Incentives and Efficiency in Constrained Allocation Mechanisms*

Joseph Root† David S. Ahn‡

June 11, 2020

Abstract

We study private-good allocation mechanisms where an arbitrary constraint delimits the set of feasible joint allocations. This generality provides a unified perspective over several prominent examples that can be parameterized as constraints in this model, including house allocation, roommate assignment, and social choice. We first characterize the set of two-agent strategy-proof and Pareto efficient mechanisms, showing that every mechanism is a "local dictatorship." For more than two agents, we leverage this result to provide a new characterization of group strategy-proofness. In particular, an N-agent mechanism is group strategy-proof if and only if all its two-agent marginal mechanisms (defined by holding fixed all but two agents' preferences) are individually strategy-proof and Pareto efficient. To illustrate their usefulness, we apply these results to the roommates problem to discover the novel finding that all group strategy-proof and Pareto efficient mechanisms are generalized serial dictatorships, a new class of mechanisms. Our results also yield a simple new proof of the Gibbard–Satterthwaite Theorem.

1 Introduction

Many market design problems involve constraints. School choice assignments must ensure quotas of low-income students are satisfied at high-performing schools. Medical residency assignments must place enough doctors in rural areas. The allocation of radio frequency in spectrum auctions must satisfy a large number of complicated engineering conditions to ensure minimal cross-channel interference.

Although successful ad hoc approaches have been tailored for particular problems, to date there is little general understanding of how constraints affect efficiency and incentives, the two classic criteria for implementation. Theoretically, a unified approach would enable analytical insights to be shared between contexts. Practically, a flexible theory of constraints for market design would greatly expand applicability. Real-world problems involve many considerations that are difficult to anticipate. The tools of market design should be general enough to accommodate these considerations.

*We thank Simon Board, Ben Brooks, Haluk Ergin, Satoshi Fukuda, Thomas Gresik, Yuhta Ishii, Yuichiro Kamada, Timothy Kehoe, Rohit Lamba, Jacob Leshno, Jay Lu, Moritz Meyer-ter-Vehn, Michèle Müller, Roger Myerson, Farzad Pourbabaee, Doron Ravid, Phil Reny, Tomasz Sadzik Chris Shannon, David Rahman, Ron Siegel, Ran Shorrer, Hugo Sonnenschein, Wenfeng Qiu, Bill Zame and seminar participants at Berkeley, Bocconi, Chicago, Duke, McGill, Minnesota, Notre Dame, Penn State, UCLA, Washington University in St. Louis, and Yale for helpful feedback.

†Department of Economics, University of California, Berkeley, 530 Evans Hall, Berkeley, CA 94720-3880. Email: jroot@econ.berkeley.edu

‡Department of Economics, University of California, Berkeley, 530 Evans Hall, Berkeley, CA 94720-3880. Email: dahn@econ.berkeley.edu

We develop a model of object allocation with private values for completely general constraints. A finite number of objects are allocated to a finite number of agents and an arbitrary constraint circumscribes the set of feasible social allocations. Each agent has strict preferences over the objects assigned to her, but is indifferent to others' assignments.

While other agents' assignments have no direct effect on one's well-being, those assignments do limit the profiles of allocations that are jointly feasible. Obviously, the assignment of a house to another agent precludes my consumption of that house. So even with purely private values, constraints introduce linkage across agents' allocations. Each agent i is indirectly concerned with any other j's assignment, not because i cares about j's consumption, but rather because j's assignment will limit the set of objects for i that are jointly feasible with the j's assignment. Our goal is to study the set of incentive compatible and efficient mechanisms for a fixed arbitrary constraint. In addition, we aim to study how different features of a constraint make it amenable for implementation, that is, to understand what kinds of constraints yield what kinds of truthful and efficient mechanisms. For any constraint on the set of feasible allocations, our main findings characterize the entire class of mechanisms that are immune to manipulation by any group of agents yet still yield Pareto efficient outcomes.

Beyond its practical benefits, a general theory of constrained allocation yields some surprising theoretical insights. Several prominent problems which at first glance may appear unconstrained and unrelated can be neatly expressed as special constraints of our model. For example, the classical social choice problem corresponds to the special constraint of our model where all agents are constrained to consume the same object.[1] From this perspective, the social choice problem presents itself as a special constrained private-goods allocation problem. In fact, a corollary application of our results is the Gibbard–Satterthwaite Theorem: that all strategy-proof social choice mechanisms are dictatorial. With this novel presentation of social choice as a constraint, we can now sensibly formulate and prove a converse to Gibbard–Satterthwaite: under what conditions does the constraint admit any non-dictatorial mechanism?

Another prominent application of our theory is to house allocation, where a finite number of indivisible objects must be assigned to agents with unit-demand. Expressed this way, the house allocation problem is almost the opposite of the social choice problem: no two agents can be assigned the same object. Recently, Pycia and Ünver (2017) provided a full characterization of the group strategy-proof[2] and Pareto efficient house allocation mechanisms, building on earlier work by Papái (2000). In an earlier version of this paper, we show how to use our results to replicate Pycia and Ünver (2017) for a small number of agents.[3]

A third prominent problem that can be expressed as a constraint is the roommates problem, where an even number of agents need to match into pairs. In this case, the "objects" are the other agents and the constraint requires that: first, no agent is matched to herself; and second, if i is assigned to j then j is commensurately assigned to i. In contrast to the previous two applications, to our knowledge no general characterization of the incentive compatible, efficient mechanisms had yet been discovered. As an application of our results, we provide such a characterization. We show that

[1] The term "object" is figurative. In social choice, the objects are usually policy choices or political candidates.

[2] Roughly, a mechanism is group strategy-proof if no coalition of agents can jointly misreport their preferences, without harming anyone in the group and making at least one agent strictly better off.

[3] The argument constructs a tedious change of variables to parameterize the Pycia and Ünver (2017) as a special case of our general formulae in the three-agent case. Details are available from the authors on request.

all group strategy-proof and Pareto efficient roommates mechanisms are "generalized serial dictatorships," a class of mechanisms we will formally introduce later.[4] The fact that our results are useful in understanding and proving results across some well-known problems is a fortunate side-effect of the model's generality.

These examples illustrate a key conceptual contribution of our paper: to provide a novel framework to unify positive and negative results across these applications, tying together seemingly disparate environments and results by viewing them as different constraints on the image rather than through restrictions of preferences on the domain. Traditionally, positive results in specific environments are seen as escaping the impossibility of the Gibbard–Sattertwhaite Theorem by restricting preferences in the *domain* of the mechanism to convenient special cases, such as assuming single-peaked rankings or quasi-linear preferences. In our model, we can provide a different reconciliation of these positive results by interpreting these environments as relaxing constraints in the *image* of the mechanism: outside of the Arrovian social choice problem, all agents need not consume the same object and instead there is room for compromise to yield mechanisms beyond dictatorship. The "diagonal" constraint implicit in the social choice problem generates maximal tension between efficiency and incentives, while other constraints allow more scope for their coexistence. Our model explicitly exposes this tension, and our results characterize the scope for positive incentive-compatible implementation of efficient outcomes when this tension is relaxed. This provides a deeper understanding of why certain environments like social choice admit so few good mechanisms while other environments like house allocation admit a broad variety of good mechanisms.

Despite allowing for complete generality in the constraint, we fully characterization all mechanisms that satisfy standard incentive and efficiency desiderata. We start by considering two-agent environments. This case admits a surprisingly parsimonious characterization of the set of individually strategy-proof and Pareto efficient mechanisms for all constraints. We show that all individually strategy-proof and Pareto efficient mechanisms are "local dictatorships" in which the set of infeasible allocations is partitioned into two regions and each region is assigned a local dictator. For a given preference profile, the agents' top choices determine some (possibly infeasible) social allocation. If this allocation is feasible, the mechanism assigns it. Otherwise, it is infeasible and there is a local dictator assigned to the allocation. The non-dictator is assigned their favorite object compatible with the dictator's top object. However, not all partitions will maintain efficiency and incentive compatibility. Instead, some structure is required of the partition to ensure these desiderata are maintained. We show that every constraint can have its infeasible allocations "block diagonalized" to yield an immediate characterization of the partitions that do yield desirable mechanisms. Every block must be assigned to a single agent as the local dictator. So the number of strategy-proof and Pareto efficient mechanisms is determined entirely by the number of blocks allowed by the constraint.

With three or more agents, the set of individually strategy-proof and Pareto efficient mechanisms no longer admits such a straightforward characterization. Indeed, even for the classic house allocation setting, the collection of all such mechanisms is still unknown. Nevertheless, if we strengthen our

[4] In common with standard serial dictatorship, there is a sequence of dictators and each dictator picks her favorite object among those that are possibly feasible with the choices of earlier dictators. In contrast to standard serial dictatorship, our generalized version allows the order of subsequent dictators to depend on the choices of earlier dictators, rather than being locked in a fixed order.

incentive compatibility condition to group strategy-proofness, we can leverage the two-agent results to get a novel recursive characterization for the multi-agent case. Group strategy-proofness requires that no group of agents can ever collectively misreport their preferences so that all agents in the group are weakly better off and at least one agent is strictly better off. Our central observation is that group strategy-proof mechanisms have the convenient property that we can restrict attention to a subset of agents, fixing a preference profile of everyone else, to get a new group strategy-proof mechanism for the subset. We call these the "marginal mechanisms." Importantly, the properties of just the two agent marginal mechanisms are enough to capture the group incentives of the entire mechanism: if all two-agent marginal mechanisms are Pareto efficient and individually strategy-proof, then the full mechanism is group strategy-proof. This discovery is especially useful given our explicit characterization of two-agent mechanisms. The two-agent mechanisms of our first result are therefore the "building blocks" of all group strategy-proof mechanisms with many agents.

Beyond its analytical power, group strategy-proofness is substantively natural for a number of reasons. First, we show that, for any constraint, group strategy-proofness is equivalent to individual strategy-proofness and a classic normative condition called "nonbossiness"[5]. In bossy mechanisms, agents can manipulate the outcome of other agents without affecting their own allocation. Therefore, the marginal power of restricting attention to group strategy-proofness, relative to requiring only individual strategy-proofness, is simply to rule out such bossy mechanisms. So the gap between group and individual incentives boils down to whether one agent is allowed to alter another's outcome while not changing her own outcome. Second, in practice, incentive problems have been highly detrimental to the practical appeal of mechanisms. Violations in strategy-proofness of the Boston mechanism lead to severe inequality between "sophisticated" agents who knew how to game the system and "naive" agents who didn't. Ultimately, the mechanism was replaced in favor a strategy-proof mechanism (Abdulkadiroğlu, Pathak, Roth, and Sonmez 2006). The Vickrey-Clarke-Groves mechanism, despite its attractive individual incentives, has largely not been implemented in practice, in part because of its susceptibility to group manipulation (Rothkopf 2007). We therefore believe that mechanisms with strong group incentives are especially useful for practical considerations. In addition, group strategy-proofness is among the most demanding incentive conditions in the literature, and this benchmark should be established to understand the gains to efficiency from demanding weaker incentive conditions like Bayesian implementation. Finally, group strategy-proofness has been long studied in other environments, and especially for the house allocation problem, so using this as our incentive condition facilitates comparisons with earlier results. That all said, our focus on Pareto-efficiency and group strategy-proofness rules out some practical mechanisms. Deferred acceptance, for example, is not Pareto efficient, is not individually strategy-proof for the accepting side, and is not group strategy-proof for the proposing side.

1.1 Literature Review

To our knowledge, this paper is the first to identify the entire set of mechanisms that satisfy criteria regarding incentives and efficiency for an arbitrary constraint in our general allocation problem. However, several papers study mechanisms for specific constraints in particular environments. One

[5]To our knowledge, nonbossiness was first introduced by (Satterthwaite and Sonnenschein 1981).

such environment is the two-sided matching problem with distributional constraints, where for example there is a cap on the number of medical residents assigned to hospitals in a certain area. The two-sided matching problem can be expressed as a constraint in our more general model, and distributional constraints can be expressed as a further sharpening of that constraint.[6] A series of papers summarized Kamada and Kojima (2017a) study the two-sided matching problem with distributional constraints, with a primary focus on understanding stability.[7] In the two-sided matching problem, stability is the primary normative concern since the ubiquitous deferred-acceptance mechanism is known to be neither strategy-proof nor Pareto efficient. While specific mechanisms are shown to work well for specific classes of constraints, a general accounting for the class of all mechanisms is still outstanding. In principle, our results applied to this problem would characterize the set of all group strategy-proof and Pareto efficient mechanisms. That said, our results are exclusively about incentives and efficiency, and we have little to directly say about stability. This is partly because, as a concept, stability is only sensible and well-defined in particular examples of our environment such as two-sided matching.

Another example of a particular environment with a constraint on allocations is the house allocation problem, although it is not often thought of as a constrained problem. Abdulkadiroğlu and Sönmez (1999) and Papái (2000) construct classes of group strategy-proof and Pareto efficient mechanisms that are strictly larger than two classic examples of group strategy-proof and Pareto efficient mechanisms for house allocation: top trading cycles, attributed to David Gale by Shapley and Scarf (1974) and shown to have these desirable features by Bird (1984), and serial dictatorship, analyzed comprehensively by Svensson (1994) and Svensson (1999), which obviously has these features. A general characterization had remained a long-standing problem until Pycia and Ünver (2017) recently provided an impressive full description of all group strategy-proof and Pareto efficient mechanisms mechanisms. These are exactly the normative criteria explored in this paper, and in fact Pycia and Ünver (2017) helped inspire this paper by demonstrating a general characterization of these criteria is even attainable for an important problem like house allocation. House allocation problems are a special constraint in our model, where $a_i \neq a_j$ is required whenever $i \neq j$. That is, our characterization when applied to this constraint also provides another parameterization of mechanisms in Pycia and Unver. We explicitly verify the connection between the two characterizations in the three-house case, and believe the general change of variables between the two formulations is feasible but would be very tedious.

While incentives and efficiency are relatively well-understood for two-sided matching and for house allocation, one-sided matching such as in the classic problem of pairing roommates into dormitory rooms has demonstrated itself to be much more intractable. This is in large part because one-sided environments may fail to yield a stable match, as originally observed by Gale and Shapley (1962) in the same article introducing their eponymous algorithm for stable two-sided matching. Since then, a very large literature in operations research and computer science, starting with Irving (1985), tries to find efficient algorithms to find stable matchings when they exist. This specific computational

[6] More precisely, the two-sided matching problem can be modeled by making the set of objects equal to the union of agents from both sides of the market with the constraint that each agent is assigned to an agent in the opposite side and that, if agent i is matched to agent j then j should also be matched to i.

[7] Work in this literature includes contributions by Hafalir, Yenmez, and Yildirim (2013), by Ehlers, Hafalir, Yenmez, and Yildirim (2013), by Kamada and Kojima (2015), by Kamada and Kojima (2017b), and by Kamada and Kojima (2018).

problem has become so well-studied that it is now called the "stable roommates problem." In contrast, there seems to be almost no discussion of incentives and efficiency for the roommates problem.[8] An application of our main results yields a characterization of group strategy-proof and Pareto efficient mechanisms for the roommates problem, which turn out to be the family of generalized serial dictatorships that we introduce in this paper. To our knowledge, this is a new observation and, analogous to the characterization theorem by Pycia and Ünver (2017) for house allocation or to the Gibbard–Satterthwaite Theorem for social choice, establishes the characterization of group strategy-proofness and Pareto efficiency for the roommates problem.

A final notable special constraint in our environment is the classic Arrovian social choice model. The first result studying incentives and efficiency was the celebrated negative finding by Gibbard (1973) and Satterthwaite (1975), which initiated the field of implementation theory. Here, the classic Arrovian social choice environment in which the Gibbard–Satterthwaite Theorem is cast corresponds to the case where all agents must be assigned the same common outcome. That is, social choice corresponds to the constraint that $a_i = a_j$ for all agents i, j. Viewed in this way, the social choice constraint is almost the opposite of the house allocation constraint. We derive the Gibbard-Sattherhwaite Theorem as a corollary of our main characterization. This provides a novel perspective on the classic result by casting light on the implications of constraining allocations so that all agents consume a common object. Our perspective allows us to understand the Gibbard–Satterthwaite Theorem as a consequence of the restrictiveness of the constraint. Correspondingly, our perspective also offers a novel escape from the assumptions of the Gibbard–Satterthwaite Theorem, namely relaxing the social choice constraint. This escape is meaningful only when Arrovian social choice is framed as a special case of private good economies. In fact, this framing allows us to generalize the Gibbard–Satterthwaite Theorem in our environment: we completely characterize the constraints where only serial dictatorships are group strategy-proof, finding the social choice constraint as a particular example. It is interesting that social choice can be cast as a special case of our model with the particular diagonal restriction on allocations, since private-goods economies are usually viewed as a special case of social choice with a particular restriction on preferences.

Our general environment with private goods was also recently studied by Barberà, Berga, and Moreno (2016) from a social choice perspective. They work focuses on the richness of preferences for a social choice function, that is, it focuses on the richness of the *domain* of preference. Throughout our paper, by contrast, we allow no restrictions on preferences and assume that mechanisms will find allocations for all preference profiles. Instead of considering restrictions on the domain, we complement Barberà, Berga, and Moreno (2016) by considering different constraints on the *image* of allocations that are feasible for a mechanism.

Our different focus on constraints on allocations, rather than on restrictions over preferences, stems partly from our different objectives. Barberà, Berga, and Moreno (2016) are primarily concerned with the relationship between group and individual incentives. Their main result reveals an important connection between group and individual strategy-proofness when the space of admissible preferences is sufficiently rich.[9] In contrast, our aim is not to relate different axioms for strategy-proofness, but

[8] The one exception we found was a working paper by Abraham and Manlove (2004) that studies the computational hardness of finding Pareto optimal matches for the roommates problem.

[9] This complements a similar connection between group and invidiual incentives for classic Arrovian environments,

rather to characterize the entire space of mechanisms that satisfy the fixed axiom of group strategy-proofness. Our main results examine the structure of the constraint to describe the structure of the group strategy-proof mechanisms. That is, our objective is not to relate strategy-proofness to other normative conditions like nonbossiness or monotonicity, but rather to relate the structure of group strategy-proof mechanisms to the structure of the constraint. Our results address concerns like how the space of strategy-proof mechanisms changes when constraints are relaxed or tightened. Of course, an improved understanding of how group strategy-proofness relates to other natural conditions can only be helpful. In fact, a key lemma in proving our characterization is to observe a tight relationship between group strategy-proofness, individual strategy-proofness and nonbossiness, and Maskin monotonicity. So our development owes a debt to these earlier realizations. However, our lemma is still distinct from these earlier observations in both substance and message, as we will explain after formally introducing the result.

Finally, a more distant body of work on random allocation tests whether a random allocation is a convex combination of deterministic allocations satisfying a fixed constraint (Balbuzanov 2019, Budish, Che, Kojima, and Milgrom 2013), extending the fairness gains of the random assignment mechanisms introduced by Bogomolnaia and Moulin (1990) to constrained environments. We focus on deterministic mechanisms, so as far as we can see our results have no direct relationship to this literature.

2 Model

We begin by introducing primitives. Let N be a finite set of **agents** and $\mathcal{O}$ be a finite set of **objects**. We use the term "object" because of our leading examples, but note that these are not necessarily physical objects, but can be political candidates, roommates, and so on. Define $\mathcal{A} = \mathcal{O}^N$ to be the set of all possible allocations of objects to agents. Equivalently, $\mathcal{A}$ is also the set of maps $\mu : N \to \mathcal{O}$ and we switch to this perspective when it is more useful. A **suballocation** is a map $\sigma : M \to \mathcal{O}$ where $M \subset N$. Let $\mathcal{S}$ denote the set of suballocations. Our task is to assign objects to agents in a way that is consistent with an exogenous *constraint* which reflects the set of feasible allocations for a particular application. Importantly, the constraint is exogenous to the problem. It is given to the mechanism designer as a fixed set of feasible outcomes. Formally, we are given a nonempty **constraint** $C \subset \mathcal{A}$ and $(a_i)_{i\in N} \in C$ means that it is feasible to allocate each agent i the object a_i simultaneously. Notice that since we place no restrictions on the constraint, it is without loss of generality to have a common set of objects for all agents because if each agent has her own set of objects then one could add the constraint that all feasible allocations cannot assign these objects to other agents.[10] Agents have strict preferences over the *objects* and are assumed to be indifferent between any two allocations in which they receive the same object. We will use P to denote the set of strict preferences (i.e. linear orders) on $\mathcal{O}$ and $\mathscr{P} = P^N$ to denote the set of preference profiles.[11] Our primary object of interest in this paper is a **feasible mechanism**, which is simply a map $f : \mathscr{P} \to C$. Our task will be to find feasible mechanisms satisfying desirable conditions regarding incentives and efficiency, to be formally introduced in the sequel.

discovered by the same authors (Barberà, Berga, and Moreno 2010) and by Le Breton and Zaporozhets (2009).

[10] More precisely, let $\mathcal{O} = \sqcup\mathcal{O}_i$ and define C_{new} by $(a_i)_{i\in N} \in C_{new}$ if and only if $(a_i)_{i\in N} \in C$

[11] A binary relation $B \subset \mathcal{O} \times \mathcal{O}$ is a linear order if it is complete, transitive, and antisymmetric

Some well-known problems can be expressed as special constraints in this model:

- House Allocation: A finite number of houses must be distributed to a finite number of agents. The houses cannot be shared so no two agents can be allocated the same one. This gives rise to the constraint
$$C = \{(a_i)_{i\in N} \,|\, a_i \neq a_j \text{ when } i \neq j\}.$$
This setting has been the subject of considerable interest since at least Shapley and Scarf (1974). Two prominent mechanisms used in practice are Gale's top trading cycles algorithm and Gale and Shapley's deferred acceptance algorithm (with priorities for houses).

- Roommates Problem: Universities are often tasked with assigning students into shared dormitory rooms. Assuming N is even, this problem can be captured in our environment by setting $\mathcal{O} = N$ and imposing the constraint
$$C = \{\mu : N \to N \,|\, \mu^2 = id \text{ and } \mu(i) \neq i \text{ for all } i\}.$$
The first condition requires that if i is assigned roommate j then j is also assigned i and the second condition requires that all agents are assigned a roommate.

- Social Choice: If the constraint specifies that all agents receive the same object (without specifying ex-ante which object will be chosen) we get the classical version of the social choice problem[12]. Specifically, if
$$C = \{(a_i)_{i\in N} \,|\, a_i = a_j \text{ for all } i, j\}$$
the constraint requires that all agents be given the same social choice, but which outcome is chosen is a function of the mechanism.

Our model is able to accommodate these examples as special cases because of its generality in admitting arbitrary constraints. We will have more explicit analyses of these examples later in the paper.

Before moving on, we record here some notation used throughout the paper. For any subset $M \subset N$, given a preference profile $\succsim = (\succsim_i)_{i\in N} \in \mathscr{P}$ and a profile of alternative preferences for agents in M, $(\succsim'_j)_{j\in M}$, we will write $(\succsim'_M, \succsim_{-M})$ to refer to the profile in which an agent j from M reports $\succsim'_j$ and any agent i from M^c reports $\succsim_i$. We will often want to consider how a mechanism f changes when a few agents change their preferences, that is the difference between $f(\succsim)$ and $f(\succsim'_M, \succsim_{-M})$. When the initial preference profile $\succsim$ is clear, we will sometimes write $\succsim_-$ instead of $\succsim_{-M}$. Given a constraint $C \subset \mathcal{A}$ and a subset of agents $M \subset N$, let $C^M = \{\mu : M \to \mathcal{O} \,|\, \exists b \in C \text{ s.t. } b_i = \mu(i)\, \forall i \in M\}$ which we will call the **projection of** C **on** M. An element of C^M will be referred to as a **feasible suballocation** for agents in M. If $\mu : M \to \mathcal{O}$ and $\mu' : M' \to \mathcal{O}$ are suballocations with $M \subset M'$ which agree on their shared domain, μ' is called a **extension** of μ. If μ' is a feasible suballocation (which of course implies that μ is) then μ' is called a **feasible extension** of μ. If μ' assigns an object to each agent, it is called a **complete extension** of μ. Given a feasible suballocation μ, we will let $C(\mu)$ denote the set of complete and feasible extensions of μ. For any agent i, let $\pi_i : \mathcal{A} \to \mathcal{O}$ be the projection map so that given an allocation $(a_j)_{j\in N}$, $\pi_i a = a_i$ and for a set of allocations $B \subset \mathcal{A}$, we

[12] See Barberà (2001) for a general statement of the social choice problem with restricted domains.

have $\pi_i B = \{a \in \mathcal{O} \mid \text{ there is a } b \in B \text{ with } \pi_i b = a\}$. For $x \in \mathcal{O}$ and $\succsim_i \in P$, define $LC_{\succsim_i}(x) = \{y \in \mathcal{O} \mid y \prec_i x\}$ be the (strict) **lower contour set** of x at $\succsim_i$. Likewise, $UC_{\succsim_i}(x) = \{y \in \mathcal{O} \mid y \succ_i x\}$ is the (strict) **upper contour set** of x at $\succsim_i$. For a preference $\succsim_i$, define $\tau_n(\succsim_i)$ as the nth top choice under $\succsim_i$. Likewise, for any preference profile $\succsim$, define $\tau_n(\succsim)$ as the allocation in which each agent gets their nth top choice. To save on notation, we will often omit the subscript when referring to the top choice (i.e. writing $\tau(\succsim)$ to mean $\tau_1(\succsim)$). We will use $\bar{C}$ to denote the set of infeasible allocations.

In practice, mechanisms are often designed to satisfy efficiency and incentive properties. Here are several well-known desiderata for allocation mechanisms.

Definition 1. A mechanism $f : \mathscr{P} \to C$ is

1. **strategy-proof** if, for every $i \in N$ and every $\succsim \in \mathscr{P}$,

$$f_i(\succsim) \succsim_i f_i(\succsim'_i, \succsim_{-i})$$

for all $\succsim'_i \in P$. That is, truth-telling is a weakly dominant strategy.

2. **group strategy-proof** if, for every $\succsim \in \mathscr{P}$ and every $M \subset N$, there is no $\succsim'_M$ such that
   (a) $f_j(\succsim'_M, \succsim_{-M}) \succsim_j f_j(\succsim)$ for all $j \in M$;
   (b) $f_k(\succsim'_M, \succsim_{-M}) \succ_k f_k(\succsim)$ for at least one $k \in M$.

3. **weakly group strategy-proof** if, for every $\succsim \in \mathscr{P}$ and every $M \subset N$, there is no $\succsim'_M$ such that

$$f_j(\succsim'_M, \succsim_{-M}) \succ_j f_j(\succsim) \text{ for all } j \in M.$$

4. **Pareto efficient** if there is no allocation $a \in C$ and preference profile $\succsim$ such that $a \neq f(\succsim)$ and $a_j \succsim_j f(\succsim)$ for all j.

5. **nonbossy** if, for all $\succsim \in \mathscr{P}$,

$$f_i(\succsim'_i, \succsim_{-i}) = f_i(\succsim) \implies f(\succsim'_i, \succsim_{-i}) = f(\succsim).$$

6. **Maskin monotonic** if, for all $\succsim, \succsim' \in \mathscr{P}$,

$$LC_{\succsim'_i}[f_i(\succsim)] \supset LC_{\succsim_i}[f_i(\succsim)] \text{ for all } i \implies f(\succsim') = f(\succsim).$$

Strategy-proofness requires that for every agent i and every possible profile of preferences for the other agents, i cannot improve her outcome by misreporting her preference. Group strategy-proofness is similar except that it requires that no group can collectively misreport their preferences without hurting anyone while strictly benefiting at least one agent. This is often called "strong group strategy-proofness" to contrast it with weak group strategy-proofness which requires that any deviating coalition make all its agents strictly better off. Pareto efficiency might also be called "constrained efficiency" since it requires that for every preference profile f selects a feasible allocation such that no other feasible allocation can improve (at least weakly) all agents outcomes. Pareto efficiency is also sometimes called

"unanamity" in the literature. Nonbossiness simply requires that no agent can exert influence on another agent without affecting her own outcome. Finally, Maskin monotonicity is the seemingly weak condition that whenever an allocation is chosen at a given preference profile, if all agents instead report a different profile in which their respective allocations have improved relative to all other allocations, then f should maintain the same outcome. This condition was famously shown to be necessary for Nash implementation by Maskin (1999).

A useful observation in building our results is the following equivalence across these conditions. We present this lemma explicitly because it is of some independent interest and to explain how this part of our argument relates to earlier observations.

Proposition 1. *If $f : \mathscr{P} \to \mathcal{A}$ the following are equivalent:*

1. *f is group strategy-proof.*

2. *f is strategy-proof and nonbossy.*

3. *f is Maskin monotonic.*

The connection between individual and *weak* group-strategy proofness was examined in social choice environments by Le Breton and Zaporozhets (2009) and by Barberà, Berga, and Moreno (2010) and in private-goods environments such as ours by Barberà, Berga, and Moreno (2016), who prove that, when the domain of preference is sufficiently rich, weak group strategy-proofness is equivalent to individual strategy-proofness for a broad class of social choice functions satisfying generalizations of nonbossiness and Maskin monotonicity. An immediate difference is our use of strong rather than weak group strategy-proofness, which follows the literature on house allocation that also studies strong group strategy-proofness.[13] While perhaps a seemingly technical distinction, this is quite a substantively important departure from the weak concept. For example, deferred acceptance is only weakly group-strategyproof on the proposing side, but is not group strategy-proof in our stronger sense. Even ignoring the difference between weak and strong incentives, the theorem of Barberà, Berga, and Moreno (2016) bears no obvious relation to Proposition 1. The two results have very different aims and messages. Barberà, Berga, and Moreno (2016) take generalizations of Maskin monotonicity (that they call "joint monotonicity") and nonbossiness (that they call "respectfulness") as *assumptions* in their results and ask how large the domain of preferences must be to ensure group and individual incentives align. Our result generates nonbossiness and Maskin monotonicity as *implications* of group strategy-proofness for full preference domains, which is important in subsequent applications where we verify that a mechanism is group strategy-proof by testing that it is Maskin monotonic. On the other hand, we assume the domain of all strict preferences throughout this paper, and have nothing to say here about the consequences of restrictions on preferences.

The relationship between group strategy-proofness and Maskin monotonicity was first revealed by the proof of the Muller–Satterwthwaite Theorem, which proceeds by showing that either group or individual strategy-proofness is equivalent to Maskin monotonicity for the social choice problem (Muller and Satterthwaite 1977).[14] This equivalence between group strategy-proofness and Maskin

[13] For the specific problem of house allocation, the equivalence between (1) and (2) was first observed by Papái (2000).
[14] Recall the Muller–Satterwhaite Theorem: all Maskin monotonic and surjective social choice functions are dictatorial.

monotonicity was then further demonstrated to hold for other problems as well, including for house allocation by Svensson (1999) and for two-sided matching by Takamiya (2001). Takamiya (2003) unified these observations in a general statement for all indivisible-good economies without externalities that also applies to our model, and should be credited for the equivalence between (1) and (3) in Proposition 1.

Group strategy-proofness requires that no group of agents can collectively misreport their preferences and benefit at least one agent without making anyone in the group worse off. One possible coalition is the grand coalition. Thus if f is group strategy-proof and $f(\succsim) = a$ for some profile $\succsim$, then a can never Pareto dominate $f(\succsim')$ for any other profile $\succsim'$, since all agents would collectively report $\succsim$.

Lemma 1. *If $f : \mathscr{P} \to \mathcal{A}$ is group strategy-proof then it is Pareto efficient on its image.*[15]

Having established this, the goal of this paper is to understand the correspondence between the primitives (the set of agents, objects, and the constraint) and the set of group strategy-proof, Pareto efficient mechanisms. We will denote the set of feasible group strategy-proof mechanisms which map into C, $GS(C)$.

3 Characterization Results

We begin by considering the two-agent case where we find an explicit characterization of the set of strategy-proof and Pareto efficient mechanisms for an arbitrary constraint. Each mechanism with these properties turns out to be a "local dictatorship." We then turn to the n-agent case where we show that an n-agent mechanism is group strategy-proof if and only if each 2-agent marginal mechanism is group strategy-proof.

3.1 Two Agents

Given just two agents, we will show that for every constraint the set of strategy-proof and Pareto efficient mechanisms corresponds exactly to the set of "local dictatorships" in which the set of infeasible allocations $\bar{C}$ is partitioned into two disjoint subsets and each agent is assigned a set. After the agents announce their preferences, if the allocation in which both agents get their top choice is feasible, the mechanism must pic this allocation by Pareto efficiency. Otherwise, it is infeasible to give both agents their top choices and one agent must compromise and consume a less-favored object. The agent who does not have to compromise is the "local dictator" and gets her top choice, and the "local compromiser" receives her favorite object among those that are jointly feasible with the local dictator's top choice.

One possible complication with this procedure is that there may be no object for the local compromiser that is jointly feasible with the local dictator's top choice. For example, if the local dictator at (x, y) is agent 1, and $(x, y') \notin C$ for all objects $y' \in \mathcal{O}$, then there is no choice for agent 2 that will allow agent 1 to consume her favorite object x. On the other hand, since agent 1 can never feasibly be assigned object x, it would seem that her preference for x is immaterial to the social choice.

[15] That is, if the constraint C is exactly $im(f)$.

This turns out to be true, and we can ignore objects that are never assigned to an agent without loss of generality. To make this precise, for any constraint $C \subset \mathcal{O}^2$ let $R_1 = \{x \in \mathcal{O} \,|\, (x,y) \notin C \text{ for all } y \in \mathcal{O}\}$ and $R_2 = \{y \in \mathcal{O} \,|\, (x,y) \notin C \text{ for all } x \in \mathcal{O}\}$. In words, R_i is the set of objects which are always infeasible for agent i because there is no object a_{-i} for the other agent that will make the joint allocation (a_i, a_{-i}) feasible. More generally, we can likewise define R_i for any number of agents as the set of objects which are always infeasible to agent i no matter what objects are assigned to everyone else. Since these objects are immaterial to the agents, it would seem natural and would certainly be convenient if the ranking of always infeasible objects should have no effect on the outcome of a mechanism. The following lemma says exactly that.

Lemma 2. *Let C be a constraint for n agents. If $f : \mathscr{P} \to C$ is group strategy-proof and Pareto efficient and if $\succsim$ and $\succsim'$ are preference profiles in which for every i the relative ordering of elements in $\mathcal{O} \setminus R_i$ is unchanged then $f(\succsim) = f(\succsim')$*

Let $\bar{C}^* = \{(x,y) \,|\, (x,y) \notin C \text{ and } x \notin R_1, y \notin R_2\}$. That is, $\bar{C}^*$ is the set of infeasible allocations in which both agents could get the associated object for some choice of the other agents' object. As mentioned, all Pareto efficient mechanisms will assign top choices to both agents when doing so is feasible. The main job of a mechanism is to adjudicate the outcome when one agent must give up on her top choice. It turns out that strategy-proofness will demand a local dictator is determined as a function of only the agents' top objects. We prove this claim by taking an approach to strategy-proofness originally developed by Barberà (1983). This approach begins with the simple but deep observation that strategy-proof social choice functions can always be written as if an "option set" is available to player i as a function of everyone else's ($j \neq i$) report, and then i's allocation maximizes agent i's reported preference over that option set. We explicitly restate Barberà's observation for our environment of private goods, because we feel it is not as generally well-known as it should be and to acknowledge the role it plays in our argument. Let $P^{N-1} = \times_{j \neq i} P$ denote the space of preference profiles for all players beside agent i.

Lemma 3 (Barberà (1983))**.** *A mechanism $f : \mathcal{P} \to C$ is strategy-proof if and only if there exist nonempty correspondences $g_i : P^{N-1} \rightrightarrows \mathcal{O}$ such that, for all agents i,*

$$f_i(\succsim) = \max_{\succsim_i} g_i(\succsim_{-i})$$

With some work, Barberà's Lemma can be used to show that all strategy-proof and Pareto efficient two-agent mechanisms assign a local dictator who gets her top choice, and the assignment of dictatorship can depend only on the top choice for each agent. So such mechanisms can be described by coloring the set $\bar{C}^*$ with one color for the top-choice pairs where agent 1 is the local dictator and the other color for the top-choice pairs where j is the local dictator.

However, not all such colorings will be strategy-proof. For example, if agent 1 is the local dictator when (a,b) are the top choices and agent 2 is the local dictator at (a,b'), then agent 2 may want to misreport her top choice as b' even in situations where b is actually her top choice because she gets dictatorship power by misreporting. The coloring of the infeasible set $\bar{C}^*$ will have to satisfy some restrictions, which motivates the following constructions. Define the binary relation B on $\bar{C}^*$

by $(a,b)B(a',b')$ if $a = a'$ or $b = b'$. Two allocations are related by B if (at least) one agent gets the same object in both allocations. Now if $(a,b)B(a',b')$, then the example above suggests that the same agents must be assigned as the dictator in both cases, to prevent the situation where one agent can move from being the local compromiser to being the local dictator by individually misreporting her top object. This relation must hold across pairs of top choices that are even indirectly linked, so common assignment of local dictatorship must also hold transitively across B. Let T be the transitive closure of B.[16] Since B is reflexive and symmetric, it can easily be shown that T is an equivalence relation.[17] As an equivalence relation on a finite set, it can be expressed as a partition with a finite number of equivalence classes $E_1, E_2, \dots E_p$, where $(a,b)T(a',b')$ if and only if (a,b) and (a',b') are both in some E_i. We will refer to the equivalence classes of T as the **blocks** of $\bar{C}^*$.

Figure 1 illustrates an example of the relation T for a specific constraint. The top-left panel shows the constraint; grey cells are infeasible allocations. Panel (B) permutes $R_1 = \{a_4\}$ and $R_2 = \{a_4, a_6\}$ to the top and left most objects. In panel (C), a particular 4-element block of $\bar{C}^*$ consisting of (a_2, a_1), (a_2, a_3), (a_6, a_3), and (a_6, a_8) is shaded black. No element of the grey set is related by B to any member of $\bar{C}^*$ which is not also shaded black. Since the order of objects is not important, we can permute the rows and columns to display the equivalence classes more easily. Hence in panel (D), we again permute the objects. As we can now easily see there are three equivalences classes of T which are indicated as E_1, E_2 and E_3. We can then assign a dictator to each block independently as described below.

Let $C^1(b) = \{a \in \mathcal{O} \,|\, (a,b) \in C\}$ and likewise $C^2(a) = \{b \in \mathcal{O} \,|\, (a,b) \in C\}$. A mechanism $f : \mathscr{P}^2 \to C$ is called a **local dictatorship** if each block E_i of $\bar{C}^*$ is assigned a (local) dictator d_i so that for any $\succsim$ if $\tau(\succsim_1, \succsim_2) = (a,b)$

$$f(\succsim) = \begin{cases} (a,b) & \text{if } (a,b) \in C \\ \left(a, \max_{\succsim_2} C^2(a)\right) & \text{if } (a,b) \in E_k \text{ and } d_k = 1 \\ \left(\max_{\succsim_1} C^1(b), b\right) & \text{if } (a,b) \in E_k \text{ and } d_k = 2 \end{cases}$$

One can easily see that any local dictatorship is strategy-proof and Pareto efficient. The surprising fact is that the converse holds. That is, T directly indicates how to construct every mechanism.

Theorem 1. *$f : P^2 \to C$ is strategy-proof and Pareto efficient if and only if it is a local dictatorship.*

To see how this works for more familiar constraints, consider Figure 2. On the left is the house allocation constraint and on the right is the social choice constraint. Each grey square on the left is a different equivalence class of T, so every mechanism corresponds to a labeling of the grey boxes with 1's and 2's, which can be done independently. Another way to think about this is that each object is owned by one of the agents. If either agent top-ranks an object they own, they're guaranteed the ability to consume it. If both agents top-rank the other agents' object, they can trade. On the right is the social choice constraint. Clearly T has a single block for this constraint since it is possible to move from any grey square to any other grey square, only changing one coordinate at a time, and

[16] The transitive closure is the minimum binary relation containing B which is transitive.

[17] It is reflexive because B is. To see that it is symmetric, if we have $(a,b)T(a',b')$ since $\bar{C}^*$ is finite, there are $(a_1,b_1), \dots (a_n,b_n)$ such that $(a,b)B(a_1,b_1)B \dots B(a_n,b_n)B(a',b')$. By reversing all these, we see that $(a',b')T(a,b)$.

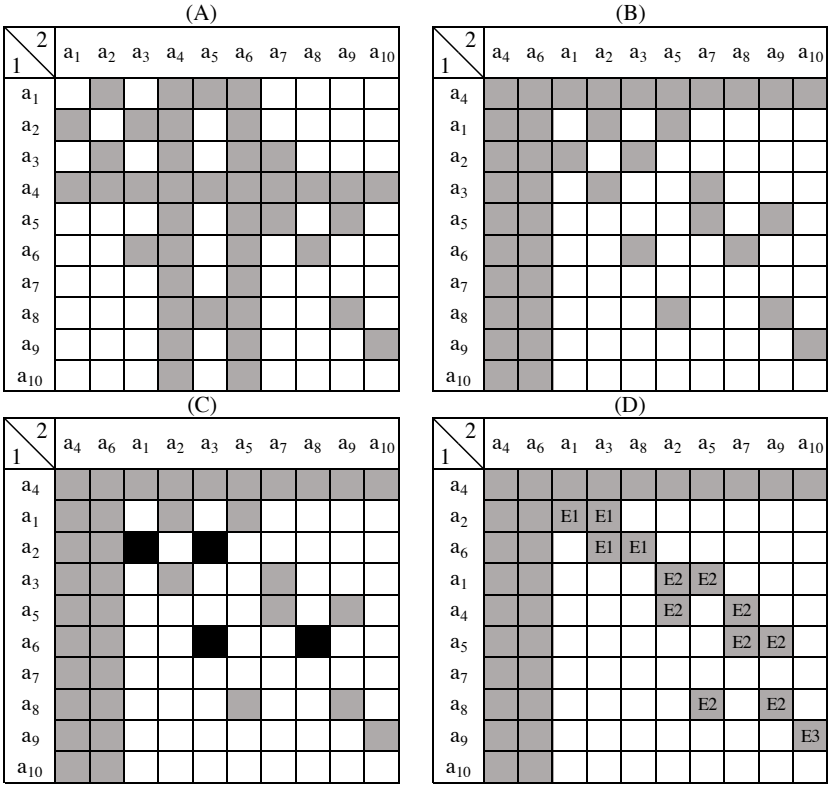

Figure 1: Two-agent Example

only passing through grey squares. Then Theorem 1 immediately yields the two-agent version of the Gibbard–Satterthwaite Theorem, that every mechanism is a dictatorship. Famously, the Gibbard–Satterthwaite Theorem requires at least three alternatives. Our analysis provide a new perspective on this cardinality requirement: observe that if the social choice constraint in Figure 2 had only two objects, the constraint would be the top-left 2×2 constraint. In this case, T has two equivalence classes corresponding to the two grey squares.

In independent and contemporaneous work, Meng (2019) provides an impressive characterization of all strategy-proof and Pareto efficient mechanisms for the two-agent social choice problem when agents are known to be indifferent between classes of alternatives that are fixed a priori. His characterization involves assigning a dictator at all profiles of preferences over announced indifference classes, where the dictator assignment must respect a cell-connected property. The structure of his result closely resembles our assignment of local dictators to the infeasible set. In fact, either result can be deduced from the other. However, these results are cast for very different questions, his for indifference and ours for constraints, so their substantive applications and contributions are quite different.

3.2 N Agents

When there are three or more agents, the approach we used for two agents fails to provide a straightforward characterization. The notion of a "local dictator" does not immediately generalize for more

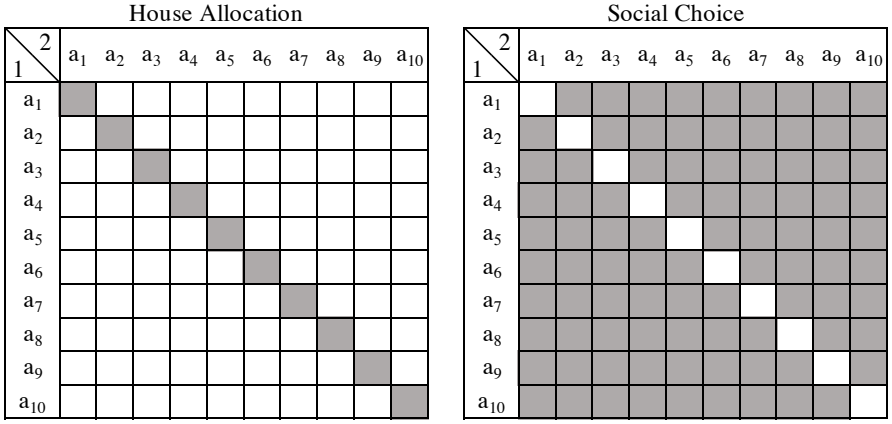

Figure 2: The social choice and house allocation constraints for two agents and 10 objects.

than two agents. One issue is that the set of compromising agents is not identified by knowing the local dictators because there are multiple agents besides the dictator. In fact, the ambiguity is deeper: not only is the identity of the compromising agent ambiguous, but the number of compromising agents is not even necessarily fixed: it may be the case that having a single compromising agent is insufficient to move to a feasible solution, and instead multiple compromising agents must move to less-preferred assignments.

However, there is a subclass of constraints for which the basic intuition does follow the two-agents case and its characterization is therefore no more difficult. A constraint C is called **single-compromising** if for every infeasible allocation $(a_i)_{i\in N}$ and every agent i there is a reassignment a_i' for agent i such that (a_i', a_{-i}) is feasible. Thus, from any infeasible allocation, any agent can unilaterally compromise to make the social allocation feasible. In this case, every group strategy-proof and Pareto efficient mechanism can be written in a simple manner analogous to the characterization of the two-agent case. The generalization again partitions the space of infeasible allocations, but now each infeasible allocation is assigned a subset of agents who must compromise. We mention this special case where the two-agent approach extends because it exposes some of the limitations in generalizing that approach to more agents. First, it will be useful to have some definitions.

A **local compromiser assignment** is a map $\alpha : \mathcal{A} \to 2^N$ such that for every infeasible $x \in \bar{C}, \alpha(x)$ is nonempty and for every feasible $y \in C$, $\alpha(y) = \emptyset$. For $x \in \bar{C}$ an agent $i \in \alpha(x)$ is referred to as a **local compromiser** at x. This definition is motivated by the following algorithm, called the **constraint-traversing algorithm** for α, which take a profile of preferences as an input and returns a feasible allocation, or, if unable to do so, returns the symbol $\emptyset$. For a given preference profile $\succsim$:

Step 0 Let $x^0 = \tau_1(\succsim)$

Step k If x^{k-1} is feasible, stop and return x^{k-1}. Otherwise, if there is any $l \in \alpha(x^{k-1})$, such that $LC(x_l^{k-1})$ is empty, stop and return $\emptyset$. If not, define $x_i^k = x_i^{k-1}$ for all $i \notin \alpha(x^{k-1})$ and let $x_j^k = \max_{\succsim_j} LC(x_j^{k-1})$ for all $j \in \alpha(x^{k-1})$.

In words, the algorithm works by starting with the allocation in which all agents get their top choice. If this is feasible, the algorithm terminates. If not, there are number of local compromisers determined by α. The algorithm next tries the allocation in which the local compromisers switch to their next-best alternative, and the other agents keep their top choice. If this is feasible, the algorithm stops. Otherwise, there are again some local compromisers and the algorithm continues in the same manner. In this way the algorithm continues down agents' preference lists. For completeness, the statement of the algorithm includes a description of what to do if the algorithm exhausts an agents objects. The assumption that the constraint is single-compromising, along with proposition 2 will ensure that this never happens. When the constraint-traversing algorithm always yields a well-defined allocation, we call the induced mechanism a **constraint-traversing mechanism**. The following proposition gives a characterization of all group strategy-proof and Pareto efficient mechanisms for single-compromising constraints, analogous to Theorem 1 for the case with just two agents.

Proposition 2. *Let n be arbitrary and let C be single-compromising. A mechanism is group strategy-proof and Pareto efficient if and only if it is a constraint-traversing mechanism such that the local compromiser assignment satisfies*

1. $|\alpha(a)| \leq 1$ *for all* a
2. $\alpha(a) = i \implies \alpha(a'_i, a_{-i}) = i$ *whenever* $(a'_i, a_{-i}) \in \bar{C}$

An earlier working version of this paper included a more comprehensive examination of constraint-traversing mechanisms in general environments beyond single-compromising constraints, and this material is currently being incorporated into another paper.[18] For more general structures of constraints, constraint-traversing mechanisms are not necessarily incentive compatible and efficient, and the main work of this additional material is finding sufficient conditions that guarantee these conditions are satisfied.

From hereon, we consider the general case of arbitrary constraints, and not just single-compromising constraints. This will force the characterization to be more involved. For the remainder of this section, we will proceed with this characterization. The key insight is to consider marginal mechanisms, defined as follows.

Definition 2. Let $f : \mathscr{P} \to C$ and let M be a proper subset of N. Let $\succsim_{M^c}$ be a profile of preferences of agents not in M. The **marginal mechanism** of f holding M^c at $\succsim_{M^c}$ is denoted $f^M_{\succsim_{M^c}} : P^M \to \mathcal{O}^M$ and is defined by

$$\succsim \mapsto [f_j(\succsim, \succsim_{M^c})]_{j \in M}$$

we will denote $I^M(\succsim_{M^c}) = im(f^M_{\succsim_{M^c}})$ which will be referred to as M's **option set** holding M^c at $\succsim_{M^c}$

Thus a marginal mechanism holds fixed some of the agents' preferences $\succsim_{M^c}$ and defines an M-agent mechanism for the remaining agents, mapping their profile of announcements $\succsim_M$ to an M-agent allocation $f^M_{\succsim_{M^c}}(\succsim) \in \mathcal{O}^M$.

Clearly, marginal mechanisms inherit the group strategy-proofness of the original grand mechanism. The main result in this section shows that, going the other direction, it is enough to check that

[18] Details are available from the authors upon request.

the two-agent marginal mechanisms are group strategy-proof to guarantee that the full mechanism is group strategy-proof.

Theorem 2. *The mechanism $f : \mathscr{P} \to C$ is group strategy-proof if and only if for every pair of agents $\{i, j\}$ and any profile $\succsim_{N\setminus\{i,j\}}$ of the other agents, the marginal mechanism of f holding $N \setminus \{i, j\}$ at $\succsim_{N\setminus\{i,j\}}$ is group strategy-proof.*

For two-agent mechanisms, there is only one group coalition—namely the grand coalition. Therefore group strategy-proofness of a two-agent mechanism is equivalent to individual strategy-proofness and Pareto efficiency on its image.

This drastically reduces the number of conditions one needs to check to ensure that a given mechanism is group strategy-proof. Rather than verifying incentives for all coalitions, it is sufficient to check that no two agents can profitably misreport their preferences. Furthermore, Theorem 2 is especially useful in conjunction with our previous characterization in Theorem 1 for all two-agent mechanisms. Application of Theorem 1 to all marginal mechanisms then provides a more explicit characterization of group strategy-proofness. We can show that the two-agent strategy-proof and Pareto efficient mechanisms form the "building blocks" of all group strategy-proof mechanisms. To do so we will need some notation. Let $F_n = \{f : P^n \to \mathcal{O}^n\}$ and $\mathscr{E}_{n,m} = \{\phi : P^n \to GS(\mathcal{O}^m)\}$. So $f \in F_n$ is just any map from the set of profiles for n agents to the set of possible allocations and any $\sigma \in \mathscr{E}_{n,m}$ provides, for each preference profile of n agents, a group strategy-proof mechanism for m other agents. Likewise, define $\mathscr{F}_{n,m} = \{\eta : P^n \to F_m\}$. We will need the following definition:

Definition 3. If $f \in F_n$ and $g \in F_m$ we may define the **direct sum** $f \oplus g : P^{n+m} \to \mathcal{O}^{n+m}$ by

$$f \oplus g(\succsim) = \left[f\left(\succsim_1, \succsim_2, \ldots, \succsim_n\right), g\left(\succsim_{n+1}, \succsim_{n+2}, \ldots, \succsim_{n+m}\right)\right]$$

This operation extends in the following way. For any $\sigma \in \mathscr{F}_{n,m}$ and $\rho \in \mathscr{F}_{m,n}$, we may define $\sigma \oplus \rho : P^{n+m} \to \mathcal{O}^{n+m}$ to be the map

$$\succsim \;\mapsto\; \left[\rho\left(\succsim_{n+1}, \cdots \succsim_{n+m}\right)\left(\succsim_1, \cdots \succsim_n\right), \sigma\left(\succsim_1, \cdots \succsim_n\right)\left(\succsim_{n+1}, \cdots \succsim_{n+m}\right)\right]$$

The final claim records these observations, explicitly providing a formula that characterizes the set of group strategy-proof mechanisms. This corollary says little other than 2, however explicitly justifies the notion the the two-agent mechanisms form the "building blocks" of arbitrary mechanisms.

Corollary 1.

$$GS(\mathcal{O}^n) = \bigcap_{\tau \in Sym(N)} \tau \circ \left[\mathscr{E}_{n-2,2} \oplus \mathscr{F}_{2,n-2}\right] \circ \tau^{-1}$$

Where $Sym(N)$ is the set of permutations of the agents N.

4 Applications

In this section, we will apply our general characterizations to specific constraints. These applications will feature a new class of mechanisms which are generalizations of serial dictatorships. In a basic

serial dictatorship, agents take turns in a fixed order choosing their favorite objects among all objects which are feasible with the objects chosen by earlier dictators. In principle, the order of future agents might depend on earlier agents' choices. Our generalization of serial dictatorship does exactly that. We begin by formally describing the class of generalized serial dictatorships. We then apply this as well as our characterization results to the social choice problem and the roommates problem.

4.1 Generalized Serial Dictatorship

First, let us recall the definition of a serial dictatorship.

Definition 4. Let $\sigma(1), \ldots, \sigma(N)$ be a strict ordering of the agents $\{1, 2, \ldots, N\}$. For any constraint C, we may define the **serial dictatorship mechanism** which for each preference profile $\succsim$ gives the allocation defined by the following algorithm:

Step 1 Agent $\sigma(1)$ chooses her favorite object a_1 from $\pi_{\sigma(1)}C$. Let μ_1 be the suballocation in which $\sigma(1)$ is assigned a_1 and all other agents are unassigned.

Step k The agent $\sigma(k)$ chooses his favorite object a_k from $\pi_{\sigma(k)}C(\mu_{k-1})$. Let μ_k be the allocation whose graph is $G(\mu_{k-1}) \cup \{(\sigma(k), a_k)\}$. If all agents have been assigned an object, stop. If not, continue to step $k+1$.

Serial dictatorships are well-defined for any constraint and are always group strategy-proof and Pareto efficient.[19] It turns out, however, that we can easily generalize this notion to allow early dictators' choices to determine who will be the subsequent dictator. The main tension here is that, in order to maintain group strategy-proofness, we will have to ensure that the mechanism is nonbossy. That is, the early dictators will not be able to determine the subsequent order arbitrarily, but will be able to determine it only through the expression of their choices.

Recall that $\mathcal{S}$ is the set of suballocations (i.e. the maps $\mu : M \to \mathcal{O}$ where $M \subset N$). Let $\mathcal{S}'$ be the set of incomplete suballocations[20]. A **GSD-ordering** is a map $\zeta : \mathcal{S}' \to N$ such that for any suballocation μ, $\zeta(\mu)$ is an agent not allocated an object under μ. For each GSD-ordering and for any constraint C we may define a **generalized serial dictatorship mechanism** whose allocation at any preference profile is determined by the following algorithm:

Step 1 The agent $d_1 \equiv \zeta(\emptyset)$ is the first dictator. She chooses her favorite object a_1 from $\pi_{d_1}C$. Let μ_1 be the suballocation in which d_1 is assigned a_1 and all other agents are unassigned.

Step k The agent $d_k \equiv \zeta(\mu_{k-1})$ chooses her favorite object a_k from $\pi_{d_k}C(\mu_{k-1})$. Let μ_k be the allocation whose graph is $G(\mu_{k-1}) \cup \{(d_k, a_k)\}$. If all agents have been assigned an object, stop. If not, continue to step $k+1$.

[19] A fact we will prove shortly.
[20] M is a proper subset of N.

Clearly, the standard serial dictatorship is the generalized serial dictatorship mechanism attained by setting $\zeta(\emptyset) = \sigma(1)$, $\zeta(\mu) = \sigma(2)$ for all suballocations μ in which a single agent is matched and so on. Unfortunately, a single mechanism can admit many GSD-orderings, that is, two different orderings might define the same mechanism. This is because the GSD-ordering ζ can be defined in any way off the "algorithm path" in the sense that, suballocations which will never be realized can be assigned any agent. For example, in the serial dictatorship mechanism, any allocation in which a single agent other than the dictator is assigned an object will never be realized, so the GSD assignment there is immaterial to the mechanism. Nevertheless, it is convenient to take $\mathcal{S}'$ as the domain of GSD-orderings. The following proposition shows that generalized serial dictatorships share the good incentive and efficiency properties of serial dictatorships.

Proposition 3. *For any constraint C, the generalized serial dictatorship mechanisms are group strategy-proof and Pareto efficient.*

Notice that this proposition demonstrates that $GS(C)$ is never empty.[21]

We can use these ideas to extend mechanisms defined on projections of the constraint. Suppose we have a constraint C and that for a proper subset $M \subset N$, we have a group strategy-proof and Pareto efficient mechanism f^M on the constraint C^M. Fix a GSD-ordering ζ. We will extend f^M to a mechanism on all of N and all of C by using a generalized serial dictatorship mechanism for agents in $N \setminus M$. In particular, define $(f^M, \zeta) : \mathscr{P} \to C$ via the following algorithm:

Step 1 Allocate $f_i^M(\succsim_M)$ to every agent i in M. Let μ_0 this suballocation. Let agent $d_1 = \zeta(\mu_0)$ choose her favorite object a_1 from among $\pi_{d_1} C(\mu_0)$ and let μ_1 be the suballocation whose graph is $G(\mu_0) \cup \{(d_1, a_1)\}$. If all agents have been allocated an object, stop. Otherwise, proceed to next step.

Step k The agent $d_k \equiv \zeta(\sigma_{k-1})$ chooses her favorite object x_k from $\pi_{d_k} C(\mu_{k-1})$. Let μ_k be the allocation whose graph is $G(\mu_{k-1}) \cup \{(d_k, x_k)\}$. If all agents have been assigned an object, stop. If not, continue to step $k+1$.

Proposition 4. *If $f^M : P^M \to C^M$ is Pareto efficient and group strategy-proof, for any GSD-ordering ζ, the mechanism (f^M, ζ) is group strategy-proof and Pareto efficient.*

4.2 The Roommates Problem

We now apply our general results to the canonical roommates problem. Our main contribution here is characterizing the group strategy-proof and Pareto efficient mechanisms for this problem.

In the roommates problem, an even number of agents who need to be paired as roommates. Each agent has a strict preference over the other agents as roommates. As discussed earlier, we can model this in our environment by letting $\mathcal{O} = N$ and using the constraint

$$C = \{\mu : N \to N \,|\, \mu(i) \neq i \text{ for all } i \text{ and } \mu^2 = id\}$$

[21] So long as the constraint is nonempty, which we assume throughout.

Any feasible mechanism for this constraint will be called a **roommates mechanism**. As mentioned in the introduction, the literature on the roommates problem has focused on the computational complexity of finding stable matching, and there is very little understanding of incentives and efficiency for one-sided matching.

Theorem 3 gives a full characterization of group strategy-proof and Pareto efficient mechanisms for the roommates problem. This is akin to the Gibbard–Satterthwaite Theorem that demonstrates all such mechanisms are dictatorships for the social choice problem and the recent result of Pycia and Ünver (2017) that characterizes all such mechanisms for the house allocation problem, but had not yet been discovered for one-sided matching. We settle this question for the roommates problem, and show that all mechanisms with these properties for the roommates problem are generalized serial dictatorships.

Theorem 3. *A roommates mechanism is group strategy-proof and Pareto efficient if and only if it is a generalized serial dictatorship.*

Although our results are generally unrelated to stability, this is one exception. As mentioned, a defining feature of the roommates problem is the lack of stable outcomes. One approach is to relax stability, with a possible direction to only require that pairs of agents where each ranks the other as her favorite must be matched. This weaker stability condition is called "mutually best" by Toda (2006) and "pairwise unanimity" by Takagi and Serizawa (2010). However, generalized serial dictatorships cannot satisfy even this very weak form of stability. So a corollary of Theorem 3 is that no group strategy-proof and Pareto efficient mechanism can satisfy mutual best or pairwise unanimity, exposing a tension between incentives and stability for the roommates problem. This negative observation for the roommates problem is not new; in fact, this corollary of our result can also be implicitly derived from Theorem 2 of Takamiya (2013) without an explicit characterization of group strategy-proofness.[22] Our constructive approach shows how this tension is related to the structure of the roommates problem as a constraint in our more general environment.

4.3 Social Choice

Here we apply the earlier theorems to provide a new proof for and insights into one of the canonical impossibility results of social choice,z by examining the structure of the social choice problem once it is expressed as a special constraint of our general model.

The first theorem in implementation theory was the celebrated negative result of Gibbard (1973) and Satterthwaite (1975) that the only strategy-proof and surjective social choice mechanisms are dictatorships. Since Pareto efficient mechanisms are necessarily surjective, this negative finding illuminates a fundamental tension between incentives and efficiency for social decisions. This tension can also be deduced as a corollary of our main result. Beyond providing a novel proof, our approach to the Gibbard–Sattertwhaite Theorem yields additional insights that help understand the theorem more deeply. First, our environment for the theorem, in a model that includes social choice as a special case, demonstrates that the reason why social choice must yield a simple dictatorship, rather than a serial dictatorship, is because the structure of the constraint forces all agents' allocations to

[22] We thank Yuichiro Kamada for pointing this out to us.

be immediately determined by fixing the dictator's allocation. If this feature is relaxed, then the dictator could consume her favorite object while still leaving flexibility in the allocation for other agents, that is, serial dictatorship is possible. So our approach shows how the dictatorship implied by the Gibbard–Satterthwaite Theorem can be seen as a special case of a more general feature of serial dictatorship.

Second and related, an immediate corollary of our main result is if all group strategy-proof mechanisms are serial dictatorships, then the marginal T relation, derived from the marginal constraint $C^{i,j}$, can have only one equivalence class. This provides a converse to the Gibbard–Sattertwhaite Theorem, showing that if all group strategy-proof mechanisms are serial dictatorships, then the constraint C must have a special structure. Again, this converse is only well-posed in a model where social choice is cast as a special case of private goods allocation, rather than vice versa as is more traditional.

One convenient feature of the diagonal social choice constraint is that, since all mechanisms are necessarily nonbossy to satisfy the constraint, there is no gap between group and individual strategy-proofness.[23]

Lemma 4. *Let C be the social choice constraint, i.e. $C = \{(a_i)_{i\in N} \mid a_i = a_j \text{ for all } i, j \in N\}$ then a map $f : \mathscr{P} \to C$ is group strategy-proof if and only if it is individually strategy-proof.*

We can then apply our main characterization results to the special case of the diagonal social choice constraint to derive that all group strategy-proof and onto mechanisms are dictatorships, which by virtue of Lemma 4 is equivalent to the Gibbard–Satterthwaite Theorem.

Theorem 4 (Gibbard–Satterthwaite)**.** *If $|\mathcal{O}| > 2$ and $f : \mathscr{P} \to C$ is surjective and strategy-proof then it is dictatorial.* [24]

As mentioned, the setup of our model enables us to sensibly ask the converse question: which types of constraints, beyond the diagonal social choice constraint, have the feature that all of the feasible, group strategy-proof mechanisms are (in some sense) dictatorial? In our context, the appropriate form of dictatorship is generalized serial dictatorship, since these always exist and specialize to dictatorship in the social choice setting. As a consequence of Proposition 4 and Theorem 1 we can show that if any two-agent projection of the constraint is such that T has two equivalence classes, then $GS^N(C)$ admits mechanisms beyond GSD.

Theorem 5. *If a constraint C is such that for some i, j, the equivalence relation T on $C^{i,j}$ admits more than one equivalence class, $GS^n(C)$ is strictly larger than the set of generalized serial dictatorship mechanisms.*

[23] This observation can also be alternatively deduced directly from the Gibbard–Sattertwhaite Theorem, since dictatorships are both individual and group strategy-proof. Since our aim is to prove that theorem, this is clearly not valid for our approach.

[24] In fact, we only need that $|im(f)| > 2$ in which case we could drop items never allowed and recover the same statement.

A Appendix

It will be convenient to introduce some additional notation for the proofs. If A and B are sets of objects and $\succsim \in P$, we say $A \succsim B$ if $a \succsim b$ for all $a \in A$ and $b \in B$. For disjoint sets of objects $A_1, A_2 \dots A_m$ we will denote

$$P[A_1, A_2 \dots A_m] = \{\succsim \in P \mid A_1 \succ A_2 \succ \cdots \succ A_m\}$$

and

$$P^{\uparrow}[A_1, A_2 \dots A_m] = \left\{ \succsim \in P \mid A_j \succ \mathcal{O} \setminus \bigcup_{i=1}^{j} A_i \text{ for all } j \right\}$$

When the A_i are singletons, we will abuse notation and drop the curly brackets, writing for example $P^{\uparrow}[a]$ to denote $P^{\uparrow}[\{a\}]$.

A.1 Proof of Propositon 1

We first need the following lemma, which is simply the forward direction of Lemma 3:

Lemma 5. *Let $f : \mathscr{P} \to \mathcal{A}$ be strategy-proof. Then for each i there is a nonempty correspondence $g_i : P^{n-1} \rightrightarrows \mathcal{O}$ such that for all $\succsim$*

$$f(\succsim) = \left(\max_{\succsim_i} g_i(\succsim_{-i}) \right)_{i \in N}$$

Proof. Define $g_i(\succsim_{-i}) = f_i(P, \succsim_{-i})$ then the result follows from strategy-proofness. □

We can now demonstrate the desired implications for the equivalence in turn:

(1) $\implies$ (2): Of course any group strategy-proof mechanism is individually strategy-proof. Suppose there is a profile $\succsim$ and an agent i with an alternative announcement $\succsim_i'$ such that $f_i(\succsim) = f_i(\succsim_i', \succsim_{-i})$ but for some j, $f_j(\succsim) \neq f_j(\succsim_i', \succsim_{-i})$. Then if $f_j(\succsim) \succ_j f_j(\succsim_i', \succsim_{-i})$, the coalition $\{i, j\}$ can improve their outcome at $(\succsim_i', \succsim_{-i})$ by announcing $(\succsim_i, \succsim_j)$. Conversely, if $f_j(\succsim) \prec_j f_j(\succsim_i', \succsim_{-i})$, the coalition $\{i, j\}$ can improve their outcome at $\succsim$ by announcing $(\succsim_i', \succsim_j)$.

(2) $\implies$ (3): Suppose we have two profiles $\succsim, \succsim' \in \mathscr{P}$ such that

$$LC_{\succsim_i'}[f_i(\succsim)] \supset LC_{\succsim_i}[f_i(\succsim)] \text{ for all } i$$

then notice that $f_1(\succsim_1', \succsim_2, \dots, \succsim_n) = f_1(\succsim)$ by Lemma 5 and by nonbossiness we have $f(\succsim_1', \succsim_2, \dots, \succsim_n) = f(\succsim)$. We can proceed, changing one preference at a time, to show that $f(\succsim') = f(\succsim)$ as desired.

(3) $\implies$ (1): Suppose f is Maskin monotonic; we will show that f is group strategy-proof. Let $\succsim \in \mathscr{P}$ and $\succsim_A'$ be a candidate violation for agents in A so that

$$f(\succsim_A', \succsim_{-A}) \succsim_j f(\succsim) \text{ for all } j \in A$$

we will show that this implies $f(\succsim_A', \succsim_{-A}) = f(\succsim)$. For each $j \in A$ construct $\succsim_j^*$ to be identical to $\succsim_j$

except that it puts $f_j(\succsim'_A, \succsim_{-A})$ first. For any $j \in A$ we have

$$LC_{\succsim_j^*}(f_j(\succsim'_A, \succsim_{-A})) \supset LC_{\succsim_j}(f_j(\succsim'_A, \succsim_{-A})) \text{ and}$$
$$LC_{\succsim_j^*}(f_j(\succsim)) \supset LC_{\succsim_j}(f_j(\succsim))$$

for all j. The first is immediate. To see the second, notice that if $f_j(\succsim'_A, \succsim_{-A}) = f_j(\succsim)$ then it holds trivially. If instead, $f_j(\succsim'_A, \succsim_{-A}) \neq f_j(\succsim)$, by assumption we have $f_j(\succsim'_A, \succsim_{-A}) \succ_j f_j(\succsim)$ and since $\succsim^*$ only moves up the position of $f_j(\succsim'_A, \succsim_{-A})$, the second statement holds. However, by Maskin monotonicity, the first statement gives $f(\succsim^*_A, \succsim_{-A}) = f(\succsim'_A, \succsim_{-A})$ and the second gives $f(\succsim^*_A, \succsim_{-A}) = f(\succsim)$, so putting them together we get

$$f(\succsim'_A, \succsim_{-A}) = f(\succsim^*_A, \succsim_{-A}) = f(\succsim)$$

as desired. □

A.2 Proof of Lemma 1

By way of contradiction, suppose that $f : \mathscr{P} \to im(f)$ is group strategy-proof and that there is a profile $\succsim$ and an allocation $(a_i)_{i\in N} \in im(f)$ such that $a_i \succsim_i f_i(\succsim)$ for all i with at least one strict. By definition, there is an alternative profile $\succsim'$ such that $f(\succsim') = (a_i)_{i\in N}$ which is a profitable deviation from $\succsim$. □

A.3 Proof of Lemma 2

Let $\{g_i\}_{i\in N}$ be as in Lemma 3. For each j the preference $\succsim'_j$ does not change the relative ranking of the objects in $g_j(\succsim_{-j})$ hence we have $f_j(\succsim'_j, \succsim_{-j}) = f_j(\succsim)$ so by nonbossiness $f(\succsim'_j, \succsim_{-j}) = f(\succsim)$. Repeating this argument one agent at a time gives the result. □

A.4 Proof of Theorem 1 (Two-agent characterization)

($\Longleftarrow$) Applying lemma 3, we see that local dictatorships are strategy-proof. They are Pareto efficient by construction.

($\Longrightarrow$) If $C = \mathcal{O}^2$ then any Pareto efficient mechanism always gives both agents their top choice, which is trivially a local dictatorship.

Suppose now that C is a nonempty, proper subset of $\mathcal{O}^2$. By Lemma 2, it is without loss to assume that for any $(a, b) \in \bar{C}$ there are a' and b' with (a', b) and (a, b') in C. Fix $f : P^2 \to C$ which is strategy-proof and Pareto efficient.[25] The proof will proceed in two steps. First we show that for any infeasible allocation (a, b) there is a local dictator who gets their top choice at every preference profile where a and b are top-ranked respectively. Then we show that the local dictator is constant within blocks.

Let $(a, b) \in \bar{C}$ and a', b' as above. Let $\succsim_1 \in P^{\uparrow}[a, a']$ and $\succsim_2 \in P^{\uparrow}[b, b']$. By Pareto efficiency, $f(\succsim_1, \succsim_2) = (a, b')$ or $f(\succsim_1, \succsim_2) = (a', b)$. Assume without loss that $f(\succsim_1, \succsim_2) = (a, b')$. We will show

[25] Serial dictatorship always is both Pareto efficient and strategy-proof (as shown in proposition 3, so the set is nonempty.

that this implies that 1 is the local dictator at (a,b). Pick any other $\succsim_2'$ which top-ranks b. By 2's strategy-proofness, $f_2(\succsim_1,\succsim_2') \neq b$, but then from Pareto efficiency, $f_1(\succsim_1,\succsim_2') = a$, since otherwise, the allocation (a',b) would Pareto dominate $f(\succsim_1,\succsim_2')$. Thus $f_1(\succsim_1,\succsim_2') = a$ whenever $\succsim_2' \in P^{\uparrow}[b]$. By 1's strategy-proofness, we have that $f_1(\succsim_1',\succsim_2') = a$ for all $\succsim_1',\succsim_2'$ with $\tau(\succsim_1',\succsim_2') = (a,b)$. Finally, by Pareto efficiency, $f(\succsim_1',\succsim_2') = (a, \max_{\succsim_2'} C^2(a))$ whenever $\tau(\succsim_1',\succsim_2') = (a,b)$. Thus we say that 1 is the local dictator at (a,b). Since (a,b), was arbitrary every other infeasible allocation has a local dictator.

Now suppose that $(a,b)B(a',b')$ and $(a,b) \neq (a',b')$. Then either $a = a'$ or $b = b'$. Without loss, assume $a = a'$. Suppose by way of contradiction that, that (a,b) and (a,b') have different local dictators. For example, suppose 1 is the local dictator at (a,b) and 2 is the local dictator at (a,b'). Consider the preference profile $(\succsim_1,\succsim_2)$ where $\succsim_1 \in P^{\uparrow}[a]$ and $\succsim_2 \in P^{\uparrow}[b,b',b'']$ where b'' is such that $(a,b'') \in C$. Then from the analysis above, we get $f(\succsim_1,\succsim_2) = (a,b'')$ since 1 is the local dictator at (a,b). However, if $\succsim_2' \in P^{\uparrow}[b']$, then $f_2(\succsim_1,\succsim_2') = b' \succ_2 b'' = f_2(\succsim_1,\succsim_2)$ since 2 is the local dictator at (a,b'), which is a violation of strategy-proofness. Thus either 1 is the local dictator at (a,b) and (a,b') or 2 is. For any two infeasible allocations (a,b) and (a',b') in an equivalence class of T, there is a sequence of infeasible allocations such that $(a,b)B(a_1,b_1)B\cdots B(a_n,b_n)B(a',b')$, so (a,b) and (a',b') have the same local dictator. □

A.5 Proof of Proposition 2

First we show that every group strategy-proof and Pareto efficient mechanism is constraint-traversing. Let C be a single-compromising constraint and fix and a group strategy-proof, Pareto efficient mechanism $f : \mathscr{P} \to C$. Let $a = (a_i)_{i\in N}$ be infeasible. For every i there is an object a_i' such that $(a_i', a_{-i}) \in C$. Let $\succsim_i \in P^{\uparrow}[a_i, a_i']$ for each i. Since f is feasible, there is at least one agent k who doesn't get their top choice at the constructed preference profile $\succsim = (\succsim_i)_{i\in N}$. However, Pareto-efficiency then implies that $f_i(\succsim) = a_i$ for all $i \neq k$ and $f_k(\succsim) = a_k'$. By Maskin monotonicity and Lemma 5 we have that for any $\succsim_{-k}'$ with $\max_{\succsim_j'} \mathcal{O} = a_j$ for all $j \neq k$, $a_k \notin g_k(\succsim_{-k})$, so that k always compromises when the top choice is a. Define $\alpha(a) = k$ (we can do this unambiguously because no other agent always compromises at a, e.g. at the profile $\succsim$). Since a was an arbitrary infeasible allocation, we can do the same for any other infeasible allocation to define α on all of $\bar{C}$. Finally, we establish inductively that f is constraint-traversing according to α. Pick any preference profile $\succsim'$. Start at $a^1 = (\max_{\succsim_i'} \mathcal{O})_{i\in N}$. If this is feasible, then f being Pareto efficient implies $f(\succsim') = a^1$. Otherwise, it is infeasible, and by the previous argument, we have an agent $k = \alpha(a^1)$ who must compromise. Replace $\succsim_k'$ with the same preference, except that it puts a_k^1 last. By Maskin monotonicity, this cannot affect the outcome of f. We therefore repeat the above process at the new profile. This is exactly how the constraint-traversing mechanism according to α works, giving the result.

Now we need to show that α has to satisfy the property that if $\alpha(a) = i$ then for any $(a_i', a_{-i}) \in \bar{C}$, we have $\alpha(a_i', a_{-i}) = \{i\}$. However this follows from similar reasoning as in the two-agent case. If, instead $k = \alpha(a_i', a_{-i})$ consider the profile $\succsim$ with $\tau(\succsim) = a$ and $\tau_2(\succsim_i) = a_i'$ and $\tau_2(\succsim_k) = a_k'$ where $(a_k', a_{-k}) \in C$. We get a violation of Pareto efficiency since the constraint-traversing algorithm would make both i and k compromise to their second-best choice, which would be Pareto dominated by (a_k', a_{-k}).

The fact that this mechanism is group strategy-proof and Pareto efficient is now a simple consequence of Maskin monotonicity and Proposition 1. □

A.6 Proof of Theorem 2 (N-agent characterization)

If f is group strategy-proof, the marginal mechanisms are group strategy-proof by definition. For the other direction, suppose that every two-agent marginal mechanism is group strategy-proof. Then f is individually strategy-proof since for any i and any profile $\succsim$ we can choose $j \neq i$ and consider the marginal mechanism $f^{i,j}_{\succsim_{-i,j}}$ then in this marginal mechanism i cannot profit from misreporting, hence she cannot in f. It remains to show that f is nonbossy. Now suppose we have $f_i(\succsim_i', \succsim_{-i}) = f_i(\succsim)$ and for some j, $f_j(\succsim_i', \succsim_{-i}) \neq f_j(\succsim)$, either $f_j(\succsim_i', \succsim_{-i}) \succ_j f_j(\succsim)$ or $f_j(\succsim_i', \succsim_{-i}) \prec_i f_j(\succsim)$. However, by assumption the marginal mechanism $f^{ij}_{\succsim_{-ij}}$ is group strategy-proof. From the two-agent characterization, no two-agent group strategy-proof mechanism can have this property. □

A.7 Proof of Corollary 1

The proof is an immediate application of Theorem 2. □

A.8 Proof of Proposition 3

Maskin monotonicity is easily seen to be satisfied, since starting from the first dictator, each agent will be given the same option set and will weakly prefer their original choice to any alternative. To see that it is Pareto efficient, by Lemma 1 it is enough to establish that its image is exactly C. By construction, the image is a subset of C. For any feasible allocation $a \in C$ let $\succsim_i$ put a_i first. Then $f(\succsim) = a$ so $im(f) = C$. □

A.9 Proof of Proposition 4

We will show that (f^M, ζ) is Maskin monotonic and Pareto efficient. Pick any $\succsim \in \mathscr{P}$ and let $\succsim'$ satisfy the conditions in the definition of Maskin monotonicity. I.e.

$$LC_{\succsim_i'}\left[(f^M, \zeta)_i(\succsim)\right] \supset LC_{\succsim_i}\left[(f^M, \zeta)_i(\succsim)\right] \text{ for all } i$$

Since f^M is group strategy-proof for the agents in M, it is Maskin monotonic. Hence we have $f^M(\succsim_M) = f^M(\succsim_M')$, then by definition, $(f^M, \zeta)_i(\succsim') = (f^M, \zeta)_i(\succsim)$ for all $i \in M$. As a consequence, the sequence of dictators is the same. Thus we have Maskin monotonicity.

By Lemma 1 it is enough to establish that the image of (f^M, ζ) is exactly C. To see this, let $(a_i)_{i \in N} \in C$, since f^M is Pareto efficient on C^M there is some profile $\succsim_M$ with $f^M(\succsim_M) = (a_i)_{i \in M}$. For agents not in M let $\succsim_j \in P^{\uparrow}(a_j)$. At this profile, we have $(f^M, \zeta) = (a_i)_{i \in N}$ as desired. □

A.10 Proof of Theorem 3 (Roommates characterization)

The "if" direction follows directly from Proposition 3.

We will prove the "only if" Theorem by mathematical induction. First, by Lemma 2, we can ignore any agents' ranking of themselves, which we will do. If $N = 2$ there is only one possible allocation, so every mechanism is trivially a generalized serial dictatorship. Furthermore, if $N = 4$ one can show that we can rewrite the problem as a social choice problem since a single agents' match determines the full outcome. In this case, the result follows from the Gibbard–Satterthwaite Theorem. Suppose that for all $m < n$ when there are $2m$ agents, all group strategy-proof and Pareto efficient roommates mechanisms are generalized serial dictatorships. We will show this for $2n$ agents. It will be enough to show that there is an agent j such that $f_j(\succsim) = \max_{\succsim_j} N$ for all $\succsim$, since, conditional on each of j's choices, the remaining $2n-2$ agents need to assigned a roommate, which itself gives a roommates mechanism guaranteed to be a generalized serial dictatorship by the induction assumption.

Let f be a group strategy-proof and Pareto efficient roommates mechanism for $2n$ agents with $n \geq 3$. We will first consider the possible two-agent marginal mechanisms. Let $i \neq j$ and fix a profile $\succsim_{-ij}$ of the other agents. Assume $(j,i) \in I^{ij}(\succsim_{-ij})$, so that it is possible for i and j to match when the other agents announce $\succsim_{-ij}$. For all $k \neq i$, $(j,k) \notin I^{ij}(\succsim_{-ij})$ since (j,k) has i matched to j but j matched to k. Likewise, for all $k \neq j$ we have $(k,i) \notin I^{ij}(\succsim_{-ij})$. Define $R_i = \{x \in N \,|\, (x,y) \notin I^{ij}(\succsim_{-ij}) \text{ for all } y \in N\}$ and $R_j = \{y \in N \,|\, (x,y) \notin I^{ij}(\succsim_{-ij}) \text{ for all } x \in N\}$. Then we get a marginal constraint like the one shown on the left of Figure 3

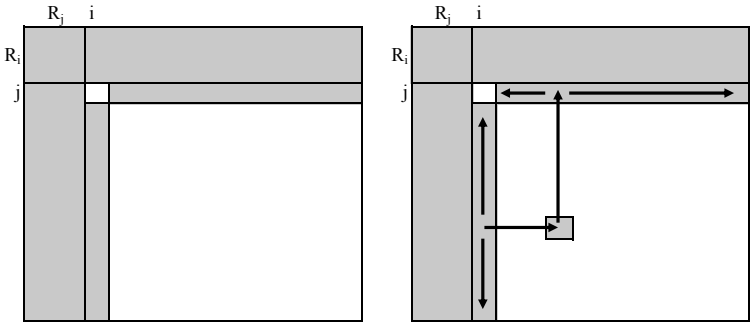

Figure 3: $I^{ij}(\succsim_{-ij})$

with the exception that some non-grey squares on the bottom right may actually be infeasible. If $[N - R_i \cup \{j\}] \times [N - R_j \cup \{i\}]$ intersects any infeasible point, then the equivalence relation T has a single equivalence class, as shown in on the right of Figure 3.[26] Therefore there must be a single dictator in the marginal mechanism $f^{ij}_{\succsim_{-ij}}$ by Theorems 1 and 2. Otherwise, every allocation in $[N - R_i \cup \{j\}] \times [N - R_j \cup \{i\}]$ is infeasible or the set is empty. In the latter case, there is of course only one marginal mechanism. In the former case, as a consequence of theorem 1 there are three possible Pareto efficient, strategy-proof marginal mechanisms as illustrated in figure 4.

Summarizing, if $(j,i) \in I^{ij}(\succsim_{-ij})$, there are four possible types of mechanisms $f^{ij}_{\succsim_{-ij}}$. (1) we could have $\{(j,i)\} = I^{ij}(\succsim_{-ij})$ so $f^{ij}_{\succsim_{-ij}}$ is constant ; (2) i could be the only dictator; (3) j could be the only dictator ;(4) we have the mechanism in panel (A) of figure 4.
We will now need the following lemma.

[26] Recall the relation T was defined immediately before the statement of Theorem 1.

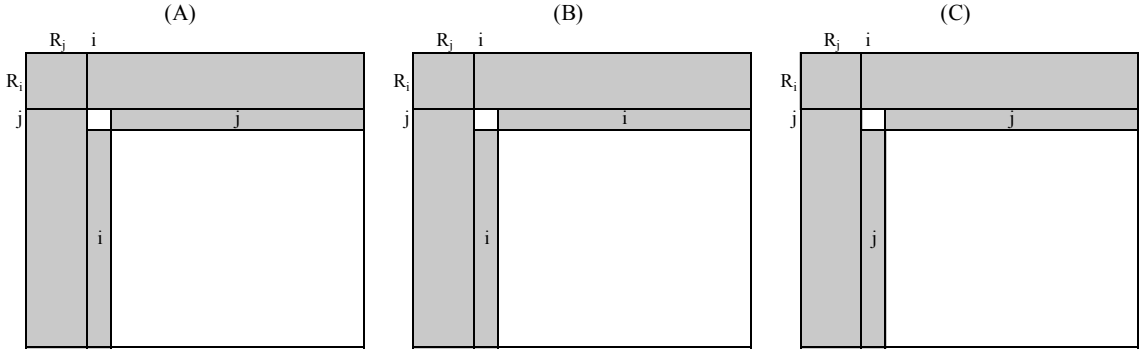

Figure 4: The three possible mechanisms $f^{ij}_{\succsim_{-ij}}$

Lemma 6. *Let A be a nonempty, proper subset of N with an even number of agents and $|A| \geq 4$. If $\succsim^*_{N\setminus A} \in \left[P^{\uparrow}(N \setminus A)\right]^{N\setminus A}$ then there is an agent $j \in A$ such that*

$$f_j(\succsim_A, \succsim^*_{N\setminus A}) = \max_{\succsim_j} N$$

whenever $\max_{\succsim_j} N \in A$

Proof. Suppose that we restrict attention to $\succsim_A \in \left[P^{\uparrow}(A)\right]^A$, then by Pareto efficiency, the agents in A are matched with one another, and the agents in $N \setminus A$ are also matched to one another. But by group strategy-proofness and Pareto efficiency of f,

$$f(\cdot, \succsim^*_{N\setminus A})|_{[P^{\uparrow}(A)]^A}$$

gives a group strategy-proof and Pareto efficient roommates mechanism for the agents in A.[27] By the induction assumption, this mechanism is a generalized serial dictatorship. Thus there is a j such that $f_j(\succsim_A, \succsim^*_{N\setminus A}) = \max_{\succsim_j} N$ whenever $\succsim_A \in \left[P^{\uparrow}(A)\right]^A$. Thus it remains to show, that j gets her top choice regardless of the announcements of the other agents in A so long as her top choice is in A and the agents in $N \setminus A$ announce $\succsim^*_{N\setminus A}$. To see this, let $\succsim_A$ be arbitrary except that $\max_{\succsim_j} N \in A$. For each $i \neq j$ in A, let $\succsim'_i \in P^{\uparrow}(A)$ put j top. Then $g_j(\succsim'_A, \succsim^*_{N\setminus A}) \supset A - \{j\}$. Consider any i, j mechanism with $i \in A$. Since $(i,j) \in I^{i,j}(\succsim'_-, \succsim^*_{N\setminus A})$ there are four options for $f^{ij}_{(\succsim'_-, \succsim^*_{N\setminus A})}$. However, the only one consistent with the fact that $g_j(\succsim'_A, \succsim^*_{N\setminus A}) \supset A - \{j\}$, (which given that $|A| \geq 4$ leaves j the option of matching with agents other than i) is that j is the only dictator in this marginal mechanism. Hence we have $g_j(\succsim_i, \succsim'_-, \succsim^*_{N\setminus A}) \supset A - \{j\}$, since i's announcement cannot affect j's option set. Repeating this argument one agent at a time gives that $g_j(\succsim_A, \succsim^*_{N\setminus A}) \supset A - \{j\}$, which is the desired result. □

We will call agent j in the lemma above, the *marginal dictator*. Having done this, the idea is to partition the agents in two ways. First we consider the partition $\{1,2\}\{3,4,\ldots,N\}$. By lemma 6 there is a marginal dictator among $\{3,4,\ldots,N\}$. Second, we consider the partition $\{1,2,3,4\}, \{5,6,\ldots,N\}$ and again lemma 6 says there is a marginal dictator among $\{1,2,3,4\}$. We show that the marginal

[27] Of course, a roommates mechanisms for the agents in A has a different domain than $f(\cdot, \succsim^*_{N\setminus A})$,but by nonbossiness the ranking of agents in $N \setminus A$ is immaterial to the mechanism when we restrict attention to $\left[P^{\uparrow}(A)\right]^A$.

dictators among these two overlapping sets determine a single dictator for the whole mechanism. If $A = \{3, 4 \ldots, N\}$, $\succsim_1^* \in P^\uparrow(2)$ and $\succsim_2^* \in P^\uparrow(1)$, there is a marginal dictator among A. Without loss, suppose this marginal dictator is 3 so that $g_3(\succsim_1^*, \succsim_2^*, \succsim_-') \supset A - \{3\}$ for all $\succsim_-'$. Now, for agents $5, \ldots N$, let $\succsim_k^* \in P^\uparrow[\{5, \ldots, N\}]$ then we also have a marginal dictator among the $\{1, 2, 3, 4\}$. Now at the profile $(\succsim_1^*, \succsim_2^*, \succsim_3, \succsim_4, \succsim_-^*)$, where 3 top ranks 4, then 3 and 4 are matched regardless of 4's announcement. Hence the marginal dictator can't be 4. This leaves three cases: the marginal dictator could be $1, 2$, or 3. Since, 1 and 2 are thus far symmetric, we will handle both of these cases at the same time.

First, however, we will start with the case where 3 is the marginal dictator. In this case, $g_3(\succsim_1', \succsim_2', \succsim_4', \succsim_-^*) \supset \{1, 2, 4\}$ for all $\succsim_1', \succsim_2'$ and $\succsim_4'$ by Lemma 6. In addition, $g_3(\succsim_1^*, \succsim_2^*, \succsim_4', \succsim_-^*) \supset N - \{1, 2, 3\}$. Putting these together we have that $g_3(\succsim_1^*, \succsim_2^*, \succsim_4', \succsim_-^*) = N - \{3\}$. Now, we know 3 is the marginal dictator in the $(1, 3), (2, 3)$ and $(3, 4)$ marginal mechanisms when all the other agents are announcing $\succsim_-^*$. Therefore, $g_3(\succsim_1'', \succsim_2^*, \succsim_4', \succsim_-^*) = N - \{3\}$ for any $\succsim_1''$. But again, 3 is the marginal dictator between 1 and 2, so $g_3(\succsim_1'', \succsim_2'', \succsim_4', \succsim_-^*) = N - \{3\}$ for any $\succsim_2''$ and so on. Hence we actually have $g_3(\succsim_1', \succsim_2', \succsim_4', \succsim_-^*) \supset N - \{3\}$ for all $\succsim_1'', \succsim_2''$ and $\succsim_4''$. Now let $\succsim_1^{**}$ be the same as $\succsim_1^*$, except that 3 is top-ranked. Likewise, let $\succsim_2^{**}$ be the same as $\succsim_2^*$, except that 3 is top-ranked. In addition, let $\succsim_4^{**}$ $P^\uparrow[3]$. From above, we have $g_3(\succsim_1^{**}, \succsim_2^{**}, \succsim_4^{**}, \succsim_-^*) \supset N - \{3\}$. Let's now consider the $(3, k)$ marginal mechanism for $k > 4$. Of course, $(3, k) \in I^{3,k}(\succsim_1^{**}, \succsim_2^{**}, \succsim_4^{**}, \succsim_-^*)$ so again we have four possible cases, however there are only two cases consistent with the fact that 3 can match k if she top-ranks her or match, say, 1 if she top-ranks 1 instead. Either 3 is the only dictator in the marginal mechanism, or the mechanism is like panel (A) of figure 4. In the latter case, we have $f_3(\succsim_1^{**}, \succsim_2^{**}, \succsim_3', \succsim_4^{**}, \succsim_-^*) = k$ whenever $\succsim_k$ top-ranks 3 for any $\succsim_3'$. However, then neither 1 or 2 are getting their top choice. By Maskin monotonicity,

$$f_3(\succsim_1^{**}, \succsim_2^{**}, \succsim_3', \succsim_4^{**}, \succsim_-^*) = f_3(\succsim_1^*, \succsim_2^*, \succsim_3', \succsim_4^{**}, \succsim_-^*)$$

Yet if $\succsim_3'$ top-ranks 4, we get a contradiction since 3 is the marginal dictator of $\{3, \ldots, N\}$ at the profile $(\succsim_1^*, \succsim_2^*)$. Hence 3 is the dictator in the in the $(3, k)$ marginal mechanism. So if we switch $\succsim_k^*$ to $\succsim_k^{**}$ which is the same, except that 3 is top-ranked, 3's option set is unchanged. Repeating the same argument for all other agents shows that

$$g_3(\succsim_{-3}^{**}) \supset N - \{3\}$$

where $\succsim_l^{**}$ top ranks 3 for all l. However, repeating our analysis from earlier, we find that in every $(3, l)$ marginal mechanism 3 is the local dictator, and finally that $g_3(\succsim_-) \supset N - \{3\}$ for all $\succsim_{-3}$. This gives the desired result and proves the theorem for this case.

For the second case in which 1 or 2 is the marginal dictator among the agents $1, 2, 3, 4$ when the other agents announce $\succsim_-^*$, the strategy will be to reduce this to case 1 by showing that this agent is the dictator among $N \setminus \{3, 4\}$ for some announcement in which 3 and 4 top-rank each other. Up to relabeling, this is the same as case 1. Since 1 and 2 are symmetric, without loss, suppose the dictator is 1. Then we have

$$f_1(\succsim_1', \succsim_2', \succsim_3', \succsim_4', \succsim_-^*) = \max_{\succsim_1'} N$$

whenever $\max_{\succsim'_1} N \in \{1,2,3,4\}$ for all $\succsim'_1, \succsim'_2, \succsim'_3, \succsim'_4$. Or, equivalently, $g_1(\succsim'_2, \succsim'_3, \succsim'_4, \succsim^*_-) \supset \{2,3,4\}$ for all $\succsim'_2, \succsim'_3, \succsim'_4$. For any k, if $\succsim_3 \in P^{\uparrow}[k]$, then by previous discussion, $f_3(\succsim^*_1, \succsim^*_2, \succsim_3, \succsim'_4, \succsim^*_-) = k$ for any $\succsim'_4$. However, since 1 is the marginal dictator, we also have that $f_1(\succsim^*_1, \succsim^*_2, \succsim_3, \succsim'_4, \succsim^*_-) = 2$. Now if we change $\succsim^*_2$ to $\succsim^{**}_2$ by putting 3 on top and we change $\succsim_3$ to $\succsim^{**}_3$ by putting 1 on top, then since 1 is the marginal dictator, neither of these changes affect the outcome of 1 and therefore 2, so by Maskin monotonicity, $f(\succsim^*_1, \succsim^{**}_2, \succsim^{**}_3, \succsim'_4, \succsim^*_-) = f(\succsim^*_1, \succsim^*_2, \succsim_3, \succsim'_4, \succsim^*_-)$. However, now consider the $(1,k)$ marginal mechanism. If $\succsim^{**}_1$ top-ranks k, but is otherwise unchanged and $\succsim^{**}_k$ top-ranks 1, but is otherwise unchanged. then if $(1,k)$ are not matched at the profile $(\succsim^{**}_1, \succsim^{**}_2, \succsim^{**}_3, \succsim'_4, \succsim^{**}_k, \succsim^*_-$, by Maskin monotonicity, the f is unchanged. However, this outcome would be Pareto dominated by the match in which 1 and k are matched, 2 and 3 are matched and all other matches are unchanged. Thus we have that $(1,k) \in I^{1,k}(\succsim^{**}_2, \succsim^{**}_3, \succsim'_4, \succsim^*_-)$. Again, this leaves us with four options. However, only 1 as the marginal dictator fits with the fact that, if 1 top-ranks 3, they are matched, and k's match is changed. In the three other cases this cannot happen. Therefore, 1 is the dictator in this marginal mechanism. As a consequence we have that $k \in g_1(\succsim^{**}_2, \succsim^{**}_3, \succsim'_4, \succsim^*_-)$. Since $2, 3$, and 4 cannot affect 1's option set, we have $k \in g_1(\succsim'_2, \succsim'_3, \succsim'_4, \succsim^*_-)$ for all $\succsim'_2, \succsim'_3, \succsim'_4$. Since k was chosen arbitrarily, we actually have $g_1(\succsim'_2, \succsim'_3, \succsim'_4, \succsim^*_-) = N - \{1\}$ for all $\succsim'_2, \succsim'_3, \succsim'_4$. However, going back to Lemma 6, if 3 and 4 top-rank each other, there is a dictator among the other agents. The fact that 1's option set is $N - \{1\}$ is only compatible with 1 being that dictator. This gets us back to case 1, and repeating the argument there, we see that 1 always gets her top choice. By the induction assumption, we are done. □

A.11 Proof of Lemma 4

Nonbossiness is immediate. Then the result follows from the observation that strategy-proofness and nonbossiness are equivlent to group strategy-proofness, recorded in Propostion 1. □

A.12 Proof of Theorem 4 (Gibbard–Satterthwaite Theorem)

Let C be the diagonal and $|\mathcal{O}| \geq 3$.

From Proposition 1, it suffices to show that any group strategy-proof mechanism is dictatorial. We will show this in two steps. First, we will show that for some i, j and some profile $\succsim_{-ij} = (\succsim_k)_{k \neq i,j}$ we have $|I^{ij}(\succsim_{-ij})| \geq 3$. From the characterization of two-agent mechanisms, we will see that $f^{ij}_{\succsim_{-ij}}$ is dictatorial. We will then show that this implies the entire mechanism is dictatorial.

1. Suppose by way of contradiction that for all i, j and all $\succsim_{-ij}$ we have $|I^{ij}(\succsim_{-ij})| < 3$. First, note that if for all i, j and all $\succsim_{-ij}$ we have $|I^{ij}_{\succsim_{-ij}}| = 1$ then f is single-valued[28] which contradicts the surjectivity of f. Hence there is at least one pair of agents i, j and $\succsim_{-ij}$ such that $|I^{ij}(\succsim_{-ij})| \geq 2$. For simplicity and without loss, let $i = 1$ and $j = 2$. By assumption then $|I^{ij}(\succsim_{-ij})| = 2$ and without loss assume $I^{ij}(\succsim_{-ij}) = \{a, b\}$. Then there must be a local dictator assigned to the incompatible pairs (a, b) and (b, a). This leaves (up to symmetry) two marginal mechanisms ϕ_1

[28] To see that $f(\succsim) = f(\succsim')$, change one preference at a time. No single change can alter f, so we get the result.

and ϕ_2 where

$$\phi_1(\succsim_1, \succsim_2) = \begin{cases} a & \text{if } a \succ_1 b \\ b & \text{if } a \prec_1 b \end{cases}$$

and

$$\phi_2(\succsim_1, \succsim_2) = \begin{cases} a & \text{if } a \succ_1 b \text{ and } a \succ_2 b \\ b & \text{otherwise} \end{cases}$$

In the first, agent 1 is a dictator. In the second, b is chosen by default and a is only chosen if both agents prefer it to b. Let c be another object in $\mathscr{O}$. If we let $\succsim_2^* \in \mathscr{P}^{\uparrow}[c, a, b]$ then in either case we have $f(\succsim_1, \succsim_2^*, \succsim_{-1,2}) = a$ if $a \succ_1 b$ and $f(\succsim_1, \succsim_2^*, \succsim_{-1,2}) = b$ if $b \succ_1 a$. We then have that a and b are in $I^{1,3}(\succsim_2^*, \succsim_4, \ldots, \succsim_n)$. As before we have two possible mechanisms and in either one, if $\succsim_3^* \in \mathscr{P}^{\uparrow}[c, a, b]$ we have $f(\succsim_1, \succsim_2^*, \succsim_3^*, \succsim_4, \ldots, \succsim_n) = a$ if $a \succ_1 b$ and $f(\succsim_1, \succsim_2^*, \succsim_3^*, \succsim_4, \ldots, \succsim_n) = b$ if $b \succ_1 a$. Continuing in this way, we get a profile of preferences in which all agents prefer c, but c is not chosen. Since any group strategy-proof map is efficient on its image we must either have that $c \notin im(f)$ or f is not group strategy-proof. Either way we have a contradiction.

2. From the characterization of two-agent mechanisms, if $|I^{1,2}(\succsim_{-1,2})| \geq 3$ we have a single dictator in the marginal mechanism $f^{ij}_{\succsim_{-ij}}$. For simplicity let $i = 1, j = 2$ and assume 1 is the dictator. We will show that for any $\succsim'$, $f(\succsim') = \max_{\succsim'_1} I^{1,2}(\succsim_{-1,2})$. Begin with $f(\succsim'_1, \succsim_2, \ldots, \succsim_n)$. The statement holds by assumption. Now since 1 is the marginal dictator, changing $\succsim_2$ to $\succsim'_2$ cannot change the outcome. Hence the statement holds for $f(\succsim'_1, \succsim'_2, \ldots, \succsim_n)$. Now we have that $I^{1,3}(\succsim'_2, \succsim_4, \ldots, \succsim_n)$ contains $I^{1,2}(\succsim_{-1,2})$ as a subset. Hence there either 1 or 3 is a local dictator. Clearly it must be 1. Therefore 3's announcement cannot change the outcome, so we have $f(\succsim'_1, \succsim'_2, \succsim'_3, \succsim_4, \ldots, \succsim_n) = \max_{\succsim'_1} I^{1,2}(\succsim_{-1,2})$. Continuing in this way gives the desired result. The assumption that f is surjective implies that 1 is a dictator. □

A.13 Proof of Theorem 5

If $C^{i,j}$ admits more than one equivalence class we may assign a different local dictator to each class as in Theorem 1. We can then extend this mechanism via any GSD-ordering as in Proposition 4. □

References

Abdulkadiroğlu, A., P. Pathak, A. E. Roth, and T. Sonmez (2006): "Changing the Boston school choice mechanism," Discussion paper, National Bureau of Economic Research.

Abdulkadiroğlu, A., and T. Sönmez (1999): "House Allocation with Existing Tenants," *Journal of Economic Theory*, 88, 233–260.

Abraham, D. J., and D. F. Manlove (2004): "Pareto Optimality in the Roommates Problem," Discussion paper, Department of Computing Science, University of Glasgow.

Balbuzanov, I. (2019): "Constrained Random Matching," Working paper, University of Melbourne.

Barberà, S. (1983): "Strategy-Proofness and Pivotal Voters: A Direct Proof of the Gibbard- Satterthwaite Theorem," *International Economic Review*, 24, 413–418.

——— (2001): "An Introduction to Strategy-proof Social Choice Functions," *Social Choice and Welfare*, 18, 619–653.

Barberà, S., D. Berga, and B. Moreno (2010): "Individual versus Group Strategy-Proofness: When Do They Coincide?," *Journal of Economic Theory*, 145, 1648–1674.

——— (2016): "Group Strategy-Proofness in Private Good Economies," *American Economic Review*, 106, 1071–1099.

Bird, C. G. (1984): "Group Incentive Compatibility in a Market with Indivisible Goods," *Economics Letters*, 14(4), 309–313.

Bogomolnaia, A., and H. Moulin (1990): "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 100, 295–328.

Budish, E., Y.-K. Che, F. Kojima, and P. Milgrom (2013): "Designing Random Allocation Mechanisms: Theory and application," *American Economic Review*, 103, 585–623.

Ehlers, L., I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim (2013): "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds," *Journal of Economic Theory*, 153, 648–683.

Gale, D., and L. Shapley (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9–14.

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

Hafalir, I. E., M. B. Yenmez, and M. A. Yildirim (2013): "Effective Affirmative Action in School Choice," *Theoretical Economics*, 8, 325–363.

Irving, R. W. (1985): "An Efficient Algorithm for the 'Stable Roommates' Problem," *Journal of Algorithms*, 6, 577–595,.

Kamada, Y., and F. Kojima (2015): "Efficient Matching under Distributional Constraints: Theory and Applications," *American Economic Review*, 105, 67–99.

——— (2017a): "Recent Developments in Matching with Constraints," *American Economic Review Papers and Proceedings*, 107, 200–204.

——— (2017b): "Stability Concepts in Matching under Distributional Constraints," *Journal of Economic Theory*, 168, 107–142.

——— (2018): "Stability and Strategy-Proofness for Matching with Constraints: A Necessary and Sufficient Condition," *Theoretical Economics*, 13, 1761–794.

Le Breton, M., and V. Zaporozhets (2009): "On the Equivalence of Coalitional and Individual Strategy-Proofness Properties," *Social Choice and Welfare*, 33, 287–309.

Maskin, E. (1999): "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 66, 23–38.

Meng, D. (2019): "Dictatorship and Connectedness for Two-Agent Mechanisms with Weak Preferences," Working paper, Southwest Baptist University.

Muller, E., and M. Satterthwaite (1977): "The Equivalence of Strong Positive Association and Strategy-proofness," *Journal of Economic Theory*, 14, 412–418.

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

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

Rothkopf, M. H. (2007): "Thirteen reasons why the Vickrey-Clarke-Groves process is not practical," *Operations Research*, 55(2), 191–197.

Satterthwaite, M. (1975): "Strategy-proofness and Arrow's conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10, 187–217.

Satterthwaite, M. A., and H. Sonnenschein (1981): "Strategy-proof allocation mechanisms at differentiable points," *The Review of Economic Studies*, 48(4), 587–597.

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

Svensson, L.-G. (1994): "Queue Allocation of Indivisible Goods," *Social Choice and Welfare*, 11, 323–330.

——— (1999): "Strategy-Proof Allocation of Indivisible Goods," *Social Choice and Welfare*, 16, 557–567.

Takagi, S., and S. Serizawa (2010): "An impossibility theorem for matching problems," *Social Choice and Welfare*, 35(2), 245–266.

Takamiya, K. (2001): "Coalition Strategy-Proofness and Monotonicity in Shapley-Scarf Housing Markets," *International Journal of Game Theory*, 41, 115–130.

——— (2003): "On Strategy-Proofness and Essentially Single-Valued Cores: A Converse Result," *Social Choice and Welfare*, 20, 77–83.

——— (2013): "Coalitional Unanimity versus Strategy-proofness in Coalition Formation Problems," *Mathematical Social Sciences*, 42, 201–213.

Toda, M. (2006): "Monotonicity and Consistency in Matching Markets," *International Journal of Game Theory*, 34, 13–31.