Obvious Dominance and Random Priority

Marek Pycia and Peter Troyan*

October 2016

Abstract

In environments without transfers, the paper shows that random priority is the unique mechanism that is obviously strategy-proof, ex-post Pareto efficient, and symmetric; hence providing an explanation for the popularity of this mechanism. The paper also constructs the full class of obviously strategy-proof and efficient mechanisms, and explores a natural strengthening of obvious strategy-proofness.

1 Introduction

The central concerns in designing mechanisms for school choice, organ transplantation, course allocation, voting, and other social choice problems without transfers are incentives, efficiency, and fairness.[1] One of the main in-

*UCLA and University of Virginia, respectively. The present work is a result of a merger of two independent single-author papers. First presentation: April 2016. First posted draft: June 2016. We are still polishing the exposition; updates coming soon. For their comments, we would like to thank Ed Green, Shengwu Li, Giorgio Martini, Stephen Morris, Utku Unver, and the Eco 514 students at Princeton. Simon Lazarus provided excellent research assistance. Marek would also like to acknowledge the financial support of the William S. Dietrich II Economic Theory Center at Princeton.

[1]See e.g. Abdulkadiroğlu and Sönmez (2003) for school choice, Roth, Sönmez, and Ünver (2004) for transplants, Sönmez and Ünver (2010) for course allocation, and Arrow (1963) for voting and social choice.

centive constraints commonly imposed is dominant strategy incentive compatibility, also known as *strategy-proofness.* While using a strategy-proof mechanism guarantees that truthtelling is a dominant strategy theoretically, this is useful only to the extent that actual participants fully understand it in practice, which may be evident to economic theorists, but inconspicuous to non-experts. Li (2016) proposes a refinement of strategy-proofness called *obvious strategy-proofness (OSP)* that characterizes the mechanisms that are 'easy' to understand as strategy-proof. While strong incentives are desirable ceteris paribus, they can lead to tradeoffs on other dimensions, which raises the question of precisely which social choice functions are implementable under OSP.

We show that there is effectively a unique mechanism that is obviously strategyproof, efficient, and symmetric: the well-known (extensive-form) Random-Priority mechanism. Thus, our results give an explanation for the popularity of this mechanism. We also construct the full class of obviously strategy-proof and efficient mechanisms for social choice environments without transfers. Last, we argue that this class still contains mechanisms for which strategy-proofness is not necessarily immediately clear, and so we introduce a further refinement we call *strong obvious strategy-proofness (SOSP).* Such mechanisms are simple to play. In particular, strongly obviously strategy-proof and efficient mechanisms are basically sequential dictatorships.

Informally, an imperfect information extensive-form game is obviously strategy-proof if, whenever an agent is called to play, there is an action such that the even *worst* possible final outcome from that action is (weakly) better than the *best* possible outcome from any other action (where what is possible may depend on the actions of other agents in the future). For example, the Random-Priority mechanism, which randomly orders the agents and at their turn gives them their favorite good that remains, can be implemented in two ways: (i) sequentially present each agent with the remaining objects at her

turn and allow them to pick their favorite or (ii) ask each agent to submit an entire preference list to 'the mechanism', which then outputs the same assignment as in (i). While both of these mechanisms are strategyproof in the classical sense, the former is obviously strategyproof, while the latter is not.[2]

Our first main result, Theorem 1, provides an explanation for the popularity of the Random-Priority mechanism in a myriad of allocation problems, both formal and informal, including house allocation, school choice, and course allocation, among others. Random Priority is obviously strategyproof, as recognized by Li; it is also well-known to be efficient and to treat agents symmetrically (that is any two agents who follow the same strategies obtain the same distribution over allocations).[3] Theorem 1 shows that no other mechanism satisfies these three properties: an extensive-form game is obviously strategy-proof, efficient, and symmetric if and only if it is random priority.[4] This insight resolves positively the quest to establish random priority as the unique mechanism with good incentive, efficiency, and fairness

[2]The observation that the extensive-form of random priority is obviously strategy-proof is due to Li (2016). Another extensive form game also called random priority is as follows: nature draws an agent uniformly at random, this agent picks an object, then nature draws an agent who has not moved yet, this agent picks an object, etc. These two games, while not identical, are equivalent in the following sense: for each profile of agents' preferences, each game admits a profile of obviously dominant strategies such that the two games lead to the same outcomes when the obviously dominant strategies are played. For simplicity, in this introduction we ignore differences between equivalent games.

[3]Random Priority as the extensive-form game introduced above is obviously strategy-proof. The static direct mechanism in which agents report their types and then the random priority game is played by proxies is not obviously strategy-proof, but when agents' type reports instruct the proxies to play the obviously dominant strategy, the static random priority leads to the same outcomes as its extensive-form counterparts. Intuitively, describing a direct-revelation game in terms of an obviously strategy-proof game played by proxies makes the direct-revelation game simpler to play than alternative descriptions. For discussion of efficiency and symmetry see, e.g., Abdulkadiroğlu and Sönmez (1998), Bogomolnaia and Moulin (2001), and Che and Kojima (2010).

[4]Note that efficiency assumes that agents report truthfully. If agents do not report truthfully, which can often be the case for mechanisms that are strategyproof, but not obviously so, efficiency may be lost.

properties.[5]

To prove this result, it is first necessary to understand the class of OSP and efficient mechanisms, which we present as Theorem 2. The extensive-form (deterministic) serial dictatorship, in which players take turns choosing their most preferred outcome (or set of outcomes) from those that remain is OSP because it requires no contingent reasoning: each agent moves only once and can easily select her most preferred outcomes from those that are presented to her. As we show in Theorem 2, serial dictatorships are not the only OSP and efficient games, but are a special case of what we call *millipede games*. In a millipede game, there are (at most) two agents who are *active* at any given point. One agent, i, is first presented with one or several subsets of outcomes that she can 'clinch'. If i clinches some subset of outcomes X, then the final outcome is guaranteed to lie in X, and i leaves the game; alternatively, i can 'pass' to the other active agent, j, who is similarly presented with one or several subsets of outcomes she can 'clinch'. Agents i and j continually pass back and forth, with their possible clinchable outcomes growing monotonically, until one of them, say i, chooses a clinching action and leaves the game, at which point a new agent k enters. The new active agents are j and k, and the process is repeated until a unique final outcome is determined.[6]

The key assumption behind these results is that outcomes an agent is not indifferent between can be ranked in any order. This key assumption

[5]In single-unit demand allocation with at most three agents and three objects, Bogomolnaia and Moulin (2001) proved that random priority is the unique mechanism that is strategy-proof, efficient, and symmetric. In markets in which each object is represented by many copies, Liu and Pycia (2013) and Pycia (2011) proved that random priority is the asymptotically unique mechanism that is symmetric, asymptotically strategy-proof, asymptotically ordinally efficient (a more demanding efficiency criterion than ex post Pareto efficiency). While these earlier results looked at either very small or very large markets, Theorem 1 is true for any number of agents and objects.

[6]In the well-known *centipede game*, two agents pass back and forth until one of them 'takes' the pot and leaves the game. Our games have a similar take-or-pass structure, but with many more options of what to take; i.e., they look like centipede games with more "legs" (see Figure 4). Hence, we call these games *millipede games*.

is trivially satisfied in standard models without transfers, but it excludes environments with transfers.[7] The results are robust to how agents' evaluate lotteries over outcomes, and they are derived in a very general model encompassing the standard allocation and social choice environments.

Li (2016) formalizes the desirability of OSP as an incentive property by showing that (i) boundedly rational agents who may not be able to engage in contingent reasoning can still see that truthful reporting is optimal in OSP games and (ii) it is easier for a principal with only partial commitment power to commit to a mechanism if it is OSP. Further, experimentally, subjects play dominant strategies significantly more often in games that are OSP compared to games that are strategyproof, but not obviously so. While Li's main focus is on environments with transfers,[8] in environments without transfers, he shows that the extensive-form random priority is obviously strategy-proof, while other standard mechanisms—top trading cycles and the static version random priority—are not.[9] A key contribution of Li (2016) is to show why some extensive forms of a direct mechanism are more often encountered in practice than others, despite the revelation principle. Our analysis of the strong obvious strategy-proofness is a next step in the agenda of understanding differences among extensive forms.

Following up on Li's work, but preceding ours, Ashlagi and Gonczarowski (2016) study whether stable mechanisms such as deferred acceptance (DA) are obviously strategy-proof. They show that the answer is generally no, ex-

[7]Li (2016) constructed the class of obviously strategy-proof mechanisms in bilateral choice environments with transfers and quasilinear utilities, and he showed that in these environments the obviously strategy-proof mechanisms resemble ascending auctions.

[8]In environments with transfers, Loertscher and Marx (2015) construct a prior-free obviously strategy-proof mechanism that becomes asymptotically optimal as the number of buyers and sellers grows.

[9]Li considers the original construction of top trading cycles of Shapley and Scarf (1974) in which each agent starts by owning exactly one object, and shows it is not obviously strategyproof. Other papers (e.g., Abdulkadiroğlu and Sönmez (2003)) have extended TTC to more general "ownership structures", some of which will be a special case of our class of millipede games, and thus will be obviously strategyproof.

cept in special environments where DA has a similar 'clinch or pass' structure as a millipede game.

More generally, this paper adds to our understanding of incentives, efficiency, and fairness in settings without transfers. In addition to Gibbard (1973, 1977) and Satterthwaite (1975), and the allocation papers mentioned above, the literature on mechanisms satisfying these key objectives includes Pápai (2000), Ehlers (2002) and Pycia and Ünver (2009) who characterized efficient and group strategy-proof mechanisms in settings with single-unit demand, and Pápai (2001) and Hatfield (2009) who provided such characterizations for settings with multi-unit demand.[10] Morrill (2014) and Hakimov and Kesten (2014) characterized deterministic mechanisms that satisfy incentive, efficiency, and fairness objectives.

2 Preliminaries

2.1 Model

Let $\mathcal{N}$ be a set of agents, and $\mathcal{X}$ a finite set of outcomes.[11] Each agent has a preference ranking over outcomes. The domain of preferences of agent $i \in \mathcal{N}$ is denoted $\mathcal{P}_i$, and a generic preference ranking in $\mathcal{P}_i$ is denoted by $\succ_i$. We allow for indifferences, and write $x \sim_i y$ if neither $x \succ_i y$ nor $y \succ_i x$. For any $x \in \mathcal{X}$, we write $I_i^{\succ_i}(x) = \{y \in \mathcal{X} : x \sim_i y\}$ to denote the indifference class for x under preferences $\succ_i$. For brevity, we sometimes suppress the dependence on $\succ_i$ and write $I_i(x)$ to represent an indifference class.

[10] Pycia and Ünver (2016) characterized individually strategy-proof and Arrovian efficient mechanisms. For an analysis of these issues under additional feasibility constraints, see also Dur and Ünver (2015).

[11] The assumption that $\mathcal{X}$ is finite simplifies the exposition and it is satisfied in such examples of our setting as voting and the no-transfer allocation environments listed in the introduction. This assumption can be relaxed. For instance, our analysis goes through with no substantive changes if we allow infinite $\mathcal{X}$ endowed with a topology such that agents' preferences are continuous in this topology and the relevant sets of outcomes are compact.

We assume that (a) for any agent i there is a partition of $\mathcal{X}$ such that for every $\succ_i \in \mathcal{P}_i$ agent i is indifferent between any two outcomes in the same element of the partition. In particular, each $\succ_i \in \mathcal{P}_i$ might be identified with a ranking on the elements of the partition. We further assume that (b) every strict ranking of the elements of the partition is in $\mathcal{P}_i$. Finally, we assume that (c) for every subset of agents $J \subsetneq \mathcal{N}$ and outcome $x \in \mathcal{X}$, if there is another outcome y such that all agents in J weakly prefer y over x and for at least one of them the preference is strict, then there exists a profile of preferences in $\times_{i \in \mathcal{N}-J} \mathcal{P}_i$ and outcome z such that all agents weakly prefer z over x and for at least one of them the preference is strict. The set of all strict preferences and the set of all preferences both satisfy this assumption. The assumption is also satisfied in allocation problems in which each agent cares only about his or her allocation. In allocation environments, each element of the partition in (a) can be identified with the allocation of agent i. Assumption (b) then means that each agent can strictly rank his or her allocations in every possible way (note that weak rankings are also allowed, but are not required). Assumption (c) is satisfied, for instance, if agents have outside options, or if there are at least as many objects as agents collectively may demand. The above assumptions fail in settings with transfers in which each agent always prefers having more money to less.[12]

When dealing with lotteries, we are agnostic as to how agents evaluate them, as long as the following property holds: an agent prefers lottery μ over ν if (i) the agent is indifferent among all outcomes in the intersection of the supports $X = \text{supp}\,(\mu) \cap \text{supp}\ \ (\nu)$, and (ii) the agent strictly prefers every outcome in $\text{supp}\,(\mu) - X$ to every outcome in $\text{supp}\,(\nu) - X$. This mild assumption is satisfied for expected utility agents; it is also satisfied for

[12]In formulating these assumptions, we balance simplicity and generality. Our results and proofs do not require the full strength of these assumptions. For instance, Appendix C makes the point that we do can relax these assumptions when agents outside options. For the analysis of obvious dominances in quasilinear environments with transfers, see the foundational paper of Li (2016).

agents who prefer μ to ν as soon as μ first-order stochastically dominates ν.

2.2 Obvious strategy-proofness

The standard notion of strategy-proofness is defined for normal form games, while obvious strategy-proofness applies to extensive-form games. An imperfect-information *extensive-form games* with perfect recall is defined in the standard way, as a collection of partially ordered *histories* (sequences of moves), where at every history h, one agent $i \in \mathcal{N}$ has a (finite) set of actions $A(h)$ from which to choose (for a formal definition, see Li (2016)). Slightly abusing notation, we write (h, a) to denote the history h followed by the move a. Each terminal history is associated with an outcome in $\mathcal{X}$, and agents receive payoffs at each terminal history that are consistent with their preferences over outcomes $\succ_i$. Without loss of generality we assume that at all information sets at which an agent moves, she has at least two actions to choose from. This simplifying assumption is only imposed on agents: we allow histories at which nature has only one move. While we do not consider incomplete information, our insights remain valid for situations in which agents' information is incomplete: we allow imperfect information, and the standard interpretation of incomplete information games as imperfect information games can be used.

A *strategy* for a player i is a function that specifies an action for agent i at each one of her information sets. A strategy profile $S = (S_i)_{i \in N}$ is a list of strategies, one for each agent. In addition, for any history at which nature makes a move, let $\omega(h) \in A(h)$ and ω be a function specifying an action for nature at each history where nature is to move.[13] We sometimes write (S, ω) to denote a strategy profile that includes nature's moves. Let $z(h, S, \omega)$ be the terminal history reached when starting from h and proceeding according to (S, ω).

[13] ω can be thought of as describing one particular realization of "chance moves" throughout the game at every history at which nature is called to play.

Following Li (2016), for a game Γ, a strategy S_i *obviously dominates* another strategy S_i' for player i if, at the first information set at which these two strategies diverge, the worst outcome the agent can obtain playing S_i is at least as good as the best outcome the agent can obtain playing the other strategy, ranging over all possible (S_{-i}, ω). A profile $(S_i(\cdot))_{i\in\mathcal{N}}$ of strategies is *obviously dominant* if for any player i and type $\succ_i$ the strategy $S_i(\succ_i)$ obviously dominates any other strategy. When there exists a profile of strategies $(S_i(\cdot))_{i\in N}$ that is obviously dominant, we say Γ is *obviously strategyproof*. An example of an obviously strategy-proof game is the well-known random priority game (see the definition and discussion in the introduction).[14]

An *extensive-form mechanism*, or simply a *mechanism,* is an extensive form game Γ together with a profile of strategies $(S_i(\succ_i))_{i\in N}$ for each preference (or type) profile $\succ$ over outcomes. A mechanism is *obviously strategy-proof* when strategies $(S_i(\succ_i))_{i\in N}$ are obviously dominant. A mechanism is *Pareto efficient* if, for all type profiles $(\succ_i)_{i\in N}$, when the agents follow strategies $(S_i(\succ_i))_{i\in N}$, every possible outcome (for all possible chance moves) is Pareto efficient, i.e., for all possible outcomes x, there is no other y such that $y \succeq_i x$ for all i, and $y \succ_i x$ for some i.

Two obviously strategy-proof extensive-form mechanisms $\left(\Gamma, (S_i(\cdot))_{i\in N}\right)$ and $\left(\Gamma', (S_i'(\cdot))_{i\in N}\right)$ are *equivalent* if, for every profile of types $(\succ_i)_{i\in\mathcal{N}}$ the distribution of outcomes in Γ when agents play $(S_i(\succ_i))_{i\in N}$ is the same as in Γ' when agents play $(S_i'(\succ_i))_{i\in N}$. Two games Γ and Γ' are *equivalent* when there are profiles of strategies $(S_i(\cdot))_{i\in\mathcal{N}}$ and $(S_i'(\cdot))_{i\in\mathcal{N}}$ such that mechanisms $\left(\Gamma, (S_i(\cdot))_{i\in\mathcal{N}}\right)$ and $\left(\Gamma', (S_i'(\cdot))_{i\in\mathcal{N}}\right)$ are obviously strategy-proof and equivalent. Lemma 1 shows that for each obviously strategy-proof game, there is an equivalent obviously strategy-proof game that has perfect information.[15]

[14] We consider pure strategies, but the analysis can be extended to mixed strategies.

[15] The insight on the sufficiency of perfect information holds true and our proof remains valid for all domains of preferences, not only domains of preferences studied in this paper.

3 Main Result on Allocation: Obvious Dominance and random priority

We now study the special case of our problem in which outcomes consist of allocation of objects to agents. There is a set of objects $\mathcal{O}$, each object $o \in \mathcal{O}$ is represented by $|o|$ copies, where every $|o|$ is a positive integer.[16] The set of feasible one agent's allocations is $Q \subseteq \times_{o\in\mathcal{O}} \{0, 1, ..., |o|\}$. We assume that if $q_i \in Q$ then Q contains all $\left(q^1, ..., q^{|\mathcal{O}|}\right) \neq (0, ..., 0)$ such that for each coordinate j we have $q^j \leq q_i^j$. For example, in the school choice problem each agent demands at most one object and $Q = \left\{q \in \times_{o\in\mathcal{O}} \{0, 1, ..., |o|\} \mid \sum_{o\in\mathcal{O}} q^o \leq 1\right\}$.[17] Each outcome $x \in \mathcal{X}$ consists of a profile of allocations $q_i \in Q$ for agents $i \in \mathcal{N}$ such that $\sum_{i\in\mathcal{N}} q_{i,o} \leq |o|$ for each $o \in \mathcal{O}$. Each agent has strict preferences over allocations from Q_i and each agent's preferences over outcomes are determined by this agent's allocation; that is each agent is indifferent among allocations of other agents.

Our main result in this setting is showing that the popular random priority extended-form mechanism—described in the Introduction—is characterized by obvious strategy-proofness together with Pareto efficiency and symmetry. A game Γ is symmetric if for all agents i and j game Γ is isomorphic to Γ with the roles of i and j exchanged. Exchanging the roles of i and j in game Γ creates game Γ' such that there is a bijection σ of histories between Γ and Γ' satisfying the following properties: terminal histories are mapped into terminal histories; if h is a subhistory of H in Γ then $\sigma(h)$ is a subhistory of $\sigma(H)$ in Γ'; and whenever i moved at history h in Γ then j moves at history $\sigma(h)$ in Γ', and vice versa; whenever an agent $k \neq i, j$ moved at h in Γ then

[16] The set of objects may include a special object called the outside option, but its presence play no role in the main-text results. Appendix C discusses how, by explicitly taking outside options into account, we can slightly extend our results beyond the class of preference domains studied in the main text.

[17] Our insights remain true beyond this environment, for instance, when each agent demands exactly one object, $Q = \left\{q \in \times_{o\in\mathcal{O}} \{0, 1, ..., |o|\} \mid \sum_{o\in\mathcal{O}} q^o = 1\right\}$; see the end of Appendix B for details.

the same agent moves at $\sigma(h)$ in Γ'; and the outcome of i after terminal history h in $\Gamma$is the same as the outcome of j after terminal history $\sigma(h)$ in Γ', and vice versa, while payoffs of other agents at terminal history h is the same as at $\sigma(h)$.[18]

A mechanism is symmetric if the underlying game is symmetric. Symmetry and efficiency are well-known to be properties of random priority, and Li (2016) shows that random priority is obviously strategyproof.

Theorem 1. *An extensive-form game is is obviously strategy-proof, symmetric, and efficient if and only if it is equivalent to random priority,*

This result explains why random priority is a very popular allocation mechanism: not only it is efficient, fair, and simple to play thanks to its obvious strategyproofness; it is the only mechanism with these properties.

We prove this theorem in the appendix. The proof relies on our characterization of obviously strategyproof and efficient mechanisms, to which we turn next. Knowing the structure of these mechanisms allows us to extend the bijective approach of Abdulkadiroğlu and Sönmez (1998) (cf. also Pathak and Sethuraman (2010) and Carroll (2010)) to our setting.[19]

4 Obvious Dominance and Efficiency in Environments without Transfers: Millipede games

In this section, we characterize the entire class of obviously strategyproof and efficient mechanisms as a class of games that we call *millipede games* (for reasons that will become clear in a moment).

[18]More generally, take a permutation of agents σ, and game Γ, and and define game $\sigma(\Gamma)$ recursively as follows: at every node at which nature moves, it has same moves in $\sigma(\Gamma)$ as in Γ; at every history at which i moves in Γ agent $\sigma(i)$ moves in $\sigma(\Gamma)$and has the same moves as i in Γ, at every terminal history the outcome of i in Γ is the same as the outcome of $\sigma(i)$ in the corresponding terminal history of $\sigma(\Gamma')$.

[19]We also discuss relaxing the symmetry assumption in the appendix.

To define millipede games, let us call an agent *active* at history h of a perfect information extensive-form game if (i) the agent moves at h, or (ii) the agent moved prior to history h, and for every terminal history following h, this agent moves between h and this terminal history. Then, a *millipede game* is an extensive-form perfect information game with the following properties. Nature moves once, at the empty history. At any non-empty history h in the game, there are at most two active agents, including the agent who moves at h, say agent i. All but at most one move at h are *clinching* moves; the remaining move (if there is one) is a *passing* move. Each clinching move is associated with a payoff of some outcome $x \in \mathcal{X}$; there may be several clinching moves associated with the same payoff. Further, millipede games satisfy the following, depending on whether i chooses a clinching move or a passing move.

(a) If i choses a clinching move, then this agent's payoff from the game is given by x and this agent never moves again. Furthermore, if there is an active agent other than i at h, say j, and if i's move a clinches an outcome inconsistent with an outcome j could have clinched at some subhistory of h , then at the following history (h, a), agent j moves and can clinch any payoff that is consistent with whatever prior agents have clinched; in such a case j has no passing moves at (h, a).

(b) If i chooses a passing move, then some other agent is called to move; if there were two active agents at j (i and j), then the next mover must be j. [20]

This procedure is continued until a unique final outcome is determined.

A simple example of the millipede game is a sequential dictatorship in

[20] Our results wouldn't change if we furthermore assumed that j can clinch any payoff she could have clinched at any subhistory of h and possibly more. The results wouldn't change also if we assumed that the sets of payoffs agents can clinch are disjoint except after the other active agent clinched a payoff, in which case all previously clinchable payoffs consistent with the clinching choice of i become clinchable by i. Notice also that except for this consistency condition after one of the active agents clinched a payoff, we impose no restrictions on the relation between the sets of payoffs agents i and j can clinch.

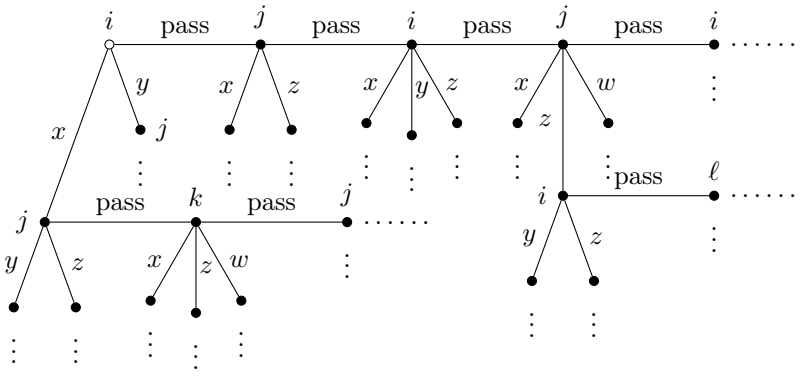


Figure 1: An example of a millipede game.

which no agent ever passes and there is only one active agent at every node. The first more complex example of a millipede we know of is due to Ashlagi and Gonczarowski (2016), who showed that for some preference domains deferred acceptance is obviously strategy-proof by implementing it as an instance of our more general millipede class of games. Yet, another example of a millipede game is given in Figure 4. Looking at the figure, the reason for the name "millipede game" becomes evident: the structure of the game resembles the well-known centipede game with more actions (i.e., "legs") at each node. Agents i and j can pass back and forth to each other. In a standard centipede game, there are only two agents; in contrast, in our millipede games when an agent clinches a payoff, and effectively "leaves" the game, the remaining active agent, say j, remains active, and we allow a third agent, say k, to become active. Agents j and k then play their own millipede game until one of them clinches a payoff and a fourth agent ℓ may become active, and so forth. This continues until there is a unique outcome that is consistent with all previously clinched payoffs.

Theorem 2. *A game Γ is obviously strategyproof and efficient if and only if it is equivalent to a millipede game.*

This theorem is applicable in many environments, including allocation en-

vironments in which agents care only about the object(s) they receive; in this case a payoff is simply the object(s) an agent receives. In these paper we focus attention on these allocation environments. Another environment Theorem 2 applies to is the standard social choice environment in which no agent is indifferent between any two outcomes. Theorem 2 has straightforward corollaries for deterministic mechanisms: in deterministic environments, we just skip nature's move, and continue with a standard millipede game.

The key to understanding Theorem 2 and its proof are the concepts of possible, guaranteeable, and clinched outcomes that we now introduce (where the general concept of clinched outcomes exactly corresponds to clinched outcomes in millipede games). In an extensive form game, every time an agent takes an action, she shrinks the set of outcomes (terminal histories) that can obtain. Determining whether an action is obviously dominant at a given node requires comparing the set of *possible* outcomes for each action with the set of outcomes that are not only possible, but are also guaranteeable. More precisely, we must compare the sets of possible vs. guaranteeable *indifference classes* that may obtain for an agent following each action.

Formally, for a game Γ, let

$$X(a, S_i) = \{x \in \mathcal{X} : z((h, a), S_i, S_{-i}, \omega) = x \text{ for some } (S_{-i}, \omega)\}$$

denote the set of outcomes that are possible if i takes action a at h and follows strategy S_i thereafter (ranging over all possible strategies of the other agents and nature). Consider an agent i of type $\succ_i$ and an indifference class $I_i(x)$. If there exists some S_i such that $y \in X(a, S_i)$ for some $y \sim_i x$, then we then we say that x is **possible for** a. If, further, there exists some S_i such that $y \sim_i x$ for all $x, y \in X(a, S_i)$, then we say x is **guaranteeable for** a. Note that we are slightly abusing notation, since when we say x is possible for a, what we really mean is some outcome in the indifference class $I_i(x)$ is possible for a. Similarly, when we say x is guaranteeable for a, we mean that there is some strategy such that agent i can ensure some outcome in the indifference class

$I_i(x)$ will obtain, no matter the actions of the other agents or nature. For brevity, we will sometimes say that the *payoff* x is possible/guaranteeable for i.[21]

Let $poss_i(a)$ be the set of outcomes that are possible for a, and $\gamma_i(a)$ be the set of outcomes that are guaranteeable for a, and define $P_i(h) = \cup_{a \in A(h)} poss_i(a)$ as the set of outcomes that are **possible at** h, and $G_i(h) = \cup_{a \in A(h)} \gamma_i(a)$ be the set of outcomes that are **guaranteeable at** h. Note that $G_i(h) \subseteq P_i(h)$, and the set $P_i(h) \setminus G_i(h)$ is the set of outcomes that are possible at h, but are not guaranteeable at h (this set will be key to many of our arguments).

Last, we define a distinction between two kinds of actions, that we call clinching actions and passing actions. We say an action $a \in A(h)$ **clinches** x if $y \sim_i x$ for all terminal histories following action a. If, at history h, there exists an action a that clinches x, we say that x is clinchable at h, and call a a clinching action (note that there can be more than one action that clinches x). If, following action a, there is some outcome x that is possible for a, but is not guaranteeable for a, then a is referred to as a **passing action**. Note the subtle difference between an object being "guaranteeable" for a and an object being "clinched" by a: x guaranteeable by a implies that there exists *some* strategy such that i can guarantee x going forward by choosing a at h; clinchable implies that for *every* strategy that chooses a at h, i receives x (or, more precisely, an outcome in the indifference class $I_i(x)$). In other words, following a clinching action, i's outcome is completely determined (modulo indifference classes), no matter what happens in the remainder of the game. Note that x clinchable at h implies x is also guaranteeable at h, but not vice-versa. In principle, an action can be neither a clinching action nor a passing action, which will happen if there are multiple possible outcomes, and each is also guaranteeable. However, we will show below that for any OSP game,

[21] In object allocation, i only cares about her allocation and is indifferent about what others receive. If at outcome x i receives object o, then for any other outcome $y \in \mathcal{X}$ such that i also receives o, we say that she receives payoff x.

we can find an equivalent game such that every action is either a clinching action or a passing action, and thus, it will be without loss of generality to assume that all actions are either clinching or passing actions; further, we will show that each history can have *at most* one passing action.

The proof of Theorem 2 builds on the above concepts. We provide the proof in the appendix, and explain its key steps below. Note that when we say every OSP game Γ satisfies some property, what we mean is that there is an OSP game Γ' that satisfies this property and is equivalent to Γ.

Step 1. *Every OSP game Γ is equivalent to a perfect information OSP game Γ' in which nature moves once, as the first mover.*

This follows because if we break any information set with imperfect information to several different information sets with perfect information, the set of outcomes that are possible shrinks. For an action a to be obviously dominant, the worst possible outcome from a must be (weakly) better than the best possible outcome from any other a'. If the set of possibilities shrinks (in the set inclusion sense), then the worst case from a only improves, and the best case from a' worsens; thus, if a was obviously dominant in Γ, it will remain so in Γ'.[22]

Step 2. *At every history, all actions except for possibly one are clinching actions.*

Step 2 is a key step, as it allows us to greatly simplify the class of OSP games to "clinch or pass" games. Recall that an action clinches x if for every possible history, agent i receives some y such that $y \sim_i x$. In addition to the clinching actions, there may be at most one action that does not clinch an outcome for i, which we call a passing action. The main insight here is that there cannot be more than one such action. Indeed, if there were two such actions a and a', then following each of a and a' there are at least two outcomes that are possible, but not guaranteeable. Thus, it will always be

[22]That every OSP game is equivalent to an OSP game with perfect information was first pointed out in a footnote by Ashlagi and Gonczarowski (2016). They also make the point that de-randomizing an OSP game leads to an OSP game.

possible to find a type of agent i for which one of the possibilities following a is at best his second choice, while one of the possibilities following a' is his first choice, which implies that a does not obviously dominate a'.

While step 2 shows that for almost all actions i's outcome is completely determined, what is perhaps more interesting is that there even *can* be one action such that i's outcome is uncertain without violating OSP. The reason for this follows from step 3.

Step 3. *If agent i passes at a history h, then she must move again between h and any terminal history, and, at this move, she will be able to clinch either: (i) everything she could have clinched previously or (ii) everything that was at possible but not clinchable for her at her first move.*

An agent may follow the passing action if she cannot clinch her favorite outcome today, and so she passes, hoping she will be able to move again in the future and get it then. Obvious strategy-proofness is retained by promising agent i that she can never be made worse off by passing: she will at least be able to clinch every payoff that she could have clinched in the past, and may possibly be able to clinch even more. Thus, there is no harm in passing, since we guarantee that she can always make herself at least as well off in the future, and possibly strictly better off.

Steps 1-3 use only obvious strategy-proofness, and not efficiency. Because of this, they only guarantee that if an agent passes, she will never be made worse off relative to what was possible for her. Efficiency implies at the start of the game, everything is possible for the first mover, and, if she passes, for the second mover as well. Efficiency is also key in to the final step of the proof.

Step 4. *There cannot be three passes in a row between distinct agents.*

The full argument of this step is somewhat subtle, and is postponed to the appendix. However, we do note here that efficiency is necessary for this to hold. Indeed, Ashlagi and Gonczarowski (2016) have an example of an OSP implementation of a deferred acceptance mechanism that allows for three ac-

tive agents. This example shows that an obviously strategyproof mechanism can still be complex, and require significant foresight by the agents. However, deferred acceptance is in general not efficient, and efficiency implies that there cannot be three passes in a row between three distinct agents.

Combining steps 1-4 implies that any OSP and efficient game Γ is equivalent to a millipede game. Steps 1-3, which rely only on OSP, give the basic structure as a "clinch-or-pass" game. Step 4 implies that two agents i and j may pass back and forth to each other, but no third agent k can become "active" until at least one of i or j has clinched some payoff. The full proof in the appendix starts by showing the theorem for a game of size $N = 3$, and then proceeds inductively on the number of agents.

5 Strong Obvious Strategy-Proofness

Our characterization of OSP and efficient mechanisms shows that they can be quite complex, and may presume a large amount of foresight (and backwards induction) by the agents. The reason is that while OSP allows agents to be hazy about other agents' choices, it presupposes that agents understand their own future choices and are confident they will get them right. How to formalize the idea that mechanisms are even simpler for the agents? Here is one idea.

A strategy $S_i(\theta_i)$ of type θ_i of agent i is *strongly obviously dominant* if for any other strategy $S_i'(\theta_i)$ at the first point (information set) of departure I between $S_i'(\theta_i)$ and $S_i(\theta_i)$ the best possible outcome from playing $S_i'(\theta_i)$, across all strategies and types of other agents, is at most as good as the worst outcome from playing $S_i(\theta_i)$ at information set I followed by by any play by other agents and any play by agent i.

Similarly to Li's OSP, the above stronger concept corresponds to a class of agents cognitive limitations. For simplicity, let us consider perfect-information games. We say that an extensive-form game Γ is outcome-set equivalent to

extensive-form game Γ' if there is a bijection ϕ between histories such that the set of possible outcomes after history h is the same as the set of possible outcomes after $\phi(h)$. We then obtain.[23]

Theorem 3. *A game Γ and strategy-profile are SOSP if and only if the corresponding strategy profile is weakly strategy-proof in any outcome-set equivalent game Γ'.*

While in the environments with transfers studied by Li, the canonical obviously strategy-proof mechanisms—ascending auctions—fails to satisfy this stronger concept, in our environment without transfers, the canonical obviously strategy-proof mechanisms—sequential dictatorships—satisfy SOSP.

The SOSP requirement basically characterizes sequential dictatorships.[24] We say that a mechanism is an *almost-sequential dictatorship* if it is a perfect-information game in which nature moves first; then the agents move in turn (the ordering being determined by nature's move, and the moves of earlier agents) with each agent moving at most once. At his move, an agent picks his objects and sends a message. As long as there are at least three objects unallocated, the moving agent can choose from all still available objects. When there are two objects remaining, an agent can be faced with either a choice between them, or, he might be given a choice between one of these object for sure or giving the next agent an opportunity to allocate the remaining two objects among the two of them.[25]

As a corollary from Theorem 2, we obtain the following

[23]While Li also shows that OSP mechanisms are precisely the mechanisms that can be implemented with bilateral commitment, this result does not extend to our setting. The reason is that bilateral commitment presumes that agents are perfectly forward looking and do not make errors in single-agent games, and SOSP relaxes this assumption.

[24]We provide an analysis below assuming both OSP and efficiency, but the characterization under SOSP alone is similar (see appendix).

[25]Pycia and Unver (2016) use the same name for deterministic mechanisms without messages that belong to the class we study; they show that this are exactly the deterministic mechanisms which are strategy-proof and Arrovian efficient with respect to a complete social welfare function. We use the name their introduced because our class is a natural extension of theirs.

Corollary 1. *A game and strategy-profile are SOSP and efficient if and only if the game is equivalent to an almost-sequential dictatorship.*

We are in the process of planning an experiment checking that SOSP extensive form games are indeed simpler to play that equivalent OSP but not SOSP games. We conjecture that the experiment will confirm the natural intuition that subjects will find it easier to play an almost-sequential dictatorship than to play a more complex millipede game.

6 Conclusion

Obvious strategyproofness is a highly desirable incentive property, as convincingly argued by Li (2016). The present paper follows up on Li's work by analyzing obvious strategyproofness in environments without transfers. The paper's main insight is that random priority is the unique obviously strategyproof mechanism that satisfies basic efficiency and fairness properties, thus providing an explanation why random priority is popular in practical allocation problems.

A Proof of Theorem 2

For clarity, the proof is organized according to the four steps described in the main text. Each step is formalized and proved using one or more lemmas. Furthermore, we couch the proof in terms of object allocation (without copies) in which each agent demands at most one object. The same proof applies to the general case if an object is understood as an element of the partition from our assumption (a) on the preference domains. In such a case, an agent clinching an object means that agents moving at continuation histories are choosing among outcomes from the element of the partition that is already clinched.

Step 1: Every OSP game Γ is equivalent to a perfect information OSP game Γ' in which nature moves once, as the first mover.

We break this result into two lemmas. The first shows that we can assume perfect information, and the second shows that we can assume nature moves only once.

Lemma 1. *(Ashlagi and Gonczarowski, 2016) Every OSP game is equivalent to an OSP game with perfect information.*

Proof. Ashlagi and Gonczarowski (2016) mention this result in a footnote; here, we provide the straightforward proof. Denote by $A(I)$ the set of actions available at information set I to the agent who moves at I. Take an obviously strategy-proof game Γ and consider its perfect-information counterpart Γ', that is the perfect information game at which at every history h in Γ the moving agent's information set is $\{h\}$ in Γ', the available actions are $A(I)$, and the outcomes in Γ' following any terminal history are the same as in Γ. Notice that the support of possible outcomes at any history h in Γ' is a subset of the support of possible outcomes at $I(h)$ in Γ. Hence, Γ' is obviously strategy-proof and equivalent to Γ'. □

Lemma 2. *Every OSP game is equivalent to a perfect information OSP game in which nature moves once, as the first mover.*

Proof. In light of the transitivity of the equivalence relationship, Lemma 1 tells us that we can assume that Γ has perfect information. We can further assume that nature moves at the empty history $\emptyset$ (possibly just having one move at this history). Let $\mathcal{H}_{\text{nature}}$ be the set of histories h at which nature moves. Consider a modified game Γ' in which at the empty history nature chooses actions from $\times_{h\in\mathcal{H}_{\text{nature}}} A(h)$ and at $h \in \mathcal{H}_{\text{nature}} - \{\emptyset\}$ nature chooses from a singleton set of actions that contains the choice from $A(h)$ that nature made at $\emptyset$. Then, the support of possible outcomes at any history h in Γ' is a subset of the support of possible outcomes at the corresponding history

in Γ, where the corresponding histories are defined by mapping the $A(h)$ component of the action taken at $\emptyset$ by nature in Γ' as an action made by nature at h in game Γ. Hence, Γ' is obviously strategy-proof, and Γ and Γ' are equivalent. □

Step 2: At every history, all actions except for possibly one are clinching actions.

The first lemma of this step points out some facts of OSP games that will be useful in the proof of the second: in particular, for a given agent i, the set of possible, but not guaranteeable outcomes monotonically shrinks, and the set of clinchable outcomes monotonically grows. We then use this to construct, from any OSP game Γ, a new game (and strategies) that is also OSP, and such that at every history, every action except possibly one is a clinching action.

Lemma 3. *Let Γ be an obviously strategyproof game. Consider a history h^i and let $P_i(h^i)$ be the set of objects that are possible for i at h^i and $C(h^i)$ be the objects that are clinchable by i at h^i.*

(a) Assume that h^i is on the path of play for type $\succ_i$. A clinching action a_x is obviously dominant at h^i for type $\succ_i$ if and only if x is the $\succ_i -$most preferred object in $P_i(h^i)$.

(b) At any later history $\tilde{h}^i$ at which i is to move, $P_i(\tilde{h}^i) \setminus G_i(\tilde{h}^i) \subseteq P_i(h^i) \setminus G_i(h^i)$.

(c) At any later history $\tilde{h}^i$ at which i is to move, either (i) $C(h^i) \subseteq C(\tilde{h}^i)$ or (ii) $(P_i(h^i) \setminus G_i(h^i)) \subseteq C(\tilde{h}^i)$ (or both).

Proof. Part (a): The if direction is immediate. For the only if direction, since the clinching action a_x clinches x, the worst case outcome from following a_x (as well as the best case) is x. If there is some other $y \in P_i(h^i)$ such that $y \succ_i x$, then there must be some other action $a' \neq a_x$ for which y is a possible outcome. The best case from a' must be weakly better than y, and so, a_x does not obviously dominate a'.

Part (b): We show the contrapositive: $x \notin P_i(h^i) \setminus G_i(h^i)$ implies $x \notin P_i(\tilde{h}^i) \setminus G_i(\tilde{h}^i)$. $x \notin P_i(h^i) \setminus G_i(h^i)$ implies that either (i) x is not possible at h^i or (ii) x is guaranteeable at h^i. It is clear by definition that if x is not possible at an earlier history h^i, then it also cannot be possible at a later history $\tilde{h}^i$. On the other hand, if x is guaranteeable at h^i, this means that there exists some strategy S_i such that if i follows this strategy from h^i forward, she is certain to receive x, no matter what the other agents do. Thus, if she continues to follow this strategy at the later history $\tilde{h}^i$, she will still be certain to receive x, which implies that x is also guaranteeable at $\tilde{h}^i$.[26] Combining these statements, we have $x \notin P_i(\tilde{h}^i) \setminus G_i(\tilde{h}_i)$.

Part (c): Assume that both (i) and (ii) are false, and let x be an object that is clinchable at h^i, but is not at $\tilde{h}^i$ and let y be an object that is possible, but not guaranteeable, at h^i, with y not clinchable at $\tilde{h}^i$.[27] Consider a type of agent i, $\succ_i$, that prefers y to x.[28] At h^i, this agent has no obviously dominant action. Since y is possible, but not guaranteeable at h^i, none of the clinching actions are obviously dominant. The passing action is also not obviously dominant. To see why, note that the worst possible outcome from passing at h^i is strictly worse than x (when history $\tilde{h}^i$ is reached, neither x nor y are guaranteeable, by assumption), while i can clinch x at h^i, which is a contradiction. □

Lemma 4. *For each obviously strategyproof game Γ, there exists an equivalent obviously strategyproof game Γ' with perfect information such that at each history h, the following hold:*

[26] It may be that the strategy S_i is such that $\tilde{h}^i$ is off-path for agent i. However, if this is the case, upon reaching $\tilde{h}^i$, no type of agent i will follow a strategy for which x is a possible outcome, and so the subgames where i receives x can be pruned away, and x is not possible at $\tilde{h}^i$.

[27] Such distinct x and y exist by construction, since $C(h^i) \cap (P_i(h^i) \setminus G_i(h^i)) = \emptyset$.

[28] History h^i must be on the path of play for some such type because, if it were not (if all types to reach this history preferred x to y), then no type of agent i would follow any strategy S_i for which y is a possible outcome, and thus the subgames where i receives y can be pruned away.

(a) At least $|A(h)|-1$ *actions at* h *are clinching actions (i.e.,* $|poss_i(a)| = 1$ *for all but at most one* $a \in A(h)$*).*

(b) For every outcome $x \in G_i(h)$*, there exists an action* a_x *that clinches* x *and, if* i *follows action* a_x*, she never moves again in the game.*

Proof. The proof of this lemma relies on the following claim:

Claim 1. *Let* Γ *be an obviously strategyproof game. Then, for any history* h*, there is at most one action* a^* *such that* $poss_i(a^*) \cap (P_i(h) \setminus G_i(h)) \neq \emptyset$*.*

The claim shows that there can be at most one action which can lead to a possible, but not guaranteeable, outcome. Thus, for every other action that can be taken, all possible outcomes are also guaranteeable. However, it may be that some action a can lead to multiple outcomes, where different outcomes are guaranteeable for i by following different actions in the future of the game. So, assume that this is the case. We will construct a new obviously strategyproof game Γ' that is equivalent to Γ and for which every action except for at most one is a clinching action.

Consider a history h, and let

$$\mathcal{S}_i(h) = \{S_i : S_i \text{ guarantees some outcome } x \text{ starting from } h\}.$$

We create a new game Γ' that is the same as Γ, except replace subgame starting from history h as follows with a new subgame defined as follows. If there is a passing action at h in the original game, then there is a passing action at h in the new game, and the subgame following the passing action is exactly the same as in the original game. Additionally, there are $M = |\mathcal{S}_i(h)|$ other actions at h, $a_1, \ldots, a_M$. Each a_m corresponds to one strategy $S_i^m \in \mathcal{S}_i(h)$, and following each a_m, we replicate the original game following the corresponding action $a = S_i^m(h)$ in the original game, except that at any future history at which i is called on to act, the node is deleted and the game proceeds as if i had chosen her action according to S_i^m. In other words, at h in the new game Γ', we are asking agent i to choose not only her current

action, but all future actions as well. By doing so, we have created a new game in which every action (except for the passing action, if it exists) at h clinches some outcome x, and further, agent i is never called upon to move again.[29] It is clear that if there is an obviously dominant profile of strategies for the original game, and the agents follow the analogous strategies in the new game,[30] the same terminal history will be reached, and so Γ and Γ' are outcome-equivalent.

We must also show that this strategy profile is obviously dominant for Γ'. This is an obviously dominant strategy for i because if her obviously dominant strategy in the original game was to guarantee some outcome x, she now is able to clinch x immediately; if her obviously dominant strategy was to pass at h, she is still able to pass, and the passing action obviously dominates any of the clinching actions (because passing in the original game obviously dominated a strategy that guaranteed some outcome). In addition, the game is also obviously strategyproof for all $j \neq i$ because, prior to h, the set of possible outcomes for j is unchanged, while for any history succeeding h where j is to move, having i make all of her choices earlier in the game only shrinks the set of possible outcomes for j, in the set inclusion sense. When the set of possible outcomes shrinks, the supremum only decreases and the infimum only increases, and so, if a strategy was obviously dominant in the original game, it will continue to be so in the new game. Repeating this process for every history h, we are left with a new game where, at each history, there are only clinching actions plus (possibly) one passing action, and, following any clinching action, an agent never acts again.

Proof of claim 1

First consider part (a) of the claim, and consider agent i. We prove

[29] Note that this only applies to the non-passing actions at h, which correspond to the guaranteeing strategies in $\mathcal{S}_i(h)$. It is still possible for i to follow the passing action at h and be called upon to move again later in the game.

[30] That is, for all $j \neq i$, they continue to follow the same action at every history as they did in the original game, and for i, at history h in the new game, she takes the action a_m that is associated with the strategy S_i^m in the original game.

this by induction on histories at which i is to move. Consider any earliest history h at which i is to move, and choose some $x \in P_i(h) \setminus G_i(h)$. If $x \in poss_i(a_1)$ for some $a_1 \in A(h)$, then for any other y that is possible but not guaranteeable at h, we must have $y \in poss_i(a_1)$ as well. To see why, assume not: $x \in poss_i(a_1)$, but $y \notin poss_i(a_1)$. There must be some a_2 such that $y \in poss_i(a_2)$. Consider the type $x \succ_i y \succ_i \cdots$.[31] Note that neither a_1 nor a_2 obviously dominate the other, because the worst case from both is strictly worse than y (because neither x nor y are guaranteeable at h), while the best case from a_1 is x and the best case from a_2 is weakly better than y. Consider any other $a' \in A(h)$ such that $a' \neq a_1, a_2$. Again, since neither x nor y are guaranteeable, the worst case from any such a' is weakly worse than x, and so a' does not obviously dominate a_1. Thus, this type has no obviously dominant action, which contradicts that the game is OSP.

To finish, assume that there were two actions a_1^* and a_2^* that could lead to possible but not guaranteeable outcomes. Consider some $x \in P_i(h) \setminus G_i(h)$, and a type of agent i that ranks x first. By the previous paragraph, we know that $x \in poss_i(a_1^*)$ and $x \in poss(a_2^*)$. However, by assumption, x is not guaranteeable at h, so, for any type who ranks x first, the worst case outcome from *any* action a' must be strictly worse than x, while the best case outcomes from a_1^* and a_2^* are both x. Therefore, no action $a' \in A(h)$ is obviously dominant, which contradicts that Γ is OSP.[32]

Thus, at the first history at which i is to move, there can be at most one passing action. Further, any obviously dominant strategy of any type $\succ_i$ that ranks any $x \in P_i(h) \setminus G_i(h)$ first must select the passing action at this history (Lemma 3, part a).

Now, consider any other later history h' for i, and assume the inductive

[31] Note that, since this is the first time agent i is to move, this history is on the path of play for all types.

[32] Note that if there were only one such action a_1, then it would still be true that there is no action that obviously dominates a_1 for this type. However, a_1 itself might be obviously dominant. When there are two such actions, a_1 does not obviously dominate a_2, nor does a_2 obviously dominate a_1, and thus there are no obviously dominant actions.

hypothesis that for all predecessor histories to h' for i there is at most one passing action. Note that $P_i(h') \setminus G_i(h') \subseteq P_i(h) \setminus G_i(h)$, i.e., the set of non-guaranteeable possibilities only shrinks (Lemma 3). Consider some $x' \in P_i(h') \setminus G_i(h')$. By the inductive hypothesis, history h' must be on the path of play for every type of agent i that ranks x' first.[33] Repeating the above argument for history h' and outcome x', we conclude that at history h', there is at most one passing action, and the strategy of any type $\succ_i$ that ranks any $x' \in P_i(h') \setminus G_i(h')$ first must be to follow the passing action at h'. □

Step 3: If agent i passes at a history h, then she must move again between h and any terminal history, and, at this move, she will be able to clinch either: (i) everything she could have clinched previously or (ii) everything that was at possible but not clinchable for her at her first move.

This step provides the key insight of why we can have a passing action while still retaining obvious strategyproofness, which is possible as long as whenever an agent passes at a history h, she knows that she will do no worse in the future than *anything* she could have clinched today.

Lemma 5. *Let Γ be an OSP game. Consider an agent i and a history $\bar{h}^i$ that is the first move of agent i (i.e., i moves at no earlier history), and for $j \neq i$, let h^j be a subsequent history such that (i) agent i has not clinched an outcome between $\bar{h}^i$ and h^j and (ii) at h^j, agent j has an action a^j_z that clinches some $z \in P_i(\bar{h}^i)$. Let $\mathcal{H}_i$ be the set of histories between $\bar{h}^i$ and h^j at which i is to move. Then, there is an equivalent OSP game such that i is the mover at (h^j, a^j_z) and:*

(a) If $z \in C(h)$ for some $h \in \mathcal{H}_i$, then i can clinch any x that was possible, but not guaranteeable at h, and at any successor history to h (and possibly

[33] All previous histories for i have had only one passing action, and x' was not guaranteeable at those histories, so any obviously dominant strategy for this type of agent i must select the passing action at any history prior to h'.

more).

(b) If $z \in P_i(h) \setminus G_i(h)$ for all $h \in \mathcal{H}_i$ then i can clinch any x that was clinchable at any $h \in \mathcal{H}_i$ (and possibly more).

Proof. Part (a): We first argue that in any OSP game, agent i must move again after h^j and before any terminal history. Consider the the immediate predecessor of h^j at which i is to move, denoted h, and let $x, y \in P_i(h) \setminus G_i(h)$,[34] and assume that there exists a terminal history h^* that can be reached from h^j such that i does not move again. Consider the types $x \succ_i z \succ_i \cdots$ and $y \succ_i z \succ_i \cdots$.[35] At any such h^*, i cannot receive z (because j has been assigned to it). If i is assigned to x at h^*, the type $y \succ_i z \succ_i \cdots$ has no obviously dominant action at h; similarly, if she is assigned to y at h^*, the type $x \succ_i z \succ_i \cdots$ has no obviously dominant action at h. If she is assigned to something other than than x or y, then neither of these types has an obviously dominant action at h. Thus, i must move again at some later history $\tilde{h}$.

We next show that if $z \in C(h)$ for some $h \in \mathcal{H}_i$, i must be able to clinch any object in $P_i(h) \setminus G_i(h)$. To see this, note that since j has clinched z, we have $C(h) \not\subseteq C(\tilde{h}^i)$. Thus, by Lemma 3, we conclude $(P_i(h) \setminus G_i(h)) \subseteq C(\tilde{h}^i)$. Further, since $P_i(h') \setminus G_i(h') \subseteq P_i(h) \setminus G_i(h)$ for all histories h' that succede h (by Lemma 3), the same holds for all successor histories h'. Note also that it is without loss of generality to assume that i moves at most once following (h^j, a_z^j), and at any history where she is to move, she has no passing action.[36] Last, we argue that any such game where this is true is equivalent to one

[34]If there is only one object x that is possible, but not guaranteeable, at h, then at $\tilde{h}^i$, it is possible that i's move is a "trivial" move for which receives x at any subsequent terminal history. However, the statement still holds.

[35]Note that by Lemma 3, $P_i(h) \setminus G_i(h) \subseteq P_i(h') \setminus G_i(h')$ for all $h' \in \mathcal{H}_i$. This implies that every obviously dominant strategy for these types must select the passing action at any $h' \in \mathcal{H}_i$, which means that history h must be on the path of play for these types of agents.

[36]This is because i is able to clinch everything in $P_i(h)$ (other than z) at either h or $\tilde{h}$. Thus, if there were a passing action at $\tilde{h}$, no type of agent i will ever follow it, and so it can be pruned.

in which i moves immediately after j clinches z. Indeed, create a new game where i moves immediately after (h^j, a_z^j), and is given all of the actions she could have taken at any earliest successor history to (h^j, a_z^j) at which she was to move in the original game. Following each of these actions, the new game continues as in the original game following (h^j, a_z^j), except at the node at which i was to act in the original game is deleted, and the game continues according to the clinching action i chose at history (h^j, a_z^j). This continues to be obviously strategyproof for i, because all of her actions are clinching actions. For any predecessor history to (h^j, a_z^j), the set of possible outcomes is the same in both games, while for any successor history following any action of i at (h_z^j, a_z^j), the set of possible outcomes of the new game for the agent j who is to move is a subset of the set of possible outcomes of the old game. Thus, the best case outcome from any strategy can only get worse, while the worst case outcome can only get better, so if a strategy was obviously dominant before, it continues to be obviously dominant in the new game.

Part (b): Agent i must move again following (h^j, a_z^j), using a similar argument as in part a.[37] Assume that there was some $h \in \mathcal{H}_i$ and $x \in C(h)$ and some path such that i could not clinch x subsequent to (h^j, a_z^j). Consider the type $z \succ_i x \succ_i \cdots$.[38] Since it is possible that i might not be able to clinch x, the worst case outcome from passing at h is strictly worse than x (since it is possible that j may clinch z, and afterwards i cannot clinch x). However, she can clinch x at h, and so passing is not an obviously dominant action. No clinching action is obviously dominant either, since z is possible after passing. Thus, this type has no obviously dominant action. This implies that i must move again at some point after (h_j, a_z^j). Constructing an equivalent OSP game in which i moves immediately at (h_j, a_z^j) and never moves again can be

[37] If there is only one object that was previously clinchable, say x, then it is possible that i's move is a trivial move in which every possible action clinches x. However, the statement still holds.

[38] All $h \in \mathcal{H}_i$ must be on the path of play for this type, since z is possible but not guaranteeable, for all $h \in \mathcal{H}_i$.

done using the same procedure from part (a). □

Step 4: There cannot be three passes in a row between distinct agents.

The first three steps showed (without invoking efficiency) that an OSP game must have a 'clinch-or-pass' structure, but still allow for many 'active' agents. Adding efficiency delivers a similar class of games, but with at most two active agents at any one time.

Lemma 6. *Let Γ be a game that is OSP and efficient. Consider a history h such that all previous actions (from the start of the game) are passing actions between at most two distinct agents, i and j, and assume (wlog) that i is to act at h. Then, for any object $x \in X$, either: (i) x was clinchable by i at some history prior to h (including h itself) or (ii) x is possible, but not guaranteeable, at h.*

Proof. Without loss of generality, assume the game starts with i passing to j. Then, it is obvious that the statement holds both at the initial history where i acts, and at the history after i passes to j for the first time. Thus, consider the next move, where j passes back to i. We will show that any x that was not clinchable at i's first move or this history must still be possible for i at this history. Indeed, assume there was some x that was not previously clinchable, nor is it possible. Let C_j be the set of objects that were clinchable for j at the previous history. Since j passed, there must be some outcome y that is possible, but not guaranteeable for her. Assume that there exists such a y that is distinct from x. Consider a preference profile where i ranks x first, j ranks y first, and all other agents $k \neq i, j$ rank some z_k first, where x, y, and z_k area all distinct. Efficiency implies that at this profile, i must be assigned to x; however, if i could not clinch x prior to this history, and x is no longer possible at this history, then we have a contradiction. Last, assume that there is no y that is distinct from x; that is, for all $y \neq x$, y was

clinchable by j at some prior history. We know that at the current history, i must be able to clinch some $z \neq x$, which implies that j must have been able to clinch z at some prior history. Consider the types $x \succ_i z \succ_i \cdots$ and $x \succ_j z \succ_j \cdots$. If x is not possible for i, and not guaranteeable for j (because j passed to i), then there must be a history $\bar{h}$ such that i and j are assigned to objects other than x. However, at most one of i or j can be assigned to z, and so one of them must be assigned to something strictly worse than z. However, this agent will then not have an obviously dominant strategy. □

Lemma 7. *If Γ is obviously strategyproof and efficient, then there cannot be three passes in a row between distinct agents (i.e, i passes to j passes to k).*

Proof. We first show this for a game of size $N = 3$, and then proceed inductively. There are several cases and subcasaes to consider, depending on what has been previously clinchable for i and j, and we will show for each case, if j could pass to a third agent $k \neq i$ we reach a contradiction.

We first must define some notation. Consider some history h^j at which the mover is j and for which all previous moves have been passing moves, and denote the immediate predecessor (at which the mover is agent i by assumption) as h^i. Recall that $\mathcal{H}_j$ is the set of histories at which j is to move, and let

$$C_j(h^j) = \cup_{\{h' \in \mathcal{H}_j : h' \preceq h^j\}} \{x \in X : x \text{ is clinchable at } h'\}$$

be the set of objects that are clinchable by j at any point between his initial move and history h^j. $C_i(h^i)$ is defined analogously. Note that for any h^j thus defined, $C_j(h^j) \cup (P_j(h^j) \setminus G_j(h^j)) = X$, by Lemma 6 (the same holds for i at her histories).

We now work through the various cases depending on the relationship between the sets $C_j(h^j)$ and $C_i(h^i)$.[39]

[39]Note also that if j (or i) passes to k, k must have at least two actions, one of which must be a clinching action (since at most one action can be a passing action).

1. $C_i(h^i) = C_j(h^j) = C$: Since both sets are nonempty, there must be at least one x that is clinchable by both i and j. If, after j passes, agent k is able to clinch some $x \in C$, then by Lemma 5, following this move, both i and j must be able to clinch any $y \notin C$; however, they cannot both end up with the same object, which is a contradiction. If agent k can clinch some $y \notin C$, then Lemma 5 implies that both i and j must be able to clinch any $x \in C$, which, again, is a contradiction.[40]

2. $C_i(h^i) \subset C_j(h^j)$ or $C_j(h^j) \subset C_i(h^i)$: We will argue for the case where $C_i(h^i) \subset C_j(h^j)$, but $C_i(h^i) \neq C_j(h^j)$; the other case is symmetric. Since $C_i(h^i) \subset C_j(h^j)$ but $C_i(h^i) \neq C_j(h^j)$, there must be some x that was previously clinchable by both i and j, and some $y \neq x$ that was previously clinchable only by j. Further, we know that there is some $z \neq x, y$ that has not been clinchable by j (or i), but is possible.[41] If k can clinch some $x \in C_i(h^i) \cap C_j(h^j)$ at her move, then Lemma 5 implies both i and j must be able to clinch $z \notin C_j(h^j)$. Similarly, if k can clinch any $z \notin C_j(h^j)$, Lemma 5 implies that both i and j must be able to clinch x. Thus, consider last the case where k can clinch some $y \in C_j(h^j) \setminus C_i(h^i)$. We argue that at every history after j passes, everything is guaranteeable for j (and thus also clinchable for j by Lemma 4), and thus this move is actually not a passing move. First, note that we have shown that k cannot clinch any $z \notin C_j(h^j) \setminus C_i(h^i)$ at this move, so she must be able to clinch some $y \in C_j(h^j) \setminus C_i(h^i)$, and must also have a passing action.[42] Fix any $z \notin C_j(h^j)$, and consider any earliest history at which some agent is able to clinch z. The only possibility is that this is agent j. To see why, assume it was i. Then, by

[40] Note that by Lemma 6, if the the set of objects that were previously clinchable for i and j are the same, then so are the sets of objects that are possible, but not guaranteeable.

[41] If not, then j has previously had the opportunity to clinch everything, and any passing action can be pruned. Since $C_i(h^i) \subset C_j(h^j)$, z was not previously clinchable by i either.

[42] If she could not pass, then k would never be able to receive x (e.g., when she ranks x first and everyone else ranks x last), and the mechanism would not be efficient.

Lemma 5, both j and k must be able to clinch y. On the other hand, if it was k who could clinch z, then both i and j again must be able to clinch x by Lemma 5. Thus, at h^j, agent j has had the opportunity to clinch all objects, and if any passing action were to exist, it can be eliminated by pruning.

3. $C_i(h^i) \cap C_j(h^j) = \emptyset$: In this case, everything that was clinchable by i has not been clinchable by j, and vice-versa. Fix some $x \in C_i(h^i)$ and $y \in C_j(h^j)$. Assume that k can clinch some object $z \in C_i(h^i)$. Then, by Lemma 5, both i and j must be able to clinch y; similarly, if $z \in C_j(h^j)$, Lemma 5 implies that both i and j must be able to clinch x. Thus, last, consider some z that is clinchable by neither i nor j. Similar to footnote 42, k must have a passing action at her move. Following the passing action, consider any earliest history at which some agent other than i is able to clinch x.[43] We argue that no matter which agent it is, the game is not OSP. First, if j could clinch x, then Lemma 5 implies that both i and k must be able to clinch z. If k could clinch x, then Lemma 5 implies that both i and j can clinch y.

4. *Neither of the previous three cases hold*: This last case consists of $C_i(h^i)$ and $C_j(h^j)$ intersecting, but neither being a strict subset of the other. We argue that in this case, j can actually clinch all objects at her move. Assume she could not. This implies that there exists some x that is clinchable by both i and j, as well as some y that is clinchable only by i and some z that is clinchable only by j (i and j may also be able to clinch other things as well). Consider the type $\succ_j$: $y \succ_j z \succ_j \cdots$ for agent j and type $\succ_i$: $z \succ_i y \succ_i \cdots$ for i. If i cannot clinch z, she must at least have a strategy that guarantees her either z or y; if j cannot clinch y, she must have a strategy that guarantees her either y or z.[44]

[43] Such an agent must exist by efficiency.

[44] By Lemma 6, y must be possible for j at h^j and z must be possible for i at h^i.

Since we assumed that j cannot clinch y, there must be a subsequent terminal history $\bar{h}$ where the type $\succ_j$ does not receive y, which implies that at $\bar{h}$, j must receive z. Further, since type $\succ_i$ of agent i must be able to guarantee either y or z, if j receives z, then there must be a history at which this type of agent j receives z and this type of agent i must receive y. However, this outcome is not Pareto efficient.

Thus, we have shown that, for a game of size $N = 3$, there cannot be three passes in a row between distinct agents. Assume that this is true for any game of size $1, \ldots, N-1$. Consider a game of size N. The above argument shows that this game cannot begin with a sequence of passes among three distinct agents; that is, before another agent moves, either i or j must clinch some object. Consider a history h such that the most recent action taken was a clinching action, but all previous actions were passing actions (i.e., the first time some agent clinches an object), and, without loss of generality, say it is i who clinches x. If x was clinchable by j at some history h' prior to h, then by Lemma 5, the next mover after h is j, and he can clinch anything that was possible, but not guaranteeable, at h' or later. Further, any object $y \notin P_j(h') \setminus G_j(h')$ must be clinchable by j at some history prior to h', by Lemma 6. Thus, after i clinches x, agent j has had the opportunity to clinch every object. So, following this history, both i and j have each clinched something, and, at the next node, we have a game of size $N-2$, which, by the inductive hypothesis, has no three passes in a row to distinct agents. On the other hand, if x was not clinchable by j at any history prior to h, after i clinches x, Lemma 5 then implies that the next mover is j, and she can clinch anything that she could clinch previously. At this node, we begin a game of size $N-1$ and again apply the inductive hypothesis.[45] □

[45]Note that at this node, everything in $X \setminus \{x\}$ is possible for j. Further, this node is "on path" for every type profile that ranks x first (because x was not clinchable by j previously), and so, when restricting to the set $X \setminus \{x\}$, every preference type for agent j is possible at this node, and so we have a complete game of size $N-1$.

Step 5: Completing the proof.

We restate Theorem 2 below, and tie all of the previous lemmas together to finish the proof.

Theorem 2. *A game Γ is OSP and efficient if and only if it is equivalent to a millipede game.*

Proof. The "if" part is straightforward. We prove the "only if" part via induction. First, consider a game of size $N = 3$. By Lemma 7, there cannot be three passes in a row between distinct agents. Thus, if at the initial node, i has a passing action to j, then j can only pass back to i, who can only pass back to j, etc. By efficiency, this chain of passes must end at some point with either i or j having been able to clinch anything they could not have previously clinched, and, at this point, the agent has no passing action. Following any clinching action by i or j, Lemma 5 implies that the other agent moves next, and can either clinch any x that was clinchable for her before (and possibly more), or can clinch any x that was possible but not guaranteeable before. Starting from this node, we thus have a game of size $N - 1$, which, via the inductive hypothesis, is a millipede game. □

B Proof of Theorem 1

Let us initially assume that each object has one copy, and each agent has single-unit demand. We relax these assumptions below.

Step 1. Let ϕ by an OSP, PE, and symmetric mechanism. By Theorem 1, there is an equivalent millipede mechanism. We may assume that this millipede mechanism is also symmetric.

Step 2. Notice that the symmetric millipede mechanism is a randomization over component mechanisms such that each one is a symmetrization (in the sense of Liu and Pycia 2016) over a deterministic millipede game. It is enough to show that each such component mechanism is equivalent to RSD.

Step 3. Notice that we can equivalently describe the component mechanism so that the nature at its initial move only selects the moving agent and his initial choice set, and then the nature moves after every move by an agent, and selects the next mover and the next choice set.[46] By symmetry, we can assume that the nature chooses the next mover uniformly at random from among the eligible agents, and that the agent's choice set does not depend on the agent drawn. Indeed, the original component mechanism is a randomization over such simpler component mechanisms, and we can just focus on one of the simpler component mechanisms. Finally, let us represent this simple component mechanism in terms of the nature initially uniformly randomizing over deterministic millipede games, in which the choice sets at corresponding nodes do not depend on the identity of agents at corresponding paths to these nodes.

Step 4. Let us now fix a preference profile, and construct a bijection from the deterministic millipede games of Step 3, to serial dictatorships. The first agent in the serial-dictatorship ordering is the agent who first received in his choice set the first object to be clinched. (To be sure: this agent is not necessarily the agent who was the first to clinch an object). The second agent is this among the remaining agents who first received in his choice set the second object to be clinched, etc. It is a well-defined mapping, and running the resulting serial dictatorship for the fixed preference profile results in the same outcome as running the millipede game, because in the millipede game the agent whose object is clinched obtains his most preferred object among objects that were not clinched earlier. (To be sure: the two mechanisms might differ when run on other preference profiles). Furthermore, the mapping is one-to-one. Because the simple component game uniformly randomizes over $|I|!$ deterministic millipede games, and there are also $|I|!$ serial dictatorship, this one-to-one mapping is onto, and hence a bijection.

[46]In fact, any millipede mechanism can be equivalently described so that the nature at its initial move only selects the moving agent and his initial choice set, and then the nature moves after every move by an agent, and selects the next mover and the next choice set

Step 5. Given Step 4, for any fixed preference profile the outcome of the simple component millipede game is the same as the outcome of uniformly randomizing over serial dictatorships, that is the same as the outcome of random priority. QED

Let us now extend the proof to multiple copies and multi-unit demands. Above, we proved Theorem 1 under additional simplifying assumptions that each object has only one copy and each agent demands at most one object. The proof in the general case follows the same steps, with the following differences. To accommodate multi-object (and multi-unit) demands, it is sufficient to change the terminology and substitute agents' allocations $q \in Q$ for objects. To additionally accommodate multiple-copies supply, we need to take additional care. With multiple copies it is possible that after the first several agents moved, irrespective of their choices, the next agent to move still is able to choose any allocation in Q. In such a case, the probabilities the relevant agents are selected for first, or second (etc.) move need not be the same; however, we obtain an equivalent mechanism by making the probabilities equal for all agents who did not move yet. On the other hand, if after some agents moved, the next agent to move might be unable to choose at least one allocation in Q, then the same argument as in Step 3 of the baseline proof suffices to say that at such a history, all agents who are yet to move have equal chances of having the move. QED

Remark. As mentioned in footnote 17, the insight of Theorem 1 remains true in some interesting environments beyond its assumptions. Suppose, for instance, that $Q = \left\{q \in \times_{o \in \mathcal{O}} \{0, 1, ..., |o|\} \,|\, \sum_{o \in \mathcal{O}} q^o = 1\right\}$. All the steps of the proof go through unchanged, except that in Step 2 we proceed as follows. Efficiency implies that each unassigned object is in the set of objects available to the agent who moves first following history h. Indeed, by way of contradiction, suppose that with positive probability i moves first after h and object o isn't in the set of objects available. By efficiency, all copies of object o need to be allocated to other agents; let j be one of these agents.

Let o' be an object in agent i's choice set. Consider the preference profile in which agent i most prefers object o and o' is his second-most preferred object, agents who moved already most prefer the objects they obtained, and other agents most prefer object is o' and o is their second-most preferred object. Then i chooses o', and j obtains o. But, then, the resulting allocation is not efficient; a contradiction that proves the claim of this step, and hence the analogue of Theorem 1 for the environment of footnote 17.

C Extensions: Outside Options

Consider the allocation model of Section 3, and suppose that each agent has an outside option.

C.1 Individual Rationality

We say that a game is individually rational if each agent can obtain at least his outside option. The analogues of our results hold true for individually rational games as soon as the domain of each agent's preferences satisfy the domain condition from Section 2 restricted to sets $X \subseteq \mathcal{X}$ that do not contain the outside option of this agent. Our proofs remain valid in this setting.

C.2 Restricted Domains

Our results hold true also in allocation domains in which any agent prefers any object to the outside option. Our proofs remain valid also in this setting.[47]

[47]Note that in this setting every game is individually rational, hence this observation is contained in the previous one.

References

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.

Arrow, K. J. (1963): *Social Choice and Individual Values*, New York: Wiley, 2nd edition ed.

Ashlagi, I. and Y. A. Gonczarowski (2016): "No stable matching mechanism is obviously strategy-proof." Working Paper.

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

Carroll, G. (2010): "A General Equivalence Theorem for Allocation of Indivisible Objects," MIT, Working paper.

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

Dur, U. M. and M. U. Ünver (2015): "Two-Sided Matching via Balanced Exchange: Tuition and Worker Exchanges," NC State University and Boston College, Working Paper.

Ehlers, L. (2002): "Coalitional Strategy-Proof House Allocation," *Journal of Economic Theory*, 105, 298–317.

Gibbard, A. (1973): "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41, 587–601.

——— (1977): "Manipulation of Schemes That Mix Voting with Chance," *Econometrica*, 45, 665–681.

Hakimov, R. and O. Kesten (2014): "The Equitable Top Trading Cycles Mechanism for School Choice," WZB and Carnegie Mellon University, Working Paper.

Hatfield, J. W. (2009): "Strategy-Proof, Efficient, and Nonbossy Quota Allocations," *Social Choice and Welfare*, 33 No. 3, 505–515.

Li, S. (2016): "Obviously Strategy-Proof Mechanisms," Working Paper.

Liu, Q. and M. Pycia (2013): "Ordinal Efficiency, Fairness, and Incentives in Large Markets," Working Paper.

Loertscher, S. and L. M. Marx (2015): "Prior-Free Bayesian Optimal Double-Clock Auctions," Working Paper.

Morrill, T. (2014): "Making Just School Assignments," *Games and Economic Behavior*, (forthcoming).

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

——— (2001): "Strategyproof and Nonbossy Multiple Assignments," *Journal of Public Economic Theory*, 3, 257–271.

Pathak, P. A. and J. Sethuraman (2010): "Lotteries in Student Assignment: An Equivalence Result," *Theoretical Economics*, forthcoming.

Pycia, M. (2011): "Ordinal Efficiency, Fairness, and Incentives in Large Multi-Unit-Demand Assignments," .

Pycia, M. and M. U. Ünver (2009): "Incentive Compatible Allocation and Exchange of Discrete Resources," UCLA and Boston College, unpublished mimeo.

——— (2016): "Arrovian Efficiency in Allocation of Discrete Resources," UCLA and Boston College, Working Paper.

Roth, A. E., T. Sönmez, and M. U. Ünver (2004): "Kidney Exchange," *Quarterly Journal of Economics*, 119, 457–488.

Satterthwaite, M. (1975): "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10, 187–216.

Shapley, L. and H. Scarf (1974): "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, 23–37.

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