Stable-dominating rules*

Samson Alva
University of Texas at San Antonio
samson.alva@gmail.com

Vikram Manjunath
University of Ottawa
vikram@dosamobile.com

September 4, 2019

Abstract

We consider a general model of indivisible goods allocation with choice-based priorities, as well as the special case of school choice. Stability is the main normative consideration for such problems. However, depending on the priority structure, it may be incompatible with Pareto-efficiency. We propose a new criterion: an allocation is *stable-dominating* if it weakly Pareto-improves *some* stable allocation.

We show that if an allocation Pareto-improves on a particular non-wasteful (and therefore stable) allocation, then it matches the same agents and matches the same number of agents to each object. This is much like the conclusion of the Rural Hospitals Theorem. In fact, we connect the existence of a stable-dominating and strategy-proof rule and the Rural Hospitals Theorem on one hand with the existence of the agent-optimal stable-dominating rule on the other.

For the school choice model, we also characterize the weak priority structures that ensure every Pareto-efficient and stable-dominating rule is stable.

Finally, for the school choice model, we show that if a rule is Pareto-efficient, stable-dominating and strategy-proof, then it is actually stable. We also show an alternative version of this result where we replace Pareto-efficiency with a mild regularity condition.

Keywords: stable-dominating, rural hospitals theorem, school choice, stability,
JEL Codes: C78; D47; D71; D82

1 Introduction

We study the allocation of indivisible goods to agents, where the terms of an agent's assignment to an object may vary and each object is associated with a choice correspondence describing priorities. The school choice problem is a special case with a single

*Some results in this paper were initially reported in an early working paper that included Alva and Manjunath [2019].

assignment-term and each object has a defined capacity and a ranking of agents (possibly with ties) as its priority criterion.

The typical notion of fairness in the school choice setting is that of *no justified envy* [Abdulkadiroğlu and Sönmez, 2003]. An agent envies another, at a particular allocation, if she would rather be assigned the object that the other agent is assigned instead of what she is actually assigned . Her envy is "justified" if she has "higher priority" at the object than the agent she envies. A stable allocation is one that is fair in this sense as well as non-wasteful[1] and individual rational.[2] For the general choice-based model, stability is defined with respect to the profile of choice correspondences of objects, in a manner that is consistent with its definition for the special setting of school choice. Stability has been the main normative consideration in the literature on such problems. However, depending on agents' preferences and the priority structure (or choice profile), preventing justified envy (or instability) may keep the allocation off of the Pareto frontier [Balinski and Sönmez, 1999].

We propose a less demanding criterion for fairness: rather than insist an allocation be stable itself, it need only (weakly) Pareto-improve a stable allocation. We call such allocations *stable-dominating* allocations. Consider the following thought experiment. Suppose that an allocation μ is stable, yet another allocation ν Pareto-improves it. Starting at μ as a reference point, if we move to ν, then no agent is worse off but some agents are better off. The only situation where such a move may not be acceptable, given the standard normative criteria we have listed above, is when there is an agent i who is not harmed by the move yet justifiably envies someone at ν. However, since μ was stable, i did not justifiably envy anyone at μ. Our proposal is that such ν is not so bad after all. Agent i would not object to the allocation μ, which is *inferior* in the Pareto sense. In particular, μ offers a remedy to the situation that i may complain about ν, but a remedy that does not actually benefit i. In other words, by offering the remedy of μ, one can make it unappealing for i to complain about ν. Allocations like ν are what we call stable-dominating.

In this paper, we try to better understand what allocations are stable-dominating and how to reconcile this requirement with that of good incentives.

First, we investigate some structural properties of the set of stable-dominating allocations in the general choice-based model. We show that if an allocation Pareto-improves a non-wasteful allocation, then A) an agent is assigned to an object at the dominating allocation if and only if she is assigned to an object at the improved allocation, B) every object is assigned to the same number of agents by both allocations, and C) if an object

[1] No object is left unassigned while some agent would prefer it to their assignment.

[2] No agent is assigned an object they find worse than being unassigned.

is not filled to capacity by the improved allocation, then it is assigned to precisely the same agents by the dominating allocation. Using this result, we characterize the priority structures in school choice problems for which the stable-dominating set coincides with the stable set. The condition on priorities that guarantees this is slightly stronger than the acyclicity condition of Ehlers and Erdil [2010].

We then turn to incentives. For the school choice setting, when priorities are strict (that is, no two agents are tied in the ranking of any object), there exists a unique agent-optimal stable-dominating rule [Alva and Manjunath, 2019]. For such priorities, the set of stable allocations satisfies the "Rural Hospitals Theorem," which states that an agent is assigned to a school at one stable allocation if and only if she is assigned to a school at *every* stable allocation [Roth, 1986]. Finally, the rule that selects the agent-optimal stable allocation is strategy-proof (that is, it provides incentives for every agent to truthfully report her preferences) [Dubins and Freedman, 1981]. We show for the general choice-based model that, given a profile of choice correspondences of objects satisfying two conditions,[3] the following statements are equivalent:

1. There exists an agent-optimal stable allocation for every profile of preferences at the given choice profile.

2. At the given profile of choice correspondences, A) there is a unique stable-dominating and strategy-proof rule and B) the Rural Hospitals Theorem holds.

In particular, the agent-optimal stable rule is the unique stable-dominating and strategy-proof rule if the second statement above holds.

Finally, for the school choice model, we show that a stable-dominating rule can be Pareto-efficient strategy-proof only if it is actually stable. Consequently, amongst the broad class of known strategy-proof and Pareto-efficient rules [Pápai, 2000, Pycia and Ünver, 2017], those rules that can be justified on the basis of being stable-dominating for some priority structure are only those that are actually stable. We also show that subject to a mild regularity condition, the only strategy-proof and stable-dominating rules are actually stable.

The lesson here is that even though weakening stability to stable-domination affords the designer some latitude to improve agents' welfare for cyclic priorities, the requirement of strategy-proofness impedes the realization of such gains.

Related Literature: Given the incompatibility of stability and efficiency [Balinski and Sönmez, 1999], there is a growing interest in notions of fairness weaker than stability.

[3] The two conditions are idempotence and size monotonicity.

Dur et al. [forthcoming] identify interesting Pareto-efficient rules that satisfy partial fairness, which weakens stability to allow instances of justified envy from an exogenously-specified collection. There exist allowable priority violations such that a partially fair allocation is not stable-dominating and vice versa. Tang and Zhang [2017] define an allocation as weakly stable if the only instances of justified envy that remain are those that cannot be resolved without creating new instances of justified envy, and relate weakly stable allocations to the outcomes of a generalized version of the efficiency-adjusted deferred acceptance algorithm [Kesten, 2010, Tang and Yu, 2014]. In some problems, there are weakly stable and Pareto-efficient allocations that are not stable-dominating, and stable-dominating allocations that are not weakly stable. Troyan et al. [2018] identify some instances of justified envy as vacuous. Roughly, an instance is vacuous if resolving it, and iteratively resolving any subsequently-created instance, results in an allocation that leaves the original envying agent worse off. They study the structure of the set of essentially stable allocations, those that admit only vacuous instances of justified envy, and show that the efficiency-adjusted deferred acceptance algorithm is essentially stable. They show that every essentially stable allocation is stable-dominating, but that the converse is not true. Ehlers and Morrill [2018] define the notion of legality, a property of a set of allocations, rather than of an allocation in isolation. They show, for the matching with contracts setting with substitutable and size-monotonic choice functions, that there is a unique legal set, that it contains the stable set and has a lattice structure, and that the efficiency-adjusted deferred acceptance algorithm produces the Pareto-optimal allocation in the set.

Finally, our results imply that even weakening the requirement of stability to stable-domination does not alleviate the tension between Pareto-efficiency and strategy-proofness for arbitrary priorities. However, Troyan and Morrill [2019] have recently proposed a weakening of strategy-proofness based on manipulations that are in a certain sense obvious. They show that, subject to a mild regularity condition, no stable-dominating rule is obviously manipulable. Subject to an even weaker regularity condition, we show that the only stable-dominating and strategy-proof rules are also stable.

The remainder of the paper is organized as follows. We introduce a general choice-based version of our matching model in Section 2.1 and the classical object allocation model, which is a special case, in Section 2.2. All of our results are in Section 3: structural results in Section 3.1 and those on incentives in Section 3.2.

2 Models

We start with a very general choice-based matching model. In Section 2.2, we specialize this to the classical object allocation model and to school choice since the additional structure allows us to prove stronger results.

2.1 Choice-based matching

Let $\boldsymbol{N}$ be a finite and nonempty set of **agents**, $\boldsymbol{O}$ be a finite and nonempty set of **objects**, $\boldsymbol{T}$ be a nonempty set of **terms** under which an agent may be assigned an object, $\boldsymbol{X} \subseteq N \times O \times T$ be a nonempty set of possible **contracts**, $\boldsymbol{\mathcal{F}_N}$ be a nonempty set of **feasible allocations for agents**, $\boldsymbol{\mathcal{F}_O}$ be a nonempty set of **feasible allocations for objects**, and $\boldsymbol{C_O}$ be a **combinatorial choice correspondence**. A choice-based matching model is a tuple $(N, O, T, X, \mathcal{F}_N, \mathcal{F}_O, C_O)$.

Each contract $x \in X$ is a triple $(i, o, t) \in N \times O \times T$ that represents "i consumes o under the terms t." Let $\boldsymbol{N(x)}$ be the agent associated with x and $\boldsymbol{O(x)}$ be the object associated with x. For each $Y \subseteq X$, let $\boldsymbol{N(Y)}$ be the set of agents associated with triples in Y. For each $i \in N$, let $\boldsymbol{Y(i)}$ be the triples in Y associated with i. For each $o \in O$, let $\boldsymbol{Y(o)}$ be the contracts in Y associated with o.

An **allocation** μ is a subset of X. If $\mu(i)$ is empty for agent i, he consumes his **outside option**, $\emptyset$. The **participants** at allocation μ, $\boldsymbol{N(\mu)}$, are the agents associated with some contract in μ.

Both $\mathcal{F}_N$ and $\mathcal{F}_O$ are collections of allocations, i.e. subsets of X. Let $\boldsymbol{\mathcal{F}_i}$ be the set of all $\mu \in \mathcal{F}_N$ at which i participates. From $\mathcal{F}_O$ we define, for each $o \in O$, the **feasible sets for $\boldsymbol{o}$** $\boldsymbol{\mathcal{F}_o}$, which is the collection of subsets of $X(o)$ that are included in some allocation μ in $\mathcal{F}_O$. The set of **feasible allocations** is $\boldsymbol{\mathcal{F}} = \mathcal{F}_N \cap \mathcal{F}_O$.

The combinatorial choice correspondence models information about how sets of contracts are prioritized at and across objects. The correspondence $C_O : 2^X \rightrightarrows 2^X$ satisfies (1) for each $Y \subseteq X$, $C_O(Y) \subseteq 2^Y$, and (2) the range of C_O is a subcollection of $\mathcal{F}_O$.[4] Condition (1) is the defining condition of a choice correspondence, requiring that, from any given set of contracts, C_O picks only subsets of it. Condition (2) requires any chosen set to be feasible for objects.

Each agent i has **preferences** R_i. Let $\boldsymbol{\tilde{\mathcal{P}}_i}$ be the set of all possible preferences of i. A **problem** is defined by a profile of preferences R from a **domain** of preference profiles $\mathcal{P} \subseteq \times_{i \in N} \tilde{\mathcal{P}}_i$. A **rule** φ is a map from the domain $\mathcal{P}$ to feasible allocations $\mathcal{F}$. If for $R \in \mathcal{P}$ and $\mu \in \mathcal{F}$ is such that $\mu = \varphi(R)$, then for each $i \in N$ denote by $\varphi_i(R)$ the set $\mu(i)$.

[4] The range of $C_O : 2^X \rightrightarrows 2^X$ is $\{Z : Y \subseteq X, Z \in C_O(Y)\}$.

Assumptions on model primitives In any allocation feasible for agents ($\mu \in \mathcal{F}_N$), no agent has more than one contract (for all $i \in N$, $|\mu(i)| \leq 1$). In any allocation feasible for objects ($\mu \in \mathcal{F}_O$), no object has more than one contract with any given agent (for all $o \in O$ and $i \in N$, $|\mu(o) \cap \mu(i)| \leq 1$). Finally, $\mathcal{F}_N$ and $\mathcal{F}_O$ are such that $\mathcal{F}$ is nonempty.

The preference domain $\mathcal{P}$ is Cartesian, that is, for each $i \in N$, there exists $\mathcal{P}_i \subseteq \tilde{\mathcal{P}}_i$ such that $\mathcal{P} = \times_{i \in N} \mathcal{P}_i$. Each agent i's preference $R_i \in \mathcal{P}_i$ depends only upon his own contract in an allocation and are *strict*. These preferences are represented by linear orders over $X(i) \cup \{\emptyset\}$, and $\mathcal{P}_i$ is a collection of such linear orders.

For each $i \in N$, each $R_i \in \mathcal{P}_i$, and each $x \in X(i)$ such that $x \; P_i \; \emptyset$, if $R_i' \in \mathcal{P}_i$ is such that (1) $x \; P_i' \; \emptyset$, and (2) for each $z \in X(i)$, $x \; P_i \; z$ if and only if $\emptyset \; P_i' \; z$. , then R_i' is a **weak truncation of $\boldsymbol{R_i}$ at $\boldsymbol{x}$**. If, in addition, for each pair $y, z \in X(i)$, $z \; P_i \; y \; P_i \; x$ if and only if $z \; P_i' \; y \; P_i' \; x$, then R_i' is a **truncation of $\boldsymbol{R_i}$ at $\boldsymbol{x}$**. We assume that $\mathcal{P}$ satisfies the following condition, which we call **truncatability**: For each $i \in N$, each $R_i \in \mathcal{P}_i$, and each $x \in X(i)$ such that $x \; P_i \; \emptyset$, the truncation of R_i at x is in $\mathcal{P}_i$. A weaker requirement, which we call **weak truncatability** only requires the $\mathcal{P}_i$ to contain a *weak* truncation of R_i at x.

Truncatability is much weaker than the typical assumption in the literature that $\mathcal{P}_i$ contains all strict preferences.

We impose the following restrictions on choice correspondences. Say C_O is **size monotonic** if, for each $Y \subseteq X$, each finite $Y' \subseteq Y$, each $Z \in C_O(Y)$, and each $Z' \in C_O(Y')$, $|Z| \geq |Z'|$.[5,6] Say C_O is **idempotent** if, for each $Y \in \text{range}(C_O)$, $Y \in C_O(Y)$.[7]

Feasibility conditions Define the Cartesian hull of $\mathcal{Y} \subseteq 2^X$ across O, denoted $\bigsqcup_O \mathcal{Y}$, to be $\{Y \subseteq X : \forall o \in O, \exists Z \in \mathcal{Y}, Y(o) \subseteq Z\}$. We say that $\mathcal{F}_O$ is **Cartesian** if $\mathcal{F}_O = \bigsqcup_O \mathcal{F}_O$. When $\mathcal{F}_O$ is not Cartesian, there are cross-object constraints. We say that $\mathcal{F}_O$ is **capacity-based** if for each $o \in O$, there exists $\boldsymbol{q_o} \in \mathbb{Z}_+$, its capacity, such that for every $Y \subseteq X(o)$ such that for each $i \in N$, $|Y(i)| \leq 1$, $Y \in \mathcal{F}_o$ if and only if $|Y| \leq q_o$.

2.2 Classical object allocation and school choice

The *classical* object allocation model is a tuple $(N, O, (q_o)_{o \in O})$, where N is a finite and nonempty set of agents, O is a finite and nonempty set of objects, and for each $o \in O$, q_o is a non-negative integer **capacity** of o. An allocation μ is a function from N to $O \cup \{\emptyset\}$ such that $|\mu^{-1}(o)| \leq q_o$ for each $o \in O$, where $\emptyset$ represents receiving no object. Since it

[5] 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].

[6] For finite Y, setting $Y' = Y$, size monotonicity implies that for each pair $Z, Z' \in C_O(Y)$, $|Z| = |Z'|$.

[7] This rules out, for instance, C_O such that $C_O(\{x, y, z\}) = \{\{x, y\}\}$ but $C_O(\{x, y\}) = \{\{x\}\}$.

should not cause confusion, denote the set $\mu^{-1}(o)$ by $\mu(o)$. Each agent has a preference representable by a linear order over $O \cup \{\varnothing\}$. For each agent i, let $\tilde{\mathcal{P}}_i$ be the set of all linear orders on $O \cup \{\varnothing\}$. Let P_i be the strict component of $R_i \in \tilde{\mathcal{P}}_i$. A domain is $\mathcal{P} = \times_{i \in N} \mathcal{P}_i$, where $\mathcal{P}_i \subseteq \tilde{\mathcal{P}}_i$ for each $i \in N$.

Every classical object allocation model can be represented in the general model as follows: T is a singleton set $\{t\}$, X is $N \times O \times \{t\}$, and $\mathcal{F}_O = \bigcup_{o \in O} \{Y \subseteq X(o) : |Y| \leq q_o\}$, that is, $\mathcal{F}_O$ is Cartesian and capacity-based. Clearly, each agent's preference over objects in a classical object allocation model can be uniquely represented as a preference over triples.

School choice and priorities A school choice model $(N, O, (q_o, \succsim_o)_{o \in O})$ is a classical object allocation model $(N, O, (q_o)_{o \in O})$ with **priorities $(\succsim_o)_{o \in O}$**, where each $\succsim_o$ is a complete, reflexive, and transitive binary relation on N. We refer to agents as students and objects as schools. Denote the strict component of $\succsim_o$ by $\succ_o$. Let $\succsim \equiv (\succsim_o)_{o \in O}$ be a list of priorities. For each $o \in O$, if o's capacity is q_o and priority is $\succsim_o$,

Each school choice model can be represented in the general model just as a classical object allocation with $C_O : 2^X \rightrightarrows 2^X$ defined as follows: for each $Y \subseteq X$, $C_O(Y) = \bigcup_{o \in O} C_o(Y)$, where

$$C_o(Y) = \begin{cases} \{Y(o)\} & \text{if } |Y(o)| \leq q_o, \\ \{Z \subseteq Y(o) : |Z| = q_o \text{ and for each } z \in Z \text{ and } y \in Y(o) \setminus Z, N(z) \succsim_o N(y)\} & \text{otherwise.} \end{cases}$$

Note that C_O is size monotonic and idempotent.

2.3 Properties of Allocations and Rules

In what follows, rules inherit properties defined for allocations in a pointwise fashion. That is, rule φ has a particular property if for each $R \in \mathcal{P}$, $\varphi(R)$ has the property. We extend relations defined amongst allocations to rules in a pointwise fashion as well. That is, given a relation on allocations, a rule φ is related to another rule φ' if for each $R \in \mathcal{P}$, $\varphi(R)$ is related to $\varphi'(R)$.

Pareto-improvement One allocation Pareto-improves another if each agent finds the first at least as desirable as the second. That is, for each $R \in \mathcal{P}$ and each pair $\mu, \nu \in \mathcal{F}$, μ Pareto-improves ν at R if, for each $i \in N$, $\mu(i)\ R_i\ \nu(i)$.[8] If μ Pareto-improves ν at R and

[8] In this case, some authors say that μ *weakly* Pareto-improves ν. However, since this is the primary Pareto-improvement relation that we consider, we drop the qualifier.

there is $i \in N$ such that $\mu(i)\ P_i\ \nu(i)$, then α **strictly Pareto-improves** β at R. If $\mu \in \mathcal{F}$ is such that no allocation strictly Pareto-improves it at R, then μ is **Pareto-efficient** at R.

Individual Rationality An allocation is individually rational at preference profile R if there is no participating agent who prefers the outside option to their assignment. That is, $\mu \in \mathcal{F}$ is individually rational at R if for each $i \in N$, $\mu(i)\ R_i\ \varnothing$.

Non-wastefulness For the classical object allocation model, Balinski and Sönmez [1999] define non-wastefulness as follows: $\mu \in \mathcal{F}$ is non-wasteful at $R \in \mathcal{P}$ 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 our general model, we present the following extension of non-wastefulness from Alva and Manjunath [2019]. It says that an allocation is wasteful if there is a way to beneficially increase the number of agents who are assigned a particular object without making any other agent worse off. That is, given $R \in \mathcal{P}$, $\mu \in \mathcal{F}$ is *wasteful at* R 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 $j \in N \setminus \{i\}$, $\nu(j)\ R_i\ \mu(j)$, so that no agent is worse off at ν compared to μ. If it is not wasteful at R, then μ is *non-wasteful at* R.

Stability An allocation is stable (with respect C_O) if it is feasible and individually rational and if no set of agents prefers to drop their assigned contracts in favor of being assigned to new objects under some terms that the objects would prioritize. That is, allocation μ is *stable* at R (given C_O) if $\mu \in \mathcal{F}$, μ is individually rational at R,[9] and there is no $Y \subseteq X \setminus \mu$ 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*) $\mu \notin C_O(\mu \cup Y)$, (*4*) there is $\mu' \in C_O(\mu \cup Y)$ such that $Y \subseteq \mu'$, and (*5*) $\mu' \in \mathcal{F}$.[10] This definition is equivalent to the standard one if C_O is the union of a profile of single-valued combinatorial choice functions for each object and $\mathcal{F}_O$ is Cartesian. For each $R \in \mathcal{P}$, we denote by $\Sigma(\mathbf{R})$ the set of stable allocations at R.

[9] 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.

[10] 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 triple in 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 the availability of Y prevents the existing allocation μ from being chosen. The second, Condition (*4*), says that there is some chosen set, μ, that contains "blocking" set Y. That is, Condition (*3*) and Condition (*4*) together say that Y is contained in some μ' that is *revealed* by C_O to have a "higher priority" than μ. 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*) ensures that the chosen rearrangement of assignments is feasible, particularly relevant if $\mathcal{F}_O$ is not Cartesian.

Remark 1. *If choice correspondences are size monotonic and idempotent, every stable allocation is non-wasteful (Lemma 2 of Alva and Manjunath [2019]).*

For the school choice model, stability is equivalent to the combination of three properties: non-wastefulness, individually rationality, and the requirement that no agent "justifiably" envies another: if $i, j \in N$ are such that i is assigned o and j prefers o to his own assignment, then j ought not have higher priority at o than i.

Given $R \in \mathcal{P}$, if a stable allocation μ Pareto-improves every other stable allocation at R, then it is the **agent-optimal stable** allocation at R. If a stable allocation μ is Pareto-improved by every other stable allocation at R, then it is the **agent-pessimal stable** allocation at R. Under only the assumptions of size monotonicity and idempotence, such stable allocations are not guaranteed to exist. However, if C_O is such that an agent-optimal stable allocation *does* exist for each $R \in \mathcal{P}$, then we denote by $\boldsymbol{\varphi^{AOS}}$ the rule that selects this allocation. Similarly, let $\boldsymbol{\varphi^{APS}}$ select the agent-pessimal stable allocation at each $R \in \mathcal{P}$.

Stable-dominating The constraints imposed by stability may keep a stable allocation below the Pareto frontier. A less demanding requirement is that it Pareto-improves *some* stable allocation. As discussed in the introduction, the existence of a stable allocation Pareto-improved by a given allocation provides a viable defense against the objections of agents whose priorities are violated. We call any such allocation a **stable-dominating** allocation and, for each $R \in \mathcal{P}$, denote the set by $\boldsymbol{\Sigma^{\uparrow}(R)}$. Since we do not insist on strict Pareto-improvement, every stable allocation is stable-dominating. That is, for each $R \in \mathcal{P}$, $\Sigma(R) \subseteq \Sigma^{\uparrow}(R)$.

Non-manipulability A rule is **strategy-proof** if it is a weakly dominant strategy for every agent to truthfully report his preferences. That is, φ is strategy-proof if for each $R \in \mathcal{P}$, each $i \in N$, and each $R'_i \in \mathcal{P}_i$, $\varphi_i(R)\ R_i\ \varphi_i(R'_i, R_{-i})$. A rule is **strongly truncation-proof** if for each $R \in \mathcal{P}$, each $i \in N$, and each $R'_i \in \mathcal{P}_i$, if R'_i is a weak truncation of R_i at $x \in X(i)$, then $\varphi_i(R)\ R_i\ \varphi_i(R'_i, R_{-i})\ R'_i\ \varphi_i(R)$. A rule is **truncation-proof** if for each $R \in \mathcal{P}$, each $i \in N$, and each $R'_i \in \mathcal{P}_i$, if R'_i is a truncation of R_i at $x \in X(i)$, then $\varphi_i(R)\ R_i\ \varphi_i(R'_i, R_{-i})\ R'_i\ \varphi_i(R)$.

Note that by strategy-proofness, if $R \in \mathcal{P}$ and $R'_i \in \mathcal{P}_i$, then $\varphi_i(R'_i, R_{-i})\ R'_i\ \varphi_i(R)$. So, both strong truncation-proofness and truncation-proofness weaken strategy-proofness, by requiring the conclusion only for certain R'_i.

3 Results

We first present our results on the structure of stable-dominating allocations in Section 3.1. In Section 3.2 we present our results on non-manipulability.

3.1 Structure of the Stable-dominating Set

In general, the existence of a stable allocation is not guaranteed, let alone the existence of an agent-optimal stable allocation. Another property of the set of stable allocations that is not guaranteed is the **Rural Hospitals Theorem**. The Rural Hospitals Theorem states that the set of assigned agents as well as the number of agents assigned to each object is invariant across all stable allocations.

Should the choice correspondences be such that stability implies non-wastefulness and an agent-optimal stable allocation exists, then the Rural Hospitals Theorem is a consequence of the following lemma.

Lemma 1 (Structure Lemma)**.** *For each $R \in \mathcal{P}$ and each pair $\mu, \nu \in \mathcal{F}$, if μ is individually rational and non-wasteful and ν Pareto-improves μ at R, then $N(\mu) = N(\nu)$ and for each $o \in O$, $|\mu(o)| = |\nu(o)|$.*

Proof. Since μ is individually rational and non-wasteful, by Remark 4 and Lemma 1 of Alva and Manjunath [2019], $N(\mu) = N(\nu)$. Then, $|\mu| = |\nu|$.

Since $\mu(o)$ and $\mu(o')$ are disjoint for distinct $o, o' \in O$, $\sum_{o \in O} |\mu(o)| = |\mu|$. By similar reasoning, $\sum_{o \in O} |\nu(o)| = |\nu|$. Then, $\sum_{o \in O} |\mu(o)| = \sum_{o \in O} |\nu(o)|$. Since ν Pareto-improves μ and μ is non-wasteful, by the definition of non-wastefulness there are two possibilities: (1) ν does not strictly Pareto-improve μ or (2) for each $o \in O$, $|\mu(o)| \geq |\nu(o)|$. If ν does not strictly Pareto-improve μ, then strict preferences implies for each $i \in N$, $\nu(i) = \mu(i)$. That is, $\nu = \mu$. If ν strictly Pareto-improves μ, case (2) applies. Since $\sum_{o \in O} |\mu(o)| = \sum_{o \in O} |\nu(o)|$, for each $o \in O$, $|\mu(o)| = |\nu(o)|$. □

This result sheds light on what drives the Rural Hospitals Theorem. The first is non-wastefulness. The Structure Lemma says that, regardless of stability, at any pair of Pareto-comparable allocations in the individually rational and non-wasteful set, the conclusion of the Rural Hospitals Theorem holds. The only additional thing required is that every pair of stable allocations be either directly or indirectly Pareto-comparable. This is guaranteed, for instance, if the agent-optimal stable allocation exists and C_O is size monotonic and idempotent.

For models with capacity-based $\mathcal{F}_O$, like the classical object allocation model, we can further strengthen the conclusion of the Structure Lemma: every object that is not allocated to capacity by μ is allocated to the *same* agents by both μ and ν.

Lemma 2 (Capacity-based Structure Lemma)**.** *Let $\mathcal{F}_O$ be Cartesian and capacity based. For each $R \in \mathcal{P}$ and each pair $\mu, \nu \in \mathcal{F}$, if μ is individually rational and non-wasteful and ν Pareto-improves μ at R, then $N(\mu) = N(\nu)$, for each $o \in O$, $|\mu(o)| = |\nu(o)|$, and for each $o \in O$ such that $|\mu(o)| < q_o$, $N(\mu(o)) = N(\nu(o))$.*

Proof. The first two claims in the conclusion follow from the Structure Lemma. We prove the last one by contradiction.

Let $o \in O$ such that $|\mu(o)| < q_o$, and suppose $N(\nu(o)) \neq N(\mu(o))$. Since $|\nu(o)| = |\mu(o)|$, there exists $i \in N(\nu(o)) \setminus N(\mu(o))$. Let $x = \nu(i)$. Given that $N(\nu) = N(\mu)$, there is $o' \neq o$ such that $y \in \mu(o)$ and $N(y) = i$. By strict preferences, $\nu(i) = x \ P_i \ y = \mu(i)$, so ν is a strict Pareto-improvement of μ.

Since $\mathcal{F}_O$ is capacity-based and $|\mu(o)| < q_o$, $\mu(o) \cup \{x\} \in \mathcal{F}_o$ and $\mu(o') \setminus \{y\} \in \mathcal{F}_{o'}$. Let $\gamma \equiv (\mu \setminus \{y\}) \cup \{x\}$. Since $\mathcal{F}_O$ is Cartesian, $\gamma \in \mathcal{F}_O$. For each $j \in N \setminus \{i\}$, $\gamma(j) = \mu(j)$, and $\gamma(i) = x \ P_i \ y = \mu(i)$, so γ is a strict Pareto-improvement of μ. But $|\gamma(o)| > |\mu(o)|$, contradicting the assumption that μ is non-wasteful. □

Relative to the Rural Hospitals Theorem, the **strong Rural Hospitals Theorem** additionally states that the set of agents matched to an object not filled to capacity is invariant across stable allocations. It is straightforward to see for capacity-based $\mathcal{F}$ that Lemma 2 drives the strong Rural Hospitals Theorem. It also facilitates analogous variants of Propositions 2, 3, and 4, below.

Lemma 2 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 2, 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.

Now consider the school choice model. 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, the agent-optimal stable rule is not Pareto-efficient. 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 whether priorities are weak and whether they are heterogeneous across objects.

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, then it is actually Pareto-efficient.[11] 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 *a*) (loop) $i \succsim_x j \succ_x k$ and $k \succsim_y i$, and *b*) (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 $l \in N_x$, $l \succsim_x j$, each $l \in N_y$, $l \succsim_y i$, $|N_x| = q_x - 1$, 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**.

Proposition 1. *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 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 $\{i_1, \ldots, i_n\} \subseteq N$ that satisfies, for each pair $i, j \in S$, $\mu(i) \neq \mu(j)$, and for each $l \in \{1, \ldots, n\}$, $\mu(i_{l+1})\ P_{i_l}\ \mu(i_l)$, where we identify $i_{n+1} \equiv i_1$ and $i_0 \equiv i_n$. It is clear that $n \geq 2$. For each $i \in N$,

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

[11] Ehlers and Westkamp [2018] provide conditions on priorities that guarantee the existence of a strategy-proof rule that selects such a stable allocation.

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 $l \in \{1,\dots,n\}$, $i_l \succsim_{o_l} i_{l-1}$ and $|\mu(o_l)| = q_{o_l}$. 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 $k \in N_{o_l}$, $k \succsim_{o_l} i_{l-1}$. 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 and individually rational, 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 $k \in \mu(o_2)$, $k \succsim_{o_2} j$. In particular, for each $k \in N_{o_2} \subseteq \mu(o_2)$, $k \succsim_{o_2} j$, and $i_2 \succsim_{o_2} j$. Thus, $i_2 \succsim_{o_2} j \succ_{o_2} i_1$. 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 $i_3 \succsim_{o_3} i_2 \succ_{o_3} i_1$. 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 $i \in N_{o_4}$, 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 $i_4 \succsim_{o_4} i_3 \succ_{o_4} i_1$. Repeating the argument, we have $i_n \succsim_{o_n} i_{n-1} \succ_{o_n} i_1$. However, since $i_1, i_n, i_{n-1} \notin N_{o_n}$, $i_1, i_n \notin N_{o_1}$, and $i_1 \succsim_{o_1} i_n$, 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 $o_1, o_2 \in O$, $i \succsim_{o_1} j \succ_{o_1} k \succsim_{o_2} i$, where $N_{o_1} \subseteq N \setminus \{i,j,k\}$, $|N_{o_1}| = q_{o_1} - 1$, $N_{o_1} \subseteq \{m \in N : m \succsim_{o_1} j\}$ and $N_{o_2} \subseteq N \setminus \{i,k\}$, $|N_{o_2}| = q_{o_2} - 1$, $N_{o_2} \subseteq \{m \in N : m \succsim_{o_2} j\}$.

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 $j \succsim_{o_2} i$. 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 $\{i,j,k\} \cup N_{o_1} \cup N_{o_2}$ ranks $\varnothing$ 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. $\square$

The following example shows that Proposition 1 does not hold under the slightly weaker condition of Ehlers and Erdil [2010]. In the proof of the proposition, we see that each "improving cycle" is a "stable improving cycle" [Erdil and Ergin, 2008] under the stronger condition. Under the weaker condition, one can only show that whenever there is an improving cycle, there is at least one stable improving cycle.

Example 1. *Proposition 1 does not hold if $\succsim$ is only strongly acyclic.*

Let $O \equiv \{o_1, o_2\}$ and $N \equiv \{i_1, i_2, i_3\}$. Let $q_{o_1} = 1$ and $q_{o_2} = 2$. Define $\succsim$ as follows:

$\succsim_{o_1}$	$\succsim_{o_2}$
i_1, i_2	i_1, i_2, i_3
i_3	

Since there are only three agents, the scarcity condition for a weak cycle is never met. Thus, $\succsim$ is strongly acyclic despite the loop condition being satisfied. However, the scarcity condition for a weak* cycle is met, so $\succsim$ is not strongly* acyclic. Consider $P \in \mathcal{P}$ as follows:

P_{i_1}	P_{i_2}	P_{i_3}
o_2	o_1	o_1
o_1	o_2	o_2

Let $\mu \in \mathcal{F}$ be such that $\mu(i_1) = o_1$, $\mu(i_2) = o_2$, and $\mu(i_3) = o_2$. Though μ is stable, it is not Pareto-efficient. There are two Pareto-improving cycles: i_1 and i_2 trade their assignments or i_1 and i_3 trade their assignments. The latter leads to an unstable allocation.

An implication of Proposition 1 is that the requirement that a rule be stable-dominating is equivalent to the requirement that it be stable if and only if priorities are strongly* acyclic. Thus, as long as priorities are not strongly* acyclic, weakening stability to stable-domination yields welfare gains.

3.2 Incentives, Stable-dominating Rules, and the Rural Hospitals Theorem

Corollary 5 of Alva and Manjunath [2019] establishes that when priorities are strict, φ^{AOS} is the only stable-dominating rule that is strategy-proof. As we have discussed following the Structure Lemma, the existence of φ^{AOS} is a sufficient condition for the Rural Hospitals Theorem to hold. Indeed, we show in this section that the following are intimately related: (*1*) the existence of φ^{AOS}, (*2*) the existence of a stable-dominating and strategy-proof rule, and (3) the Rural Hospitals Theorem.[12]

We begin with a lemma, similar to Corollary 1 of Alva and Manjunath [2019]. In relation to that result, we weaken strategy-proofness to truncation-proofness but strengthen the domain condition to strict preferences satisfying truncatability. A pair of rules φ and φ' are defined to be **participation-equivalent** if for each $R \in \mathcal{P}$, $N(\varphi(P)) = N(\varphi'(P)))$.

Lemma 3 (Participation-equivalence Lemma)**.** *Let the preference domain satisfy truncatability. If a pair of truncation-proof and individually rational rules, φ and φ', are participation-equivalent, then they are identical.*[13]

Proof. Suppose φ and φ' are not identical. We will obtain a contradiction.

Let $R \in \mathcal{P}$ such that $\varphi(R) \neq \varphi'(R)$. By strict preferences, there are $i \in N$ and $x \in X(i)$ such that, without loss of generality, $x = \varphi'_i(R)\ P_i\ \varphi_i(R)$. By participation-equivalence and individual rationality, $\varphi_i(R)\ P_i\ \emptyset$. By truncatability, there is $R'_i \in \mathcal{P}_i$ such that R'_i is a truncation of R_i at x and $R'_i \neq R_i$. By truncation-proofness, $\varphi'_i(R'_i, R_{-i})\ R'_i\ \varphi'_i(R) = x\ R_i\ \varphi'_i(R'_i, R_{-i})$. By definition of a truncation, $\varphi'_i(R'_i, R_{-i})\ R_i\ x$. So, by strict preferences, $\varphi'_i(R'_i, R_{-i}) = x$.

By individual rationality, $\varphi_i(R'_i, R_{-i})\ R'_i\ \emptyset$. By participation-equivalence, there exists $y \in X(i)$ such that $\varphi(R'_i, R_{-i}) = y$. So by strict preferences and individual rationality, $y\ P'_i\ \emptyset$. By definition of a truncation, $y\ R_i\ x$. However, this means $\varphi_i(R'_i, R_{-i}) = y\ R_i\ x\ P_i\ \varphi_i(R)$, contradicting truncation-proofness of φ. □

First, we show existence of φ^{AOS} ensures that φ^{AOS} is the *unique* stable-dominating and truncation-proof rule and that the Rural Hospitals Theorem holds, given size monotonicity and idempotence of C_O.[14]

Proposition 2. *If φ^{AOS} exists, then it is the unique stable-dominating and truncation-proof rule and the Rural Hospitals Theorem holds.*

Proof. A stable allocation is also non-wasteful (Remark 1). By definition, for each $R \in \mathcal{P}$, $\varphi^{AOS}(R)$ Pareto-improves each stable allocation at P. Thus, by the 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.

[12] Hatfield and Kojima [2010] provide conditions on single-valued C that guarantee the first two statements.

[13] The lemma also holds if truncatability is weakened to weak truncatability and truncation-proofness is replaced by strongly truncation-proofness.

[14] Though Hirata and Kasuya [2017] show that whenever φ^{AOS} exists it is the only candidate for a stable and strategy-proof rule, Alva and Manjunath [2019] show that this is not so for our 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, and so truncation-proof, in our setting.

Next, if any other rule φ is stable-dominating, there exists a stable $\underline{\varphi}$ that both φ and φ^{AOS} Pareto-improve, where $\underline{\varphi}$ is also non-wasteful (Remark 1). Then, from the Structure Lemma, for each $R \in \mathcal{P}$, $N(\varphi(R)) = N(\underline{\varphi}(R)) = N(\varphi^{AOS}(R))$. By Participation-equivalence Lemma, if φ is truncation-proof as well, then $\varphi = \varphi^{AOS}$. Thus φ^{AOS} is the only stable-dominating and truncation-proof rule. □

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 rule that is truncation-proof. This follows from the Structure Lemma and the Participation-equivalence Lemma.

Proposition 3. *If the Rural Hospitals Theorem holds, then there is at most one stable-dominating and truncation-proof rule.*

Proof. Let φ and φ' be stable-dominating and truncation-proof rules. For each $R \in \mathcal{P}$, there exist stable $\mu, \mu' \in \mathcal{F}$ such that $\varphi(R)$ Pareto-improves μ and $\varphi'(R)$ Pareto-improves μ'. By Remark 1 and the Structure Lemma, $N(\varphi(R)) = N(\mu)$ and $N(\varphi'(R)) = N(\mu')$. By the Rural Hospitals Theorem, $N(\mu) = N(\mu')$, so $N(\varphi(R)) = N(\varphi'(R))$. Since stable-dominating rules are individually rational, by the Participation-equivalence Lemma, $\varphi = \varphi'$. □

Third, we show that if the Rural Hospitals Theorem holds, then a stable-dominating and truncation-proof rule Pareto-improves *every* stable rule.

Proposition 4. *If the Rural Hospitals Theorem holds and φ is a stable-dominating and truncation-proof rule, then φ Pareto-improves every stable rule.*

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

Since $\mathcal{P}$ satisfies truncatability, there exists $R'_i \in \mathcal{P}$ such that R'_i is a truncation of R_i at $\nu(i)$. Note that $R'_i \neq R_i$, since $\nu(i) \; P_i \; \varphi_i(R) \; P_i \; \emptyset$. Let $R' = (R'_i, R_{-i})$.

We first show, by contradiction, that ν is stable at R'. If ν is not stable at R', there is $Y \subseteq X \setminus \nu$ such that, for each $i \in N$, $|Y(i)| \leq 1$, for each $z \in Y$, $z \; P'_{N(z)} \; \nu(N(z))$, $\nu \notin C_O(\nu \cup Y)$, and there is $\gamma \in C_O(\nu \cup Y)$ such that $Y \subseteq Z$ and $\gamma \in \mathcal{F}$. Since, for each $j \in N \setminus \{i\}$, $R'_j = R_j$ and ν is stable at R, there is $z \in Y$ such that $N(z) = i$. However, since $z \; P'_i \; \nu(i)$, we have $z \; P_i \; \nu(i)$, by definition of R'_i. But then Y would undermine the stability of ν at R, a contradiction.

By the Rural Hospitals Theorem, since $\nu(i) \neq \emptyset$, for each $\mu \in \Sigma(R')$, $\mu(i) \neq \emptyset$. Thus, by the Structure Lemma, for each $\mu \in \Sigma^{\uparrow}(R)$, $\mu(i) \neq \emptyset$. So $\varphi_i(R') \neq \emptyset$. Since φ is individually rational and the definition of R'_i, $\varphi_i(R')\, P'_i\, \emptyset\, P'_i\, \varphi_i(R)$. Then, by definition of a truncation, $\varphi_i(R')\, R_i\, \nu(i)\, P_i\, \varphi_i(R)$. However, this contradicts the truncation-proofness of φ, which requires that $\varphi_i(R)\, R_i\, \varphi_i(R')$. $\square$

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

We consider specifically the school choice model. While we cannot pin down all stable-dominating and strategy-proof rules for general priorities, we are able to say more if they are also Pareto-efficient. For the following result, we again assume that $\mathcal{P}$ is closed under truncation.

Proposition 5. *Consider a Pareto-efficient rule in the school choice model. If it is stable-dominating and truncation-proof, then it is stable.*

Proof. If priorities are strict then, by Proposition 2 the agent-optimal stable rule is the unique truncation-proof and stable-dominating rule [Alva and Manjunath, 2019]. Thus, for strict priorities every truncation-proof and stable-dominating, efficient or not, is stable.

If priorities satisfy strong* acyclicity, then every stable-dominating allocation is stable. Thus, every stable-dominating rule, truncation-proof and efficient or not, is stable.

This leaves the case of weak priorities that contain a weak* cycle. By the scarcity condition in the definition of a weak* cycle, it is without loss of generality to assume that $O = \{a, b\}$, $q_a = q_b = 1$, and $N = \{1, 2, 3\}$—since we consider stable-dominating rules, we may consider profiles of preferences where agents other than 1, 2, and 3 find only one object acceptable and agents 1,2, and 3 find all objects but a and b to be unacceptable.

Given three agents and two objects, each with capacity of one, suppose that the agents and objects are labeled such that the following weak* cycle appears in the priorities: $1 \succsim_a 2 \succ_a 3$, $3 \succ_b 1$, and $3 \succ_b 2$.

Fix $P \in \mathcal{P}$ as follows:

P_1	P_2	P_3
b	b	a
a	a	b
$\emptyset$	$\emptyset$	$\emptyset$

Let φ be an efficient and stable-dominating rule. We distinguish two cases.

Case 1: $1 \sim_a 2$. There are two stable allocations: $(a,\varnothing,b)$ and $(\varnothing,a,b)$.[15] Since φ is efficient and stable-dominating, $\varphi(P) \in \{(b,\varnothing,a),(\varnothing,b,a)\}$. Suppose $\varphi(P) = (b,\varnothing,a)$ and consider the truncation of P_3 at a:

$$\begin{array}{c} P_3' \\ \hline a \\ \varnothing \\ b \end{array}$$

There are only two stable allocations at (P_1,P_2,P_3'): $(a,b,\varnothing)$ and $(b,a,\varnothing)$. Since each of these assigns $\varnothing$ to 3 and φ is stable-dominating, by the Structure Lemma, φ is $\varphi_3(P_1,P_2,P_3') = \varnothing$. Then, $a = \varphi_3(P_1,P_2,P_3)\ P_3'\ \varphi(P_1,P_2,P_3') = \varnothing$, so φ is not truncation-proof.

If, on the other hand, $\varphi(P) = (\varnothing,b,a)$, the proof is analogous.

Case 2: $1 \succ_a 2$ The unique stable allocation is $(a,\varnothing,b)$. By the Structure Lemma, since φ is efficient and stable-dominating, $\varphi(P) = (b,\varnothing,a)$. By an analogous argument to Case 1, we conclude that φ is not truncation-proof.

□

The broad class of known truncation-proof and Pareto-efficient rules [Pápai, 2000, Pycia and Ünver, 2017] for the subdomain of $\mathcal{P}$ where agents rank $\varnothing$ below each $o \in O$ are readily extended to $\mathcal{P}$. Proposition 5 says that unless the priorities are such that there is a truncation-proof and Pareto-efficient rule that is also stable, none of these rules—including the *top trading cycles* rules [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 [2018] characterizes the priority structures that admit rules that are group strategy-proof, Pareto-efficient, and stable. By Proposition 5, stability can be weakened to stable-dominating in this characterization.[16]

For strict priorities, in light of Proposition 5, the agent-optimal stable rule is the unique candidate for a truncation-proof, Pareto-efficient, and stable-dominating rule. Thus, there exists a truncation-proof, Pareto-efficient, and stable-dominating rule if and only if the priority structure is Ergin-acyclic [Ergin, 2002]. Moreover, the top trading cycles rule is stable-dominating (or stable) if and only if it coincides with the agent-optimal

[15] We represent $\mu \in \mathcal{M}$ in the format of $(\mu(1),\mu(2),\mu(3))$.

[16] Han [2018] and Ehlers and Westkamp [2018] also give necessary and sufficient conditions for existence of a strategy-proof rule that is stability-constrained Pareto-efficient.

stable rule. In turn, the top trading cycles rule coincides with the agent-optimal stable rule if and only if the priority structure satisfies Kesten-acyclicity [Kesten, 2006].

Proposition 5 tells us that among truncation-proof rules, stable-domination is compatible with Pareto-efficiency only when stability itself is compatible with Pareto-efficiency. However, it leaves open the possibility that weakening stability to stable-domination may expand the set of truncation-proof rules. Indeed, Troyan and Morrill [2019] have recently shown that stable-dominating rules have some desirable strategic properties. Rather than strategy-proofness, they consider a strategic requirement on rules that allows manipulations, but only if they are not "obvious." They show that, under a regularity condition, no stable-dominating rule, regardless of the priority structure, is obviously manipulable.[17]

We weaken their condition to the following. A rule φ satisfies **truncation non-bossiness** if for each $R \in \mathcal{P}$, each $i \in N$, and each $R'_i \in \mathcal{P}_i$, if R'_i is a truncation of R_i at $\varphi_i(R)$ and $\varphi_i(R'_i, R_{-i}) = \varphi_i(R)$, then $\varphi(R'_i, R_{-i}) = \varphi(R)$.

It turns out that, among rules satisfying truncation non-bossiness, if we insist on truncation-proofness, we do not expand the set of admissible rules by weakening the requirement of stability to stable-domination.

Proposition 6. *Consider a truncation non-bossy rule in the school choice model. If it is stable-dominating and truncation-proof, then it is stable.*

Proof. Suppose that φ satisfies truncation non-bossiness, truncation-proofness, and stable-domination.

First, for each $i \in N$ and each $R_i \in \mathcal{R}_i$, let $\lambda(R_i) = |\{o \in O : o\ P_i\ \varnothing\}|$, the cardinality of the strict upper level set of R_i at $\varnothing$. Let $R, R' \in \mathcal{P}$. Define $R \preccurlyeq R'$ if and only if for each $i \in N$, $\lambda(R_i) \leq \lambda(R'_i)$. Denote the strict part of $\preccurlyeq$ by $<$.

For each $R \in \mathcal{P}$, if there exists $i \in N$ and $o \in O$ such that $\varphi_i(R)\ P_i\ o\ P_i\ \varnothing$, then there exists $R'_i \in \mathcal{P}_i$ such that R'_i is a truncation of R_i at $\varphi_i(R)$ and $R'_i \neq R_i$, since $\mathcal{P}$ satisfies truncatability. But then $(R'_i, R_{-i}) < R$.

Note that if, for each $i \in N$, $\lambda(R_i) = 0$, then the unique stable allocation assigns $\varnothing$ to each agent, and is also Pareto efficient. So for such R, $\varphi(R)$ is stable at R. This establishes that, for any sufficiently truncated $R \in \mathcal{P}$, $\varphi(R)$ is stable.

Suppose there is $R \in \mathcal{P}$ such that $\varphi(R)$ is not stable at R. By the above observation regarding sufficiently truncated preferences, it is without loss of generality to assume

[17] A preference $R'_i \in \mathcal{P}_i$ is a lower reshuffling of a preference $R_i \in \mathcal{P}_i$ at allocation $\mu(i)$ if for every $o, o' \in \mathcal{O}$, $o\ R_i\ o'\ R_i\ \mu(i)$ if and only if $o\ R'_i\ o'\ R'_i\ \mu(i)$. The regularity condition of Troyan and Morrill [2019] is independence of irrelevant rankings of φ: for each $R \in \mathcal{P}$, each $i \in N$, and each $R'_i \in \mathcal{P}_i$, if R'_i is a lower reshuffling of R_i at $\varphi_i(R)$, then $\varphi(R) = \varphi(R'_i, R_{-i})$. That is, the rule is not sensitive to an agent's preferences below the object that it assigns her.

that for each $R' \in \mathcal{P}$ such that $R > R'$, $\varphi(R')$ is stable. That is, we may assume that R is minimal with respect to $\leqslant$ among preference profiles where φ is not stable. Let $\nu = \varphi(R)$. Since φ is stable-dominating, there is $\mu \in \Sigma(R)$ such that ν Pareto-improves μ. By the Capacity-based Structure Lemma, ν is non-wasteful and individually rational. Since it is not stable but Pareto-improves μ, there are $o \in O$ and a pair $j, k \in N$ such that $j \succ_o k$ and $\nu(k) = o \; P_j \; \nu(j) \; R_j \; \mu(j)$. Since μ is stable and Pareto-improved by ν, there is $i \in N$ such that $i \gtrsim_o j$ and $\nu(i) \; P_i \; \mu(i) = o$.

Let $R'_i \in \mathcal{P}_i$ be the truncation of R_i at $\nu(i)$ and let $\nu' = \varphi(R'_i, R_{-i})$. By truncation-proofness, $\nu(i) = \nu'(i)$. By individual rationality, $\nu'(i) \; R'_i \; \varnothing$. This implies, by truncation non-bossiness, that $\nu' = \nu$. Since $R > (R'_i, R_i)$, by definition of R, ν is stable at (R'_i, R_{-i}). However, only i's preferences are different at R compared to (R'_i, R_{-i}) and i's preferences above $\nu(i)$ are the same at both profiles. Thus, ν is stable at R, which contradicts the premise that φ is not stable at R. $\square$

Note that Propositions 5 and 6 are independent since Pareto-efficiency and truncation non-bossiness are logically independent.

References

A. Abdulkadiroğlu and T. Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, June 2003. [2], [18]

A. Alkan. A class of multipartner matching markets with a strong lattice structure. *Economic Theory*, 19(4):737–746, June 2002. [6]

A. Alkan and D. Gale. Stable schedule matching under revealed preference. *Journal of Economic Theory*, 112(2):289–306, October 2003. [6]

S. Alva and V. Manjunath. Strategy-proof Pareto-improvement. *Journal of Economic Theory*, 181:121–142, 2019. [1], [3], [8], [9], [10], [14], [15], [17]

M. Balinski and T. Sönmez. A tale of two mechanisms: Student placement. *Journal of Economic Theory*, 84(1):73–94, January 1999. [2], [3], [8]

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. Yılmaz. School choice under partial fairness. *Theoretical Economics*, forthcoming. [4]

L. Ehlers and A. Erdil. Efficient assignment respecting priorities. *Journal of Economic Theory*, 145(3):1269–1282, May 2010. [3], [12], [13], [14]

L. Ehlers and T. Morrill. (Il)legal assignments in school choice. Working paper, North Carolina State University, 2018. [4]

L. Ehlers and A. Westkamp. Strategy-proof tie-breaking in matching with priorities. *Theoretical Economics*, 13(3):1009–1042, Sep 2018. [12], [18]

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. [14]

H. I. Ergin. Efficient resource allocation on the basis of priorities. *Econometrica*, 70(6): 2489–2497, 2002. [11], [18]

T. Fleiner. A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research*, 28:103–126, February 2003. [6]

X. Han. Stable and efficient resource allocation under weak priorities. *Games and Economic Behavior*, 107:1–20, January 2018. [18]

J. W. Hatfield and F. Kojima. Substitutes and stability for matching with contracts. *Journal of Economic Theory*, 145(5):1704–1723, September 2010. [15]

J. W. Hatfield and P. Milgrom. Matching with contracts. *American Economic Review*, 95 (4):913–935, September 2005. [6], [15]

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. [15]

O. Kesten. On two competing mechanisms for priority-based allocation problems. *Journal of Economic Theory*, 127(1):155–171, March 2006. [19]

O. Kesten. School choice with consent. *Quarterly Journal of Economics*, 125(3):1297–1348, August 2010. [4]

S. Pápai. Strategyproof assignment by hierarchical exchange. *Econometrica*, 68(6):1403–1434, November 2000. [3], [18]

M. Pycia and M. U. Ünver. Incentive compatible allocation and exchange of discrete resources. *Theoretical Economics*, 12(1):287–329, January 2017. [3], [18]

A. E. Roth. On the allocation of residents to rural hospitals: A general property of two-sided matching markets. *Econometrica*, 54(2):425–427, March 1986. [3]

Q. Tang and J. Yu. A new perspective on Kesten's school choice with consent idea. *Journal of Economic Theory*, 154:543–561, 2014. [4]

Q. Tang and Y. Zhang. Weak stability and Pareto efficiency in school choice. Working paper, Shanghai University of Finance and Economics, 2017. [4]

P. Troyan and T. Morrill. Obvious manipulations. Working paper, North Carolina State University, 2019. [4], [19]

P. Troyan, D. Delacrétaz, and A. Kloosterman. Essentially stable matchings. Working paper, University of Virginia, 2018. [4]