

Strategy-proof Pareto-improvement*

Samson Alva

University of Texas at San Antonio
samson.alva@utsa.edu

Vikram Manjunath

Texas A&M University
vmanjunath@tamu.edu

March 27, 2017

Abstract

We consider a general framework, where each agent has an outside option of privately known value, that encompasses object allocation, matching with contracts, provision of excludable public goods, and more. First, we show that if a strategy-proof mechanism is required to (weakly) Pareto-improve an individually rational and *participation-maximal* benchmark mechanism, there is at most one choice. A consequence is that many known mechanisms, especially in market design applications, are on the Pareto frontier of strategy-proof mechanisms. Second, we characterize the Pareto-improvement relation over strategy-proof and individually rational mechanisms: one of these mechanisms Pareto-improves another if and only if it weakly expands the set of participants at each profile of preferences. Third, for settings with transfers, we obtain a characterization of the pivotal mechanism and a revenue equivalence theorem.

Keywords: strategy-proofness, participation-maximality, Pareto-improvement, school choice, matching with contracts, revenue equivalence

JEL Codes: C78; D47; D71; D82

1 Introduction

Mechanism design has been successful in providing desirable alternatives to allocation mechanisms used in the real world.¹ If a particular mechanism is a status quo or has

*We thank Inácio Bó, Alexander Brown, Rossella Calvi, Lars Ehlers, Ben Golub, Guillaume Haeringer, Patrick Harless, Eun Jeong Heo, Sean Horan, Fuhito Kojima, Scott Kominers, Silvana Krasteva, Timo Mennle, Debasis Mishra, Thayer Morrill, Mallesh Pai, William Phan, Santanu Roy, Erling Skancke, Tayfun Sönmez, Alexander Teytelboym, William Thomson, Guoqiang Tian, Ryan Tierney, Juuso Toikka, Utku Ünver, Rodrigo Velez, Dan Walton, Alexander Westkamp, Alexander Wolitzky, and Bumin Yenmez for helpful comments and discussions. We also thank seminar audiences at the University of Rochester, the University of Cologne, the University of Bonn, Southern Methodist University, the University of Windsor, Harvard/MIT, the University of Ottawa, North Carolina State University, and numerous conference participants for their feedback.

¹ The placement of medical students to residencies [Roth and Peranson, 1999], assignment of students to schools [Abdulkadiroğlu et al., 2005, Abdulkadiroğlu et al., 2005], exchange of kidney donors amongst

been identified based on normative considerations, then it forms a benchmark for agents' welfare. When designing a new mechanism, a natural requirement is that no agent be worse off than he would be under the benchmark, regardless of other agents' privately known preferences. Otherwise, an agent who is made worse off has reason to oppose such a change.

We study strategy-proof Pareto-improvements of a benchmark mechanism, in a general framework that unifies several standard models, from object allocation [Hylland and Zeckhauser, 1979], to school choice [Abdulkadiroğlu and Sönmez, 2003], to matching with contracts [Hatfield and Milgrom, 2005], to the provision of excludable public goods [Cantala, 2004], and more. While many mechanisms may Pareto-improve a given benchmark,² our main contribution is to show that at most one of them is strategy-proof if the benchmark is *individually rational* and *participation-maximal* (Theorem 1). In our framework, each allocation involves the presence of a certain subset of agents. We call them the *participants* at that allocation. The remaining agents are non-participants, and consume their outside option. Individual rationality says that no agent is asked to participate in an allocation that he finds worse than his outside option. Participation-maximality says that bringing a non-participating agent into the fold would make someone worse off. The uniqueness of the strategy-proof Pareto-improvement, should one exist, is in terms of welfare since we do not assume strict preferences. A corollary is that participation-maximality and individual rationality are jointly sufficient for a strategy-proof mechanism to be on the Pareto frontier of the set of strategy-proof mechanisms.

The generality of the framework permits myriad applications, which we discuss below. However, to help fix ideas, consider the *school choice problem* with strict preferences. Each school has its own priority ranking of students. An allocation assigns each student to a school or leaves him unassigned, subject to the constraint that no school is assigned beyond its capacity. It is individually rational if each assigned student prefers his assignment to receiving nothing. In this context, an allocation is participation-maximal if it is impossible to place a student assigned nothing at an acceptable school without exceeding capacities or making someone worse off.³ Many normatively appealing requirements on allocations imply individual rationality and participation-maximality. For instance, the most widely considered requirement in school choice is *stability*: if student i is assigned to school s , no student with higher priority than i at s envies i .⁴ An ideal list of desider-

patients [Roth et al., 2004, 2005], and allocation of the electromagnetic spectrum [Cramton, 1995, McAfee and McMillan, 1996, Milgrom, 2000] are just a few examples of success stories.

² By Pareto-improvement we mean each agent is at least as well off.

³ This is weaker than the standard criterion of *non-wastefulness* [Balinski and Sönmez, 1999].

⁴ Stability additionally requires an allocation to be non-wasteful and individually rational.

ata for a mechanism would include strategy-proofness, stability, and Pareto-efficiency. However, stability is typically incompatible with Pareto-efficiency [Balinski and Sönmez, 1999]. So, consider weakening the normative criterion from stability to the following: an allocation is *stable-dominating* if it Pareto-improves *some* stable allocation—no student has a reason to object to a violation of his priority if doing so would result in a move to this Pareto-inferior stable allocation.⁵ Theorem 1 implies that the *student-optimal stable mechanism*⁶ is the only stable-dominating and strategy-proof mechanism.^{7,8} We can draw similar novel conclusions for other normative concepts as well. Consider, for instance, the recently proposed notion of *legality* [Morrill, 2016]. The legal set contains a Pareto-worst allocation that is individually rational and participation-maximal. Furthermore, this set contains every stable allocation. Consequently, Theorem 1 says that the student-optimal stable mechanism is the unique legal and strategy-proof mechanism.⁹

When the benchmark mechanism is not participation-maximal, there may exist many strategy-proof Pareto-improvements. Nevertheless, considering the sets of participants at the allocations chosen by mechanisms, maximal or not, sheds light on the structure of the set of strategy-proof and individually rational mechanisms. Our second contribution is to characterize the Pareto-improvement relation over such mechanisms on the basis of comparing these sets of participants: one strategy-proof and individually rational mechanism Pareto-improves another if and only if it weakly expands, in the sense of set inclusion, the set of participants at each profile of preferences (Theorem 2).

To understand the value of Theorem 2, we consider its implications for a specific problem: the assignment of doctors to residency positions at hospitals. A long standing concern has been the shortage of doctors in rural areas [Roth, 1986]. One way of mitigating the shortage is to cap the number of doctors matched to other regions. Given caps, the *flexible deferred acceptance* mechanism [Kamada and Kojima, 2015] is strategy-proof for the doctors and, subject to the caps, Pareto-efficient. Loosening the caps leads to a Pareto-improvement, from the doctors’ point of view, in the allocation chosen by this mechanism. Thus, Theorem 2 says that loosening the caps for some or all of the regions

⁵ This is similar in spirit to the proposal of obtaining consent for priority violations that would not affect a consenting student’s welfare but improve another’s [Kesten, 2010].

⁶ The student-optimal stable mechanism selects the allocation determined by the student-proposing deferred acceptance algorithm [Gale and Shapley, 1962], and is strategy-proof [Dubins and Freedman, 1981].

⁷ This follows from Theorem 1 with the student-*pessimal* stable mechanism as the benchmark.

⁸ This is stronger than the existing result that there is no strategy-proof mechanism that strictly Pareto-improves the student-optimal stable mechanism [Abdulkadiroğlu et al., 2009].

⁹ In general, if a normative property refines individual rationality and participation-maximality, and the set of allocations satisfying this property contains a Pareto-best or Pareto-worst allocation for every preference profile, then there is at most one strategy-proof selection from it (Proposition 1).

without reducing those at the others affects the sets of matched doctors: for every profile of preferences, a doctor matched at the tighter caps is matched at the looser caps and there are profiles of preferences where a strictly larger (by inclusion) set of doctors is matched under the looser caps. This has implications for the proportion of doctors who are assigned to some hospital, the *match rate*, which is a simple yet policy-relevant measure of a mechanism’s performance. When designing a mechanism, the relevant statistic is actually the *expected* match rate with regards to a distribution over preference profiles. The implication of Theorem 2 is that changing the caps in a way that leads to a Pareto-improvement leads to a *strict* increase in the expected match rate.¹⁰ Thus, while it is possible to increase the match rate for rural hospitals by tightening caps on other regions, Theorem 2 provides a cautionary message: an expected increase of one doctor matched to a rural hospital leads to an expected decrease of more than one doctor matched to other areas.

Our framework also accommodates problems with transferable utility. In such settings, a mechanism can be decomposed into a pair consisting of a *decision* rule and a *payment* rule, and the participants are determined by the decision rule. Our third contribution is to relate our results to *revenue equivalence*.¹¹ Suppose a decision rule can be implemented in an individually rational way.¹² Every pair of payment rules that implement it differ, for each agent, only by a function of other agents’ preferences (Proposition 5).

We make certain assumptions on agents’ preferences over their outside options. Throughout the paper we assume that there are no externalities on non-participants. That is, each agent is indifferent between any pair of allocations that he does not participate in. So the welfare of a non-participant is completely determined by his outside option. We make two further assumptions on each agent’s preferences vis-à-vis his outside option. The first is an assumption of richness: the (upward) movement of the agent’s outside option in his preference ranking is essentially unrestricted. The second is that he is not indifferent between his outside option and an allocation in which he participates. Each of the two steps in proving Theorem 1 relies on one of these.

Richness of outside options says that the range of values that an agent’s outside option may take are on the order of what the mechanism designer may offer him, which is natural in many real-world applications. It is enough to prove a striking lemma about strategy-proof and individually rational mechanisms. For each profile of preferences, by

¹⁰ The conclusion that the increase is strict holds if the distribution has full support.

¹¹ There is an extensive literature on revenue (also known as payoff) equivalence [Holmström, 1979, Myerson, 1981, Krishna and Maenner, 2001, Chung and Olszewski, 2007, Heydenreich et al., 2009].

¹² A decision rule is implementable if there is a payment rule with which it forms a strategy-proof mechanism.

selecting an allocation, a mechanism selects a list of participants as well—those agents who participate at the allocation that it selects. Pinning down who participates at each profile per a strategy-proof and individually rational mechanism pins down the entire mechanism in terms of welfare (Lemma 1). Our results for transferable utility problems rely on this lemma, since richness of outside options is compatible with quasilinearity.

No indifference with the outside option requires that an agent is never indifferent between participating and not participating. It rules out quasilinear preferences, but is commonly assumed in many settings, such as those with discrete goods. With this assumption alone, the set of individually rational and participation-maximal allocations has a very particular structure. There is a unique finest partition of this set so that no two allocations in separate components can be compared in the Pareto sense. Every pair of allocations in the same component have the same participants (Lemma 2). This is much like the *Rural Hospitals Theorem* [Roth, 1986], which, in matching models, draws the same conclusion for every pair of *stable* allocations. In fact, the Rural Hospitals Theorem is a consequence of the entire stable set residing within a single component of the above described partition (Proposition 2).

Applications Since the framework is very general, our results are best appreciated through their corollaries for various special cases.

Many of the salient applications of Theorem 1 fall under market design. Consider first an object allocation model with strict preferences, augmented with a choice correspondence for each object. A choice correspondence conveys exogenous information about how to prioritize the different ways the object can be allocated, a generalization of weak priority relations in the school choice problem. Under certain conditions,¹³ each stable allocation is participation-maximal (Lemma 3).¹⁴ Thus, a corollary of Theorem 1 is that at most one strategy-proof mechanism Pareto-improves a stable benchmark. Consequently, if a stable mechanism is strategy-proof, it is strategy-proofness-constrained Pareto-efficient. As described above, for the school choice model with strict priorities, the agent-optimal stable mechanism is the unique strategy-proof mechanism to Pareto-improve a stable mechanism. For matching with contracts settings like the matching of military cadets to branches of the US Army [Sönmez and Switzer, 2013, Sönmez, 2013], medical residents to hospitals [Roth, 1984, Hatfield and Milgrom, 2005], or students to schools with diversity considerations [Kominers and Sönmez, 2016], the cumulative offer

¹³ The two conditions are *size monotonicity* (we use the terminology of Alkan and Gale [2003], whereas Hatfield and Milgrom [2005] refer to it as the *law of aggregate demand*) and *idempotence* (this is weaker than irrelevance of rejected contracts [Aygün and Sönmez, 2013]).

¹⁴ In fact, given these choice conditions, stability implies the stronger property of non-wastefulness.

mechanism¹⁵ is stable and strategy-proof. Even if stability does not imply participation-maximality, this mechanism is typically participation-maximal. Hence, Theorem 1 implies that it is strategy-proofness-constrained Pareto-efficient.

Going beyond matching problems, as an example, consider a model of locating an excludable public facility along an interval. Suppose that agents are located at either end of the interval and the only private information is how far each is willing to travel. An allocation here consists of a location along the interval as well as a set of users. If we insist on Pareto-efficiency alongside strategy-proofness and individual rationality, only dictatorial mechanisms remain: no compromise is possible. However, there are non-dictatorial strategy-proof mechanisms that are individually rational and participation-maximal. Theorem 1 tells us that these Pareto-inefficient mechanisms are all strategy-proofness-constrained Pareto-efficient.

For transferable utility problems, Lemma 1 implies that a given decision rule can be implemented by at most one payment rule in an individually rational way. If we consider an (utilitarian) efficient decision rule, Groves schemes implement it [Groves, 1973]. Pivotal payment rules [Vickrey, 1961, Clarke, 1971] are the most well known members of this family—for instance, they amount to the second price auction in the allocation of a single object. Lemma 1 implies a characterization of pivotal payment rules as the only ones to implement efficient decision rules in an individually rational way. Our framework requires that non-participants receive no payments. Relaxing this, we are able to adapt the argument of Lemma 1 to prove a revenue equivalence theorem. It says that if a decision rule can be implemented in an individually rational way, then any two payment rules that implement it, individually rationally or not, differ, for each agent, only by a function of others' preferences (Proposition 5). This illuminates a bridge between mechanism design with and without transfers: that the sets of participants at each profile determine a strategy-proof mechanism (Lemma 1) can be thought of as an analog for models without transfers of the revenue equivalence theorem.

Related Literature The uniqueness of a strategy-proof Pareto-improvement over a participation-maximal and individually rational benchmark (Theorem 1) is a novel result. Previous results have only dealt with Pareto-improvements over an already strategy-proof benchmark. The first one along these lines shows, for the school choice problem, that the student-optimal stable mechanism has no strategy-proof strict Pareto-improvement [Abdulkadiroğlu et al., 2009].¹⁶ A more recent result of this sort is for

¹⁵ This mechanism generalizes the agent-proposing deferred acceptance mechanism.

¹⁶ While outside options are crucial in our analysis, this particular result can be shown without it when school seats exceed the number of students [Kesten, 2010, Kesten and Kurino, 2015]. Indeed, if the bench-

the matching with contracts setting with choice functions. A strategy-proof, individually rational, and non-wasteful benchmark mechanism has no strategy-proof strict Pareto-improvement [Hirata and Kasuya, 2017].¹⁷ This is also true for the probabilistic allocation of indivisible goods [Erdil, 2014]. The more general question of Pareto-improving a benchmark that may not, itself, be strategy-proof is novel to the current paper.

In addition to the above-mentioned results, there have been others that add to our understanding of the Pareto frontier of strategy-proof mechanisms for various problems. These include the allocation with transfers of a single object [Sprumont, 2013] or multiple identical objects [Ohseto, 2006], the division of multiple perfectly divisible goods with single-peaked preferences [Anno and Sasaki, 2013], and the allocation of multiple types of objects [Anno and Kurino, 2016]. Our results tell us that for the general framework that we work in, individual rationality and participation-maximality are sufficient conditions for a mechanism to be on this frontier.

The generality of our framework and our assumptions on preferences are reminiscent of Sönmez [1999], who shows that single-valuedness of the core (in welfare terms) is a necessary condition for a strategy-proof, individually rational, and Pareto-efficient mechanism to exist. In his model, the individual rationality constraint is with regards to an endowment rather than an outside option. Nonetheless, his assumptions on the preference domain play a similar role in showing essential uniqueness as in our proofs.

We are heretofore aware of only one result from this literature in the spirit of Theorem 2. One of the directions has been shown to hold for probabilistic allocation of indivisible goods when each agent has unit demand and strict preferences: Pareto-improvement in the stochastic dominance sense implies a participation-expansion relation [Erdil, 2014]. This result is limited by the model's restriction to allocation of objects. However, since we work with deterministic allocations only, Theorem 2 does not imply this result either.

The pivotal payment rule has been characterized for both indivisible private goods [Chew and Serizawa, 2007] and pure public goods [Moulin, 1986]. We provide a new characterization for the provision of excludable public goods (Corollary 8). Characterizations of the pivotal payment rule are closely related to revenue equivalence [Holmström, 1979]. Proposition 5 demonstrates that revenue equivalence holds for every decision rule

mark mechanism is student-optimal stable, then each student has similar “rights” to an “under-demanded” school as he might have to an outside option. Intuitively, the under-demanded school serves as a pseudo-outside option, and so the logic behind the result remains the same. Nevertheless, the identification of under-demanded schools requires care.

¹⁷ This result requires choice functions satisfy irrelevance of rejected contracts [Aygün and Sönmez, 2013].

that can be implemented in an individually rational way as long as our rich outside options assumption is met. Since [Heydenreich et al. \[2009\]](#) characterize the decision rules and valuation spaces where revenue equivalence holds, our contribution is to provide easy to check conditions (individual rationality and richness of outside options) that ensure that their theorem applies.

Our results also contribute to the large literature on mechanism design in quasilinear settings with voluntary participation.¹⁸ Our analysis departs from this literature by focusing on dominant strategy, rather than Bayesian-Nash, incentive compatibility, where ex-post individual rationality is natural. Moreover, we provide insights into problems without quasilinearity in transfers.

The remainder of the paper is organized as follows. We introduce our framework in Section 2. We define several properties of allocations and mechanisms in Section 3. Our main results are in Section 4. Applications are in Section 5. We conclude in Section 6. All proofs are in Appendix A. Supporting material are in Appendices B through E.

2 The Framework

We first consider the general framework in which we prove our results. We then add structure to show how to accommodate the problem of object allocation, studied in Sections 5.1 and 5.2. We illustrate in Section 5.3 how to model excludable public goods, and in Section 5.4 how to work with transferable utility. We embolden notation only when first introduced.

Let N be a finite and nonempty set of agents. Let \mathcal{F} be the nonempty set of allocations. Given $\alpha \in \mathcal{F}$, let nonempty $N(\alpha) \subseteq N$ be the **participants** at α . For each $i \in N$, let $\mathcal{F}_i \subseteq \mathcal{F}$ be the nonempty set of allocations that i participates in. If $\alpha \in \mathcal{F}$ is chosen and i is not a participant at α , then i consumes his **outside option**. We denote consumption of the outside option by \emptyset . For each $\alpha \in \mathcal{F}$, we denote by $\alpha(i)$ either α if $i \in N(\alpha)$ or \emptyset otherwise. Note that the participation of each $i \in N(\alpha)$ is a part of the definition of α . That is, α *requires* the participation of i . Thus, participation is anterior to i 's preferences—it is not a choice that i makes.

For each $i \in N$, his preferences are a complete, reflexive, and transitive binary relation on $\mathcal{F}_i \cup \{\emptyset\}$.¹⁹ We denote it by R_i . For each pair $\alpha, \beta \in \mathcal{F}$, we write $\alpha(i) R_i \beta(i)$ to mean

¹⁸ See, for instance, [Jullien \[2000\]](#), [Jackson and Palfrey \[2001\]](#), and [Compte and Jehiel \[2007, 2009\]](#).

¹⁹ Since i 's preferences are over $\mathcal{F}_i \cup \{\emptyset\}$ rather than \mathcal{F} , i is indifferent between any pair of allocations that he does not participate in. Consequently, his welfare from such allocations is fully determined by his outside option. This rules out externalities on non-participants.

that i finds $\alpha(i)$ to be at least as good as $\beta(i)$. We use P_i to denote strict preference and I_i to denote indifference, the asymmetric and symmetric components of R_i , respectively. Let \mathcal{R}_i be a set of preference relations for i . The *preference domain* is $\mathcal{R} \equiv \times_{i \in N} \mathcal{R}_i$.

Our first assumption on \mathcal{R} is **richness of outside options**: for each $i \in N$, each $R_i \in \mathcal{R}_i$, and each pair $\alpha, \beta \in \mathcal{F}_i$ such that $\alpha(i) P_i \beta(i) R_i \emptyset$, there is $R'_i \in \mathcal{R}_i$ such that (1) $\alpha(i) P'_i \beta(i)$, and (2) for each $\gamma \in \mathcal{F}_i$, if $\gamma(i) R'_i \emptyset$ then $\gamma(i) P_i \beta(i)$. All of our results, except for Lemma 2, rely on this assumption. Our second assumption on \mathcal{R} is **no indifference with \emptyset** : for each $i \in N$ and each $R_i \in \mathcal{R}_i$, there is no $\alpha \in \mathcal{F}_i$ such that $\alpha(i) I_i \emptyset$. We rely on it for all but Lemma 1 and its applications to transferable utility problems.²⁰

Let $\mathcal{P}_i \subseteq \mathcal{R}_i$ be the subset of all strict (antisymmetric) preferences in \mathcal{R}_i . For each $R_i \in \mathcal{P}_i$, I_i is trivial, so P_i completely identifies R_i . When working with strict preference, we refer to $P_i \in \mathcal{P}_i$. The *strict subdomain* of \mathcal{R} is $\mathcal{P} \equiv \times_{i \in N} \mathcal{P}_i$. Notice that the set of all strict preferences satisfies both the no indifference with \emptyset and the richness of outside options conditions.

Our analysis is for fixed N , \mathcal{F} , and \mathcal{R} . Thus, an economy is entirely described by $R \in \mathcal{R}$. A **mechanism**, $\varphi : \mathcal{R} \rightarrow \mathcal{F}$, associates each economy with an allocation. For each $R \in \mathcal{R}$ and each $i \in N$, instead of $\varphi(R)(i)$, we write $\varphi_i(R)$.

While this framework covers many applications, we next show how to accommodate a more structured model of object allocation. This facilitates the relation of our results to existing literature on object allocation.

The object allocation model The object allocation model, in addition to a set of agents N , consists of a finite and nonempty set O of objects, a nonempty set T of terms under which an agent may be assigned an object, and a nonempty set $X \subseteq N \times O \times T$ of possible triples. The triple $(i, o, t) \in N \times O \times T$ represents “ i consumes o under the terms t .”²¹ For each $x \in X$, let $N(x)$ be the agent associated with x . For each $Y \subseteq X$, let $N(Y)$ be the set of agents associated with triples in Y , let $Y(i)$ be the triples in Y associated with $i \in N$, and let $Y(o)$ be the triples in Y associated with $o \in O$. Each object may only be allocated in certain ways. These constraints define, for each $o \in O$, the **feasible sets** for o , which is a collection of subsets of $X(o)$. We denote it by \mathcal{F}_o . In an allocation, each agent has one triple from $X(i)$ or consumes his outside option, \emptyset . Only one’s own triple matters, so he has a preference relation over $X(i) \cup \{\emptyset\}$.

This object allocation model can be embedded into the general framework as follows.

²⁰ Sönmez [1999] makes two similar assumptions. Also see Erdil and Ergin [2015] for another instance of no indifference with \emptyset .

²¹ These are *contracts* in Hatfield and Milgrom [2005].

An allocation $\mu \in \mathcal{F}$ is a subset Y of X such that no two triples name the same agent and each object's assignment is a feasible set. If $Y(i)$ is empty for agent i , he consumes his outside option, \emptyset . That is, \mathcal{F} is a subset of 2^X such that for each $\mu \in \mathcal{F}$, each $i \in N$, and each $o \in O$, $|\mu(i)| \leq 1$ and $\mu(o) \in \mathcal{F}_o$. Thus, the participants at μ , $N(\mu)$, are the agents associated with some triple in μ . Each agent's preference relation in the object allocation model defines a preference relation over \mathcal{F} . The definitions of our two assumptions—richness of outside options and no indifference with \emptyset —map to assumptions on preferences for the object allocation model in the obvious way.

When T is a singleton, each $x \in X$ is fully identified by the associated agent and object. In such cases, for each $i \in N$, each triple in $X(i)$ is identified by an element of O , while for each $o \in O$, each triple in $X(o)$ is identified by an element of N . Also, each object's feasible set is identified by a collection of subsets of N , while each agent's preference relation is identified by an ordering of $O \cup \{\emptyset\}$.

As an example, consider the classical object allocation model, where each $o \in O$ is an object with capacity $q_o \in \mathbb{Z}_+$. In this model, there is only one term under which an agent can be assigned to an object. For each $o \in O$, \mathcal{F}_o consists of all subsets of $X(o)$ containing no more than q_o elements. In such cases, where, for each $o \in O$, there is $q_o \in \mathbb{R}_+$ such that for each $Y \subseteq X(o)$, $Y \in \mathcal{F}_o$ if and only if $|Y| \leq q_o$, we say that \mathcal{F}_o is **capacity-based**.

Our framework accommodates problems with cross-object constraints [Kamada and Kojima, 2015, Goto et al., forthcoming]: \mathcal{F} can be any subset of 2^X so long as each allocation contains at most one triple associated with each agent. In the absence of such cross-object constraints, the only constraints are that the triples chosen for each object ought to be feasible. That is, for each $\mu \subseteq X$, $\mu \in \mathcal{F}$ if and only if (1) for each $i \in N$, $|\mu(i)| \leq 1$, and (2) for each $o \in O$, $\mu(o) \in \mathcal{F}_o$. In such cases, we say that \mathcal{F} is **Cartesian**. If, in addition to \mathcal{F} being Cartesian, for each $o \in O$, \mathcal{F}_o is capacity-based, then we say that \mathcal{F} is **capacity based**.

3 Properties of Allocations and Mechanisms

Individual rationality An allocation is **individually rational** if each agent finds it to be at least as good as not participating. That is, for each $R \in \mathcal{R}$ and each $\alpha \in \mathcal{F}$, α is individually rational at R if, for each $i \in N$, $\alpha(i) R_i \emptyset$.²² A mechanism, φ , is *individually rational* if, for each $R \in \mathcal{R}$, $\varphi(R)$ is individually rational at R . Individual rationality ensures that no such agent has an incentive to exercise his outside option when the allocation chosen by the mechanism relies on his presence.

²² Since each $i \notin N(\alpha)$ consumes \emptyset , it suffices to verify that for each $i \in N(\alpha)$, $\alpha(i) R_i \emptyset$.

Participation-expansion An allocation **participation-expands** another if participation in the latter entails participation in the former. That is, for each pair $\alpha, \beta \in \mathcal{F}$, α participation-expands β if $N(\alpha) \supseteq N(\beta)$. If they have the same participants, then they are **participation-equivalent**. That is, α is participation-equivalent to β if $N(\alpha) = N(\beta)$. Given a pair of mechanisms φ and φ' , φ *participation-expands* φ' if for each $R \in \mathcal{R}$, $\varphi(R)$ participation-expands $\varphi'(R)$. They are *participation-equivalent* if for each $R \in \mathcal{R}$, $\varphi(R)$ and $\varphi'(R)$ are participation equivalent.

Participation-maximality An allocation is **participation-maximal** at a given $R \in \mathcal{R}$ if there is no other allocation that strictly expands the set of participants without harming anyone. That is, for each $R \in \mathcal{R}$, $\alpha \in \mathcal{F}$ is participation-maximal at R if there is no $\beta \in \mathcal{F}$ such that (1) $N(\alpha) \subsetneq N(\beta)$, so that at least one additional agent participates at β compared to α and (2) there is no $i \in N$, such that $\alpha(i) P_i \beta(i)$, so that nobody is worse off at β than at α . A mechanism, φ , is *participation-maximal* if, for each $R \in \mathcal{R}$, $\varphi(R)$ is participation-maximal at R .²³

Pareto-improvement One allocation **Pareto-improves** another if each agent finds the first to be at least as desirable as the second. That is, for each $R \in \mathcal{R}$, and each pair $\alpha, \beta \in \mathcal{F}$, α Pareto-improves β at R if for each $i \in N$, $\alpha(i) R_i \beta(i)$.²⁴ If α Pareto-improves β at R and there is $i \in N$ such that $\alpha(i) P_i \beta(i)$, then α **strictly Pareto-improves** β at R . If $\alpha \in \mathcal{F}$ is such that no allocation strictly Pareto-improves it at R , then α is **Pareto-efficient** at R . Every Pareto-efficient allocation is participation-maximal, but the converse is not true.

The Pareto-improvement relation is reflexive and transitive but not complete. Two allocations are **Pareto-comparable** if one Pareto-improves the other. Two individually rational and participation-maximal allocations are **Pareto-connected** (within the individually rational and participation-maximal set) if there is a sequence of individually rational and participation-maximal allocations starting at one and ending at the other such that successive allocations are Pareto-comparable. That is, two individually rational and participation-maximal allocations, $\alpha, \beta \in \mathcal{F}$, are Pareto-connected if there is a sequence $(\alpha_k)_{k=0}^K$, with $\alpha_0 \equiv \alpha$ and $\alpha_K \equiv \beta$, such that for each $k \in \{1, \dots, K\}$, α_k is individually rational and participation-maximal and Pareto-comparable to α_{k-1} .

For each pair of mechanisms φ and φ' , φ *Pareto-improves* φ' if, for each $R \in \mathcal{R}$, $\varphi(R)$

²³Participation-maximality is weaker than non-wastefulness, which we define for the object allocation model under strict preferences in Section 5.1.

²⁴ In this case, some authors say that α *weakly* Pareto-improves β . However, since this is the main form of Pareto-improvement that we consider, we drop the qualifier.

Pareto-improves $\varphi'(R)$ at R . If φ Pareto-improves φ' and for some $R \in \mathcal{R}$, $\varphi(R)$ strictly Pareto-improves $\varphi'(R)$ at R , then φ *strictly Pareto-improves* φ' . If they are both individually rational and participation-maximal, then φ and φ' are *Pareto-connected* if, for each $R \in \mathcal{R}$, $\varphi(R)$ and $\varphi'(R)$ are Pareto-connected at R . If, for each $R \in \mathcal{R}$ and each $i \in N$, $\varphi_i(R) \succeq_i \varphi'_i(R)$, then φ and φ' are **welfare-equivalent**. If, for each $R \in \mathcal{R}$, $\varphi(R)$ is Pareto-efficient at R , then φ is *Pareto-efficient*.

Strategy-proofness A mechanism, φ , is **strategy-proof** if no agent can benefit by misreporting his preferences, no matter what other agents do. That is, for each $R \in \mathcal{R}$, each $i \in N$, and each $R'_i \in \mathcal{R}_i$, $\varphi_i(R) \succeq_i \varphi_i(R'_i, R_{-i})$.

Extending this concept to groups of agents, φ is **group strategy-proof** if no group of agents can misreport their preferences in a way that at least one member is better off while each member is at least as well off. That is, for each $R \in \mathcal{R}$ and each $S \subseteq N$, there is no $R'_S \in \times_{i \in S} \mathcal{R}_i$, such that for each $i \in S$, $\varphi_i(R'_S, R_{-S}) \succeq_i \varphi_i(R)$ and for some $i \in S$, $\varphi_i(R'_S, R_{-S}) \succ_i \varphi_i(R)$.

4 Strategy-proof Pareto-improvement

We present here our fundamental results. First, given an individually rational and participation-maximal benchmark mechanism, we are interested in strategy-proof mechanisms that Pareto-improve it. When a designer's choice of a strategy-proof mechanism is constrained in this way, Theorem 1 says that his problem has a unique solution (in terms of welfare) if it has one at all.

Theorem 1. *Consider an individually rational and participation-maximal benchmark mechanism. In welfare terms, there is at most one strategy-proof mechanism that Pareto-improves it.*

We prove Theorem 1 by way of two lemmas of independent interest. The first is the Participation-equivalence Lemma, which implies that two welfare-distinct individually rational mechanisms that are participation-equivalent cannot both be strategy-proof.

Lemma 1 (Participation-equivalence Lemma). *If a pair of strategy-proof and individually rational mechanisms are participation-equivalent, then they are welfare-equivalent.*

The Participation-equivalence Lemma is interesting in its own right. By selecting an allocation for each profile of preferences, a mechanism selects a list of participants as well—those agents who participate at the allocation that it selects. The lemma says that

pinning down who participates at each profile pins down, in welfare terms, the entire mechanism. It does not require \mathcal{R} to satisfy no indifference with \emptyset , and so is relevant for problems with transferable utility. We return to this fact in Section 5.4.

The second lemma is the Structure Lemma. It says that the set of individually rational and participation-maximal allocations has a particularly nice structure. The Pareto-connectedness relation over this set is reflexive and symmetric. Moreover, since it is transitive, it is an equivalence relation over the set of individually rational and participation-maximal allocations. Therefore, it partitions the set of individually rational and participation-maximal allocations into Pareto-connected components. The following lemma says that every allocation in the same component involves the same participants. Since the statement of the lemma is specific to a fixed profile of preferences, it does not rely on richness of outside options.²⁵

Lemma 2 (Structure Lemma). *For each profile of preferences, if a pair of individually rational and participation-maximal allocations are Pareto-connected, then they are also participation-equivalent.*

In fact, these lemmas actually deliver more than Theorem 1. The next proposition, a consequence of these lemmas, states that among a set of individually rational, participation-maximal, and Pareto-connected mechanisms, at most one can be strategy-proof.

Proposition 1. *If a pair of distinct strategy-proof, individually rational, and participation-maximal mechanisms are Pareto-connected, then they are welfare-equivalent.*

Next, it is clear that Theorem 1 does not generally hold for benchmark mechanisms that are not participation-maximal. If, for instance, the trivial mechanism that assigns \emptyset to each agent is the benchmark, then each individually rational mechanism Pareto-improves it. However, depending on the structure of \mathcal{F} , many of these may be strategy-proof: in object allocation problems, for instance, each serial dictatorship is a strategy-proof Pareto-improvement.

Nevertheless, we can still say something about strategy-proof mechanisms that Pareto-improve one another. By Theorem 2, the participation-expansion relation over strategy-proof and individually rational mechanisms is equivalent to the Pareto-improvement relation. Since the two relations are equivalent, this means that a strategy-proof mechanism strictly Pareto-improves another if and only if it strictly expands participation.

²⁵ We show, in Section 5.1, that by adding structure to the model, we can draw stronger conclusions. Indeed, we show that the Structure Lemma is closely related to the Rural Hospitals Theorem for matching problems [Roth, 1986].

Theorem 2. *The Pareto-improvement relation coincides with the participation-expansion relation on the set of strategy-proof and individually rational mechanisms.*

For individually rational mechanisms, it is clear that no indifference with \emptyset ensures that Pareto-improvement implies participation-expansion. What about the converse? A proof similar to that of the Participation-equivalence Lemma shows that for mechanisms that are both strategy-proof and individually rational, participation-expansion implies Pareto-improvement even if agents may be indifferent between participating and \emptyset .

The **strategy-proofness-constrained** Pareto frontier consists of each strategy-proof mechanism that is not strictly Pareto-improved by another strategy-proof mechanism. Theorem 1 tells us that a strategy-proof mechanism is on this frontier if it is individually rational and participation-maximal.

Corollary 1. *A strategy-proof mechanism is strategy-proofness-constrained Pareto-efficient if it is participation-maximal and individually rational.*

Theorem 2 says that each strategy-proof mechanism below this frontier is strictly participation-expanded by some mechanism on the frontier. We might ask whether a mechanism on the frontier must be participation-maximal. This is not so. Depending on the structure of \mathcal{F} , there may be mechanisms on this frontier that are not participation-maximal.

Remark 1. *Strategy-proofness-constrained Pareto-efficiency does not imply participation-maximality.*

We provide an example, in Appendix A.1, of a strategy-proof and individually rational mechanism not on the strategy-proofness-constrained frontier with no strategy-proof Pareto-improvement that is participation-maximal. It is noteworthy that the setting is not particularly special. It is the classical object allocation problem with strict preferences. The example actually demonstrates more than Remark 1: there is a *group* strategy-proof mechanism that is not Pareto-improved by a strategy-proof mechanism, yet is not participation-maximal. Thus, even among group strategy-proof mechanisms, participation-maximality is not a necessary condition for a mechanism to be strategy-proofness-constrained Pareto-efficient.

5 Applications

We consider several applications to demonstrate our results. We start, in Section 5.1, with the object allocation model, augmented with choice correspondences, which gen-

eralizes the matching with contracts model. We describe applications of our theorems as well as some further results in this setting. In Section 5.1.1, we discuss their implications for recent developments in market design. We further specialize the model to school choice in Section 5.1.2. We consider the reallocation of objects from an endowment in Section 5.2, and a model of excludable public goods in Section 5.3. We apply Participation-equivalence Lemma to transferable utility problems in Section 5.4.

5.1 Object Allocation and Matching

We consider the *object allocation* model in this section. We also assume that each agent has *strict preferences*, as is typical in the much of this literature.²⁶ That is, we restrict attention to profiles in \mathcal{P} .

In many applications, there is more information available about each object than just the feasible sets. These might be priorities over agents as in school choice, objectives of the army in cadet-branching, and so on. We model the extra information about how feasible sets are prioritized by associating each $o \in O$ with a **choice correspondence**, $C_o : 2^{X(o)} \rightrightarrows 2^{X(o)}$, such that (1) for each $Y \subseteq X(o)$, $C_o(Y) \subseteq 2^Y$, and (2) $\text{range}(C_o) = \mathcal{F}_o$.²⁷ Condition (1) says that from any set, C_o picks only subsets of it. Condition (2) says that the feasible sets are exactly those that are chosen from *some* set. To satisfy Condition (2), it would suffice, for instance, to select each feasible set from itself.²⁸ Let $C \equiv (C_o)_{o \in O}$.

Our model is more general than the matching with contracts model [Hatfield and Milgrom, 2005] since we associate each object with a choice correspondence rather than a *choice function*. Since applications like school choice with weak priorities [Erdil and Ergin, 2008, Abdulkadiroğlu et al., 2009] are better modeled with choice correspondences, we adopt this more general approach.²⁹ Furthermore, since we do not require \mathcal{F} to be Cartesian, our results may be applied even where there are distributional constraints [Kamada and Kojima, 2015].

The constraints imposed by this information sometimes keep the benchmark mechanism below the Pareto frontier.

Stability An allocation is stable if no set of agents prefer to drop their assignments in favor of being assigned to a new object under some terms that the object would “choose.”

²⁶ See Bogomolnaia et al. [2005] and Erdil and Ergin [2015] for more on the problems that arise when modeling indifference in such contexts.

²⁷ The range of $C_o : 2^{X(o)} \rightrightarrows 2^{X(o)}$ is $\bigcup_{Y \subseteq X(o)} C_o(Y)$.

²⁸ An alternative approach is to start with C_o as the primitive and define \mathcal{F}_o to be its range.

²⁹ To our knowledge, the first and only other analysis of matching to consider general choice correspondences as a primitive is by Erdil and Kumano [2014].

That is, for each $\mu \in \mathcal{F}$ and $P \in \mathcal{P}$, μ is **stable** at P if it is individually rational³⁰ and there are no $o \in O$ and $Y \subseteq X(o) \setminus \mu(o)$ such that (1) for each $i \in N, |Y(i)| \leq 1$, (2) for each $y \in Y, y P_{N(y)} \mu(N(y))$, (3) $\mu(o) \notin C_o(\mu(o) \cup Y)$, (4) there is $Z \in C_o(\mu(o) \cup Y)$ such that $Y \subseteq Z$, and (5) $(\mu \setminus (\mu(o) \cup \mu(N(Y)))) \cup Z \in \mathcal{F}$. Condition (1) says that Y contains at most one triple per agent. Condition (2) says that every agent associated with a triple in Y finds it preferable to his assignment at μ . These are familiar conditions from the definition of stability for choice functions. Since we are concerned with choice correspondences, the next part of the definition needs to be broken into two parts. The first, Condition (3), says that $\mu(o)$ is not among what is chosen by \emptyset when Y is available. The second, Condition (4), says that there is some chosen set, Z , that contains Y . That is, Condition (3) and Condition (4) together say that Y is contained in some Z that is *revealed* by C_o to have a higher priority than $\mu(o)$. The standard definition of stability typically does not include Condition (3) since it is implied by Condition (4) when choice correspondences are single-valued. Condition (5) is only relevant if \mathcal{F} is not Cartesian. It requires that replacing the assignments of o and $N(Y)$ at μ with Z is feasible. A mechanism, φ , is *stable* if, for each $P \in \mathcal{P}$, $\varphi(P)$ is stable.³¹

Stability is relevant if the choice correspondences represent more than feasibility constraints: they may represent the rights of agents with regards to the objects or particular design goals of the policy maker. A stable mechanism is thus a natural candidate for a benchmark that the mechanism designer may need to Pareto-improve. This is particularly the case because stability is often incompatible with Pareto-efficiency. The designer's choice of mechanism might then be constrained to those that select, at every preference profile, an allocation that Pareto-improves *some* stable allocation. Such a mechanism is **stable-dominating**.³² The Pareto-improvement need not be *strict*, so every stable mechanism is stable-dominating.

Since Theorem 1 only applies to participation-maximal benchmark mechanisms, we need stability to imply participation-maximality to invoke it here. We will, in fact, consider a stronger property than participation-maximality called non-wastefulness.

Non-wastefulness For the classical object allocation model, where feasibility is capacity-based, a natural requirement is that an agent ought not to prefer an object that has remaining capacity to his assignment. If he were to, we could allow him to consume

³⁰ Individual rationality accounts for agents' preferences while feasibility, along with the requirement that, for each $o \in O$, \mathcal{F}_o be the range of C_o , accounts for objects' choice correspondences.

³¹ If, for each $o \in O$, C_o is single-valued, this definition is equivalent to the standard definition of stability.

³² We define this property for allocations as well: $\mu \in \mathcal{F}$ is *stable-dominating* if there is $\nu \in \mathcal{F}$ such that ν is stable and μ Pareto-improves ν .

this available resource at no expense to the other agents.

We define non-wastefulness for our general object allocation model as follows: Given $P \in \mathcal{P}$, $\mu \in \mathcal{F}$ is *wasteful* if there are $o \in O$, $i \in N$, and $\nu \in \mathcal{F}$, such that (1) $|\nu(o)| > |\mu(o)|$, so that ν allocates o to more agents than μ does, (2) $\nu(i) P_i \mu(i)$, so that i prefers his assignment at ν to that at μ , and (3) for each $j \in N \setminus \{i\}$, $\nu(j) R_j \mu(j)$, so that no agent is worse off at ν compared to μ . If it is not wasteful, then μ is **non-wasteful**. A mechanism, φ , is *non-wasteful* if, for each $P \in \mathcal{P}$, $\varphi(P)$ is non-wasteful. Non-wastefulness is a stronger requirement on an allocation than participation-maximality.

For the classical object allocation model, [Balinski and Sönmez \[1999\]](#) define non-wastefulness as follows: $\mu \in \mathcal{F}$ is non-wasteful at $P \in \mathcal{P}$ if there is no $o \in O$ such that $|\mu(o)| < q_o$ and $i \in N$ such that $o P_i \mu(i)$. For this narrower setting, our definition of non-wastefulness is equivalent to this.³³ For more on non-wastefulness, see [Appendix B](#).

In many applications, like school choice, where priority rankings define the choice correspondences, it is easy to see that stability implies non-wastefulness, and thus implies participation-maximality. However, this may not be the case without any restrictions on choice correspondences, even if they are single-valued ([Example 1](#) in [Appendix A.2](#)).

We place two restrictions on choice correspondences to address this. The first says that the choices from each set should be at least as large as each choice from each of its subsets. That is, C is **size monotonic** if, for each $o \in O$, each $Y \subseteq X(o)$, each finite $Y' \subseteq Y$, each $Z \in C_o(Y)$, and each $Z' \in C_o(Y')$, $|Z| \geq |Z'|$.^{34,35} The second restriction is a mild consistency requirement. It says that if a set is among those chosen from a larger set, it ought to be among what is chosen from itself. That is, C is **idempotent** if, for each $o \in O$, and each $Y \in \text{range}(C_o)$, $Y \in C_o(Y)$.³⁶ Unlike most of the literature on matching with contracts, we do not assume that choice correspondences satisfy a condition like *irrelevance of rejected contracts* (IRC) [[Aygün and Sönmez, 2013](#)].³⁷

The assumptions that preferences are strict and that choice correspondences are size monotonic and idempotent ensure that every stable allocation is non-wasteful. The complexity of the proof, in [Appendix A.2](#), is because we have not assumed IRC.

³³ For capacity-based \mathcal{F} and singleton T , [Ehlers and Klaus \[2014\]](#) define an even weaker property that they call *weak non-wastefulness*. However, even in that specific setting, a result like the [Structure Lemma](#) does not hold for such a weak version of non-wastefulness.

³⁴ This is an extension to correspondences of a condition defined for choice functions [[Alkan, 2002](#), [Alkan and Gale, 2003](#), [Fleiner, 2003](#), [Hatfield and Milgrom, 2005](#)].

³⁵ For finite Y , setting $Y' = Y$, size monotonicity implies that for each pair $Z, Z' \in C_o(Y)$, $|Z| = |Z'|$.

³⁶ This rules out, for instance, C_o such that $C_o(\{x, y, z\}) = \{\{x, y\}\}$ but $C_o(\{x, y\}) = \{\{x\}\}$.

³⁷ [Aygün and Sönmez \[2013\]](#) define IRC for choice functions. [Alva \[2016\]](#) extends the definition to choice correspondences and shows its equivalence to the weak axiom of revealed preference.

Lemma 3. *Suppose that C is size monotonic and idempotent. For each profile of preferences, if an allocation is stable, then it is non-wasteful.*

Lemma 3 allows us to link our fundamental results to the matching with contracts setting. The assumptions of size monotonicity and idempotence make this possible but are not necessary. It bears emphasizing that to apply Theorem 1, in this setting one only needs to ensure that the benchmark mechanism is *participation-maximal*. In particular, we have the following corollary of Theorem 1 and Lemma 3 that there is at most one strategy-proof Pareto-improvement from a stable benchmark mechanism.

Corollary 2. *If C is size monotonic and idempotent, for each stable-dominating benchmark mechanism, there is at most one strategy-proof mechanism that Pareto-improves it.*

An implication of Corollary 2 is that if a stable-dominating mechanism is also strategy-proof, then it is not Pareto-improved by any other strategy-proof mechanism. Hirata and Kasuya [2017] show that this result holds for stable rules and single-valued choice functions that satisfy IRC. In fact, they show that for such choice functions, there is at most one stable and strategy-proof mechanism. In Appendix D, we show that under our assumptions, there may be more than one such mechanism.³⁸ This highlights the difference between their approach and ours. Nevertheless, we demonstrate below the uniqueness of a stable and strategy-proof mechanism under further conditions.

Kamada and Kojima [2015], in a matching model with distributional constraints, define a strategy-proof mechanism that Pareto-improves on a deferred acceptance mechanism. This does not contradict the corollary, which takes \mathcal{F} as fixed. By making flexible the application of the constraints that define \mathcal{F} , these authors obtain a strategy-proof Pareto-improvement on a benchmark that is no longer participation-maximal under the redefined \mathcal{F} .³⁹ In a similar setting with more general constraints, Goto et al. [forthcoming] show that an adaptation of the deferred acceptance mechanism is strategy-proofness-constrained Pareto-efficient.

³⁸ The example in Appendix D consists of an object associated with a choice function that endogenously selects a “tie breaker” over the agents. Consequently, it violates IRC. While this may appear strange, for the school choice model with weak priorities, Ehlers and Erdil [2010] provide an example where fixed tie-breaking entails a loss of efficiency while there is a mechanism with endogenous tie-breaking that does not. Hatfield and Kominers [2014] show that in a *slot-specific priorities* setting, if the order of precedence in which slots are filled depends on the set of agents that are being considered, then the choice function may violate IRC. Thus, it may be worthwhile considering such choice functions.

³⁹ In particular, the benchmark mechanism in their study is the deferred acceptance mechanism where objects have target capacities that satisfy distributional constraints. Taking these targets as given, this benchmark has no strategy-proof Pareto-improvement. However, by allowing a mechanism the flexibility to exceed these targets while satisfying the distributional constraints, which changes the set of feasible allocations \mathcal{F} , Kamada and Kojima [2015] manage to obtain a strategy-proof Pareto-improvement.

Given a profile of preferences P , if an allocation μ Pareto-improves every stable allocation and is stable at P , then μ is the **agent-optimal stable** allocation—strict preferences implies it is unique. If C is such that an agent-optimal stable allocation exists for each profile of preferences, then we denote by φ^{AOS} the mechanism that selects this allocation. Since we do not make assumptions about choice correspondences beyond idempotence and size monotonicity, the existence of a stable allocation is not guaranteed, let alone an agent-optimal stable allocation. Another property of the set of stable allocations that our assumptions do not guarantee is the **Rural Hospitals Theorem**. It states that the set of participants and the number of assigned agents at each object are the same across all stable allocations.

When choice correspondences are such that stability implies non-wastefulness and an agent-optimal stable allocation exists, the Rural Hospitals Theorem is a consequence of the following variant of the Structure [Lemma](#). The lemma strengthens participation-maximality to non-wastefulness and draws the stronger conclusion that Pareto-comparable allocations assign each object to the same number of agents.⁴⁰

Lemma 4 (Non-wasteful Structure Lemma for Object Allocation). *For each profile of preferences, if one allocation Pareto-improves another one that is individually rational and non-wasteful, then the two are participation-equivalent and assign each object to the same number of agents.*

This result sheds light on what drives the Rural Hospitals Theorem. The first is non-wastefulness. The Non-wasteful Structure [Lemma](#) says that, regardless of stability, at any pair of Pareto-connected allocations from the individually rational and non-wasteful set, the conclusion of the Rural Hospital Theorem holds. The only additional thing required is that the stable set be Pareto-connected. This is guaranteed, for instance, if the agent-optimal stable allocation exists and C is size monotonic and idempotent.

In fact, it turns out that (1) the existence of φ^{AOS} , (2) the existence of a stable-dominating and strategy-proof mechanism, and (3) the Rural Hospitals Theorem are intimately connected.⁴¹

First, using [Proposition 1](#), we show that size monotonicity and idempotence of C and existence of φ^{AOS} ensures that φ^{AOS} is the *unique* stable-dominating and strategy-proof

⁴⁰ In [Appendix B](#), we further strengthen the conclusions of the Non-wasteful Structure [Lemma](#) by assuming that \mathcal{F} is capacity-based, and relate it to a stronger version of the Rural Hospitals Theorem. This version requires that the set of agents assigned an object not filled to capacity be the same across all stable allocations.

⁴¹ [Hatfield and Kojima \[2010\]](#) provide conditions on single-valued C that guarantee the first two statements.

mechanism and that the Rural Hospitals Theorem holds.⁴²

Proposition 2. *Suppose that C is size monotonic and idempotent. If φ^{AOS} exists, then it is the unique stable-dominating and strategy-proof mechanism and the Rural Hospitals Theorem holds.*

Second, towards a converse of Proposition 2, we first show that if the Rural Hospitals Theorem holds, then there can be at most one stable-dominating mechanism that is strategy-proof. This follows from the Non-wasteful Structure Lemma and the Participation-equivalence Lemma.

Proposition 3. *Suppose that C is size monotonic and idempotent. If the Rural Hospitals Theorem holds, then there is at most one stable-dominating and strategy-proof mechanism.*

Third, under an additional restriction on \mathcal{P} , we show that if the Rural Hospitals Theorem holds, then a stable-dominating and strategy-proof mechanism Pareto-improves every stable mechanism. For each $i \in N$, \mathcal{P}_i is **closed under truncation** if for each $P_i \in \mathcal{P}_i$ and each $x \in X(i)$, there is $P'_i \in \mathcal{P}_i$ that ranks the same triples in $X(i)$ above x as P_i and ranks \emptyset immediately below x . That is, for each $y \in X(i)$, (1) $y P_i x$ if and only if $y P'_i x$ and (2) if $x P_i y$ then $x P'_i \emptyset P'_i y$. If, for each $i \in N$, \mathcal{P}_i is closed under truncation, then \mathcal{P} is closed under truncation.⁴³

Proposition 4. *Suppose that \mathcal{P} is closed under truncation. If the Rural Hospitals Theorem holds and φ is a stable-dominating and strategy-proof mechanism, then it Pareto-improves every stable mechanism.*

Propositions 3 and 4 together imply that, when the Rural Hospitals Theorem holds, if a mechanism is *stable* and strategy-proof, then it coincides with φ^{AOS} and is the unique stable-dominating and strategy-proof mechanism.

5.1.1 Recent Developments in Market Design

Having developed our main results in the general matching with contracts setting of Section 5.1, we consider their application to some recent developments in the literature on market design.

⁴² Hirata and Kasuya [2017] show that whenever φ^{AOS} exists, it is the only candidate for a stable and strategy-proof mechanism. Their proof, unlike ours, does not use the Rural Hospitals Theorem, which fails in their setting even when φ^{AOS} exists. Hatfield et al. [2015] show, in a very general setting, that the cumulative offer mechanism is the only candidate for a stable and strategy-proof mechanism.

⁴³ This is the same type of restriction that Sönmez [1999] imposes on preferences.

School choice with diversity constraints Public school districts are often concerned not only about parental preferences but also about the composition of the student body at each school, which they typically express using quotas and reserves. One can encode “soft” quotas and reserves in size monotonic and *substitutable* choice functions [Hafalir et al., 2013, Ehlers et al., 2014]. Since the student-proposing deferred acceptance mechanism is actually the agent-optimal stable mechanism, our results from Section 5.1 apply. The *school-proposing deferred acceptance mechanism* with these constructed choice functions always picks the stable matching that is closest to satisfying the diversity constraints amongst all stable allocations [Ehlers et al., 2014, Bó, 2016]. For some profiles, it violates strictly fewer constraints than φ^{AOS} , but at the cost of strategy-proofness. Corollary 2 answers a natural question: If we are willing to give up on stability in order to obtain fewer diversity constraint violations than φ^{AOS} , is there a strategy-proof mechanism that does at least as well as the benchmark of the school-proposing deferred acceptance mechanism in terms of student welfare? Since φ^{AOS} is strategy-proof and Pareto-improves the school-proposing deferred acceptance mechanism, by Corollary 2 it is the *only* candidate. The conclusion is that any attempt to maintain the incentives of students without increasing diversity constraint violations requires a redesign of the diversity objectives themselves.

Slot-specific priorities The *slot-specific priorities* framework [Kominers and Sönmez, 2016] is one way to handle applications where there are multiple policy objectives to reconcile.⁴⁴ In this framework, for capacity-based feasibility, each unit of an object can be thought of as a slot and these slots can be apportioned among the various objectives (or priorities). These priorities along with the *precedence order* over the slots—the order in which the slots are filled—generate choice functions. These do not always permit an agent-optimal stable mechanism, nor are they necessarily size monotonic. Nonetheless, Kominers and Sönmez [2016] show that the *cumulative offer* mechanism is stable and strategy-proof. Since the choice functions are not necessarily size monotonic, stability need not imply non-wastefulness. So Corollary 2 does not apply immediately. Yet, the proof technique that the authors use can be adapted to show that their mechanism is, indeed, non-wasteful. So Theorem 1 applies.⁴⁵ Therefore, the cumulative offer mech-

⁴⁴ Take, for instance, the US Army’s problem of assigning cadets to branches [Sönmez and Switzer, 2013, Sönmez, 2013]. There are two objectives to balance: prioritizing cadets on the basis of an “order-of-merit” list on one hand, and increasing the years of service by prioritizing cadets willing to serve longer terms on the other.

⁴⁵ Their proof associates each instance of their model with a representative instance of the matching with contracts model satisfying *substitutability* and size monotonicity. Since the cumulative offer mechanism corresponds to a stable mechanism in the representation, and the representation preserves the feasible set \mathcal{F} , the cumulative offer mechanism is non-wasteful in the original economy.

anism of [Kominers and Sönmez \[2016\]](#) is non-wasteful and is thus strategy-proofness-constrained Pareto-efficient.⁴⁶

Dynamic reserves Quotas and reserves discussed above are one way of accounting for policy objectives—in particular, diversity—in matching problems. An alternative approach is to target a specific distribution of the various types of agents among the objects, while specifying how to redistribute unused type-specific capacity when it is left over [[Westkamp, 2013](#), [Aygün and Turhan, 2016](#)]. These models are similar to the model with slot-specific priorities, except that each slot’s capacity is endogenous. If there is only one term under which an agent may be matched, under certain conditions on how slots’s capacities are redistributed, the choice function associated with each object is substitutable and size monotonic [[Westkamp, 2013](#)]. Since this ensures existence of φ^{AOS} , [Corollary 2](#) says that it is the unique stable-dominating and strategy-proof mechanism. If there are multiple terms, under similar conditions on capacity redistribution, the cumulative offer mechanism is stable and strategy-proof [[Aygün and Turhan, 2016](#)], even though φ^{AOS} need not exist. As with the discussion of slot-specific priorities, [Theorem 1](#) implies that it is strategy-proofness-constrained Pareto-efficient.

Distributional constraints In some applications, there are distributional constraints on allocations that are across objects, in addition to object specific constraints. An example of such constraints features prominently in the way graduating medical students are matched to residencies in Japan [[Kamada and Kojima, 2015](#)]: the country is divided into various regions and the number of residents that can be matched to hospitals in each region is capped independently of the actual number of positions at each hospital. [Kamada and Kojima](#) define a new strategy-proof and individually rational mechanism called *flexible deferred acceptance* and show that it Pareto-improves the strategy-proof mechanism that is currently in use, without violating the caps. They also consider the consequences of varying the regional caps: tightening the cap for at least one region without loosening any of the others leads to a Pareto-worsening in the allocation made by the flexible deferred acceptance mechanism. Thinking of flexible deferred acceptance with different profiles of caps as different mechanisms, [Theorem 2](#) thus tells us that such a tightening of the caps leads, for every profile of preferences, to an expansion (in terms of set inclu-

⁴⁶ [Hatfield and Kominers \[2014\]](#) provide *completeness* conditions on choice functions that ensure that the cumulative offer mechanism is stable and strategy-proof. Their proof technique is similar to that of [Kominers and Sönmez \[2016\]](#). So by an argument similar to the one above, if choice functions are completable to being substitutable and size monotonic, the cumulative offer mechanism is strategy-proofness-constrained Pareto-efficient.

sion) of the set of unmatched residents. Indeed, it tells us that this expansion is strict for some profiles. This has an important implication for the *match rate*: the proportion of medical students matched to *some* hospital. If the prior distribution over preference profiles, from a policy maker’s point of view, has full support, then Theorem 2 allows us to conclude that tightening the caps leads to a *strict* reduction in the expected match rate. The purpose of these caps is to increase the number of residents matched to rural areas by tightening the caps on other regions. So the above reasoning says that tightening the caps in a way that increases the expected number of residents matched to a rural area is accompanied by a *greater* decrease in the overall expected number of residents matched.

5.1.2 School Choice

We consider here the more specialized school choice model, which is the classical object allocation model augmented with weak priorities. Recall that in this model, T is a singleton and \mathcal{F} is capacity-based. Additionally, each $o \in O$ is associated with a **priority** over N , denoted by \succsim_o , which is a complete, transitive, and reflexive relation. Let $\succsim \equiv (\succsim_o)_{o \in O}$.

Given priorities, \succsim , for each $o \in O$, we define C_o^{\succsim} as follows. For each $Y \subseteq N$,

$$C_o^{\succsim}(Y) \equiv \begin{cases} \{Y\} & \text{if } |Y| \leq q_o, \\ \{Z \subseteq Y : |Z| = q_o \text{ and for each } i \in Z \text{ and each } j \in Y \setminus Z, i \succsim_o j\} & \text{otherwise.} \end{cases}$$

That is, for each subset of agents, if it contains no more than q_o elements, then the entire set is the only one that is chosen. If not, then all subsets that contain exactly q_o elements are chosen, except for ones that include agents of *strictly* lower priority than an excluded agent. This C_o^{\succsim} is size monotonic and idempotent.

Suppose that an agent prefers a particular object o to the one that he is assigned. Under the interpretation of priorities as consumption “rights”, if o is assigned to someone else who has *strictly* lower priority, then he has the right to protest this allocation. For each $\mu \in \mathcal{F}$, μ **respects priorities** if no agent can protest on such grounds. That is, there is no pair $i, j \in N$ and $o \in O$ such that $o P_j \mu(j)$, $\mu(i) = o$, and $j \succ_o i$.

Stability and stable-domination as fairness Interpreting respect for the priorities as a fairness constraint [Balinski and Sönmez, 1999, Abdulkadiroğlu and Sönmez, 2003], we are interested in mechanisms that are individually rational, non-wasteful, and fair. Using the choice correspondence for each object defined from these priorities, respect for priorities alongside individual rationality and non-wastefulness is equivalent to stability

with respect to these correspondences.

Remark 2. *For each profile of preferences, an allocation is stable with respect to C^{\succsim} if and only if it is individually rational, non-wasteful, and respects \succsim .*

Remark 2 says that, among individually rational and non-wasteful mechanisms, the requirement that a mechanism or allocation respect priorities is equivalent to the requirement that it be stable. Since they are equivalent, we speak of an allocation being stable rather than it respecting priorities and being individually rational and non-wasteful.

When priorities are strict (that is, they contain no ties between agents), φ^{AOS} exists [Gale and Shapley, 1962]. However, when priorities contain ties, there may not exist a single stable allocation that Pareto-improves every other stable allocation. Since φ^{AOS} may not exist, a common approach to handling weak priorities is to use a tie breaker to form strict priorities. Let $\tau \equiv (\tau_o)_{o \in O}$ be a profile of linear orders over N , one for each object. For each such τ , let \succsim^τ be the priorities **tie broken by τ** . That is, for each distinct pair $i, j \in N$, $i \succ_o^\tau j$ if either (1) $i \succ_o j$ or (2) $i \sim_o j$ and $i \tau_o j$. Let \mathcal{T} be the set of all profiles of tie breakers. Since there is an agent-optimal stable mechanism for strict priorities, given \succsim and $\tau \in \mathcal{T}$, we define the agent-optimal stable mechanism for the priorities tie broken by τ as $\varphi^{AOS\tau}$. These are the mechanisms studied by Abdulkadiroğlu et al. [2009].⁴⁷

Ergin [2002] shows that, for strict priorities, unless they satisfy a restrictive condition that he calls acyclicity, stability and efficiency are at odds. That is, unless priorities are acyclic, φ^{AOS} is not Pareto-efficient. When priorities are weak, it is therefore clear that arbitrarily breaking ties could cause Pareto-inefficiency. In fact, no matter how ties are broken, the agent-optimal stable mechanism with tie broken priorities may select an allocation that is Pareto-improvable by another stable allocation [Erdil and Ergin, 2008]. Furthermore, there are examples of priorities that permit a stable, group strategy-proof, and Pareto-efficient mechanism, while for no $\tau \in \mathcal{T}$ is $\varphi^{AOS\tau}$ Pareto-efficient [Ehlers and Erdil, 2010].

Since $\varphi^{AOS\tau}$ may not be Pareto-efficient, Abdulkadiroğlu et al. [2009] consider the following question: for each $\tau \in \mathcal{T}$, is it possible to find a strategy-proof mechanism that Pareto-improves $\varphi^{AOS\tau}$? They show that the answer is negative. However, even if the answer were positive, the allocations chosen by the Pareto-improving mechanism would not be stable. The reason such a mechanism would pass muster here is that it Pareto-improves a *stable* mechanism: $\varphi^{AOS\tau}$. But as we have explained above, there is nothing special about $\varphi^{AOS\tau}$, other than strategy-proofness, when priorities are weak. In fact, they may even select allocations that are Pareto-improved by other stable allocations.

⁴⁷ These $\varphi^{AOS\tau}$ correspond to what Abdulkadiroğlu et al. [2009] call deferred acceptance with multiple tie-breaking.

Therefore, the exercise is to justify the choice of a mechanism on the grounds that it Pareto-improves *some* stable allocation at each profile of preferences. Then if an agent were to protest the violation of his priority at some object, we offer a move to the Pareto-improved stable allocation so this protest would be moot: the agent would not be better off at *this* stable allocation.⁴⁸ That is, by requiring the chosen mechanism to just be *stable-dominating* rather than stable, we may enlarge the options for a strategy-proof Pareto-improvement.

For strict priorities, since φ^{AOS} exists, Proposition 2 implies the following, which is stronger than the known result that φ^{AOS} is the only stable and strategy-proof mechanism [Alcalde and Barberà, 1994].

Corollary 3. *If \succsim consists solely of strict priorities, then φ^{AOS} is the unique stable-dominating and strategy-proof mechanism.*

On the other hand, for weak priorities, there may be more than one stable-dominating and strategy-proof mechanism. Nevertheless, Proposition 1 yields the following corollary.

Corollary 4. *No stable-dominating and strategy-proof mechanism, including, for each $\tau \in \mathcal{T}$, $\varphi^{AOS\tau}$, is Pareto-improved by any other strategy-proof mechanism. Further, no two stable-dominating and strategy-proof mechanisms are Pareto-connected.*

While, for each $\tau \in \mathcal{T}$, $\varphi^{AOS\tau}$ is strategy-proofness-constrained Pareto-efficient, $\varphi^{AOS\tau}$ s are not the only stable and strategy-proof mechanisms, as Ehlers and Erdil [2010] have shown. Corollary 4 extends the result of Abdulkadiroğlu et al. [2009] to all of these.

We present further results on stable-dominating mechanisms in Appendix C.

Beyond stability as fairness While we have focused on stability and stable-domination as notions of fairness, our results allow us to draw conclusions about other fairness concepts as well. Take for instance the *legal set* of allocations under strict priorities [Morrill, 2016], where only harmful and *redressable* priority violations are ruled out.⁴⁹ The legal set includes the stable set and always has a Pareto-worst member. Consequently, Theorem 1 implies that φ^{AOS} is the unique strategy-proof selection from the legal set.

Corollary 5. *If \succsim consists solely of strict priorities, then φ^{AOS} is the unique strategy-proof mechanism selecting from the legal set.*

⁴⁸ See Dur et al. [2015] for recent work on algorithms that select from the stable-dominating set in a way that only certain priorities are violated.

⁴⁹ We refer the reader to Morrill [2016] for a formal definition.

5.2 Reallocating Objects From an Endowment

Consider the model of [Shapley and Scarf \[1974\]](#). There is a classical object allocation model where $|O| = |N|$ and for each $o \in O$, $q_o = 1$. Additionally, there is a reference allocation ω , where $\omega(i)$ is agent i 's **endowment**.

The literature following [Shapley and Scarf \[1974\]](#) focuses on the subdomain of strict preferences profiles such that each agent ranks each object above \emptyset . We denote this subdomain by $\underline{\mathcal{P}}$. It also takes individual rationality to mean that each agent receives an object that he finds at least as desirable as his endowment. In this model, for each $P \in \underline{\mathcal{P}}$, the *core* consists of a single allocation [[Roth and Postlewaite, 1977](#)]. The core mechanism, which selects this unique allocation, is actually the only mechanism that is strategy-proof, Pareto-efficient, and individually rational in the sense of Pareto-improving ω [[Ma, 1994](#), [Sönmez, 1999](#)].

On the subdomain $\underline{\mathcal{P}}$, since \emptyset is at the bottom of each preference relation, individual rationality according to our definition is vacuous: every allocation Pareto-improves \emptyset . Since $|O| = |N|$, the set of all permutations of the endowment across agents is non-wasteful. Thus, on $\underline{\mathcal{P}}$, the **no-trade** mechanism, which always selects ω , is strategy-proof, non-wasteful, and individually rational yet the strategy-proof core mechanism strictly Pareto-improves it. This does not contradict [Theorem 1](#) because $\underline{\mathcal{P}}$ violates the richness of outside options.

We expand our focus to a domain with rich outside options. That is, where an agent may rank \emptyset above any of the other objects, including his own endowment. On this larger domain, Pareto-efficiency implies that each agent considers his assignment at least as good as \emptyset —otherwise, disposing of the object leads to a Pareto-improvement. For this domain, there are many strategy-proof and Pareto-efficient mechanism that Pareto-improve ω .⁵⁰ For the remainder of this section, we take individual rationality to mean Pareto-improving both \emptyset as well as ω . As we show below, there are many rules that are strategy-proof, individually rational, and participation maximal, and therefore strategy-proofness-constrained Pareto-efficient, but not fully Pareto-efficient.

Because the no-trade benchmark may not Pareto-improve \emptyset on \mathcal{P} , it is not individually rational. A better benchmark would allow an agent to discard his endowment. Of course, this could be wasteful: i 's trash may be j 's treasure. Often, in such situations, there are institutional rules that determine how these discarded resources are distributed

⁵⁰ For instance, consider the *hierarchical exchange* mechanisms [[Pápai, 2000](#)] where each agent is at the root of his endowment's inheritance tree, modified by including branches labeled by \emptyset , and where agents are allowed to point to \emptyset . See [Pycia and Ünver \[2016\]](#) for a more detailed discussion.

among other agents. For instance, fantasy sports leagues have “waivers”⁵¹ systems where a waiver order over the “fantasy owners” is fixed beforehand and any players that the owners discard (or “waive”) can be picked up by the remaining owners in this order. In other situations—for instance, in the workplace—seniority serves such a purpose.

Suppose that we have a linear order over the agents, \triangleright , along with an endowment, ω . The process described above can be formalized as Algorithm 1 in Appendix E. It takes the endowment, preference profile, and order as arguments and returns an individually rational and non-wasteful allocation that Pareto-improves the endowment. If the status quo is to use such a procedure, then any change to a new mechanism would have to Pareto-improve it. Thus, we have a benchmark mechanism $\underline{\varphi}^{\omega, \triangleright}$.

It turns out that there *is* a strategy-proof mechanism, φ , that Pareto-improves this benchmark mechanism. We define it as follows. For each $o \in O$, let \succsim_o be a linear order over N that agrees with \triangleright on all agents but $\omega(o)$ and ranks $\omega(o)$ above all other agents. Consider the school choice problem associated with \succsim . For each $P \in \mathcal{P}$, let $\varphi(P)$ be the agent-optimal stable allocation with regards to \succsim .

Since $\underline{\varphi}^{\omega, \triangleright}$ can be computed using the *object-proposing deferred acceptance* algorithm with regards to \succsim , φ Pareto-improves it. However, $\underline{\varphi}^{\omega, \triangleright}$ is individually rational and non-wasteful. So Theorem 1 says that φ is actually the *only* strategy-proof mechanism that Pareto-improves $\underline{\varphi}^{\omega, \triangleright}$. But these \succsim do not satisfy acyclicity [Ergin, 2002], so φ is not Pareto-efficient. From this and Theorem 1, we have the following corollary.

Corollary 6. *No Pareto-efficient and strategy-proof mechanism Pareto-improves $\underline{\varphi}^{\omega, \triangleright}$.*

Having extended the lower bound on agents’ welfare from the no-trade mechanism to $\underline{\varphi}^{\omega, \triangleright}$, we ended up with an impossibility result on \mathcal{P} , which is in contrast with the results on Pareto-improving ω for $\underline{\mathcal{P}}$.

Corollary 6 shows a novel type of application of Theorem 1: by showing the existence of a strategy-proof but Pareto-*inefficient* mechanism that Pareto-improves the benchmark, we establish the impossibility of a Pareto-efficient strategy-proof improvement.

5.3 Excludable Public Goods with Voluntary Participation

In this section, we show how to accommodate excludable public goods in the general framework through a simple example.

Suppose a public facility is to be located on the interval $[0, 1]$ and a set of users chosen. The set of agents is partitioned into two sets: agents in N_L live at 0 and agents in N_R live

⁵¹ See, for instance, <http://www.espn.com/fantasy/football/ffl/story?page=fflruleswaiverwalk>.

at 1. Each $i \in N$ prefers to have the facility located as close to his own home as possible. Further, he is unwilling to travel beyond a certain **threshold** $t_i \in [0, 1]$. That is, if i lives at 0, then he is unwilling to travel to the right of t_i and if he lives at 1, then he is unwilling to travel to the left of t_i . Since we know exactly how each agent ranks the locations, the only private information for i is t_i . So a preference profile is identified by the threshold profile t . To satisfy no indifference with \emptyset , each $i \in N$ prefers to travel to t_i and enjoy the facility over opting out. However, he prefers to opt out rather than travel beyond t_i . An *allocation* consists of two parts: the location of a public facility in the interval $[0, 1]$ and the set of users. That is, $\mathcal{F} \equiv [0, 1] \times 2^N$. A mechanism is **dictatorial** if it ignores the thresholds and always locates the facility at 0 or always locates the facility at 1.

Even in this stark model, if we insist on strategy-proofness, individually rationality, and Pareto-efficiency, then the only two options are the dictatorial mechanisms.⁵² Thus, the requirement of Pareto-efficiency alongside strategy-proofness and individual rationality precludes the possibility that any compromise is ever reached.

If we weaken this requirement to strategy-proofness-constrained Pareto-efficiency, compromises are possible. To show this, we define a mechanism that selects a compromise at some threshold profiles and yet is strategy-proof, individually rational, and participation-maximal: Given a candidate for the compromise $x \in (0, 1)$, consider the mechanism that offers x to N_R . If they *unanimously* accept x , then the mechanism selects the location x . Otherwise, it selects the location 0. The set of users is the set of agents willing to travel to the chosen location. This mechanism is *participation-maximal*: when the location is x , every agent in N_R participates. Any agent in N_L who does not participate would only participate if the location were moved to the left, which would make every member of N_R worse off. On the other hand, if the location is 0, then any other location would make members of N_L worse off. Though it is not Pareto-efficient, by Theorem 1, this mechanism is strategy-proofness-constrained Pareto-efficient.

5.4 Transferable Utility

Here we consider the problem of making a social decision and assigning payments to agents. We assume that preferences are quasilinear in payments. Consequently, there is indifference with \emptyset . Nonetheless, the Participation-equivalence Lemma applies. Below we study its implications.

Suppose \mathcal{D} is a set of **social decisions**. At each $\delta \in \mathcal{D}$, the agents participating in δ are $N(\delta)$. Let $\mathcal{T} \subseteq \mathbb{R}^N$ be the set of **possible payment profiles**. At each $\tau \in \mathcal{T}$, for each $i \in N$,

⁵²See Appendix A.3 for a proof of this statement.

τ_i is the payment that i makes. Like \mathcal{F}_i , let \mathcal{D}_i be the set of decisions that i participates in.

To fit this into our general model, an allocation in \mathcal{F} is a pair $(\delta, \tau) \in \mathcal{D} \times \mathcal{T}$. Further, $N(\delta, \tau) = N(\delta)$. Since a non-participant consumes his outside option, we require that for each $i \notin N(\delta)$, $\tau_i = 0$ —we relax this later.

We assume that each agent’s preferences are quasilinear in the payment that he makes. Thus, for each $i \in N$, each $R_i \in \mathcal{R}_i$ is identified by a **valuation** $v_i \in \mathbb{R}^{\mathcal{D}_i \cup \{\emptyset\}}$. Let i ’s **valuation space** \mathcal{V}_i be the set of all possible valuations for i . Thus, i ’s preference relation is represented by $v_i(\delta) - \tau_i$, where $v_i(\delta)$ is the δ th coordinate of v_i if $i \in N(\delta)$ and the \emptyset th coordinate otherwise. The set of all **valuation profiles** is $\mathcal{V} \equiv \times_{i \in N} \mathcal{V}_i$.

A mechanism in this context consists of two parts: a decision rule, $d : \mathcal{V} \rightarrow \mathcal{D}$, and a payment rule, $t : \mathcal{V} \rightarrow \mathcal{T}$. If (d, t) is strategy-proof, we say that t **implements** d . If there is a payment rule that implements d , then d is **implementable**. As defined, implementation says nothing about individual rationality. We define parallel concepts with this requirement added: t **IR-implements** d if (d, t) is not only strategy-proof but also individually rational. In this case, we say that d is **IR-implementable**. If there is a unique payment rule that IR-implements d , we say that d is **uniquely IR-implementable**.

Richness of outside options The richness condition that we described in Section 2 was for preference relations over $\mathcal{F}_i \cup \{\emptyset\}$. This condition would be met if, for instance, \mathcal{V}_i were such that for each $v_i \in \mathcal{V}_i$ there were $v'_i \in \mathcal{V}_i$, such that the valuations of each of the decisions were unchanged, but the valuation of the outside option were increased. That is, for each $\kappa > v_i(\emptyset)$, there is v'_i such that $v'_i(\emptyset) = \kappa$ and for each $\delta \in \mathcal{D}_i$, $v'_i(\delta) = v_i(\delta)$.⁵³

With these concepts in hand, we have the following corollary of the Participation-equivalence [Lemma](#):

Corollary 7. *Every IR-implementable decision rule is uniquely IR-implementable.*

In fact we can say more than Corollary 7. For any pair of decision rules d and d' that are participation-equivalent, if t IR-implements d and t' IR-implements d' , then (d, t) and (d', t') are welfare-equivalent.

The decisions in \mathcal{D} , since they may exclude some of the agents, may be thought of as excludable public goods. An **efficient** decision rule maximizes the relative value of the decision to its participants. Given $v \in \mathcal{V}$, denote, for each $i \in N$ and each $\delta \in \mathcal{D}$, by $u_i^v(\delta)$ value of δ to i relative to being excluded. That is, $u_i^v(\delta) \equiv v_i(\delta) - v_i(\emptyset)$. Of

⁵³ We could state this with a bound on κ if there were a bound on valuations of the decisions in \mathcal{D}_i and on the payments that i may make in \mathcal{T} .

course, if $i \notin N(\delta)$, this means that $u_i^v(\delta) = 0$. Interpreting $v_i(\emptyset)$ as i 's opportunity cost of participating, $u_i^v(\delta)$ is i 's valuation of δ net of this cost.

A decision rule d is efficient if, for each $v \in \mathcal{V}$,

$$d(v) \in \operatorname{argmax}_{\delta \in \mathcal{D}} \sum_{i \in N} u_i^v(\delta).$$

In a context where the status quo is that each agent enjoys his outside option, an efficient decision maximizes the total surplus over the status quo.

A **pivotal** payment rule is one that assigns a payment to each agent equal to the externality that his participation imposes on the other agents. That is, given an efficient decision rule d , t is a pivotal payment rule if, for each $v \in \mathcal{V}$ and each $i \in N$,

$$t_i(v) \equiv \max_{\delta \in \mathcal{D} \setminus \mathcal{D}_i} \left\{ \sum_{j \neq i} u_j^v(\delta) \right\} - \sum_{j \neq i} u_j^v(d(v)).$$

A pivotal payment rule is a particular instance of a Groves scheme [Groves, 1973]. So, if d is an efficient decision rule and t is a corresponding pivotal payment rule, then (d, t) is strategy-proof. For each $v \in \mathcal{V}$ and each $i \in N$, if $i \notin N(d(v))$, then t prescribes a zero payment for i . Therefore, (d, t) is feasible. Finally, the efficiency of d ensures that (d, t) is also individually rational. Thus, we have the following corollary.

Corollary 8. *If d is efficient, then t IR-implements d if and only if it is a pivotal payment rule.*

When we have neither the restriction that non-participants receive zero payments nor the requirement of individual rationality, Groves schemes are the only ones to implement efficient decision rules [Green and Laffont, 1977, Holmström, 1979]. For private goods problems, pivotal rules are the only Groves schemes that are individually rational without making payments to non-participants [Chew and Serizawa, 2007]. For pure public goods, the counterpart of individual rationality requires that agents have no incentive to free-ride. Substituting this property for individual rationality similarly characterizes pivot rules [Moulin, 1986]. Corollary 8 neither implies nor is implied by existing characterizations of the pivotal rule since our assumption that \mathcal{V} satisfy richness of outside options is logically independent from the domain assumptions made in existing results.

Corollary 7 is similar to revenue equivalence results which pin down the payment rule for each agent, up to a function of others' reported valuations [Holmström, 1979, Chung and Olszewski, 2007, Heydenreich et al., 2009]. Unlike these results, it pins down

a unique payment rule. This uniqueness is driven by individual rationality and the requirement that no payments be assigned to non-participants. We make this assumption here in order to invoke the Participation-equivalence [Lemma](#). While the assumption is reasonable for many practical settings, we next relax it and show a revenue equivalence result.

Allowing payments to non-participants Even if non-participants may make a payment, the requirement that agents must participate voluntarily rather than exercising their outside option is unchanged. So the definition of individual rationality is unchanged: $(\delta, \tau) \in \mathcal{F}$ is individually rational if for each $i \in N$, $v_i(\delta) - \tau_i \geq v_i(\emptyset)$. Of course, this implies that a non-participant never make a *positive* payment. That is, if (δ, τ) is individually rational, then for each $i \notin N(\delta)$, $\tau_i \leq 0$.

We now define revenue equivalence as a property of a decision rule. **Revenue equivalence** holds for d if two payment rules that both implement d (individually rationally or not) differ, for each agent, by a function of others' valuations. That is, if t and t' both implement d , then for each $i \in N$ there is a function $h_i : \mathcal{V}_{-i} \rightarrow \mathbb{R}$ such that for each $v \in \mathcal{V}$, $t_i(v) - t'_i(v) = h_i(v_{-i})$.

Revenue equivalence is not guaranteed for every implementable decision rule. However, [Proposition 5](#) states that it *does* hold for every IR-implementable decision rule.

Proposition 5. *Revenue equivalence holds for every IR-implementable decision rule.*

[Chung and Olszewski \[2007\]](#) provide a sufficient (and in some cases necessary) condition on the valuation space to ensure revenue equivalence for every implementable decision rule. Even when their condition is violated, \mathcal{V} may still satisfy richness of outside options. Thus, [Proposition 5](#), by strengthening the hypothesis to IR-implementability, still ensures revenue equivalence. On the other hand, [Heydenreich et al. \[2009\]](#) provide a joint condition on the decision rule and the valuation space that is necessary and sufficient for revenue equivalence to hold. [Proposition 5](#) says that individual rationality and richness of outside options, which are easy to check, are sufficient for their condition to hold.

6 Conclusion

We proposed a general framework where agents have rich outside options and each allocation relies on the participation of a particular set of agents. Individually rational allocations are those where no participant has an incentive to walk away when it is chosen.

We introduced the property of participation-maximality, an important consideration in many applications. If an allocation with this property admits an enlargement of the set of participants, this is at the cost of harming some agent. For instance, in the problem of school choice, it would be unacceptable to leave students unmatched while seats at desirable schools are unfilled.

Taking individual rationality and participation-maximality as baseline requirements, we considered the set of strategy-proof mechanisms. For our first result (Theorem 1), we showed that there is at most one strategy-proof mechanism that meets or exceeds, in the Pareto sense, the welfare level provided by a given individually rational and participation-maximal benchmark mechanism. Thus, given such a benchmark, the requirement of Pareto-improving it in a strategy-proof way eliminates all degrees of freedom from the design problem.

Digging deeper into the relationship between participation and welfare, we showed that a Pareto comparison between two strategy-proof and individually rational mechanisms amounts to comparing the sets of participants (Theorem 2). One mechanism strictly Pareto-improves another if and only if it strictly expands the set of participants at some preference profiles and leaves them unchanged at the others. By connecting the welfare ordering of strategy-proof and individually rational mechanisms with the participation-expansion ordering, we identified that participation-maximality, with individual rationality, is a sufficient condition for such a mechanism to be second best.

To show these results, we relied on the Participation-equivalence [Lemma](#), which says that the only freedom a designer has is in the selection of the participants at each profile of preferences when choosing a strategy-proof and individually rational mechanism—once that is done, the entire mechanism is pinned down in welfare terms. Interestingly, this has consequences for the problem of making a social decision in the presence of transfers and quasilinear preferences. For such problems, a mechanism can be decomposed into an decision rule and a payment rule. The decision rule determines the set of participants independently of the payment rule. Adapting the argument of the lemma, we showed that any two payment rules that each combine with a given decision rule to form a strategy-proof and individually rational mechanism are identical, up to the payments of non-participants (Proposition 5).

As noted by [Compte and Jehiel \[2007, 2009\]](#), outside options have strong implications even when strategy-proofness is weakened to Bayesian incentive compatibility. We leave for future research an extension of the ideas in this paper to Bayesian implementation.

References

- A. Abdulkadiroğlu, P. A. Pathak, and A. E. Roth. The New York City high school match. *American Economic Review*, 95(2):364–367, May 2005. [1]
- A. Abdulkadiroğlu, P. A. Pathak, A. E. Roth, and T. Sönmez. The Boston public school match. *American Economic Review*, 95(2):368–371, May 2005. [1]
- A. Abdulkadiroğlu, P. A. Pathak, and A. E. Roth. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match. *American Economic Review*, 99(5):1954–1978, December 2009. [3], [6], [15], [24], [25], [52]
- A. Abdulkadiroğlu and T. Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, June 2003. [2], [23], [53]
- J. Alcalde and S. Barberà. Top dominance and the possibility of strategy-proof stable solutions to matching problems. *Economic Theory*, 4(3):417–35, May 1994. [25]
- A. Alkan. A class of multipartner matching markets with a strong lattice structure. *Economic Theory*, 19(4):737–746, 2002. [17]
- A. Alkan and D. Gale. Stable schedule matching under revealed preference. *Journal of Economic Theory*, 112(2):289–306, 2003. [5], [17]
- S. Alva. WARP and combinatorial choice. Working paper, University of Texas at San Antonio, 2016. [17]
- H. Anno and M. Kurino. On the operation of multiple matching markets. *Games and Economic Behavior*, 100:166–185, 2016. [7]
- H. Anno and H. Sasaki. Second-best efficiency of allocation rules: Strategy-proofness and single-peaked preferences with multiple commodities. *Economic Theory*, 54(3):693–716, 2013. [7]
- O. Aygün and T. Sönmez. Matching with contracts: Comment. *American Economic Review*, 103(5):2050–51, August 2013. [5], [7], [17]
- O. Aygün and B. Turhan. Dynamic reserves in matching markets: Theory and applications. Working paper, ITAM, 2016. [22]
- M. Balinski and T. Sönmez. A tale of two mechanisms: Student placement. *Journal of Economic Theory*, 84(1):73–94, January 1999. [2], [3], [17], [23], [47], [48]
- I. Bó. Fair implementation of diversity in school choice. *Games and Economic Behavior*, 97:54–63, 2016. [21]
- A. Bogomolnaia, R. Deb, and L. Ehlers. Strategy-proof assignment on the full preference domain. *Journal of Economic Theory*, 123(2):161–186, Aug 2005. [15]

- D. Cantala. Choosing the level of a public good when agents have an outside option. *Social Choice and Welfare*, 22(3):491–514, 06 2004. [2]
- S. H. Chew and S. Serizawa. Characterizing the Vickrey combinatorial auction by induction. *Economic Theory*, 33:393–406, 2007. [7], [30]
- K.-S. Chung and W. Olszewski. A non-differentiable approach to revenue equivalence. *Theoretical Economics*, 2(4):469–487, 2007. [4], [30], [31]
- E. H. Clarke. Multipart pricing of public goods. *Public Choice*, 11(1):17–33, 1971. [6]
- O. Compte and P. Jehiel. On quitting rights in mechanism design. *American Economic Review, Papers and Proceedings*, 97(2):137–141, May 2007. [8], [32]
- O. Compte and P. Jehiel. Veto constraint in mechanism design: Inefficiency with correlated types. *American Economic Journal: Microeconomics*, 1(1):182–206, February 2009. [8], [32]
- P. C. Cramton. Money Out of Thin Air: The Nationwide Narrowband PCS Auction. *Journal of Economics & Management Strategy*, 4(2):267–343, Summer 1995. [2]
- L. E. Dubins and D. A. Freedman. Machiavelli and the Gale-Shapley algorithm. *American Mathematical Monthly*, 88(7):485–494, 1981. [3]
- U. Dur, A. A. Gitmez, and O. Yilmaz. School choice under partial fairness. Working paper, North Carolina State University, 2015. [25]
- L. Ehlers and A. Erdil. Efficient assignment respecting priorities. *Journal of Economic Theory*, 145(3):1269–1282, May 2010. [18], [24], [25], [50], [52]
- L. Ehlers and B. Klaus. Strategy-proofness makes the difference: Deferred-acceptance with responsive priorities. *Mathematics of Operations Research*, 39(4):949–966, November 2014. [17]
- L. Ehlers and A. Westkamp. Strategy-proof tie-breaking. Working paper, CIREQ, Feb 2016. [50], [53]
- L. Ehlers, I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim. School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory*, 153: 648 – 683, 2014. [21]
- A. Erdil. Strategy-proof stochastic assignment. *Journal of Economic Theory*, 151:146 – 162, 2014. [7]
- A. Erdil and H. I. Ergin. What’s the matter with tie-breaking? Improving efficiency in school choice. *American Economic Review*, 98(3):669–89, June 2008. [15], [24], [52]
- A. Erdil and H. I. Ergin. Two-sided matching with indifferences. Working paper, University of California, Berkeley, 2015. [9], [15]

- A. Erdil and T. Kumano. Prioritizing diversity in school choice. Working paper, University of Cambridge, 2014. [15]
- H. I. Ergin. Efficient resource allocation on the basis of priorities. *Econometrica*, 70(6): 2489–2497, 2002. [24], [27], [50], [53]
- T. Fleiner. A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research*, 28:103–126, Feb 2003. [17]
- D. Gale and L. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69:9–15, 1962. [3], [24]
- M. Goto, F. Kojima, R. Kurata, A. Tamura, and M. Yokoo. Designing matching mechanisms under general distributional constraints. *American Economic Journal: Microeconomics*, forthcoming. [10], [18]
- J. Green and J.-J. Laffont. Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica*, 45:417–438, 1977. [30]
- T. Groves. Incentives in teams. *Econometrica*, 41(4):617–631, 1973. [6], [30]
- I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim. Effective affirmative action in school choice. *Theoretical Economics*, 8(2):325–363, 2013. [21]
- X. Han. Priority-augmented house allocation. Working paper, Southern Methodist University, 2015. [53]
- J. W. Hatfield and F. Kojima. Substitutes and stability for matching with contracts. *Journal of Economic Theory*, 145(5):1704–1723, September 2010. [19]
- J. W. Hatfield and S. D. Kominers. Hidden substitutes. Working paper, University of Texas at Austin, 2014. [18], [22]
- J. W. Hatfield and P. Milgrom. Matching with contracts. *American Economic Review*, 95(4):913–935, September 2005. [2], [5], [9], [15], [17], [43], [55]
- J. W. Hatfield, S. D. Kominers, and A. Westkamp. Stability, strategy-proofness, and cumulative offer mechanisms. Working paper, University of Texas at Austin, 2015. [20]
- B. Heydenreich, R. Müller, M. Uetz, and R. V. Vohra. Characterization of revenue equivalence. *Econometrica*, 77(1):307–316, 2009. [4], [8], [30], [31]
- D. Hirata and Y. Kasuya. On stable and strategy-proof rules in matching markets with contracts. *Journal of Economic Theory*, 168:27–43, March 2017. [7], [18], [20]
- B. Holmström. Groves’ scheme on restricted domains. *Econometrica*, 47(5):1137–1144, 1979. [4], [7], [30]
- A. Hylland and R. Zeckhauser. The efficient allocation of individuals to positions. *Journal of Political Economy*, 87(2):293–314, April 1979. [2]

- M. O. Jackson and T. R. Palfrey. Voluntary implementation. *Journal of Economic Theory*, 98(1):1–25, May 2001. [8]
- B. Jullien. Participation constraints in adverse selection models. *Journal of Economic Theory*, 93(1):1–47, July 2000. [8]
- Y. Kamada and F. Kojima. Efficient matching under distributional constraints: Theory and applications. *American Economic Review*, 105(1):67–99, January 2015. [3], [10], [15], [18], [22]
- O. Kesten. On two competing mechanisms for priority-based allocation problems. *Journal of Economic Theory*, 127(1):155–171, March 2006. [53]
- O. Kesten. School choice with consent. *Quarterly Journal of Economics*, 125(3):1297–1348, August 2010. [3], [6]
- O. Kesten and M. Kurino. Do outside options matter in matching? A new perspective on the trade-offs in school choice. Working paper, Carnegie Mellon University, 2015. [6]
- S. D. Kominers and T. Sönmez. Matching with slot-specific priorities: Theory. *Theoretical Economics*, 11(2):683–710, 2016. [5], [21], [22]
- V. Krishna and E. Maenner. Convex potentials with an application to mechanism design. *Econometrica*, 69(4):1113–1119, 2001. [4]
- J. Ma. Strategy-proofness and the strict core in a market with indivisibilities. *International Journal of Game Theory*, 23(1):75–83, 1994. [26]
- R. P. McAfee and J. McMillan. Analyzing the airwaves auction. *The Journal of Economic Perspectives*, 10(1):159–175, 1996. [2]
- P. Milgrom. Putting auction theory to work: The simultaneous ascending auction. *Journal of Political Economy*, 108(2):245–272, 2000. [2]
- T. Morrill. Which school assignments are legal? Working paper, North Carolina State University, 2016. [3], [25]
- H. Moulin. Characterizations of the pivotal mechanism. *Journal of Public Economics*, 31: 53–78, 1986. [7], [30]
- R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. [4]
- S. Ohseto. Characterizations of strategy-proof and fair mechanisms for allocating indivisible goods. *Economic Theory*, 29:111–121, 2006. [7]
- S. Pápai. Strategyproof assignment by hierarchical exchange. *Econometrica*, 68(6):1403–1434, November 2000. [26]

- M. Pycia and M. U. Ünver. Outside options in neutral allocation of discrete resources. Working paper, Boston College, 2016. [26]
- M. Pycia and M. U. Ünver. Incentive compatible allocation and exchange of discrete resources. *Theoretical Economics*, 12(1):287–329, January 2017. [53]
- A. E. Roth. The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy*, 92(6):991–1016, December 1984. [5]
- A. E. Roth. On the allocation of residents to rural hospitals: A general property of two-sided matching markets. *Econometrica*, 54(2):425–427, 1986. [3], [5], [13]
- A. E. Roth and E. Peranson. The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American Economic Review*, 89(4):748–780, September 1999. [1]
- A. E. Roth and A. Postlewaite. Weak versus strong domination in a market with indivisible goods. *Journal of Mathematical Economics*, 4(2):131–137, August 1977. [26]
- A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange. *Quarterly Journal of Economics*, 119(2):457–488, 2004. [2]
- A. E. Roth, T. Sönmez, and M. U. Ünver. A kidney exchange clearinghouse in New England. *American Economic Review*, 95(2):376–380, 2005. [2]
- L. Shapley and H. Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1: 23–37, 1974. [26]
- T. Sönmez. Strategy-proofness and essentially single-valued cores. *Econometrica*, 67(3): 677–690, May 1999. [7], [9], [20], [26]
- T. Sönmez. Bidding for army career specialties: Improving the ROTC branching mechanism. *Journal of Political Economy*, 121(1):186–219, March 2013. [5], [21]
- T. Sönmez and T. B. Switzer. Matching with (branch-of-choice) contracts at the United States Military Academy. *Econometrica*, 81(2):451–488, 2013. [5], [21]
- Y. Sprumont. Constrained-optimal strategy-proof assignment: Beyond the Groves mechanisms. *Journal of Economic Theory*, 148:1102–1121, 2013. [7]
- W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, March 1961. [6]
- A. Westkamp. An analysis of the German university admissions system. *Economic Theory*, 53(3):561–589, 2013. [22]

A Proofs

A.1 Proofs from Section 4

Proof of Participation-equivalence Lemma. Let φ and φ' be a pair of strategy-proof and individually rational mechanisms that are participation-equivalent. If they are not welfare-equivalent, then there are $R \in \mathcal{R}$ and $i \in N$ such that $\varphi_i(R) P_i \varphi'_i(R)$. Let $\alpha \equiv \varphi(R)$ and $\beta \equiv \varphi'(R)$.

Since both φ and φ' are individually rational, we have $\alpha(i) R_i \emptyset$ and $\beta(i) R_i \emptyset$. Since $\alpha(i) P_i \beta(i) R_i \emptyset$, we deduce that $i \in N(\alpha)$. By participation-equivalence of α and β , $i \in N(\beta)$.

Since $\alpha(i) P_i \beta(i) R_i \emptyset$, by richness of outside options, there is $R'_i \in \mathcal{R}_i$ such that (1) $\alpha(i) P'_i \emptyset P'_i \beta(i)$, and (2) for each $\gamma \in \mathcal{F}_i$, if $\gamma(i) R'_i \emptyset$ then $\gamma(i) P_i \beta(i)$. Let $\gamma \equiv \varphi(R'_i, R_{-i})$ and $\gamma' \equiv \varphi(R'_i, R_{-i})$.

Since φ is strategy-proof, $\gamma(i) R'_i \alpha(i)$. Otherwise, i would have an incentive to misreport R_i if his true preferences relation were R'_i . Thus, by definition of R'_i , $\gamma(i) P'_i \emptyset$. So $i \in N(\gamma)$. Again, since φ and φ' participation-equivalent, $i \in N(\gamma')$. Since φ' is individually rational, $\gamma'(i) R'_i \emptyset$. By definition of R'_i , we know that $\gamma'(i) P_i \beta(i)$. This contradicts the strategy-proofness of φ' since i has an incentive to misreport his preference as R'_i if his true preference were R_i . Thus, φ and φ' are welfare-equivalent. \square

It is noteworthy that the proof of the Participation-equivalence Lemma is at the level of a single agent.

Proof of the Structure Lemma. We begin with the following claim.

Claim: For each pair $\alpha, \beta \in \mathcal{F}$, if α is individually rational and participation-maximal and β Pareto-improves α , then $N(\alpha) = N(\beta)$ and β is individually rational and participation-maximal as well.

Proof of claim: Since β Pareto-improves α , which is individually rational, for each $i \in N$, $\beta(i) R_i \alpha(i) R_i \emptyset$. Thus β is individually rational.

For each $i \in N(\alpha)$, by no indifference with \emptyset , $\alpha(i) P_i \emptyset$. Since $\beta(i) R_i \alpha(i)$, $\beta(i) P_i \emptyset$, and so $i \in N(\beta)$. Thus, $N(\alpha) \subseteq N(\beta)$. If $N(\alpha) \subsetneq N(\beta)$, then we contradict the assumption that α is participation-maximal, since for each $i \in N(\beta) \setminus N(\alpha)$, by no indifference with \emptyset , $\beta(i) P_i \alpha(i) = \emptyset$. Therefore, $N(\beta) = N(\alpha)$.

Next, we show that β is participation-maximal. If it is not, there is $\gamma \in \mathcal{F}$ such (1) $N(\beta) \subsetneq N(\gamma)$ and (2) there is no $i \in N$, such that $\beta(i) P_i \gamma(i)$. Since $N(\beta) = N(\alpha)$, $N(\alpha) \subsetneq N(\gamma)$. Since preferences are transitive and β Pareto-improves α , there is no $i \in N$ such that $\alpha(i) P_i \gamma(i)$. This contradicts the participation-maximality of α . \diamond

Suppose $\alpha, \beta \in \mathcal{F}$ are individually rational, participation-maximal, and Pareto-connected allocations. Then, there is a sequence of individually rational and participation-maximal allocations, $(\alpha^k)_{k=0}^K$, with $\alpha^0 = \alpha$ and $\alpha^K = \beta$, such that for each $k \in \{1, \dots, K\}$, either α^k Pareto-improves α^{k-1} or vice versa. In either case, by the claim, $N(\alpha^k) = N(\alpha^{k-1})$. \square

Proof of Proposition 1. Consider a pair, φ and φ' , of strategy-proof, individually rational, and participation-maximal mechanisms. If they are Pareto-connected, then by the Structure Lemma, they are participation-equivalent. Thus, by the Participation-equivalence Lemma, they are welfare-equivalent. \square

Proof of Theorem 1. Follows from Proposition 1. \square

Proof of Theorem 2. We start with a pair of individually rational mechanisms φ and φ' such that φ' Pareto-improves φ . Thus, for each $R \in \mathcal{R}$, and each $i \in N$, $\varphi'_i(R) R_i \varphi_i(R)$. By individual rationality of φ and no indifference with \emptyset , if $i \in N(\varphi(R))$, then $\varphi_i(R) P_i \emptyset$, so $\varphi'_i(R) P_i \emptyset$. Thus, $i \in N(\varphi'(R))$. Since this holds at each R for each $i \in N(\varphi(R))$, φ' participation-expands φ .

We now prove the converse for strategy-proof and individually rational mechanisms even if there is indifference with \emptyset . If φ' participation-expands φ but does not Pareto-improve it, then there are $i \in N$ and $R \in \mathcal{R}$ such that $\alpha(i) \equiv \varphi_i(R) P_i \varphi'_i(R) \equiv \beta(i)$. By individual rationality of φ' , $\beta(i) R_i \emptyset$.

By richness of outside options, there is $R'_i \in \mathcal{R}_i$ such that (1) $\alpha(i) P'_i \emptyset P'_i \beta(i)$, and (2) for each $\gamma \in \mathcal{F}_i$, if $\gamma(i) R'_i \emptyset$ then $\gamma(i) P_i \beta(i)$. Let $\gamma \equiv \varphi(R'_i, R_{-i})$ and $\gamma' \equiv \varphi'(R'_i, R_{-i})$.

The reasoning from the proof of the Participation-equivalence Lemma shows that $i \in N(\gamma)$. Since $N(\gamma') \supseteq N(\gamma)$, $i \in N(\gamma')$. The remainder of the proof proceeds just as that of the Participation-equivalence Lemma. \square

Example demonstrating Remark 1. Consider an object allocation problem with $N \equiv \{i_1, i_2, i_3, \dots\}$ and $O \equiv \{a, b, c, \dots\}$. Suppose that T is a singleton, $|N| \geq 3$, $|O| \geq 3$, for each $o \in O$, $\mathcal{F}_o = \{\{i\} : i \in N\} \cup \{\emptyset\}$, \mathcal{F} is Cartesian, and the preferences are strict.

Consider the benchmark mechanism, $\underline{\varphi}$, defined by setting, for each $P \in \mathcal{P}$,

$$\underline{\varphi}_{i_1}(P) = P_{i_1}\text{-max}(O \setminus \{a\}).^{54}$$

$$\underline{\varphi}_{i_2}(P) = \begin{cases} P_{i_2}\text{-max}(O \setminus \underline{\varphi}_{i_1}(P)) & \text{if } \underline{\varphi}_{i_1}(P) \neq c \text{ or} \\ P_{i_2}\text{-max}(O \setminus (\underline{\varphi}_{i_1}(P) \cup \underline{\varphi}_{i_3}(P))) & \text{otherwise} \end{cases}$$

$$\underline{\varphi}_{i_3}(P) = \begin{cases} P_{i_3}\text{-max}(O \setminus (\underline{\varphi}_{i_1}(P) \cup \underline{\varphi}_{i_2}(P))) & \text{if } \underline{\varphi}_{i_1}(P) = c \text{ or} \\ P_{i_3}\text{-max}(O \setminus \underline{\varphi}_{i_1}(P)) & \text{otherwise} \end{cases}$$

and for $k > 3$,

$$\underline{\varphi}_{i_k}(P) = P_{i_k}\text{-max}(O \setminus (\underline{\varphi}_{i_1}(P) \cup \dots \cup \underline{\varphi}_{i_{k-1}}(P)))$$

In words, this mechanism assigns to i_1 his most preferred object except for a . The remaining objects are distributed among the remaining agents sequentially in the order i_2, i_3, i_4, \dots if i_1 is not assigned c . The places of i_2 and i_3 are swapped if i_1 is assigned c . Since i_1 is barred from receiving a , this mechanism is not participation-maximal: at each $P \in \mathcal{P}$ such that, for each $i \in N \setminus \{i_1\}, \emptyset P_i a$ and, for each $o \in O \setminus \{a\}, a P_{i_1} \emptyset P_{i_1} o, \underline{\varphi}_{i_1}(P) = \emptyset$ and a is not assigned to anyone, so $\underline{\varphi}$ is not participation-maximal.

While it may be possible to find a strategy-proof mechanism that Pareto-improves $\underline{\varphi}$, we show that no such mechanism is participation-maximal. Thus, there *is* a strategy-proof mechanism that cannot be Pareto-improved by another strategy-proof mechanism but is not participation-maximal.

To prove this claim, suppose that φ is participation-maximal and Pareto-improves $\underline{\varphi}$. Consider $P \in \mathcal{P}$ as follows:

P_{i_1}	P_{i_2}	P_{i_3}	and for $k > 3,$	P_{i_k}
a	b	b		\emptyset
\emptyset	a	\emptyset		
	\vdots			

By definition of $\underline{\varphi}$, we have $\underline{\varphi}_{i_2}(P) = b$, and for each $i \in N \setminus \{i_2\}, \varphi_i(P) = \emptyset$. Since φ Pareto-improves $\underline{\varphi}$ and φ is participation-maximal, $\varphi_{i_1}(P) = a, \varphi_{i_2}(P) = b$, and, for each $i \in N \setminus \{i_1, i_2\}, \varphi_i(P) = \emptyset$.

⁵⁴ Given $P_i \in \mathcal{P}_i$ and $A \subseteq O$, we denote the best element of A according to P_i by $P_i\text{-max}(A)$.

Now consider $P'_{i_1} \in \mathcal{P}_{i_1}$ as follows:

$$\begin{array}{c} P'_{i_1} \\ a \\ c \\ \vdots \end{array}$$

By definition of $\underline{\varphi}$, we have $\underline{\varphi}_{i_1}(P'_{i_1}, P_{-i_1}) = c$, $\underline{\varphi}_{i_2}(P'_{i_1}, P_{-i_1}) = a$, $\underline{\varphi}_{i_3}(P'_{i_1}, P_{-i_1}) = b$, and for each $i \in N \setminus \{i_1, i_2, i_3\}$, $\underline{\varphi}_i(P'_{i_1}, P_{-i_1}) = \emptyset$. Since this allocation is Pareto-efficient at (P'_{i_1}, P_{-i_1}) and $\underline{\varphi}$ Pareto-improves $\underline{\varphi}$, $\underline{\varphi}(P'_{i_1}, P_{-i_1}) = \underline{\varphi}(P'_{i_1}, P_{-i_1})$. But then $\varphi_{i_1}(P_{i_1}, P_{-i_1}) = a P'_{i_1} c = \varphi_{i_1}(P'_{i_1}, P_{-i_1})$, so $\underline{\varphi}$ is not strategy-proof.

We conclude that no strategy-proof mechanism that Pareto-improves $\underline{\varphi}$ is participation-maximal. \square

Actually, $\underline{\varphi}$ in the above example is *group strategy-proof*. Since no participation-maximal strategy-proof mechanism Pareto-improves it, there is a *group strategy-proof* mechanism on the strategy-proofness-constrained Pareto frontier that is not participation-maximal.

While $\underline{\varphi}$ does not satisfy participation-maximality, it does satisfy a range-based non-wastefulness condition: there is no allocation in its range that Pareto-improves on the chosen allocation in a way that assigns an object to more agents. However, this is a very weak property—even the constant mechanism that always selects \emptyset satisfies it—that does not guarantee that a mechanism is strategy-proofness-constrained Pareto-efficient.

A.2 Proofs from Section 5.1

We first provide an example showing that without any assumptions on choice correspondences, even if they are single-valued, stability may not imply non-wastefulness.

Example 1. *There may be a stable allocation that is not participation-maximal.*

Let $N \equiv \{i_1, i_2\}$, $T \equiv \{t_1, t_2\}$, $O \equiv \{o\}$, and $X(o) \equiv \{(i_1, o, t_1), (i_1, o, t_2), (i_2, o, t_1)\}$. Let C_o be such that for each $Y \subseteq X(o)$, if $(i_1, o, t_1) \in Y$ then $C_o(Y) = \{(i_1, o, t_1)\}$ and otherwise $C_o(Y) = \{Y\}$. Let $R \in \mathcal{R}$ be such that $(i_1, o, t_2) P_{i_1} (i_1, o, t_1) P_{i_1} \emptyset$ while $(i_2, o, t_1) P_{i_2} \emptyset$. Let $\mu \equiv \{(i_1, o, t_1)\} \in \mathcal{F}$. Then μ is stable at R . However, it is not participation-maximal as there is $\nu \equiv \{(i_1, o, t_2), (i_2, o, t_1)\} \in \mathcal{F}$, which makes every agent better off and $i_2 \notin N(\nu) \setminus N(\mu)$.

Notice that C_o in Example 1 is not size monotonic.

Before showing that stability implies non-wastefulness under the assumptions of size monotonicity and idempotence, we start with a definition and a lemma. For each $P \in \mathcal{P}$,

each $\mu \in \mathcal{F}$, and each $o \in O$, let $\mathbf{Y}_o^\mu(\mathbf{P})$ be the triples in $X(o)$ that are associated with an agent who prefers it to what he is assigned at μ . That is, $Y_o^\mu(P) \equiv \{x \in X(o) : x P_{N(x)} \mu(N(x))\}$.

Lemma 5. *Suppose that C is size monotonic and idempotent. For each $P \in \mathcal{P}$, each stable $\mu \in \mathcal{F}$, each $o \in O$, each finite $Y \subseteq Y_o^\mu(P)$, and each $Z \in C_o(\mu(o) \cup Y)$, $|Z| = |\mu(o)|$.*

Proof. We proceed by induction over subsets of $Y_o^\mu(P)$. Let $Y \subseteq Y_o^\mu(P)$.

For the base case, where $Y = \emptyset$, since $\mu(o) \in \text{range}(C_o)$, by idempotence of C , $\mu(o) \in C_o(\mu(o) \cup Y)$ and by size monotonicity of C , for each $Z \in C_o(\mu(o))$, $|Z| = |\mu(o)|$.

As an induction hypothesis, assume that for each $Y' \subsetneq Y$ and each $Z \in C_o(\mu(o) \cup Y')$, $|Z| = |\mu(o)|$. Equivalently, for each $T \subseteq \mu(o) \cup Y$ such that $Y \not\subseteq T$, for each $Z \in C_o(\mu(o) \cup T)$, $|Z| = |\mu(o)|$.

The induction step is to show that, for each $Z \in C_o(\mu(o) \cup Y)$, $|Z| = |\mu(o)|$. Let $Z \in C_o(\mu(o) \cup Y)$. By idempotence of C , $Z \in C_o(Z)$. Thus, since $Z \subseteq \mu(o) \cup Z$, by size monotonicity of C ,

$$\text{for each } Z' \in C_o(\mu(o) \cup Z), |Z| \leq |Z'|. \quad (1)$$

By stability of μ , either $Y \not\subseteq Z$ or $\mu(o) \in C_o(\mu(o) \cup Y)$. If $\mu(o) \in C_o(\mu(o) \cup Y)$, then by size monotonicity of C , for each $Z \in C_o(\mu(o) \cup Y)$, $|Z| = |\mu(o)|$, concluding the proof.

Thus, we consider the case where $Y \not\subseteq Z$. By the induction hypothesis, since $Y \not\subseteq Z$,

$$\text{for each } Z' \in C_o(\mu(o) \cup Z), |Z'| = |\mu(o)|. \quad (2)$$

By (1) and (2), $|Z| \leq |\mu(o)|$. Since $\mu \in \mathcal{F}$, by idempotence of C , $\mu(o) \in C_o(\mu(o))$. Thus, since $Z \in C_o(\mu(o) \cup Y)$, by size monotonicity of C , $|\mu(o)| \leq |Z|$. Then, we conclude that $|Z| = |\mu(o)|$. \square

Proof of Lemma 3. Suppose that μ is wasteful. Then there are $o \in O$ and $\nu \in \mathcal{F}$ such that $|\nu(o)| > |\mu(o)|$ and, for each $y \in \nu(o) \setminus \mu(o)$, $y P_{N(y)} \mu(N(y))$. Let $Y \equiv \nu(o) \setminus \mu(o)$. Since $Y \subseteq Y_o^\mu(P)$ and Y is finite, by Lemma 5, for each $Z' \in C_o(\mu(o) \cup Y)$, $|Z'| = |\mu(o)|$. However, $\nu(o) \subseteq \mu(o) \cup Y$. So by size monotonicity of C , for each $Z \in C_o(\nu(o))$ and each $Z' \in C_o(\mu(o) \cup Y)$, $|Z| \leq |Z'| = |\mu(o)|$. Since $\nu \in \mathcal{F}$, $\nu \in \mathcal{F}_o = \text{range}(C_o)$. So by idempotence of C , $\nu(o) \in C_o(\nu(o))$. Thus, $|\nu(o)| \leq |\mu(o)|$. This contradicts the definition of ν . \square

Proof of the Non-wasteful Structure Lemma. Since μ is non-wasteful, it is participation-maximal as well, by Lemma 10 in Appendix B. So by the Structure Lemma, $N(\mu) = N(\nu)$, so $|\mu| = |\nu|$.

Since $\mu(o)$ and $\mu(o')$ are disjoint for distinct $o, o' \in O$, $\sum_{o \in O} |\mu(o)| = |\mu|$. By similar reasoning, $\sum_{o \in O} |\nu(o)| = |\nu|$. Since $|\mu| = |\nu|$, $\sum_{o \in O} |\mu(o)| = \sum_{o \in O} |\nu(o)|$.

Since ν Pareto-improves μ and μ is non-wasteful, by the definition of non-wastefulness there are two possibilities: (1) for each $o \in O$, $|\mu(o)| \geq |\nu(o)|$ or (2) for each $i \in N$, $\nu(i) I_i \mu(i)$.

If ν strictly Pareto-improves μ , case (1) applies. Since $\sum_{o \in O} |\mu(o)| = \sum_{o \in O} |\nu(o)|$, for each $o \in O$, $|\mu(o)| = |\nu(o)|$. The other case is trivial, since strict preferences implies $\nu = \mu$. \square

Proof of Proposition 2. Since C is size monotonic and idempotent, by Lemma 3, a stable allocation is also non-wasteful. By definition, for each $P \in \mathcal{P}$, $\varphi^{AOS}(P)$ Pareto-improves each stable allocation at P . Thus, by the Non-wasteful Structure Lemma, the Rural Hospitals Theorem holds.

The proofs of Theorems 10 and 11 of Hatfield and Milgrom [2005] use only the conclusion of the Rural Hospitals Theorem to show that φ^{AOS} is strategy-proof in their setting. Since they work on the entire strict preference domain and \mathcal{F} is a subset of their set of feasible allocations, as long as the Rural Hospitals Theorem holds and φ^{AOS} exists, it is strategy-proof in our setting.

Next, if any other mechanism φ is stable-dominating, there exists a stable $\underline{\varphi}$ that both φ and φ^{AOS} Pareto-improve. Since C is size monotonic and idempotent, by Lemma 3, $\underline{\varphi}$ is non-wasteful. Then, $\underline{\varphi}$ is Pareto-connected to φ^{AOS} . By Proposition 1, if $\varphi \neq \varphi^{AOS}$, then it is not strategy-proof. Thus φ^{AOS} is the only stable-dominating and strategy-proof mechanism. \square

Proof of Proposition 3. Let φ and φ' be stable-dominating and strategy-proof mechanisms. For each $P \in \mathcal{P}$, there exist stable $\mu, \mu' \in \mathcal{F}$ such that $\varphi(P)$ Pareto-improves μ and $\varphi'(P)$ Pareto-improves μ' . Since C is size monotonic and idempotent, by the Non-wasteful Structure Lemma, $N(\varphi(P)) = N(\mu)$ and $N(\varphi'(P)) = N(\mu')$. By the Rural Hospitals Theorem, $N(\mu) = N(\mu')$, so $N(\varphi(P)) = N(\varphi'(P))$. By the Participation-equivalence Lemma, φ and φ' are welfare-equivalent. Since preferences are strict, $\varphi = \varphi'$. \square

Proof of Proposition 4. For the sake of contradiction, suppose that there are $P \in \mathcal{P}$ and $\nu \in \mathcal{F}$ such that ν is stable at P and $\varphi(P)$ does not Pareto-improve ν . So there is $i \in N$ such that $\nu(i) P_i \varphi_i(P)$. So $\nu(i) \neq \emptyset$. By the Rural Hospitals Theorem, $\varphi_i(P) \neq \emptyset$. Let $x \equiv \varphi_i(P)$.

Since \mathcal{P} is closed under truncation, there is $P'_i \in \mathcal{P}_i$ such that for each $y \in X(i)$, (1) $y P'_i x$ if and only if $y P_i x$ and (2) if $x P_i y$ then $x P_i \emptyset P_i y$.

We first show, by contradiction, that ν is stable at P' . If ν is not stable at P' , there are $o \in O$ and $Y \subseteq X(o)$ such that, for each $z \in Y$, $z P'_{N(z)} \nu(N(z))$, $\nu(o) \notin C_o(\nu(o) \cup Y)$ and there is $Z \in C_o(\nu(o) \cup Y)$ such that $Y \subseteq Z$. Since, for each $j \in N \setminus \{i\}$, $P'_j = P_j$ and ν is stable at

P , there is $z \in Y$ such that $N(z) = i$. However, since $z P'_i \nu(i) = x$, we have $z P_i x$. This contradicts the stability of ν at P .

By the Rural Hospitals Theorem, since $\nu(i) \neq \emptyset$ and since $\varphi(P')$ and ν are Pareto-connected at P' , $\varphi_i(P') \neq \emptyset$. Since φ is individually rational, $\varphi_i(P') P'_i \emptyset$. By definition of P'_i , $\varphi_i(P') P_i x$. However, this contradicts the strategy-proofness of φ . \square

Proof of Remark 2. Let μ be stable with respect to C . By definition of stability it is individually rational. Since C is size monotonic and idempotent, by Lemma 3, μ is non-wasteful. If it violates priorities, there are a pair $i, j \in N$ and $o \in O$ such that $\mu(i) = o$, $o P_j \mu(j)$, and $j >_o i$. Since μ is non-wasteful, $|\mu(o)| = q_o$. Since $j >_o i$, $C_o(\mu(o) \cup \{j\}) = \{(\mu(o) \setminus \{i\}) \cup \{j\}\}$. This contradicts the stability of μ .

Suppose that μ is individually rational, non-wasteful, and respects priorities. If it is not stable, then there are $o \in O$ and $Y \subseteq N \setminus \mu(o)$ such that for each $i \in Y$, $o P_i \mu(i)$ and (1) $Y \subseteq Z$ for some $Z \in C_o(\mu(o) \cup Y)$ and (2) $\mu(o) \notin C_o(\mu(o) \cup Y)$. Since, for each $i \in Y$, $o P_i \mu(i)$, and μ is non-wasteful, $|\mu(o)| = q_o$. Thus, for each $Z \in C_o(\mu(o) \cup Y)$, $|Z| = |\mu(o)|$. Since $\mu(o) \notin C_o(\mu(o) \cup Y)$, there are $i \in Y$ and $j \in \mu(o)$ such that $i >_o j$. This contradicts the assumption that μ respects priorities. \square

A.3 Proofs from Section 5.3

Proposition 6. *For the problem of locating of a public facility on a line, if a mechanism is strategy-proof, individually rational, and Pareto-efficient, then it is dictatorial.*

Proof. If the highest threshold for an agent in N_L is strictly lower than the lowest threshold for an agent in N_R , no compromise is possible. Let l_1, \dots, l_k be the agents in N_L and let r_1, \dots, r_{n-k} be the agents in N_R .

Suppose that φ is strategy-proof, individually rational, and Pareto-efficient. If no compromise is possible at t , then Pareto-efficiency and individual rationality require that $\varphi(t)$ is either $(0, N_L)$ or $(1, N_R)$. Suppose that it is $(0, N_L)$. The argument is symmetric if it is $(1, N_R)$.

Step 1: *For every t' such that no compromise is possible, $\varphi(t') = (0, N_L)$.*

For each $i \in N_L$, let $\underline{t}_i = 0$. Since $\varphi(t) = (0, N_L)$, by strategy-proofness, Pareto-efficiency, and individual rationality, $\varphi(\underline{t}_{l_1}, t_{-l_1}) = (0, N_L)$. Otherwise, if $\varphi(\underline{t}_{l_1}, t_{-l_1}) = (x, S)$ for some $x > 0$ or if $l_1 \notin S$, then l_1 benefits by reporting t_1 when his true threshold is \underline{t}_{l_1} . Furthermore, by Pareto-efficiency, $N_L \subseteq S$ and by individual rationality, since 0 is below the threshold for each member of N_R , $S = N_L$. Repeating the argument, we replace, for each member of N_L , his threshold by 0 to conclude that $\varphi(\underline{t}_{N_L}, t_{N_R}) = (0, N_L)$. By a similar argument we replace, one at a time, the threshold of each agent in N_R at the profile $(\underline{t}_{N_L}, t_{N_R})$

by his threshold at t' to see that $\varphi(\underline{t}_{N_L}, t'_{N_R}) = (0, N_L)$. Again, by a similar argument we replace, one at a time, the threshold of each agent in N_L at the profile $(\underline{t}_{N_L}, t'_{N_R})$ by his threshold at t' to conclude that $\varphi(t') = (0, N_L)$.

Step 2: *If t is such that the threshold of each member of N_L is 0, then $\varphi(t) = (0, S)$ where $N_L \subseteq S$.*

By Step 1, if \hat{t}_{N_R} is such that the threshold of each agent in N_R is greater than 0, then $\varphi(t_{N_L}, \hat{t}_{N_R}) = (0, N_L)$.

Let \underline{t}_{N_R} be such that the threshold of each agent in N_R is 0. We will first show that that replacing, one by one, the threshold of each member of N_R at the profile (t_{N_L}, \hat{t}_{N_R}) with his threshold at \underline{t}_{N_R} does not move the chosen location from 0. Starting with r_1 , we claim that by strategy-proofness, individual rationality, and Pareto-efficiency, $\varphi(t_{N_L}, \underline{t}_{r_1}, \hat{t}_{N_R \setminus \{r_1\}}) = (0, N_L \cup \{r_1\})$. Suppose that $\varphi(t_{N_L}, \underline{t}_{r_1}, \hat{t}_{N_R \setminus \{r_1\}}) = (x, S)$. By Pareto-efficiency, $r_1 \in S$. If $x > 0$, let $t_{r_1}^x \in (0, x)$. Then, by Step 1, $\varphi(t_{N_L}, t_{r_1}^x, \hat{t}_{N_R \setminus \{r_1\}}) = (0, N_L)$ since there is no compromise. Then, when r_1 's true threshold is $t_{r_1}^x$, he gains by reporting \underline{t}_{r_1} instead, contradicting strategy-proofness. Thus, $x = 0$. By individual rationality and Pareto-efficiency, $S = N_L \cup \{r_1\}$. Repeating this argument for each agent in N_R , we reach the desired conclusion of this step of the proof.

Step 3: *For every t , $\varphi(t) = (0, N_L \cup \{i \in N_R : t_i = 0\})$*

Let \underline{t}_{N_L} be such that the threshold of each agent in N_L is 0. By Step 2, $\varphi(\underline{t}_{N_L}, t_{N_R}) = (0, S)$ where $N_L \subseteq S$. Replacing, one by one, the thresholds of each agent in N_L at $(\underline{t}_{N_L}, t_{N_R})$ with his threshold at t as in the argument of Step 1, we reach the desired conclusion. \square

A.4 Proofs from Section 5.4

We start with a piece of notation: Given $v_i \in \mathcal{V}_i$ and $\kappa > v_i(\emptyset)$, let $\chi(v_i, \kappa) \in \mathcal{V}_i$ be such that it increases the valuation of \emptyset from $v_i(\emptyset)$ to κ , leaving the other valuations unchanged. By the richness of outside options, $\chi(v_i, \kappa)$ exists in \mathcal{V}_i .

Our first lemma is to show that if an implementable decision rule is IR-implementable, then regardless of other agents' valuations, there is a report for each agent that ensures his non-participation.

Lemma 6. *If d is IR-implementable, then for each $v \in \mathcal{V}$ and each $i \in N$ there is $v'_i \in \mathcal{V}_i$ such that $i \notin N(d(v'_i, v_{-i}))$.*

Proof. Suppose that t IR-implements d . If $i \notin N(d(v))$, we are done. Otherwise, let $\kappa > v_i(d(v)) - t_i(v)$. Let $v'_i \equiv \chi(v_i, \kappa)$. By individual rationality, $v'_i(d(v'_i, v_{-i})) - t_i(v'_i, v_{-i}) \geq \kappa$. If $i \in$

$N(d(v'_i, v_{-i}))$, then by definition of v'_i , $v'_i(d(v'_i, v_{-i})) = v_i(d(v'_i, v_{-i}))$. This violates strategy-proofness since $v_i(d(v'_i, v_{-i})) - t_i(v'_i, v_{-i}) \geq \kappa > v_i(d(v)) - t_i(v)$. Thus, $i \notin N(d(v'_i, v_{-i}))$. \square

The next lemma states that if t implements d then, given the valuations of other agents, then t assigns an agent the same payment for each of his possible reports that result in his non-participation. Note that this is vacuously true if every agent participates at every profile of valuations, in which case the decision rule would not be IR-implementable.

Lemma 7. *If t implements d , then for each $v \in \mathcal{V}$, each $i \in N$, and each $v'_i \in \mathcal{V}_i$, such that $i \notin N(d(v))$ and $i \notin N(d(v'_i, v_{-i}))$, we have $t_i(v'_i, v_{-i}) = t_i(v)$.*

Proof. Suppose that $t_i(v'_i, v_{-i}) \neq t_i(v)$. Without loss of generality, say $t_i(v'_i, v_{-i}) < t_i(v)$. Combining this with $v_i(d(v)) = v_i(\emptyset) = v_i(d(v'_i, v_{-i}))$, we have $v_i(d(v)) - t_i(v) < v_i(d(v'_i, v_{-i})) - t_i(v'_i, v_{-i})$. This contradicts the strategy-proofness of (d, t) . \square

The remainder of the proof is similar to the proof of the Participation-equivalence [Lemma](#). The main difference is that we do not directly appeal to individual rationality. Instead, we appeal to the above stated lemmas. The first step is to prove that if we pin down the payments of non-participants at a strategy-proof mechanism, then we have pinned down the payment rule entirely. The logic is very similar to Participation-equivalence [Lemma](#).

Lemma 8. *If both t and t' implement d and for each $v \in \mathcal{V}$ and each $i \notin N(d(v))$, $t_i(v) = t'_i(v)$, then $t = t'$.*

Proof. If $t \neq t'$ then there are $v \in \mathcal{V}$ and $i \in N(d(v))$ such that $t_i(v) \neq t'_i(v)$. Without loss of generality, suppose that $t_i(v) < t'_i(v)$. By [Lemma 6](#), there is $\hat{v}_i \in v_i$ such that $i \notin N(d(\hat{v}_i, v_{-i}))$. Let $\hat{t}_i \equiv t_i(\hat{v}_i, v_{-i}) = t'_i(\hat{v}_i, v_{-i})$. Define $\kappa \equiv v_i(d(v)) - \frac{t_i(v) + t'_i(v)}{2} + \hat{t}_i$ and let $v'_i \equiv \chi(v_i, \kappa)$.

Consider $d(v'_i, v_{-i})$. If $i \notin N(d(v'_i, v_{-i}))$ then by [Lemma 7](#), $t_i(v'_i, v_{-i}) = \hat{t}_i$. So $v'_i(d(v'_i, v_{-i})) - t_i(v'_i, v_{-i}) = \kappa - \hat{t}_i$. However, by definition of κ and v'_i , $\kappa - \hat{t}_i < v'_i(d(v)) - t_i(v)$. This contradicts strategy-proofness of (d, t) , since $v'_i(d(v'_i, v_{-i})) - t_i(v'_i, v_{-i}) = \kappa - \hat{t}_i < v'_i(d(v)) - t_i(v)$. Thus, $i \in N(d(v'_i, v_{-i}))$.

By strategy-proofness of (d, t') , we have $v'_i(d(v'_i, v_{-i})) - t'_i(v'_i, v_{-i}) \geq \kappa - \hat{t}_i = v'_i(d(\hat{v}_i, v_{-i})) - t'_i(\hat{v}_i, v_{-i})$. By definition of κ and v'_i , $v_i(d(v'_i, v_{-i})) - t'_i(v'_i, v_{-i}) = v'_i(d(v'_i, v_{-i})) - t'_i(v'_i, v_{-i}) = \kappa - \hat{t}_i > v_i(d(v)) - t'_i(v)$. This violates the strategy-proofness of (d, t') . Thus, t' does not implement d . \square

Finally, we show that adding a function of others' payments to each agent's payment does not compromise strategy-proofness.

Lemma 9. *Suppose that t implements d . For each $i \in N$, let $h_i : \mathcal{V}_{-i} \rightarrow \mathbb{R}$. If t' is such that for each $v \in \mathcal{V}$ and $i \in N$, $t'_i(v) = t_i(v) + h_i(v_{-i})$, then t' also implements d .*

Proof. If (d, t') is not strategy-proof, then there are $v \in \mathcal{V}$, $i \in N$, and $v'_i \in \mathcal{V}_i$ such that $v_i(d(v)) - t'_i(v) < v_i(d(v'_i, v_{-i})) - t'_i(v'_i, v_{-i})$. Adding $h_i(v_{-i})$ to either side of this inequality, $v_i(d(v)) - t_i(v) < v_i(d(v'_i, v_{-i})) - t_i(v'_i, v_{-i})$. This contradicts the strategy-proofness of (d, t) . \square

Proof of Proposition 5. Let \underline{t} be the payment rule that IR-implements d with the least payment to non-participants. By Lemmas 8 and 9, for any other t that implements d , the difference in payments to the non-participants is a function of others' valuations and is the difference between t_i and \underline{t}_i . \square

B More on Non-wastefulness

Extending non-wastefulness We present two examples that illustrate challenges to extending the non-wastefulness definition of [Balinski and Sönmez \[1999\]](#) beyond the school choice model.

Example 2. *Feasibility not capacity-based.*

Let T be a singleton, $O \equiv \{o\}$, $N \equiv \{i_1, i_2, i_3\}$, and $\mathcal{F}_o \equiv \{\emptyset, \{i_1\}, \{i_2\}, \{i_3\}, \{i_2, i_3\}\}$. Consider $P \in \mathcal{P}$ as follows:

P_{i_1}	P_{i_2}	P_{i_3}
o	o	o
\emptyset	\emptyset	\emptyset

What is the “capacity” of o ? The largest set of agents that may consume o contains two elements while the smallest non-trivial set contains only one. If we naïvely extend non-wastefulness by setting the capacity of o to be two, then allocating it to i_1 would be wasteful, even though this is the only allocation where i_1 receives his top choice. On the other hand if we set the capacity of o to be one, then allocating it to i_2 alone would *not* be wasteful even though o could be assigned to i_3 as well. Neither of these is sensible. Thus, non-wastefulness cannot be extended in a way that relies on a fixed capacity for each object.

Example 3. *Complementarities in feasibility.*

Let T be a singleton, $O \equiv \{o_1, o_2\}$, $N \equiv \{i_1, i_2\}$, $\mathcal{F}_{o_1} \equiv \{\emptyset, \{i_1, i_2\}\}$, $\mathcal{F}_{o_2} \equiv \{\emptyset, \{i_1\}, \{i_2\}\}$, and \mathcal{F} be Cartesian (though not capacity-based). Consider $P \in \mathcal{P}$ as follows:

P_{i_1}	P_{i_2}
o_1	o_2
o_2	o_1
\emptyset	\emptyset

There are two allocations of interest. The first assigns o_1 to both agents. This is the only allocation where i_1 receives his top choice. The second assigns \emptyset to i_1 and o_2 to i_2 . This is the only allocation where i_2 receives his top choice. At either of these allocations there is an agent who prefers the unallocated object to what he receives. However, the only way he can be assigned this unallocated object is by making the other agent worse off. A sensible definition of non-wastefulness, in the general setting, should not rule either of these allocations out.

Though there is no fixed notion of capacity, as demonstrated by Example 2, non-wastefulness should seek to ensure that each object is utilized to the greatest extent possible. Yet, as demonstrated by Example 3, it should take care to ensure that increasing the utilization of an object by allocating it to agents who prefer it does not harm other agents.

Our definition of wastefulness summarizes the discussion above. Given $P \in \mathcal{P}$, $\mu \in \mathcal{F}$ is wasteful if there are $o \in O$, $i \in N$, and $\nu \in \mathcal{F}$, such that (1) $|\nu(o)| > |\mu(o)|$, so that ν allocates o to more agents than μ does, (2) $\nu(i) P_i \mu(i)$, so that i prefers his assignment at ν to that at μ , and (3) for each $j \in N \setminus \{i\}$, $\nu(j) R_j \mu(j)$, so that no agent is worse off at ν compared to μ . As we now demonstrate, this definition is an extension of the definition of that by Balinski and Sönmez [1999] to our more general setting.

Claim 1. *Suppose that T is a singleton and \mathcal{F} is capacity-based. An allocation μ is non-wasteful by the definition of Balinski and Sönmez [1999] if and only if it is non-wasteful.*

Proof. Suppose that μ is wasteful by the definition of Balinski and Sönmez [1999]. Then there are $o \in O$ and $i \in N$ such that $o P_i \mu(i)$ and $|\mu(o)| < q_o$. Let $\nu \equiv (\mu \cup \{(i, o)\}) \setminus \mu(i)$. Then $|\nu(o)| = |\mu(o)| + 1 \leq q_o$ and, for each $o' \in O \setminus \{o\}$, $|\mu(o')| - 1 \leq |\nu(o')| \leq |\mu(o')| \leq q_{o'}$. Thus, $\nu \in \mathcal{F}$ and $|\nu(o)| > |\mu(o)|$. Furthermore, $\nu(i) P_i \mu(i)$ while for each $j \in N \setminus \{i\}$, $\nu(j) = \mu(j)$. Thus, μ is wasteful.

Suppose that μ is non-wasteful by the definition of Balinski and Sönmez [1999]. Consider $\nu \in \mathcal{F}$ such that, for each $i \in N$, $\nu(i) R_i \mu(i)$ and for some $i \in N$, $\nu(i) P_i \mu(i)$. Let $o \in O$. If $|\mu(o)| < q_o$, by the Balinski and Sönmez definition of non-wastefulness, there is no $i \in N$

such that $o P_i \mu(i)$. So $|\nu(o)| \leq |\mu(o)|$. If $|\mu(o)| = q_o$, by feasibility of ν , $|\nu(o)| \leq q_o = |\mu(o)|$. Thus, μ is non-wasteful. \square

Relationship to participation maximality We first show that if an allocation is non-wasteful, then it is participation-maximal.

Lemma 10. *Non-wastefulness implies participation-maximality.*

Proof. If $\mu \in \mathcal{F}$ is non-wasteful but not participation-maximal, then there is $\nu \in \mathcal{F}$ such that $N(\nu) \supsetneq N(\mu)$ and for each $i \in N$, $\nu(i) R_i \mu(i)$. By no indifference with \emptyset , for each $i \in N(\nu) \setminus N(\mu)$, $\nu(i) P_i \mu(i) = \emptyset$. Finally, since $|N(\nu)| > |N(\mu)|$, there is $o \in O$ such that $|\nu(o)| > |\mu(o)|$. Thus, μ is wasteful. \square

The two properties are, however, not equivalent, even for classical object allocation.

Example 4. *An allocation that is participation-maximal but wasteful.*

Let T be a singleton, $O = \{o_1, o_2\}$, $N = \{i_1\}$, $\mathcal{F}_{o_1} = \mathcal{F}_{o_2} = \{\emptyset, \{i_1\}\}$, and \mathcal{F} be Cartesian. Consider $R \in \mathcal{R}$ such that $o_1 R_{i_1} o_2 R_{i_2} \emptyset$. Let $\mu \in \mathcal{F}$ be such that $\mu(i_1) = o_2$. Since $N(\mu) = N$, it is participation-maximal. However, consider $\nu \in \mathcal{F}$ such that $\nu(i_1) = o_1$. Since $\nu(i_1) P_{i_1} \mu(i_1)$ and $1 = |\nu(o_1)| > |\mu(o_1)| = 0$, μ is wasteful.

Capacity-based Structure Lemma and Rural Hospitals Theorem For models with capacity-based \mathcal{F} , we can further strengthen the conclusion of the Non-wasteful Structure Lemma: every object that is not allocated to capacity by μ is allocated to the same agents by both μ and ν .

Lemma 11. *Let \mathcal{F} be capacity-based. For each profile of preferences, if one allocation Pareto-improves another one that is individually rational and non-wasteful, then the two are participation-equivalent, assign each object to the same number of agents, and assign each object not filled to capacity to the same set of agents.*

Proof. The first two claims in the conclusion follow from the Non-wasteful Structure Lemma, so we prove the last one. By strict preferences, if μ is Pareto-efficient, then $\nu = \mu$. So suppose μ is not Pareto-efficient. By the Non-wasteful Structure Lemma, for each $o \in O$, $|\nu(o)| = |\mu(o)|$. Suppose that $o \in O$ is such that $|\mu(o)| < q_o$ but $N(\nu(o)) \neq N(\mu(o))$. Then there is $i \in N(\nu(o)) \setminus N(\mu(o))$. Let $x \equiv \nu(i)$ and $y \equiv \mu(i)$. Since ν Pareto-improves μ , $x P_i y$. Let $\gamma \equiv (\mu \setminus \{y\}) \cup \{x\}$. Let o' be the object associated with y . Since $i \notin N(\mu(o))$, $o' \neq o$. Since \mathcal{F} is capacity-based, $\mu(o') \setminus \{y\} \in \mathcal{F}_{o'}$ and, since $|\mu(o)| < q_o$, $\mu(o) \cup \{x\} \in \mathcal{F}_o$. Thus, since \mathcal{F} is Cartesian, $\gamma \in \mathcal{F}$. Since $\gamma(i) P_i \mu(i)$, for each $j \in N \setminus \{i\}$, $\gamma(j) = \mu(j)$, and $|\gamma(o)| = |\mu(o)| + 1$, this contradicts the assumption that μ is non-wasteful. \square

We can strengthen the Rural Hospitals Theorem to state that two stable allocations also keep the same set of agents matched to objects that do not fill up to capacity. It is straightforward to see for capacity-based \mathcal{F} that Lemma 11 drives the stronger Rural Hospitals Theorem. It also facilitates analogous variants of Propositions 2, 3, and 4.

Lemma 11 narrows down the ways in which an individually rational and non-wasteful allocation can be Pareto-improved when feasibility is capacity-based and preferences are strict. Suppose that $\mu, \nu \in \mathcal{F}$ are such that μ is individually rational and non-wasteful and ν Pareto-improves μ . By Lemma 11, each agent to whom μ assigns an object that it does not allocate to capacity receives the same object from ν . So the only way that ν can change the allocation of objects is through *trading-cycles* consisting of agents to whom μ assigns objects that are allocated to capacity. That is, ν assigns to each agent the object that μ assigns to the next agent in the cycle (possibly under different terms). In particular, if T is a singleton then every Pareto-improvement from a non-wasteful allocation results from a set of disjoint trading-cycles.

C More on Stable-dominating Mechanisms

When priorities are so weak as to be degenerate (in the sense of each agent having equal priority at each object), stability reduces to the combination of individually rationality and non-wastefulness, so each individually rational and Pareto-efficient allocation is also stable. On the other hand, if priorities are strict and each object has the same priority over agents, then there is a unique stable allocation and it is Pareto-efficient. The tension between stability and efficiency is thus dependent on the priorities.

Ehlers and Erdil [2010] define the following property of the priorities to guarantee that if a stable allocation is not strictly Pareto-improved by any other stable allocation,⁵⁵ then it is actually Pareto-efficient. They say that \succsim contains a *weak cycle* if there is a distinct triple $i, j, k \in N$ and a distinct pair $x, y \in O$ such that (1) (loop) $i \succsim_x j \succ_x k$ and $k \succ_y i$, and (2) (scarcity) there exist disjoint $N_x \subseteq N \setminus \{i, j, k\}$ and $N_y \subseteq N \setminus \{i, j, k\}$ such that for each $l \in N_x$, $l \succ_x j$, each $l \in N_y$, $l \succ_y i$, $|N_x| = q_x - 1$, and $|N_y| = q_y - 1$. They say that \succsim is *strongly acyclic* if it does not contain a *weak cycle*. Under a slightly stronger condition, we can say something about all stable allocations. We define a *weak* cycle* exactly as a weak cycle except that we only require $N_y \subseteq N \setminus \{i, k\}$ rather than $N_y \subseteq N \setminus \{i, j, k\}$ in the scarcity condition. If it does not contain a weak* cycle, \succsim is *strongly* acyclic*.⁵⁶

⁵⁵ Ehlers and Westkamp [2016] provide conditions on priorities that guarantee the existence of a strategy-proof mechanism that selects such a stable allocation.

⁵⁶ When priorities are strict, strong acyclicity is equivalent to Ergin-acyclicity [Ergin, 2002], but strong* acyclicity is stronger.

Proposition 7. For any (\succsim, q) the following statements are equivalent:

1. \succsim is strongly* acyclic
2. $\mu \in \mathcal{F}$ is stable-dominating if and only if it is stable

Proof. First, we show that strong* acyclicity is sufficient for a stable-dominating allocation to be stable. By the Non-wasteful Structure Lemma, each allocation that Pareto-improves on μ reallocates objects among agents in (possibly several) cycles so that each agent obtains the object assigned by μ to the next agent in the cycle. If the same object appears twice in the same cycle, we can divide the cycle into two separate cycles. Thus, it suffices to show that, for each cycle S , $\mu^S \in \mathcal{F}$ defined below is stable, where a cycle S is a set $\{i_1, \dots, i_n\} \subseteq N$ that satisfies, for each pair $i, j \in S$, $\mu(i) \neq \mu(j)$, and for each $l \in \{1, \dots, n\}$, $\mu(i_{l+1}) P_{i_l} \mu(i_l)$, where we identify $i_{n+1} \equiv i_1$. It is clear that $n \geq 2$. For each $i \in N$,

$$\mu^S(i) = \begin{cases} \mu(i) & \text{if } i \notin S, \\ \mu(i_{l+1}) & \text{if } i = i_l \text{ where } l \in \{1, \dots, n\}. \end{cases}$$

For each $l \in \{1, \dots, n\}$, let $o_l \equiv \mu(i_l)$. Since μ is stable and $o_l = \mu^S(i_{l-1}) P_{i_{l-1}} \mu(i_{l-1})$, for each $l \in \{1, \dots, n\}$, $i_l \succ_{o_l} i_{l-1}$ and $|\mu(o_l)| = q_{o_l}$, where we identify $i_0 \equiv i_n$. Let $N_{o_l} \equiv \mu(o_l) \setminus \{i_l\}$. Since $|\mu(o_l)| = q_{o_l}$, $|N_{o_l}| = q_{o_l} - 1$. Since $\mu(i_{l-1}) \neq o_l$, $i_{l-1} \notin N_{o_l}$. Since μ is stable, for each $k \in N_{o_l}$, $k \succ_{o_l} i_{l-1}$. Thus, for each $l \in \{1, \dots, n\}$, $i_l, i_{l-1} \notin N_{o_l}$. Further, for each pair $o, o' \in \{o_1, \dots, o_n\}$, N_o and $N_{o'}$ are disjoint.

Suppose that μ^S is not stable. Since it is non-wasteful, it violates priorities. Without loss of generality, there is j such that $j \succ_{o_2} i_1$ and $o_2 P_j \mu^S(j)$. However, since $\mu^S(j) R_j \mu(j)$ and μ is stable, for each $k \in \mu(o_2)$, $k \succ_{o_2} j$. In particular, for each $k \in N_{o_2} \subseteq \mu(o_2)$, $k \succ_{o_2} j$, and $i_2 \succ_{o_2} j$. Thus, $i_2 \succ_{o_2} j \succ_{o_2} i_1$. We have that \succsim is strongly* acyclic, N_{o_2} and N_{o_3} are disjoint, $i_2, i_1, j \notin N_{o_2}$, and $i_2 \notin N_{o_3}$. Then either $i_2 \succ_{o_3} i_1$ or $i_1 \in N_{o_3}$. If $i_1 \in N_{o_3}$, then by definition of N_{o_3} , $\mu(i_1) = o_3$, and so $o_3 = o_1$. But $i_1 \notin N_{o_1} = N_{o_3}$, a contradiction. Thus $i_2 \succ_{o_3} i_1$, so $i_3 \succ_{o_3} i_2 \succ_{o_3} i_1$. Again, we have that \succsim is strongly* acyclic, N_{o_3} and N_{o_4} are disjoint, $i_3, i_2, i_1 \notin N_{o_3}$, and $i_3 \notin N_{o_4}$. Then either $i_3 \succ_{o_4} i_1$ or $i_1 \in N_{o_4}$. If $i_1 \in N_{o_4}$, then by definition of N_{o_4} , $\mu(i_1) = o_4$, and so $o_4 = o_1$. But $i_1 \notin N_{o_1} = N_{o_4}$, a contradiction. Thus $i_3 \succ_{o_4} i_1$ so that $i_4 \succ_{o_4} i_3 \succ_{o_4} i_1$. Repeating the argument, we have $i_n \succ_{o_n} i_{n-1} \succ_{o_n} i_1$. However, since $i_1, i_n, i_{n-1} \notin N_{o_n}$, $i_1, i_n \notin N_{o_1}$, and $i_1 \succ_{o_1} i_n$, contradicting the assumption that \succsim is strongly* acyclic. Thus, μ^S is stable.

Second, we show that strong* acyclicity is necessary for each stable-dominating allocation to be stable. Suppose \succsim contains the weak* cycle: there exists $o_1, o_2 \in O$,

$i \succ_{o_1} j \succ_{o_1} k \succ_{o_2} i$, where $N_{o_1} \subseteq N \setminus \{i, j, k\}$, $|N_{o_1}| = q_{o_1} - 1$, $N_{o_1} \subseteq \{m \in N : m \succ_{o_1} j\}$ and $N_{o_2} \subseteq N \setminus \{i, k\}$, $|N_{o_2}| = q_{o_2} - 1$, $N_{o_2} \subseteq \{m \in N : m \succ_{o_2} j\}$.

If $j \notin N_{o_2}$ or $q_{o_2} = 1$, then this weak* cycle is also a weak cycle. Then by [Ehlers and Erdil \[2010\]](#), there exists a constrained efficient stable allocation that is not efficient, and so there exists a stable-dominating allocation that is not stable.

So, assume $q_{o_2} > 1$ and $j \in N_{o_2}$. Notice that $j \succ_{o_2} i$. Consider the preference profile given below, where l is a generic agent in N_{o_1} , m is a generic agent in $N_{o_2} \setminus \{j\}$, and every agent not amongst $\{i, j, k\} \cup N_{o_1} \cup N_{o_2}$ ranked \emptyset at the top:

P_i	P_j	P_k	P_l	P_m
o_2	o_1	o_1	o_1	o_2
o_1	o_2	o_2		

Define allocation μ by $\mu(i) = \mu(l) = o_1$ and $\mu(j) = \mu(k) = \mu(m) = o_2$, for every $l \in N_{o_1}$ and $m \in N_{o_2} \setminus \{j\}$. Every other agent is left unmatched. Notice that μ is stable at the given preference profile. Define allocation $\hat{\mu}$ by $\hat{\mu}(i) = \hat{\mu}(j) = \hat{\mu}(m) = o_2$ and $\hat{\mu}(k) = \hat{\mu}(l) = o_1$, for every $l \in N_{o_1}$ and $m \in N_{o_2} \setminus \{j\}$. Every other agent is left unmatched. Clearly $\hat{\mu}$ Pareto-improves μ , and so it stable-dominating. However, j blocks $\hat{\mu}$ with o_1 , so $\hat{\mu}$ is unstable. \square

An immediate implication of Proposition 7 is that, if priorities are strongly* acyclic, then the requirement that a mechanism be stable-dominating is equivalent to the requirement that it be stable. In the proof of the proposition, we see that each “improving cycle” is a “stable improving cycle” [[Erdil and Ergin, 2008](#)] under the strong* acyclicity. Under the weaker condition of [Ehlers and Erdil \[2010\]](#), one can only show that whenever there is an improving cycle, there is at least one stable improving cycle.

While we cannot pin down all stable-dominating mechanisms for general priorities, we are able to say more if they are also group strategy-proof. For the following result we assume a stronger richness condition on the domain of preferences than the one stated in Section 2. If for each $i \in N$ and each $o \in O$, there is $P_i \in \mathcal{P}_i$ such that for each $o' \in O \setminus \{o\}$, $o P_i \not\geq P_i o'$, then \mathcal{P} is rich*.⁵⁷

Proposition 8. *Suppose that \mathcal{P} is rich*. If a mechanism is group strategy-proof and stable-dominating, then it is stable.*

Proof. If φ is not stable, there is $P \in \mathcal{P}$ such that $\varphi(P)$ is not stable. Let $\mu \equiv \varphi(P)$. Since φ is stable-dominating, there is $\underline{\mu} \in \mathcal{F}$ such that $\underline{\mu}$ is stable and μ Pareto-improves it. By Lemma 3, $\underline{\mu}$ is non-wasteful. So by the Non-wasteful Structure Lemma, μ is non-wasteful.

⁵⁷ This assumption is implicit in the analysis of [Abdulkadiroğlu et al. \[2009\]](#).

Since μ is not stable, it does not respect priorities. So there are a pair $i, j \in N$ and $o \in O$ such that $\mu(i) = o$, $j \succ_o i$, and $o P_j \mu(j)$ and $|\mu(o)| = q_o$. Since $j \succ_o i$, $|\{k \in N \setminus \{j\} : \mu(k) = o \text{ and } k \succ_o j\}| \leq q_o - 1$.

Let $S \equiv N \setminus \{j\}$. By the richness* assumption on \mathcal{P} , there is $\bar{P}_S \in \times_{k \in S} \mathcal{P}_k$ such that, for each $k \in S$ and each $o \in O \setminus \{\mu(k)\}$, $\mu(k) \bar{P}_k \emptyset \bar{P}_k o$. Then $|\{k \in N \setminus \{j\} : o \bar{P}_k \emptyset \text{ and } k \succ_o j\}| \leq q_o - 1$. So for each $v \in \mathcal{F}$, if v is stable at (\bar{P}_S, P_j) , then $v(j) R_j o$.

Let $\bar{\mu} = \varphi(\bar{P}_S, P_j)$. By definition of \bar{P}_S , for each $k \in S$, $\mu(k) \bar{R}_k \bar{\mu}(k)$. If there is $k \in S$ such that $\bar{\mu}(k) \neq \mu(k)$, then $\mu(k) \bar{P}_i \bar{\mu}(k)$. This contradicts the group strategy-proofness of φ , since S may beneficially report P_S when the true preferences are \bar{P}_S . Thus, for each $k \in S$, $\bar{\mu}(k) = \mu(k)$. Since $\bar{\mu}$ Pareto-improves a stable allocation, $\bar{\mu}(j) R_j o P_j \mu(j)$. Thus, $\bar{\mu}$ Pareto-improves μ . This contradicts the group strategy-proofness of φ , since N may beneficially report (\bar{P}_S, P_j) when the true preferences are P . \square

The broad class of group strategy-proof and Pareto-efficient mechanisms defined by [Pycia and Ünver \[2017\]](#) for the subdomain of \mathcal{P} where agents rank \emptyset below each $o \in O$ are readily extended to \mathcal{P} . Proposition 8 says that unless the priorities are such that there is a group strategy-proof and Pareto-efficient mechanism that is also stable, none of these mechanisms—including the *top trading cycles* mechanism [[Abdulkadiroğlu and Sönmez, 2003](#)], which interpret priorities as providing ownership and not just consumption rights—can be justified on the basis being stable-dominating. In the case of unit capacity for objects, [Han \[2015\]](#) characterizes the priority structures that admit mechanisms that are group strategy-proof, Pareto-efficient, and stable. By Proposition 8, stability can be weakened to stable-dominating in this characterization.⁵⁸

For strict priorities, in light of Proposition 8, the agent-optimal stable mechanism is the unique candidate for a group strategy-proof and stable-dominating rule (Corollary 3). Thus, there exists a group strategy-proof and stable-dominating rule if and only if the priority structure is Ergin-acyclic [[Ergin, 2002](#)]. Moreover, the top trading cycles mechanism is stable-dominating (or stable) if and only if it coincides with the agent-optimal stable mechanism. In turn, the top trading cycles mechanism coincides with the agent-optimal stable mechanism if and only if the priority structure satisfies Kesten-acyclicity [[Kesten, 2006](#)].

⁵⁸ [Han \[2015\]](#) and [Ehlers and Westkamp \[2016\]](#) also give necessary and sufficient conditions for existence of a strategy-proof mechanism that is stability-constrained Pareto-efficient.

D Multiple Stable and Strategy-proof mechanisms without IRC

We provide an example of a capacity-based setting where the ranking according to which agents are chosen depends upon the agents being compared.

Consider a situation where there are two positions for teachers at one school o . There are four candidates $N \equiv \{m_1, m_2, p_1, p_2\}$. There is only one term each teacher can be hired under, so T is a singleton. Let C_o be a single-valued choice correspondence described by the following process. Two of the teachers, m_1 and m_2 , specialize in math and the other two, p_1 and p_2 , specialize in physics. The math teachers are able to teach physics but not as well as the physics teachers and vice versa. As overall teachers, m_2 is the best, followed by p_1 , m_1 , and p_2 , in that order. If more math specialists are being considered than physics specialists, then the math faculty are more likely to weigh in, so the positions are filled according to how good the candidates are as math teachers. Vice versa if there are more physics specialists. If there are equal numbers of math and physics specialists, the candidates are compared based on their overall teaching ability.

Below, the boxed elements show the choices from each set of candidates.

$$\begin{array}{c}
 \{m_1, \boxed{m_2, p_1}, p_2\} \\
 \\
 \{\boxed{m_1, m_2}, p_1\} \quad \{\boxed{m_1, m_2}, p_2\} \quad \{m_1, \boxed{p_1, p_2}\} \quad \{m_2, \boxed{p_1, p_2}\} \\
 \\
 \{\boxed{m_1, m_2}\} \quad \{\boxed{m_1, p_1}\} \quad \{\boxed{m_2, p_1}\} \quad \{\boxed{m_1, p_2}\} \quad \{\boxed{m_2, p_2}\} \quad \{\boxed{p_1, p_2}\} \\
 \\
 \{\boxed{m_1}\} \quad \{\boxed{m_2}\} \quad \{\boxed{p_1}\} \quad \{\boxed{p_2}\}
 \end{array}$$

Though C satisfies our assumptions of size monotonicity and idempotence, it violates IRC.

For each $S \subseteq N$ such that $|S| \leq 2$, let $\mu^S \in \mathcal{F}$ be such that it assigns agents in S to o and leaves the others unassigned. That is, $\mu^S(o) = S$ and for each $i \in N \setminus S$, $\mu^S(i) = \emptyset$. For each $P \in \mathcal{P}$, let $G(P) \equiv \{i \in N : o P_i \emptyset\}$.

Consider the mechanism φ defined by setting, for each $P \in \mathcal{P}$,

$$\varphi(P) \equiv \begin{cases} \mu^{\{m_1, m_2\}} & \text{if } \{m_1, m_2\} \subseteq G(P), \\ \mu^{\{p_1, p_2\}} & \text{if } \{p_1, p_2\} \subseteq G(P) \text{ and } \{m_1, m_2\} \not\subseteq G(P), \\ \mu^{G(P)} & \text{otherwise.} \end{cases}$$

Claim 2. φ is stable and strategy-proof.

Proof. We first establish that φ is stable by considering four cases.

Case 1: $m_1, m_2 \in G(P)$. Then $\varphi(P) = \mu^{\{m_1, m_2\}}$. Regardless of whether $p_1, p_2 \in G(P)$, there is no $Y \subseteq G(P) \setminus \{m_1, m_2\}$ such that $Y \subseteq C_o(\mu^{\{m_1, m_2\}} \cup Y)$. Thus, $\varphi(P)$ is stable.

Case 2: $m_1 \notin G(P)$ but $m_2 \in G(P)$. If $p_1, p_2 \in G(P)$, then $\varphi(P) = \mu^{\{p_1, p_2\}}$. Since $C_o(\{m_2, p_1, p_2\}) = \{p_1, p_2\}$, $\varphi(P)$ is stable. Otherwise, $\varphi(P) = \mu^{G(P)}$ and each agent receives his top choice. Thus $\varphi(P)$ is stable.

Case 3: $m_1 \in G(P)$ but $m_2 \notin G(P)$. This is symmetric to Case 2.

Case 4: $m_1, m_2 \notin G(P)$. Since, $\varphi(P) = \mu^{G(P)}$ and each agent receives his top choice, $\varphi(P)$ is stable.

To show that φ is strategy-proof, we again consider the same four cases.

Case 1: $m_1, m_2 \in G(P)$. Then $\varphi(P) = \mu^{\{m_1, m_2\}}$ and neither m_1 nor m_2 can benefit by misreporting his preferences. Regardless $P_{\{p_1, p_2\}}$, φ selects $\mu^{\{m_1, m_2\}}$ so neither of p_1 or p_2 can benefit by misreporting his preference either.

Case 2: $m_1 \notin G(P)$ but $m_2 \in G(P)$. If $p_1, p_2 \in G(P)$, then $\varphi(P) = \mu^{\{p_1, p_2\}}$, so neither p_1 nor p_2 benefits by misreporting and since φ selects $\mu^{\{p_1, p_2\}}$ regardless of m_2 's preference, he has no incentive to misreport either. Otherwise, $\varphi(P) = \mu^{G(P)}$ and no agent can benefit by misreporting since he receives his top choice.

Case 3: $m_1 \in G(P)$ but $m_2 \notin G(P)$. This is symmetric to Case 2.

Case 4: $m_1, m_2 \notin G(P)$. Since $\varphi(P) = \mu^{G(P)}$, no agent can benefit by misreporting since he receives his top choice. \square

Now, consider the mechanism φ' defined by setting, for each $P \in \mathcal{P}$,

$$\varphi'(P) \equiv \begin{cases} \mu^{\{p_1, p_2\}} & \text{if } \{p_1, p_2\} \subseteq G(P), \\ \mu^{\{m_1, m_2\}} & \text{if } \{m_1, m_2\} \subseteq G(P) \text{ and } \{p_1, p_2\} \not\subseteq G(P), \\ \mu^{G(P)} & \text{otherwise.} \end{cases}$$

Since it is symmetric to φ , φ' is also stable and strategy-proof. In fact, both of these mechanisms are group strategy-proof. Neither of these mechanisms is generated by a cumulative offer algorithm [Hatfield and Milgrom, 2005], which, regardless of the order, outputs the unstable allocation $\mu^{\{m_2, p_1\}}$ for each $P \in \mathcal{P}$ such that $G(P) = N$.

E Waiver Algorithm

The procedure ‘‘WaiverOrder’’ (Algorithm 1) takes as input an endowment, ω , a preference profile, P , and a waiver order over the agents, \triangleright . The output is a non-wasteful

allocation where agents drop and pick up available objects in the order of \triangleright , starting from ω .

Algorithm 1 Procedure to waive and pick up objects on waivers

```

1: procedure WAIVERORDER( $\omega, P, \triangleright$ )
2:    $\mu = \emptyset$ 
3:   for  $i \in N$  do
4:      $\mu(i) = \begin{cases} \omega(i) & \text{if } \omega(i) P_i \emptyset \\ \emptyset & \text{otherwise} \end{cases}$ 
5:    $A = \{o \in O : \mu(o) = \emptyset\}$  ▷ Set of unassigned objects.
6:   while  $\mu$  is wasteful at  $P$  do
7:     for  $i = \text{first in } \triangleright$  to  $i = \text{last in } \triangleright$  do
8:       for  $a \in A$  do
9:         if  $a P_i \mu(i)$  then
10:            $A = (A \setminus a) \cup \mu(i)$  ▷ Remove  $a$  from and add  $\mu(i)$  to  $A$ .
11:            $\mu(i) = a$  ▷ Assign  $a$  to  $i$ .
12:           go to 7
13:   return  $\mu$ 

```
