Invariance and Matching Market Outcomes." For their comments, I would like to thank the audiences at U. Melbourne, LSE, Essex, UCLA Summer Workshop, U Zurich, 2016 CIREQ Montreal Microtheory Conference, 2016 NY Abu Dhabi Mechanism Design Conference, MATCH-UP 2017, NYU 2017 Econometrics Meets Theory Conference, Matching-in-Practice 2017, SSE, Bonn, 2017 Barcelona JOCS Talk, UPV/EHU, EUI, Oxford, FU/ESMT/TU/HU Berlin, U Zielona Gora, Chicago Econ and Booth, Aix-Marseille, as well as Estelle Cantillon, Denis Chetverikov, Jernej Copic, Federico Echenique, Lars Ehlers, Fuhito Kojima, David Levine, Shengwu Li, Jacob Leshno, Jay Liu, Anna Myjak-Pycia, Krzysztof Oleszkiewicz, Julian Teichgraeber, and Utku Unver. For their generous support of this work, I would like to thank the UCLA Department of Economics, the William S. Dietrich II Economic Theory Center, and Department of Economics at Princeton University.

1 Introduction

The allocation of school seats is a canonical example of the allocation of discrete resources.[1] In many public school districts—in countries as varied as Australia, Chile, China, Finland, France, Ghana, Hungary, Ireland, the Netherlands, Norway, Poland, Romania, Spain, Taiwan, Turkey, US, and the UK—the applicants are consulted and asked to submit their rankings of schools. School seats are then allocated taking these rankings into account. The applicants' willingness to pay is not elicited.

The choice of the mechanism used to allocate seats can be highly contentious and many allocation mechanisms have been proposed in the economic literature.[2] At the same time, the empirical analyses of school choice uncovered a puzzle: many aggregate measures of student outcomes are the same or very similar for a variety of different mechanisms. This equivalence occurs in empirical data from school districts in Amsterdam, Boston, Cambridge, New Orleans, and New York. The empirical studies and the presented analysis allows multidimensional measures. For instance, the standard reported aggregate measure of a matching mechanism performance reports the number of applicants obtaining their first choice outcome, the number of applicants receiving their second choice outcome, etc.[3]

[1]The analysis of this paper is presented in the context of the allocation of school seats but the presented results are also applicable to other problems in which objects are allocated instead of being sold, e.g. to the matching of refugees and resettlement countries (or communities).

[2]Pathak and Sethuraman (2011) discuss the controversies surrounding the choice. For examples of the lively economic literature constructing new, mostly ordinal, mechanisms, see e.g.: Balinski and Sönmez (1999), Pápai (2000), Bogomolnaia and Moulin (2001), Abdulkadiroğlu and Sönmez (2003), Kesten (2010), Pycia and Ünver (2011), Budish, Che, Kojima, and Milgrom (2014), Ashlagi and Shi (2014), Abdulkadiroğlu et al. (2015), Morrill (2015), Nguyen, Peivandi, and Vohra (2016), Hakimov and Kesten (2018), He, Miralles, Pycia, and Yen. (2018), and Abdulkadiroglu, Che, Pathak, Roth, and Tercieux (2017b).

[3]To see how close the equivalence is in empirical data, consider Abdulkadiroglu, Che, Pathak, Roth, and Tercieux (2017b) who report the number of students obtaining seats in their respective first, second, third, fourth, and fifth or worse most preferred school for four different efficient mechanisms. The rank numbers are nearly identical with the largest difference below 6 per mille of the number of students participating (and typical differences below 1 per mille). They report similar numbers for Boston. See also e.g. De Haan, Gautier, Oosterbeek, and van der Klaauw (2015) (Amsterdam), and Abdulkadiroglu et al. (2017a) (New York). Pathak (2017) survey highlights this empirical puzzle.

The present paper explains this empirical puzzle and in so doing throws light on the controversies surrounding the selection of school choice mechanisms. The main results establish that the near equivalence of the aggregate measures of student outcomes are driven by three factors: market size, the efficiency and incentive properties of the underlying mechanisms, and the anonymity of the aggregate measures.[4] An aggregate measure is anonymous (or invariant) if it is invariant with respect to permutations of students' outcomes. For instance, in counting the number of students who received their first, or second, or third choice outcome the identities of these students are not relevant, only their allocations and their preferences. Also measures such as how many students attend schools on the right side of the river (or train tracks) are anonymous. The anonymity of the aggregate measures is crucial for the equivalence insight. An example of a non-anonymous measure is the number of students whose allocation is better under a new mechanism than under the status quo. For this and other non-anonymous measures the equivalence fails.[5]

The equivalence holds in both positive and normative sense. The positive theorem relies on two assumptions: the matching market is large in the sense that there are many applicants relative to the number of schools and the mechanisms are robust in that a change of report by one agent affects the allocations only for a bounded number of other agents; this robustness assumption is satisfied by standard allocation mechanisms such as Serial Dictatorship, Top Trading Cycles based mechanisms, and the Boston mechanism. Under these assumptions, the positive theorem of the paper establishes that any two strategy-proof and Pareto efficient mechanisms result in anonymous aggregate statistics that are very similar for most possible preference

[4]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 (see Section 3.2).

[5]This failure of equivalence is also in line with regularities reported in empirical literature. For studies of the improvement over status quo see He 2011, Calsamiglia and Miralles (2012), and Agarwal and Somaini (2018). The number of violations of stability (Kesten, 2010) is also non-anonymous and the equivalence fails e.g. in the data on stability violations reported by Abdulkadiroglu, Che, Pathak, Roth, and Tercieux (2017b).

profiles.[6] Asymptotically, as the number of seats grows, the statistics become identical for all but measure zero of possible preference profiles.

From a normative perspective of a social planner who does not yet know the participants' preferences and assumes that they are exchangeable draws from any full-support distribution, the above positive equivalence implies as a corollary that in expectation any two mechanisms satisfying the above incentives and efficiency assumptions generate the same expected anonymous aggregate measures. We furthermore prove several finite-market theorems establishing that for the standard class of mechanisms introduced by Abdulkadiroğlu and Sönmez (2003) and called Top Trading Cycles, this normative corollary remains true regardless of the market size. This standard class of mechanisms contains for instance Serial Dictatorships and, on the house allocation subdomain, all Pápai (2000) fixed-endowment Hierarchical Exchange mechanisms.[7]

The equivalence of anonymous statistics of standard allocation mechanisms established in the paper does not mean that the mechanisms themselves are outcome-identical: the equivalence of the anonymous statistics is not the same as the equivalence of mechanisms. Even within the class of serial dictatorship mechanisms—in which one-by-one market participants receive seats in their most favorite still not oversubscribed school—the order in which the seats are assigned matters for individual participants: being the first to receive the seat is better than having a random position in the ordering, and the random position is better than being last. Only the aggregate statistics are the same in all these mechanisms.

At the same time the results suggest that to choose between various standard allocation mechanisms the school districts need to look beyond the usually considered

[6]For allocation mechanisms that are deterministic, the strategy-proofness assumption can be relaxed to there being a pure strategy Nash equilibrium. However, as de Haan et al. (2015) and Kapor, Neilson, and Zimmerman (2018) recently shown, participants' heterogenous beliefs may lead to welfare losses in non strategy-proof school choice mechanisms.

[7]On this subdomain, we show that all Li (2017) obviously strategy-proof mechanisms have also the same expected (and median) anonymous statistics; a result which builds on recent work by Pycia and Troyan (2019).

anonymous statistics.[8] As discussed above, the districts might compare the mechanisms to the status quo, which by definition is not anonymous. Or the district might pay attention to the simplicity of mechanisms, e.g. defined as in Li (2017) or Pycia and Troyan (2019).

Furthermore, the results suggest that school districts that primarily care about anonymous statistics cannot improve upon the standard mechanisms while eliciting only ordinal rankings of schools. To construct better mechanisms we need to elicit preference intensity. There is a small but growing literature devoted to this problem in the school choice environment starting with the seminal work by Hylland and Zeckhauser (1979) (cf. also Abdulkadiroğlu, Che, and Yasuda (2015), He, Miralles, Pycia, and Yen. (2018), Azevedo and Budish (2011), Miralles and Pycia (2014), and Nguyen, Peivandi, and Vohra (2016)).[9] And, indeed, mechanisms eliciting preference-intensity do better than purely ordinal ones as shown by Miralles (2008), Abdulkadiroğlu et al. (2011), and Featherstone and Niederle (2016).[10]

Methodologically, the results rely on a simple, yet unexpected, link between pointwise properties of symmetric mechanisms and population properties of asymmetric mechanisms. The paper formalizes this link as a duality principle between these two types of environments, and it exploits it in obtaining both large market and small market results.[11] The asymptotic equivalence results for strategy-proof mechanisms build on Liu and Pycia (2011), who proved the general large-market equivalence of symmetric, strategy-proof, and efficient mechanisms.[12] The proofs of the finite-

[8] We also refer to anonymous statistics as invariant outcome measures.

[9] This school choice literature builds on pioneering analysis of mechanisms eliciting preference intensities without transfers in other environments, e.g. Sönmez and Ünver (2010), Budish (2010), Budish and Cantillon (2010), Budish, Che, Kojima, and Milgrom (2014). (For recent examples see also e.g. Bogomolnaia et al. (2017)and Babaioff et al. (2017)).

[10] For more general analysis see also Troyan (2012), Pycia (2011a), Ashlagi and Shi (2015), Abdulkadiroglu, Agarwal, and Pathak (2017a). For multi-unit-demand assignment, Budish and Cantillon (2010) showed that strategy-proof ordinal mechanisms incur even more substantial welfare losses than those encountered in single-unit-demand school choice.

[11] The paper is also one of the first economic applications of the group concentration inequalities. For an earlier use of other concentration inequalities, see e.g. Kalai (2004) study large games.

[12] Cf. also Pycia (2011b). A narrower asymptotic equivalence for environments with copies was earlier proved by Che and Kojima (2010), who showed the asymptotic equivalence of two symmetric

market results for standard allocation mechanisms also build on the duality principle, but the finite-market analysis of symmetric standard mechanisms in many-to-one environments has been missing, and hence the current paper provides this analysis.

In particular, one of the paper's auxiliary contributions is the first finite-market equivalence for symmetric standard school choice mechanisms for environments in which schools can have more than one seat each. Our school choice symmetric equivalence builds on the rich literature studying various classes of symmetric mechanisms, but only in the environment without copies (usually referred to as house allocation). Abdulkadiroğlu and Sönmez (1998) were the first to prove an equivalence of two symmetric mechanisms: Random Priority and the Core from Random Endowment. Carroll (2014) proved the equivalence for all standard mechanisms and his result is the main basis for ours.[13]

The present paper results do not hinge on the symmetry of the mechanisms. This is important as many of the mechanisms for which the equivalence of anonymous statistics have been observed in the data are not symmetric; also the mechanisms among which school districts are choosing are often not symmetric. We show that in large markets even in the absence of symmetry the outcome measures are nearly identical for a typical preference profile in many mechanisms. While this paper is the first to address the positive question without restricting attention to symmetric assumptions, a special case of the normative question we study was analyzed by Che and Tercieux (2018). They restrict attention to payoff-distributions—which is an

mechanisms: Random Priority and Bogomolnaia and Moulin (2001) Probabilistic Serial. Miralles (2008) independently established the equivalence of these two mechanisms in the continuum economy limit; for differences between continuum limit economies and large finite market economies, see Miralles and Pycia (2015). Azevedo and Leshno (2016) proved related limit results for Deferred Acceptance and Ashlagi, Kanoria, and Leshno (2017) showed that when the market sides differ in size, the mean outcome ranks of agents on the proposing side in Random Deferred Acceptance are close to their ranks in Random Priority. Less related to our work is Immorlica and Mahdian (2005) and Kojima and Pathak (2008) who provided the first large market matching analysis in two-sided environments.

[13] Cf. also Pathak and Sethuraman (2011) and Lee and Sethuraman (2011). Pycia and Troyan (2019) proved finite-market equivalence of all symmetric mechanisms that are simple to play and efficient.

example of an anonymous statistic—and show that asymptotically in large markets the theoretical distributions converge.[14] In contrast, most of our normative results do not rely on any large market assumptions and they apply to a much broader range of market-outcome statistics.[15]

2 Model

Let A be a finite set of schools; each school $a \in A$ has $|a| > 0$ seats.[16] Let N be a finite set of agents, to whom we also refer to as applicants; each agent i demands a single seat and has a strict preference ranking $\succ_i$ over schools. Let Θ be the set of agents' rankings over schools, to which we also refer to as agents' types.

An *allocation* (or matching) μ specifies for each agent i the school $\mu(i)$ the agent is assigned. An allocation is *Pareto efficient* if no other allocation is weakly better for all agents and strictly better for at least one agent.

For each allocation μ, we code the outcome $a = \mu(i)$ of agent i of type $\succ \in \Theta$ as as an element of set $K = \{1, ..., k\}$ for some $k = 2, 3, ...$. The set of codes K is assumed to be fixed throughout the paper, except for the duality analysis of Lemmas 1 and 4. The coding is given by a function $f : N \times \Theta \times A \to K$ and we say that the coding function is *anonymous* (or invariant) if it does not depend on the individual identity, that is

$$f(i, \succ, a) = f(j, \succ, a)$$

for all agents $i, j \in N$, rankings $\succ \in \Theta$, and outcomes $a \in A$. We then write $f(\succ, a)$ for the common code. A statistics (or outcome measure) $F : (\Theta \times A)_{i \in N} \to [0,1]^K$ is an empirical distribution of individual outcome codes. We call the statistics anonymous (or *invariant*) if the coding function is anonymous.

Examples of anonymous statistics include: how many applicants obtain their top

[14] The market is large in that there are many schools relative to the number of seats per school.
[15] Che and Tercieux' complex and impressive proof does not rely on the duality principle.
[16] We may allow for a null object $\oslash$ that has many copies, $|\oslash| \geq |N|$.

outcome; how many applicants obtain their two top outcomes; how many applicants are assigned to school a or school b; how many applicants obtained an outcome they prefer to school a. An example of a non-anonymous statistics is how many applicants obtained an outcome they prefer to their local school.

A *mechanism* ϕ maps profiles of messages to allocations. We focus on mechanisms in which agents report their rankings of schools. Mechanism ϕ is *strategy-proof* if, at any profile of preferences $\succ_N$, for any agent i, reporting $\succ_i$ weakly dominates reporting any $\succ'_i$. While we focus on deterministic mechanisms, the analogues of our results for random strategy-proof mechanisms follow immediately. A random mechanism maps profiles of messages to lotteries over allocations and it is strategy-proof if reporting the truth weakly first-order stochastically dominates all other reports.[17]

A mechanism is robust with ratio $c > 0$ at preference profile $\succ$ if, starting at this profile, a change of report by one agent affects the allocations of at most c agents. A mechanism is robust with ratio $c > 0$ if it is robust with this ratio at all preference profiles.[18] For instance, deterministic mechanisms such as serial dictatorships, Abdulkadiroğlu and Sönmez (2003) top trading cycles, Boston mechanism, Papai's (2000) hierarchical exchange and Pycia and Unver's (2017; 2011) trading cycles with fixed endowments are robust with ratio $|A|$. To see it observe that a preference change by one agent affects another agent only if this agent took, or just missed, the last copy of an object. There is only one affected agent per object because if an agent who took the last copy is affected (and doesn't get the object after the change) then the agent who originally just missed the object, still cannot obtain it.[19]

[17]A profile of strategies $\hat{\succ}_N$ in mechanism ϕ is in a (pure-strategy) *Nash equilibrium* if, for any agent i, reporting $\hat{\succ}_i$ weakly dominates reporting any $\succ'_i$. For deterministic mechanisms, the analogues of our results are true for pure strategy Nash equilibria.

[18]In pre-2017 drafts and presentations, I referred to robustness as weak continuity. When extending our results to random mechanisms we can weaken robustness by allowing slight impact on the marginal allocations of the remaining agents. Furthermore, at the cost of added conceptual and notational complication, Theorem 1 and its lemmas can be extended to sequences of mechanisms such that for every $\epsilon > 0$ there is $\delta > 0$ such that changing the reports of a fraction δ of agents affects at most fraction ϵ of agents.

[19]The robustness ratio of stable mechanisms, such as Deferred Acceptance, depends on the preference profile. Under the assumptions in the analysis of incentives in Kojima and Pathak (2008),

3 Positive Equivalence

The first result of this section is formulated for finite markets with sufficiently many applicants (or agents).

Theorem 1. *Let F be an anonymous statistics. For every $\epsilon, c > 0$ and for $|N|$ sufficiently large relative to ϵ,c and $|A|$, for any two Pareto efficient and strategy-proof mechanisms ϕ and ψ that are robust with ratio c, we have:*

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

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

In this result, instead of counting preference profiles we can measure them by any distribution satisfying the assumptions of Theorem 6.

The normative analysis of this paper—in particular, Theorem 6—is key to the proof of Theorem 1. The normative analysis gives us estimates for the means of $F_\ell (\succ, \phi(\succ))$ and $F_\ell (\succ, \psi(\succ))$ and allows us to use the probabilistic concentration theory to estimate for how many preference profiles the aggregate statistics are far from the mean, and hence for how many preference profiles the bound of Theorem 1 fails. The details are in the appendix.

Theorem 1 establishes the asymptotic equivalence of many standard mechanisms including:

- **Serial Dictatorships**, in which agents are ordered, and the first agent in the ordering obtains the seat in his or her most favorite school, the second agent obtains the seat in his or her most favorite school that still has available seats, etc.

the ratio of Deferred Acceptance is low at typical preference profiles. In Section 3.3 on priorities, we formulate the our positive result for stable mechanisms.

- **Top Trading Cycles** mechanisms of Abdulkadiroğlu and Sönmez (2003). Each Top Trading Cycle mechanism takes as input a profile of orderings of agents, one at each object. We call these orderings priority lists. The allocation is then determined in rounds similarly to the original Gale's Top Trading Cycles. In each round, each agent who did not receive a school seat in previous rounds points to his most preferred school with available seats, and each such school points to the pointing agent who is highest in the school priority list. There is at least one cycle in which an agent i points to a school, which points to an agent, etc, till a school points to agent i. We then allocate to each agent in such a cycle a seat in the school he pointed to, and for each such allocation we decrease the number of the available seats in this school by one.

For these standard mechanisms, we provide bounds on the precision of the approximation in finite markets; see the next subsection. With this analysis in mind, let us denote by $\mathcal{M}_{\text{Standard}}$ the above class of Top Trading Cycles mechanisms. Notice that it is a broad class of mechanisms that contains Serial Dictatorship mechanisms as a special case in which priority lists at all schools are the same. Indeed, in the subdomain of house allocation environments—that is the class of environments delineated by the assumption $|a| = 1$ for all $a \in A$—this class of mechanisms contains all of Pápai (2000) Fixed-Endowment Hierarchical Exchange mechanisms.

3.1 Equivalence Precision Bounds

Denote by $\mathbb{P}$ any distribution on preference profiles that is iid across agents.

Theorem 2. *If ϕ and ψ are mechanisms from $\mathcal{M}_{Standard}$ and F is an anonymous statistics then*

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

and

$$\mathbb{P}\left(\left|F_{\ell}\left(\succ,\phi\left(\succ\right)\right)-F_{\ell}\left(\succ,\psi\left(\succ\right)\right)\right|>t\right)\leq 8\exp\left(-\frac{t^{2}N}{4|A|^{2}}\right),\quad \forall\ell=1,...,|K|.$$

If, for instance, $\mathbb{P}$ is the uniform distribution on all preference profiles, the estimates above tell us for how few preference profiles the anonymous statistics are further away than t.[20]

This theorem remains true if we include in $\mathcal{M}_{\text{Standard}}$ all mechanisms that are Pareto-efficient and strongly obviously strategy-proof mechanisms in the sense of Pycia and Troyan (2019). It also extends to Generalized Top Trading Cycles mechanisms discussed in the Appendix. The proof—provided in the Appendix—is based on the finite-market normative equivalence results of Section 4 and a concentration inequality of Talagrand (1995).

Example 1. The University of California system has 9 campuses that offer undergraduate education. 221,788 prospective undergraduates applied to at least one campus in 2017. For problem this size, any two mechanisms from $\mathcal{M}_{\text{Standard}}$ have anonymous aggregate statistics of enrollment different for each coding category by 10% or more of the relevant range ($t = .1$) for only $8\exp\left(-\frac{.1^2*221788}{4*9^2}\right)\approx .008$ possible preference profiles.

Example 2. In 2015, 1.2 mln first-time applicants applied for asylum in the European Union, 1.8 mln migrants were caught crossing the border illegally, and 2.2 mln migrants were discovered being illegally present in the EU.[21] Were EU to match them

[20] As discussed in appendix, our approach leads to only slightly weaker bounds for Pycia and Ünver (2011) extension of Pápai (2000) Hierarchical Exchange mechanisms (with and without fixed endowments) to the school choice environment, a much larger class of mechanisms than the standard mechanisms in $\mathcal{M}_{\text{Standard}}$. Furthermore, the proof of Theorem 1 implies weaker bounds that however are applicable to all mechanisms studied in that theorem.

[21] Source: The statistical office of the European Union (Eurostat) official statistics. There is some overlap among these numbers; e.g. someone turned away from the border might be then discovered as being illegally in the EU. In this illustrative calculation, I ignore this overlap but even under the conservative estimates of the total as 2.2 mln, the bound on the fraction of preference profiles would be .007.

to its 28 countries via one of the standard mechanisms from $\mathcal{M}_{\text{Standard}}$, the selection of the mechanism would impact any coding category of any anonymous statistics by less than 10% under all but $8\exp\left(-\frac{.1^2*5200000}{4*28^2}\right) \approx .0000005$ of possible preference profiles.[22] In the same year 2015, 749,487 refugees and asylum-seekers officially registered in Germany.[23] Were Germany to match them to its 16 lands via one of the standard mechanisms from $\mathcal{M}_{\text{Standard}}$, the selection of the mechanism would impact the coding categories of anonymous statistics by less than 10% under all but $8\exp\left(-\frac{.1^2*749487}{4*16^2}\right) \approx .005$ of possible preference profiles.

3.2 Converse

Does an analogue of Theorem 1 hold true for non-anonymous statistics? We establish a converse of Theorem 1, first however we need to recognize that the large market approximate equivalence for anonymous statistics immediately implies such an approximate equivalence for approximately anonymous statistics. We say that an aggregate statistics $F : (\Theta \times A)_{i \in N} \to [0,1]^K$ is ϵ-approximately anonymous if for every permutation σ and strategy-proof and Pareto efficient mechanism ϕ[24]

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

for no more than the fraction ϵ of preference profiles $\succ_N$. If we can take ϵ to zero along a sequence of statistics as the population N increases, we say that the sequence is asymptotically anonymous. With these terms in place, Theorem 1 implies

Corollary 1. *Let F^N be an asymptotically anonymous sequence of aggregate statis-*

[22] The closeness would be within 5% for all but .1 of possible preference profiles. I do not take stance on the question to what extent the asylum seekers had preferences over EU countries (or German lands); they might well did so e.g. because of the location of other migrants they knew.

[23] Source: The United Nations Refugee Agency (UNHCR) official statistics.

[24] Notice the slight abuse of notation in writing the argument of the aggregate statistics F as $(\succ_N, \phi(\succ_N))$. By $\succ_{\sigma(N)}$ and $\sigma(\phi(\succ_N))$ we denote respectively the permutations of the preference profile and allocation.

tics. For every $\epsilon, c > 0$ and for $|N|$ sufficiently large relative to ϵ,c, $|A|$, and the sequence of statistics, for any two Pareto efficient and strategy-proof mechanisms ϕ and ψ that are robust with ratio c, and for at least fraction $1-\epsilon$ of all preference profiles $\succ$, we have:

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

This preparation allows us to formulate the partial converse of Theorem 1.

Theorem 3. *If F^N is a sequence of aggregate statistics that is not asymptotically anonymous, then there exists $\epsilon > 0$ such that for all n^* there exists N such that $|N| > n^*$ and there are two Pareto efficient and strategy-proof mechanisms ϕ and ψ and at least $1-\epsilon$ fraction of all preference profiles $\succ$ such that*

$$\sum_{\ell=1}^{|K|} \left| F_\ell^N \left(\succ, \phi\left(\succ\right)\right) - F_\ell^N \left(\succ, \psi\left(\succ\right)\right) \right| > \epsilon. \tag{2}$$

Proof. The failure of asymptotic anonymity means that for all n^* there exists N such that $|N| > n^*$, a permutation σ, at least fraction ϵ of preference profile $\succ_N$, and a strategy-proof and Pareto efficient mechanism ϕ for which (1) holds. Setting ψ to be the σ permutation of ϕ allows us to verify that (2) holds true. Because ψ is then strategy-proof and Pareto efficient, this concludes the proof. QED

3.3 Priorities

Some school choice districts run priority-based mechanisms, such as Deferred Acceptance, in order to guarantee outcomes that are stable. Stability in the school choice context is also known also as the lack of justified envy; a more precise term.

To define and consider stability we need to assume that schools are endowed with priorities. We assume that there is a finite set K of priority types, and the agents are partitioned across the types. The priority types $K(i)$ and $K(j)$ of two agents i

and j determines their priority ranking at each school. The rankings are strict iff the two agents have different priority type. An allocation is *stable* if there is 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.

Our analysis carries through to the setting with priorities with no major changes.

Theorem 4. *Let F be an anonymous statistics. For every $c, \epsilon > 0$ and for sufficiently many agents in each priority group relative to ϵ, c, and $|A|$, for any two stable, constrained-Pareto-efficient, and strategy-proof mechanisms ϕ and ψ, we have*

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

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

It is important that Theorem 4 relies only on the robustness of mechanisms on a subset of preference profiles.[25] In contrast, to the setting without priorities—where standard mechanisms are robust with ratio equal to the number of schools $|A|$—this local assumption matters in the results stated for the settings with priorities . For instance, the robustness ratio of Deferred Acceptance, the standard stable mechanism, depends on the preference profile and at some preference profiles the ratio is larger than $|A|$.

3.4 Random Mechanisms

Our results extend to random mechanisms. Theorem 2 remains true for lotteries over mechanisms from $\mathcal{M}_{\text{Standard}}$, and its proof follows essentially the same steps. In Theorem 1, we can relax strategy-proofness to asymptotic strategy-proofness.

[25] A local analogue of Theorem 1 is a special case of Theorem 2.

4 Normative Results

We study populations of agents whose preferences are drawn from exchangeable distributions. A distribution of types is *exchangeable* if the probability of the type profile θ_N is the same as the probability of the profile $\theta_{\sigma(N)}$ for any permutation $\sigma : N \to N$. For instance, iid distributions are exchangeable. We can also construct exchangeable distributions by drawing preferences from a mixture of iid distributions: with probability π_i, $i = 1, ..., I$, the preferences are drawn iid from a distribution F_i. A natural interpretation of this procedure is that the draw of one of the iid distributions reflects an aggregate shock; the constructed distribution is then iid conditional on an aggregate shock.

4.1 Normative results for large markets

From a normative perspective the anonymous statistics of the mechanisms we study are asymptotically equivalent in large market. For the special case when the preferences of agents are drawn uniformly the asymptotic normative result takes the following form.

Theorem 5. *Let F be an anonymous statistics. For every $\epsilon, c > 0$ and for $|N|$ sufficiently large relative to ϵ, c, and $|A|$, for any two Pareto efficient and strategy-proof mechanisms ϕ and ψ that are robust with ratio c, we have:*

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

where the expectation is taken over the uniform distribution over all preference profiles.

This result follows from the stronger, but more complex, normative result for exchangeable distributions. In its formulation a special role is played by the following class of preference profiles

$$\mathcal{P}_\delta = \{\succ_N\colon\ (\forall \succ)\ |\{i \in N : \succ_i = \succ\}| > \delta\,|N|\}$$

where $\delta > 0$ is arbitrarily fixed.

Theorem 6. *Let F be an anonymous statistics. For every $\epsilon, c > 0$ and for $|N|$ sufficiently large relative to ϵ, c, δ, and $|A|$, for any two Pareto efficient and strategy-proof mechanisms ϕ and ψ that are robust with ratio c, we have:*

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

where the expectation is taken over any exchangeable distribution $\mathbb{P}$ on preference profiles such that $\mathbb{P}(\mathcal{P}_\delta) \geq 1 - \frac{\epsilon}{3}$.

The uniform distribution version of the result, stated first, follows because for sufficiently large N, the condition $\mathbb{P}(\mathcal{P}_\delta) > 1 - \frac{\epsilon}{3}$ is satisfied by the iid uniform distribution on all preference profiles. Equivalently, as we take $|N| \to \infty$ and $\delta \to 0$, the fraction of preference profiles in $\mathcal{P}_\delta$ to all preference profiles goes to 1.

The proof of this result builds on the analysis of Liu and Pycia (2011), who studied random mechanisms and showed that symmetric, asymptotically strategy-proof, asymptotically ordinarily efficient mechanisms satisfying their regularity condition are equivalent not only in terms of anonymous statistics but, more strongly, in terms of marginal outcome distributions. They also proved asymptotic ordinal efficiency of symmetric mechanisms built by uniformly randomizing over mechanisms that are Pareto efficient, provided the randomization satisfies their regularity condition. The challenge of the proof is two-fold: taking care of their regularity assumption and translating the result on symmetric mechanisms into a result on population means. We discuss the second challenge in Section 5 and provide the details on both in the Appendix.

For replica economies, the above general result can be expressed in a simpler manner. For every base economy $(N, \succ)$ the q-fold replica economy N_q has $q \in \{1, 2, ...\}$ copies of each agent from N; each copy has preferences inherited from the agent in N. We assume that all preference rankings are represented in the base economy.

Corollary 2. *Let F_q be an invariant outcome measure on N_q. For every $\epsilon > 0$ and and sufficiently large q, for any two Pareto efficient and strategy-proof mechanisms ϕ and ψ, we have:*

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

where the expectation is taken over any exchangeable distribution.

Some of our other results also take simpler form for replica economies; see Appendix C.

4.2 Normative results for finite markets

We now show that the standard mechanisms from $\mathcal{M}_{\text{Standard}}$ have exactly identical means and medians of anonymous statistics already in any size finite markets.

Theorem 7. *The population mean and median of any anonymous statistics, with respect to any exchangeable distribution, do not vary on the class of standard mechanisms, $\mathcal{M}_{Standard}$.*

This theorem remains true if we include in $\mathcal{M}_{\text{Standard}}$ all mechanisms that are Pareto-efficient and strongly obviously strategy-proof mechanisms in the sense of Pycia and Troyan (2019). It also extends to Generalized Top Trading Cycles mechanisms discussed in the Appendix.

As a step in proving this theorem, we establish the following result for symmetric random mechanisms constructed as follows. We identify a random mechanism φ with the marginal probabilities agents are allocated objects, that is the probabilities $\varphi(i, s)(\succ)$ that i obtains a. The deterministic mechanisms are naturally embedded in this notation. The *symmetrization* ϕ^S of (deterministic or stochastic) mechanism ϕ is then given by

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

The class of Symmetric Top Trading Cycles mechanisms consists of symmetrizations of Top Trading Cycles mechanisms from $\mathcal{M}_{\text{Standard}}$.

Theorem 8. *Any two Symmetric Top Trading Cycles mechanisms generate exactly the same distribution over matchings.*

The proof—provided in the appendix—builds on Carroll (2014), who proved a related result for Pápai (2000) Fixed-Endowment Hierarchical Exchange mechanisms in the house allocation environment, in which no object has a copy. The Fixed-Endowment Hierarchical Exchange are exactly Top Trading Cycles mechanisms for the house allocation subdomain of our problem.[26]

The next section focuses on the relationship between results such as Theorem 8 and Theorem 7.

5 Population–Symmetry Duality

We establish duality between population properties of a mechanism and pointwise properties of its symmetrized version. This duality enables our analysis of normative questions in school choice mechanism design. We first formulate the duality for means in finite markets, and later discuss extensions to medians as well as approximate and asymptotic results.

[26] In pre-June 2019 drafts, I erroneously credited a version of Theorem 8 to Pathak and Sethuraman (2011); however they only studied the environment in which each school has only one seat, that is the house allocation environment. To the best of my current knowledge, Pycia and Troyan (2019) result on Strong Obviously Strategy-Proof mechanisms and Theorem 8 above are the first finite market equivalence result for school choice environments with multiple-seats schools.

Lemma 1. ***(Finite Population-Symmetry Duality).***

(1) Consider two mechanisms, ϕ and ψ, such that their symmetrizations have identical marginal distributions of outcomes. If agents' types are drawn from an exchangeable distribution then the mean of any anonymous statistics under ϕ is the same as under ψ.

(2) Conversely, consider two mechanisms, ϕ and ψ, such that for any exchangeable distribution of preferences the mean of any anonymous statistics under ϕ is the same as under ψ, then the symmetrizations of ϕ and ψ have identical marginal distributions of outcomes.

To prove Lemma 1, let us first show

Claim. The distribution of anonymous statistics for ϕ^S and fixed preference profile $\succ_N$ is the same as for ϕ and the uniform distribution over all permutations of $\succ_N$.

To prove it, let us fix label $k \in K$, and agent $i \in N$. By invariance of the coding function f, for any permutation σ on N we have

$$f\left(\succ_{\sigma(i)}, \phi\left(\sigma\left(i\right), \succ_{\sigma(I)}\right)\right) = f\left(\succ_i, \sigma\left(\phi\right)\left(i, \succ_I\right)\right).$$

Averaging over all permutations on N we conclude that the probability

$$f\left(\succ_{\sigma(i)}, \phi\left(\sigma\left(i\right), \succ_{\sigma(I)}\right)\right) = k$$

where permutation σ is drawn uniformly from all permutations on N is equal to the probability

$$f\left(\succ_i, \phi^S\left(i, \succ_I\right)\right) = k$$

where ϕ^S is the symmetrization of ϕ. Hence, the corresponding aggregate statistics F (which is just the profile of probabilities of outcome codes) are the same for the population drawing its preferences from permutations of the preference profile $\succ$ and submitting them to ϕ as from the single preference profile $\succ$ submitted to the symmetrized mechanism ϕ^S.

By definition, an exchangeable distribution is a mechanical sum of improper uniform distributions over permutations of preference profiles.[27] Hence, the mean aggregate statistics of ϕ is the weighted sum of the means of aggregate statistics from ϕ^S summed up over equivalence classes of preference profiles (where two profiles are equivalent is they are each other permutations).

The same analysis can be repeated for ψ and ψ^S. If the mean aggregate statistics of ϕ^S and ψ^S are equal—which is the case under our assumption that for each agent the marginal distributions of the schools the agent obtains is the same under both symmetrizations—then the first claim of Lemma 1 obtains.

To prove the second claim take two mechanisms, ϕ and ψ. Fix a preference profile $\succ$ and consider the exchangeable distribution obtained by uniformly randomizing over all permutations of $\succ$. Fix $a^* \in A$ and an individual preference ranking $\succ^* \in \Theta$ and consider an invariant coding function such that $f(\succ, a) = 1$ iff $a = a^*$ and $\succ = \succ^*$. The equality of this outcome measure for ϕ and ψ on this population means that the marginal probability that some agent with type $\succ^*$ obtains a^* under ϕ^S and ψ^S are identical. Since under ϕ^S and ψ^S any two agents with identical type have the same marginal probabilities of receiving a^*, we can conclude that the marginal probability distributions induced by ϕ^S and ψ^S are the same. QED

In light of the last paragraph of the proof, the second claim of Lemma 1 can be strengthened: we could weaken the assumption by restricting it to uniform distributions and the coding functions considered therein.

Remark 1. When the symmetrizations of the two mechanisms, ϕ and ψ, have not only identical marginal distributions of outcomes but also identical distributions of entire allocations, then the Claim from the previous lemma tells us that for any uniform

[27] Indeed, we can take any preference profile that has positive probability and remove it and all its permutations; by exchangeability all these permutations have the same probability and the remaining improper distribution is also exchangeable. After a finite number of such steps we have represented the original distribution as a mechanical sum of uniform distributions over permutations of preference profiles.

distribution on any closed-under-permutations set of preference profiles the medians of ϕ and ψ are the same. In particular, this observation about medians obtains if agents' preferences are drawn iid.

Approximate and asymptotic analogues of Lemma 1 are presented in Appendix B.3.

6 Conclusion

The paper shows that Pareto efficient deterministic mechanisms are asymptotically equivalent—profile-by-profile—for invariant outcome statistics. An analogous result holds true for stable and constrained efficient mechanisms. The positive results provide an explanation of the empirical puzzle posed by the equivalence of various mechanisms in measures reported in empirical work. Furthermore, the paper shows that many standard mechanisms are exactly equivalent in terms of mean invariant outcome statistics.

The anonymity (also referred to as invariance) of the outcome statistics is crucial: the paper shows that the asymptotic equivalence only holds for statistics that are at least asymptotically anonymous. The prediction that all invariant statistics have these property, and only them, is testable, and the literature summarized in the introduction is consistent with this prediction.

The results highlight two dimensions of the mechanism selection. How do the mechanisms differ with respect to non-invariant outcome measures? How do the mechanisms differ with respect to non-outcome measures, such as simplicity of playing them? The results also suggest that optimizing the mechanisms with regard to these objectives comes without losses in terms of anonymous statistics as long as we restrict attention to the broad class of mechanisms studied in this paper.

Furthermore, the positive and normative results suggest that to make welfare gains we need to go beyond strategy-proof ordinal mechanisms, for which the equivalence obtains. We can do it by eliciting the intensity of preferences or by incorporating exogenous sources of information about applicants' preferences.[28]

The techniques developed to prove these results are of use in other environments.

[28]The former approach is taken e.g. in Hylland and Zeckhauser (1979); Abdulkadiroğlu et al. (2015); He et al. (2018); Nguyen et al. (2016), and the latter in Leshno and Lo (2018).

References

Abdulkadiroglu, A., N. Agarwal, and P. Pathak (2017a): "Centralized vs. Decentralized School Assignment: Evidence from NYC," *American Economic Review*, 107(12), 3635–3689.

Abdulkadiroglu, A., Y.-K. Che, P. Pathak, A. Roth, and O. Tercieux (2017b): "Minimal-Envy Efficient Assignment Mechanisms: The Design of New Orleans OneApp," .

Abdulkadiroğlu, A., Y. Che, and Y. Yasuda (2011): "Resolving Conflicting Preferences in School Choice: the Boston Mechanism Reconsidered," *American Economic Review*, 101, 399–410.

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

Abdulkadiroğlu, A. and T. Sönmez (1998): "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems," *Econometrica*, 66, 689–701.

——— (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Agarwal, N. and P. Somaini (2018): "Demand Analysis using Strategic Reports: An application to a school choice mechanism," *Econometrica*, 86, 391–444.

Ashlagi, I., Y. Kanoria, and J. D. Leshno (2017): "Unbalanced random matching markets," .

Ashlagi, I. and P. Shi (2014): "Improving community cohesion in school choice via correlated-lottery implementation, Operations Research, 62(6), 1247-1264, 2014. Ashlagi I., P. Shi, Improving community cohesion in school choice via correlated-lottery implementation," *Operations Research*, 62, 1247–1264.

——— (2015): "Optimal allocated design without money: an engineering approach," *Management Science*, 62(4), 1078–1097.

Azevedo, E. and E. Budish (2011): "Strategyproofness for "Price-Takers" as a Desideratum for Market Design," Unpublished manuscript.

Azevedo, E. and J. Leshno (2016): "A Supply and Demand Framework for Two-sided Matching Markets," .

Babaioff, M., N. Nisan, and I. Talgam-Cohen (2017): "Competitive Equilibrium with Indivisible Goods and Generic Budgets," Working Paper.

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

Bogomolnaia, A. and H. Moulin (2001): "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 100, 295–328.

Bogomolnaia, A., H. Moulin, F. Sandomirskiy, and E. Yanovskaya (2017): "Competitive Division of a Mixed Manna," *Econometrica*, 85, 1847–1871.

Budish, E. (2010): "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes," Working Paper.

Budish, E. and E. Cantillon (2010): "The Multi-unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard," *American Economic Review*, forthcoming.

Budish, E., Y.-K. Che, F. Kojima, and P. Milgrom (2014): "Implementing Random Assignments: A Generalization of the Birkhoff - von Neumann Theorem," *American Economic Review*.

Calsamiglia, C. and A. Miralles (2012): "Catchment Areas and Access to Better Schools," Working Paper.

Carroll, G. (2014): "A General Equivalence Theorem for Allocation of Indivisible Objects," *Journal of Mathematical Economics*, 51, 163–177.

Che, Y.-K. and F. Kojima (2010): "Asymptotic Equivalence of Random Priority and Probabilistic Serial Mechanisms," *Econometrica*, 78, 1625–1672.

Che, Y.-K. and O. Tercieux (2018): "Payoff Equivalence of Efficient Mechanisms in Large Matching Markets," *Theoretical Economics*, 13, 239–271.

de Haan, M., P. A. Gautier, H. Oosterbeek, and B. van der Klaauw (2015): "The Performance of School Assignment Mechanisms in Practice," Working Paper.

Featherstone, C. and M. Niederle (2016): "Ex Ante Efficiency in School Choice Mechanisms: An Experimental Investigation," *Games and Economic Behavior*, stanford University, unpublished mimeo.

Hakimov, R. and O. Kesten (2018): "The Equitable Top Trading Cycles Mechanism for School Choice," *International Economic Review*, 59, 2219–2258.

He, Y., A. Miralles, M. Pycia, and J. Yen. (2018): "A Pseudo-Market Approach to Allocating with Priorities," *American Economic Journal: Microeconomics*, 10, 272–314.

Hylland, A. and R. Zeckhauser (1979): "The Efficient Allocation of Individuals to Positions," *Journal of Political Economy*, 87, 293–314.

Immorlica, N. and M. Mahdian (2005): "Marriage, Honesty, and Stability," *SODA 2005*, 53–62.

Kalai, E. (2004): “Large Robust Games,” *Econometrica*, 72, 1631–1665.

Kapor, A., C. A. Neilson, and S. D. Zimmerman (2018): “Heterogeneous Beliefs and School Choice Mechanisms,” Working Paper.

Kesten, O. (2010): “School Choice with Consent,” *Quarterly Journal of Economics*, 125, 1297–1348.

Kojima, F. and P. A. Pathak (2008): “Incentives and Stability in Large Two-sided Markets,” *American Economic Review*, 99, 608–627.

Lee, T. and J. Sethuraman (2011): “Equivalence results in the allocation of indivisible objects: A unified view,” .

Leshno, J. D. and I. Lo (2018): “The Cutoff Structure of Top Trading Cycles in School Choice,” Working Paper.

Li, S. (2017): “Obviously Strategy-Proof Mechanisms,” *American Economic Review*, 107(11), 3257–3287.

Liu, Q. and M. Pycia (2011): “Ordinal Efficiency, Fairness, and Incentives in Large Markets,” .

Maurey, B. (1979): “Construction de Suites Symetriques,” *Comptes Rendus de l’Académie des Sciences*, 288, 679–681.

Miralles, A. (2008): “School Choice: The Case for the Boston Mechanism,” Boston University, unpublished mimeo.

Miralles, A. and M. Pycia (2014): “Prices and Efficient Assignments without Transfers,” Working Paper.

——— (2015): “Large vs. Continuum Assignment Economies,” Working Paper.

Morrill, T. (2015): “Making Just School Assignments,” *Games and Economic Behavior*, 92, 18–27.

Nguyen, T., A. Peivandi, and R. Vohra (2016): “Assignment problems with complementarities,” *Journal of Economic Theory*, 165.

Pápai, S. (2000): “Strategyproof Assignment by Hierarchical Exchange,” *Econometrica*, 68, 1403–1433.

Pathak, P. (2017): “What Really Matters in Designing School Choice Mechanisms,” in *Advances in Economics and Econometrics, 11th World Congress of the Econometric Society*, ed. by B. Honore, A. Pakes, M. Piazessi, and L. Samuelson, Cambridge University Press.

Pathak, P. A. and J. Sethuraman (2011): “Lotteries in Student Assignment: An Equivalence Result,” *Theoretical Economics*.

Pycia, M. (2011a): “The Cost of Ordinality,” Working Paper.

——— (2011b): “Ordinal Efficiency, Fairness, and Incentives in Large Multi-Unit-Demand Assignments,” Working Paper.

Pycia, M. and P. Troyan (2019): “A Theory of Simplicity in Games and Mechanism Design,” Working Paper.

Pycia, M. and M. U. Ünver (2017): “Incentive Compatible Allocation and Exchange of Discrete Resources,” *Theoretical Economicse*, 12, 287–329.

Pycia, M. and U. Ünver (2011): “Trading Cycles for School Choice,” Working Paper.

Sönmez, T. and M. U. Ünver (2010): “Course Bidding at Business Schools,” *International Economic Review*, 51, 99–123.

Talagrand, M. (1995): “Concentration of measure and isoperimetric inequalities in product spaces,” *Publications Mathématiques de l’I.H.É.S.*, 81, 73–205.

Troyan, P. (2012): “Comparing School Choice Mechanisms by Interim and Ex-Ante Welfare,” *Games and Economic Behavior*, 75(2), 936–947.

A Proofs

A.1 Proof of Theorem 1

From Theorem 5—proven without any direct or indirect reliance on Theorem 1—we know that

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

where the expectation is taken over the uniform distribution on all preference profiles.[29] The displayed inequality implies its counterpart for all $\ell=1,\ldots|K|$,

$$\mathbb{E}\left|F_{\ell}\left(\succ,\phi\left(\succ\right)\right)-F_{\ell}\left(\succ,\psi\left(\succ\right)\right)\right|<\epsilon.$$

Since ϕ and ψ are robust with ratio c, we can conclude that a change of report by one agent changes $F_{\ell}\left(\succ,\phi\left(\succ\right)\right)$ by at most $\frac{c}{|N|}$, and hence $F_{\ell}\left(\succ,\phi\left(\succ\right)\right)$ is Lipschitz

[29]For the analogue of the positive result for distributions satisfying the assumptions of Theorem 6, we base the above estimate on Theorem 6 (whose proof is also does not rely on Theorem 1).

in Hamming metrics, and the above Talagrand concentration inequality for product spaces allows us to conclude that for nearly all preference profiles, $F_\ell(\succ, \phi(\succ))$ is close to its mean. The same obtains for $F_\ell(\succ, \psi(\succ))$. On the intersection of this nearly full measure sets of profiles $F_\ell(\succ, \phi(\succ))$ and $F_\ell(\succ, \psi(\succ))$ are close, and hence so is the sum of the absolute differences. Hence, Theorem 1 obtains.[30] QED

A.2 Proof of Theorem 2

Theorem 7—proven without direct or indirect reliance on Theorem 2—tells us that for any two mechanisms in $\mathcal{M}_{\text{Standard}}$ the medians of anonymous statistics over any exchangeable population are the same. Let us fix the population and the exchangeable distribution, and denote the common median of F by F_M. Fix also a coding label ℓ.

The mechanisms in $\mathcal{M}_{\text{Standard}}$ are robust with ratio $c = |A|$; cf. the discussion after the robustness definition. Thus F_ℓ is Lipschitz with constant $L = \frac{c}{|N|}$. By Talagrand (1995) concentration inequality

$$\mathbb{P}\left(|F_\ell(\succ, \phi(\succ)) - F_M| > t\right) \leq 4 \exp\left\{-\frac{t^2}{L^2 N}\right\}.$$

Because the same holds true for ψ, we infer that

$$\mathbb{P}\left(|F_\ell(\succ, \phi(\succ)) - F_\ell(\succ, \psi(\succ))| > 2t\right) \leq 8 \exp\left\{-\frac{t^2}{L^2 N}\right\}.$$

Plugging in for L and halving t delivers the first claim of the theorem. The proof of the second claim is analogous except that $L = \frac{2c}{|N|}$ because a change of report by one agent can affect allocation of c agents and each allocation change can impact two codes in F. QED

[30]For the version of Theorem 1 formulated under the distributional assumptions of Theorem 6, instead of the product-space concentration inequalities, we use Maurey Theorem quoted above.

A.3 Proof of Theorem 6

The key to the proof is to first consider symmetric versions of mechanisms ϕ and ψ and then to apply our asymptotic duality Theorem 5. The analysis of symmetrizations of ϕ and ψ builds on Liu and Pycia (2011). They establish a general asymptotic equivalence of symmetric mechanisms—a class that includes symmetrizations—but their analysis relies on regularity conditions we did not impose. Because of this some work beyond their paper is needed.

We need to establish that the effect of reports of any small groups of agents other than j on the allocation of an agent j vanishes as $|N| \to \infty$. We will prove something stronger than they required for their analysis, but only for preferences $\succ \in \mathcal{P}_\delta$. While they imposed a regularity condition on all preference profiles, the restricted condition below is sufficient for a version of their equivalence restricted to $\mathcal{P}_\delta$. We say a mechanism ϕ is *continuous* with ratio $C > 0$ if for every $\epsilon > 0$ and any preference profiles $\succ, \succ'$ if

$$|\{i \in N | \succ'_i \neq \succ_i\}| < \frac{1}{C}\epsilon \, |N| \, , \tag{3}$$

then for all agents $j \in N$ such that $\succ'_j = \succ_j$ we have:

$$\max_{a \in A} |\phi_N(\succ_N)(j,a) - \phi_N(\succ'_N)(j,a)| < \epsilon. \tag{4}$$

Lemma 2. *Suppose $\succ \in \mathcal{P}_\delta$. If a mechanism ϕ is robust with ratio c then, its symmetrization ϕ^S is continuous with any ratio $C \geq \frac{c}{\delta}$.*

To use Liu and Pycia equivalence approach we also need to ensure approximate ordinal efficiency of the symmetrizations under our assumptions. Following their terminology, for any $\epsilon > 0$, we say that a random allocation μ is $(1-\epsilon)$-*efficient* with respect to a preference profile $\succ$ iff (i) for no $n \geq 2$ there is a cycle of agents $i_1, ..., i_n$ who can improve by trading ϵ probability shares of some objects $a_1, ..., a_n$

that is $\mu(i_k, a_k) > \epsilon$ and $a_{k+1} \succ_{i_k} a_k$ for all $k = 1, ..., n$ (modulo n), and (ii) if $a \succ_i b$ and $\mu(i, b) > \epsilon$, then no more than probability ϵ of a is unallocated.

Lemma 3. *For any $\succ \in \mathcal{P}_\delta$, if a Pareto efficient deterministic mechanism ϕ is robust with ratio c then its symmetrization ϕ^S is $(1-\epsilon)$-efficient for any $\epsilon \geq 4\sqrt{c}\sqrt[4]{\log(2|A|)}n^{-\frac{1}{4}}$.*

The above two lemmas—proved below—allow us to recover an analogue of the main equivalence result of Liu and Pycia (2011) restricted to $\mathcal{P}_\delta$. We can hence conclude that for all $\epsilon > 0$, and sufficiently large N, and any strategy-proof, Pareto-efficient, and robust deterministic mechanisms ϕ and ψ, their symmetrizations are approximately identical:

$$\left|\phi^S(i,a)(\succ) - \psi^S(i,a)(\succ)\right| < \epsilon$$

for all $i \in N$, $a \in A$, and $\succ \in \mathcal{P}_\delta$.

This approximate equivalence of symmetrized mechanisms allows us to apply our approximate duality Lemma 4 because $\mathcal{P}_\delta$ is permutation invariant. Hence for the distribution $\mathbb{P}$ conditioned on $\succ \in \mathcal{P}_\delta$, the analysis of symmetrizations ϕ^S and ψ^S in Section 3 and our asymptotic duality Theorem 5 imply that

$$\mathbb{E}_{\mathcal{P}_\delta} \sum_{\ell=1}^{|K|} |F_\ell(\succ, \phi(\succ)) - F_\ell(\succ, \psi(\succ))| < \frac{\epsilon}{3}$$

(where we substituted $\frac{\epsilon}{3}$ for ϵ in applying the analysis referenced). On the complement of $\mathcal{P}_\delta$, we have

$$\sum_{\ell=1}^{|K|} |F_\ell(\succ, \phi(\succ)) - F_\ell(\succ, \psi(\succ))| \leq 2.$$

By averaging these bounds in the worst case scenario of $\mathbb{P}(\mathcal{P}_\delta) = 1 - \epsilon$, we obtain

$$\mathbb{E} \sum_{\ell=1}^{|K|} |F_\ell(\succ, \phi(\succ)) - F_\ell(\succ, \psi(\succ))| < \frac{\epsilon}{3}(1-\epsilon) + 2\frac{\epsilon}{3} \leq \epsilon,$$

as required. QED

Proof of Lemma 3.[31] Consider a sequence of preference profiles $(\succ_N)_{N=\{1\},\{1,2\},\ldots}$. The key step in our proof is the following.

Claim. Let θ be a preference ranking, let a be an object, and let ψ^S be a symmetrization of a robust deterministic mechanism ψ with ratio c. If $(\succ_N) \in \mathcal{P}_{\delta,N}$ and there is an agent with preference ranking θ who obtains a under ψ^S with probability at least $\epsilon \in (0,1)$, then the fraction of permutations σ at which no agent of type θ is allocated object a under $\sigma\left(\psi^S\right)$ is at most $2e^{-\frac{\epsilon^4\delta^2 n}{16^2c^2}}$.

Proof of the Claim. Let $\hat{\Sigma} \subseteq \Sigma_N$ be the set of permutations σ such that no agent of type θ obtains a at $\sigma(\psi)$. Because we are bounding the size of $\hat{\Sigma}$ from above, we can without loss of generality assume that there is at least one permutation $\tau^* \in \hat{\Sigma}$.

Let $B \subseteq \Sigma_N$ be the set of permutations $\sigma \in \Sigma_N$ such that at least the fraction $\frac{\epsilon}{2}$ of agents of type θ obtains a under ψ^σ. Notice that B contains at least the fraction $\frac{\epsilon}{2}$ of all permutations. Indeed, if not then at most the fraction $\frac{\epsilon}{2}$ of the permutations of ψ contains more than fraction $\frac{\epsilon}{2}$ of agents of type θ obtaining a, and at all other permutations at most fraction $\frac{\epsilon}{2}$ of these agents obtain a. Thus the average probability an agent of type θ obtains a is bounded above by $\frac{\epsilon}{2} * 1 + \left(1-\frac{\epsilon}{2}\right)\frac{\epsilon}{2} < \epsilon$. By symmetry of the randomization, all these agents have the same probability of obtaining a, and because this probability is at least ϵ, we obtain a contradiction.

Because ψ is robust with ratio c, so is $\sigma(\psi)$ for any $\sigma \in \Sigma_N$. Because there is at least fraction δ of agents of type θ, if weakly more than the fraction $\frac{\epsilon}{2}$ of agents of type θ obtain a under $\sigma(\psi)$ then at least the fraction $\frac{\epsilon}{2c}\delta$ of agents must have changed their reports between σ and τ^* and hence $f(\sigma) = d(\sigma, \tau^*) \geq \frac{1}{|N|}\frac{\epsilon}{2c}\delta|N| = \frac{\epsilon}{2c}\delta$. From this and the bound on B we can conclude that $\mathbb{E}f = \frac{\sum_{\tau\in\Sigma_N} f(\tau)}{|\Sigma_N|} \geq \frac{\epsilon\delta}{2c} * \frac{\epsilon}{2}$.

Because f equals zero on $\hat{\Sigma}$, we conclude that all $\sigma \in \hat{\Sigma}$ belong to $\left\{\sigma \in \Sigma_N : \left|f(\sigma) - \frac{\sum_{\tau\in\Sigma_N} f(\tau)}{|\Sigma_N|}\right| \geq t\right\}$ for $t = \frac{\epsilon^2\delta}{4c}$. We can use Maurey's (1979) concentration theorem: If $f : \Sigma_N \to$

[31] The proof slightly refines the analysis of symmetric mechanisms and asymptotic efficiency in Liu and Pycia (2016).

$\mathbb{R}$ is Lipschitz with constant L with respect to the Hamming distance $d(\sigma,\sigma') = \frac{1}{|N|}|\{i \in N : \sigma(i) \neq \sigma'(i)\}|$, then:

$$\left|\left\{\sigma \in \Sigma_N : \left|f(\sigma) - \frac{\sum_{\tau\in\Sigma_N} f(\tau)}{|\Sigma_N|}\right| \geq t\right\}\right| \leq 2e^{-\frac{t^2 n}{16L^2}} |\Sigma_N| \, .$$

Setting $f(\sigma) = d(\sigma, \tau^*)$ where $\tau^* \in \hat{\Sigma}$ is fixed above observe that this function is Lipschitz with respect to the distance d with constant $L = 1$. By Maurey's Theorem $\hat{\Sigma}$ can contain at most $2e^{-\frac{\epsilon^4\delta^2 n}{16^2 c^2}} |\Sigma_N|$ permutations and the claim obtains. QED

To finish the proof of the efficiency lemma, suppose that ϵ-efficiency fails for preference profile $\succ_N$. Thus, either condition (i) or (ii) fails. The analyses of these two cases are similar; let us consider the slightly more complex case when (i) fails. Then, at each N, there is a cycle of ℓ agents $i_1, ..., i_\ell$ that have a profitable swap of size ϵ, that is $\mu(i_k, a_k) > \epsilon$ and $a_{k+1} \succ_{i_k} a_k$ for all $k = 1, ..., \ell$. Notice that the uniformity of randomization implies that the agents with the same preference type obtain the same probability shares. The resulting symmetry of allocations allows us to assume that $\ell \leq |A|$ as otherwise we could shorten the cycle.

By Pareto efficiency of ψ, at each $\sigma \in \Sigma_N$ at least one agent of type from the cycle does not obtain the corresponding object; that is this claim obtains with probability 1. Because Claim 1 puts an upper bound of $|A| \, 2e^{-\frac{\epsilon^4\delta^2 n}{16^2 c^2}}$ on the total probability this might happen , we infer that for

$$\epsilon \geq 4\sqrt{c}\sqrt[4]{\log(2|A|)} n^{-\frac{1}{4}}$$

condition (i) cannot fail. The failure of condition (ii) would give us a less tight bound on ϵ. QED

Proof of Lemma 2. We are to show that no agent i's random allocation is changed by more than ϵ when an agent $j \neq i$ changes his or her report. By way of contradiction suppose that there are two different agents i, j, object a, and such that when j changes the reported preferences from $\succ_j$ to $\succ'_j$ then

$$\left|\phi^S\left(\succ_N\right)(i,a)-\phi^S\left(\succ'_j,\succ_{N-\{j\}}\right)(i,a)\right|\geq\epsilon.$$

By symmetry, the change by j affects in the same way all agents of the same preference type as i, and by assumption there are at least $\delta\left|N\right|$ such agents.

For each σ, agent j affects at most c agents of same type as i, and by symmetry ϵ is bounded above by the probability that i is one of c uniform random draws (without replacement) from $\delta\left|N\right|$ agents, that is $\epsilon\leq\frac{c}{\delta N}$. QED

A.4 Proof of Theorem 7

The Theorem follows from Theorem 8 and the median-version of Lemma 1 (cf. Remark 1). QED

A.5 Proof of Theorem 8

Consider a school choice environment and fix a preference profile $(\succ_i)_{i\in N}$. Let ϕ^S be a Symmetric Top Trading Cycles mechanism and let ϕ be its underlying Top Trading Cycles mechanism. Consider an auxiliary house allocation environment, with the same set of agents the school choice environment we study and the set of objects O constructed as a union of all seats in all schools; let us associate each seat in school a with a unique number in $1,...,|a|$. Let us construct each agent i's preference ranking $\tilde{\succ}_i$ over these objects as follows: if $o_1,o_2\in O$ are seats in different schools a_1 and a_2 (respectively) then $o_1\tilde{\succ}_i o_2$ iff $a_1\succ_i a_2$; if $o_1,o_2\in O$ are seats in the same school then $o_1\tilde{\succ}_i o_2$ iff the number associated with seat o_1 is strictly lower than the number associated with seat o_2.

Let us construct the auxiliary mechanism $\tilde{\phi}$ on the auxiliary house allocation problem so that $\tilde{\phi}$ is Top Trading Cycles in which each seat o has the same priority ranking over agents as the school a of this seat. The key observation is that ϕ is identical to the projection of the allocation of $\tilde{\phi}$ into the space of allocation of schools to applicants. The same obtains for any permutation of roles in ϕ and correspondingly in $\tilde{\phi}$, and hence ϕ^S is equivalent to the projection of $\tilde{\phi}^S$ into the space of allocation of schools to applicants. By Carroll (2014), all $\tilde{\phi}^S$ are equivalent in the auxiliary house allocation problem, and hence we can conclude that all ϕ^S are equivalent in the original problem. QED

B Additional Results and Remarks

B.1 Subdomains of Preference Profiles

We can restrict the definition of anonymous statistics to a support of an exchangeable distribution that is to a subset of preference profiles that is closed under permutations of agents: a class $\mathcal{C}$ of preference profiles is closed under permutations if for every $\succ_N \in \mathcal{C}$ and every permutation of agents σ the permuted profile $\succ_{\sigma(N)} \in \mathcal{C}$. All the results of this paper—including Theorem 1—would remain valid.

B.2 Replica Economies

Restricting attention to replica economies would simplify some of our results. For instance, in Lemma 3 we may drop the restriction to $\succ \in \mathcal{P}_\delta$ and we do not need to replace it with any full support assumptions. In the formula for efficiency tolerance ϵ, we then substitute the share of the least frequent preference profile for δ.

B.3 Approximate and Asymptotic Duality

The proofs of the following additional duality lemmas are essentially the same as that of the finite-market duality Lemma 1.

Lemma 4. ***(Approximate Population-Symmetry Duality).*** *Suppose that class $\mathcal{C}$ of preference profiles is closed under permutations. Then:*

(1) Take two mechanisms, ϕ and ψ, such that their symmetrizations have approximately identical marginal distributions of outcomes

$$\left|\phi^{S}(i,a)(\succ)-\psi^{S}(i,a)(\succ)\right|<\epsilon \quad \text{for } \succ\in\mathcal{C}, i\in N, a\in A \tag{5}$$

If agents' types are drawn from an exchangeable distribution then the mean of any invariant measure of outcomes under ϕ is approximately the same as under ψ,

$$\mathbb{E}_{\mathcal{C}}\sum_{\ell=1}^{|K|}\left|F_{\ell}(\succ,\phi(\succ))-F_{\ell}(\succ,\psi(\succ))\right|<\epsilon. \tag{6}$$

(2) Conversely, if for two mechanisms, ϕ and ψ, and any exchangeable distribution of preferences the mean of any invariant measure of outcomes under ϕ is approximately the same as under ψ in the sense of inequality 6, then the symmetrizations of ϕ and ψ have nearly identical marginal distributions of outcomes in the sense of inequality 5.

Lemma 5. ***(Asymptotic Population-Symmetry Duality).***

(1) Take two sequences of populations N and corresponding mechanisms, ϕ_N and ψ_N such that the symmetrizations ϕ_N^S and ψ_N^S have asymptotically equivalent marginal distributions of outcomes. If agents' types are drawn from a exchangeable distribution then the mean of any invariant measure of outcomes under ϕ_N is asymptotically the same as under ψ_N.

(2) Conversely, for two sequences of populations N and corresponding mechanisms, ϕ_N and ψ_N such that for any exchangeable distribution of preferences the mean of any invariant measure of outcomes under ϕ is the same as under ψ, then the symmetrizations ϕ_N^S and ψ_N^S have asymptotically equivalent marginal distributions of outcomes.

B.4 Further Equivalences for Symmetric School Choice Mechanisms

The proof approach from Theorem 8 allows us to derive other equivalences of symmetric mechanisms in school choice. For instance, it allows us to leverage the analysis in Lee and Sethuraman (2011) to conclude that any two Pycia and Ünver (2011) extensions of Pápai (2000) Hierarchical Exchange mechanisms to school choice environment have identical marginal distributions of outcomes; that is the lottery over schools an agent obtains is the same for all these mechanisms.

Furthermore, the same proof as in Theorem 8 allows us to establish the analogue of this theorem for the following larger class of school choice mechanisms, which we call Generalized Top Trading Cycles. There are two inputs into each of these mechanisms: a profile of priority orderings separate for each seat in each school, and sets of rankings of seats in schools, one set of rankings per agent. Given these two inputs, the outcome of Generalized Top Trading Cycles is the same as the outcome of Top Trading Cycles in which each seat is treated as a school and agents' preferences among seats are constructed so that if $o_1, o_2 \in O$ are seats in different schools a_1 and a_2 (respectively) then $o_1 \tilde{\succ}_i o_2$ iff $a_1 \succ_i a_2$; if $o_1, o_2 \in O$ are seats in the same school then $o_1 \tilde{\succ}_i o_2$ iff agent i ranks seat o_1 above o_2 in line with his ranking of seats in this school.[32]

[32] As mentioned above, Top Trading Cycles restricted to this domain are identical to Pápai (2000) Fixed-Endowment Hierarchical Exchange. The strategy-proofness and Pareto efficiency properties of Generalized Top Trading Cycles are inherited from Top Trading Cycles.