Strategy-proof Allocation for Restricted Economic Domains

James Peter Schummer, Jr.

Submitted in Partial Fulfillment
of the
Requirements for the Degree
Doctor of Philosophy

Supervised by
Professor William Thomson

Department of Economics
The College
Arts and Sciences
University of Rochester
Rochester, New York

1997

Dedication

I dedicate this work to my parents.

Curriculum Vitæ

The author was born on the twenty-first day of October, 1969, in Quakertown, Pennsylvania. After finishing his secondary education there in 1988, he began attending The Pennsylvania State University. In 1992, he graduated from Penn State with a B.S. in Finance and began his studies in Economics at the University of Rochester. Focusing his research on Game Theory, he earned an M.A. in 1996.

Acknowledgements

Many people have contributed both directly and indirectly to my education at Rochester, and I am thankful to all of them. While the confines of this space do not allow me to recognize every single person who has done so, I hope to mention at least those who have contributed most.

First and foremost, I am indebted beyond words to my advisor, William Thomson. He has structured my thinking in a way that I will appreciate always. He has taught me how to present a logical argument in a clear manner. He introduced me to the areas of Game Theory and Mechanism Design which lead to my writing this work. His reputation as a great teacher is truly deserved.

I have also benefited from contact with, and encouragement from, the rest of the faculty at Rochester. I am especially grateful to David Austen-Smith, Jeff Banks, John Boyd, Hugo Hopenhayn, Lionel McKenzie, Rangarajan Sundaram, and the other participants of the Economic Theory Workshop.

Professors from outside Rochester have provided me with helpful discussions and with useful comments on earlier versions of my work. I thank Colin Campbell, Don Campbell, Luis Corchòn, John Duggan, Matt Jackson, François Maniquet, Yves Sprumont, Ennio Stacchetti, John Weymark, and Lin Zhou.

During my early years as a student, I was fortunate to receive guidance from other students of William Thomson, and I thank them all. I am particularly grateful to Stephen Ching, Tarık Kara, Hideo Konishi, and Tayfun Sönmez for many helpful interactions. I also thank Ban Chuan Cheah, Matt Mitchell, and Eiichi Miyagawa for insightful conversations.

My decision to attend Rochester — even my decision to study economics at a graduate school — was due to the suggestion and encouragement of a Rochester graduate who instructed me when I was an undergraduate at Penn State. I thank Ping Wang for pointing me toward this path.

A few close people deserve special thanks for putting up with me during both good and bad times. They are Viviane, for her caring support; Glenda for tolerance and helpfulness; and Tim, Pete, and Chris for shaping who I am. I also thank Dave B., Paul, and Dave O. for many good times.

Finally, I would like to thank my parents, James and Joyce, for their support and love, without which none of this would have been possible.

Abstract

Consider rules that allocate resources based on the preferences of agents. Facing such a rule, selfish agents may not find it in their best interest to reveal their *actual* preferences. Here, we address this concern by searching for *strategy-proof* rules: Regardless of the other agents' stated preferences, an agent should not have the incentive to announce false preferences. We conduct our search on three types of environments: 2-agent exchange economies, economies with public goods, and economies with indivisible goods.

For the class of 2-agent exchange economies in which agents have homothetic, strictly convex preferences, we show, as Zhou (1991b) did for a larger domain of preferences, that among the set of rules that are not dictatorial, *strategy-proofness* is incompatible with *efficiency*. We then show that this incompatibility holds even when preferences are restricted to the domain of linear preferences. Finally, we reach the same conclusion for any superdomain of monotonic preferences containing any one of these two small domains.

Next we consider the combination of *strategy-proofness* and *efficiency* for various classes of economies whose allocation spaces can be represented by simplices. It has long been known that if agents have von Neumann-Morgenstern preferences over lotteries, then the only *strategy-proof* and *efficient* rules choosing lotteries are, as above, dictatorial (Gibbard, 1977; Hylland, 1980). We strengthen this result by showing it to actually be a consequence of the same incompatibility on a series of much smaller domains of preferences over simplices, each of which has a standard economic interpretation.

Finally, we consider economies in which each agent consumes a different indivisible object and some amount of a divisible good. We establish another incompatibility between *strategy-proofness* and *efficiency*. Therefore we continue by dropping *efficiency* altogether.

For the simple case of two agents and two objects, we obtain a complete characterization of *strategy-proof* rules. For the case of more than two agents and objects, we argue that while there are many *strategy-proof* rules, many of them choose allocations "arbitrarily." To eliminate such rules from consideration, we strengthen *strategy-proofness* in three alternate ways. We show that by additionally requiring either *non-bossiness* or *coalitional strategy-proofness*, a rule must have a limited range, containing at most one allocation per assignment of the indivisible objects. The third strengthening of *strategy-proofness* is the addition of the requirement that no agent should be able to "bribe" another agent to change his preferences, making both agents better off after the bribe. We show that only constant rules are *bribe-proof*.

Contents

List of Figures

Chapter 1

Introduction

1.1 The General Problem

Consider the problem of allocating resources to a set of agents based on the preferences of those agents. A solution to such a problem is an (allocation) rule assigning a single allocation to any combination of preferences the agents may have. Since preferences are unobservable, they must be elicited from the agents before the rule can be used. When asked about their preferences, however, selfish agents may or may not find it in their best interest to reveal their *actual* preferences.

In this work we address this concern by searching for allocation rules that give agents the incentive to be truthful when announcing their preferences. Specifically, it should be the case that an agent, regardless of what preferences the other agents announce, should be no worse off announcing his true preferences than he would be stating any misrepresentation. In the language of the literature, we search for *strategy-proof* allocation rules. We conduct this search in essentially three types of economic environments: 2-agent exchange economies, economies with public goods, and economies with indivisible goods.

As shown in the seminal works of Gibbard (1973) and Satterthwaite (1975), *strategy-proofness* is a very demanding property in the sense that it is essentially equivalent to *dictatorship* on the unrestricted domain of preferences. The strength of the property in their model may not come as a surprise due to the fact that *strategy-proofness* is a stronger property on larger domains of preferences. Intuitively, this is because agents have more ways to misrepresent their preferences on larger domains. Therefore, one may expect that on smaller domains of preferences, such as economic domains, a rule may satisfy *strategy-proofness*, *non-dictatorship*, and perhaps even some other desirable properties, such as *efficiency*.

In general, however, this is not the case. Gibbard (1977), for example, considers only von Neumann-Morgenstern preferences over lotteries.[1] He shows, that if a *strategy-proof* and *efficient* solution for this domain is required to depend only on the agents' preferences over degenerate lotteries, then it must be *dictatorial*. Hylland's (1980) unpublished work strengthens Gibbard's result by dropping the latter requirement. In the same framework, Barberà (1977) considers *strategy-proof* solutions that choose *sets* of alternatives (after appropriately re-defining *strategy-proofness* for the choice of sets). If such a solution satisfies some other minor requirements, then it must also be *dictatorial*.

Other examples of restricted domains of preferences are those that apply when we consider choosing levels of public goods. Zhou (1991a) examines various classes of n-person pure public good economies: one with satiated preferences, and some with monotonic preferences. He shows that when the dimension of the range of a *strategy-proof* solution is at least two, the solution must be *dictatorial*. Barberà and Jackson (1994) then characterize those *strategy-proof* solutions with a one-dimensional range.

While the domains of preferences in these works are smaller than the Gibbard-Satterthwaite domain, they are still in some sense "large," at least when compared to domains in which, for example, agents care only about their own private consumption. That is, one might hope that when we consider private good economies, the resulting reduction of the set of admissible preferences would lessen the negative consequences of *strategy-proofness* seen in the public-goods-type of model.

An example of this is given by Serizawa (1996), who considers one-private-good one-public-good economies when preferences are monotonic and strictly convex. He characterizes the *strategy-proof*, *non-bossy*,[2] and *individually rational* solutions as those according to which (1) agents share the cost of producing the public good according to pre-specified cost sharing functions and, (2) the level of the public good provided is the minimum level demanded by all of the agents, given these cost sharing functions.

An even richer example is the work of Barberà and Jackson (1995), which actually provides some of the methods for Serizawa (1996). Considering exchange economies in which agents have strictly convex, monotonic preferences, they characterize the *strategy-proof*, *non-bossy*, and *anonymous* rules as "fixed-price trading" rules: agents trade only in pre-specified proportions from their endowments. While this class of rules is limited, it shows that the implications of *strategy-proofness* are not as strong on at least one interesting restricted domain

[1] Considering unrestricted preferences over lotteries would just put us back in the Gibbard-Satterthwaite model.

[2] See Section 4.3.1 for a definition of *non-bossiness*.

of preferences — one pertaining to exchange economies.

For 2-agent exchange economies with two goods, Sprumont (1995), using a domain of preferences similar to Barberà and Jackson's, characterizes *strategy-proof* rules satisfying a *continuity* condition with respect to preferences. They are rules in which a fixed agent receives his most preferred bundle from a predetermined set.

Since *strategy-proofness* apparently loses at least some of its strength on these restricted domains of preferences, one might hope that the combination of that property with *efficiency* is achievable. Unfortunately, some other works show that this is not to be the case.

One of the earliest such result is by Hurwicz (1972), who considers a certain class of 2-agent exchange economies. On this domain, he shows the incompatibility of *strategy-proofness*, *individual rationality*, and *efficiency*. Zhou (1991b) provides an elegant strengthening of that result by showing that when both agents have strictly convex, monotonic preferences, any *strategy-proof* and *efficient* rule must be *dictatorial*. For the case of more than two agents, Barberà and Jackson (1995) note that none of their fixed-price trading rules is *efficient*. In addition, the rules described by Serizawa (1996) are not *efficient*.

A natural question follows from such results: Does *strategy-proofness* retain this strength (perhaps in combination with other properties) on even smaller subdomains of preferences?[3] This type of question, applied in particular to reasonable subdomains in economic environments, is the motivation for our work.

1.2 An Overview of the Results

The purpose of Chapter 2 is to determine whether *strategy-proof*, *efficient*, and *non-dictatorial* rules exist for restricted domains of preferences for 2-agent exchange economies, even though they do not for the domain of Zhou (1991b).[4] First we restrict attention to the case in which both agents have only homothetic, strictly convex preferences. Homothetic preferences are commonly used in consumer and producer theory, international trade theory, and the theory of the aggregation of preferences (see Chipman, 1974). Examples of preferences in our

[3] Alternative questions are: "'How often' is a *strategy-proof* solution not *efficient*?" and "'How often' does an *efficient* solution violate the conditions of *strategy-proofness*?". These questions are provided with negative answers by Hurwicz and Walker (1990) for general economies, and by Beviá and Corchón (1995) for public goods economies.

[4] Nagahisa (1995) asks "how large" can a subdomain of Zhou's be if it admits such a solution. Using the closed convergence topology, he shows that such a subdomain must be nowhere dense in Zhou's domain. That result does not imply any of the results obtained here.

domain are Cobb-Douglas and CES preferences. Given the importance of the homothetic domain, it should be noted that the proof in Zhou (1991b) requires the admissibility of non-homothetic preferences.

The first result is that even on this small domain, any *strategy-proof* and *efficient* solution is necessarily *dictatorial*.

Second, we restrict attention to what intuitively appears as an even smaller domain of preferences: linear, strictly monotonic preferences. While such a small domain has been considered less often than one of homothetic preferences, it is of interest in that such preferences exhibit a high degree of substitutability across goods, whereas the proofs of Zhou's result and of the first result here require arbitrary degrees of complementarity between the goods at an arbitrary consumption bundle (*i.e.*, require that a preference relation may be arbitrarily close to a Leontieff-type preference relation).

The second result parallels the first: Even on this domain, any *strategy-proof* and *efficient* solution is necessarily *dictatorial*. Finally, we show this type of result to hold on any domain of monotonic preferences that contains at least one of our two small subdomains (*e.g.*, Zhou, 1991b).

In Chapter 3, we ask the question from Chapter 2 for economies with public goods: Do *strategy-proof*, *efficient*, *non-dictatorial* rules exist for restricted domains of preferences in such economies, even though they do not for larger ones (*e.g.*, Gibbard, 1977, and Hylland, 1980)?

Our main result for this chapter (on our "smallest" domain) shows that on the class of 2-agent economies in which agents have linear preferences over one private good and one public good (produced according to a constant-returns production function), a *strategy-proof* and *efficient* solution must be *dictatorial*.

We extend this main result to a class of economies with many public goods. We finally extend the result to the class of economies with an arbitrary (finite) number of agents with linear preferences over just public goods. This last domain coincides with the domain of von Neumann-Morgenstern preferences over lotteries, hence providing a strengthening of a result of Hylland (1980). We also discuss the connection of this work to the literature on the Clarke-Groves mechanisms.

Finally in Chapter 4, we consider the problem of allocating a set of indivisible objects and a fixed amount of a divisible good (*e.g.*, money) to a set of agents with the restriction that each agent consumes precisely one object. These indivisibles may represent positions such as jobs, offices, or housing locations — objects of which we typically consume no more than one.

For the simple case of two agents and two objects, we characterize the entire class of *strategy-proof* solutions. Such a solution either: (1) is constant, (2) allows one agent to choose among two pre-specified allocations, or (3) chooses a pre-specified ("status-quo") allocation unless both agents prefer a pre-specified second allocation. This characterization is domain-free in the sense that it holds on any (sub)domain of quasi-linear preferences for our model.

For the case of more than two agents, many *strategy-proof* solutions exist. We provide a parameterized family of them suggesting that many *strategy-proof* solutions use preference information in an arbitrary way. To weed out such solutions, we consider in turn three alternate auxiliary conditions.

First we add to *strategy-proofness* the requirement that if a change in one agent's preferences does not change what he consumes, than it also does not change what anyone else consumes (*non-bossiness*). The result we obtain is that despite the existence of a continuum of allocations, a solution satisfying this additional condition must have a small number of allocations in its range — at most one allocation per assignment of the indivisible objects to the agents. We show as corollaries that some of these solutions are dictatorial on their ranges, and none of them always treats agents with identical preferences equally. However, an interesting class of *strategy-proof* solutions satisfying our additional requirement can be derived from allocation rules provided by Roth (1982) for a related model.

Second, we consider the requirement that no *coalition* of agents can gain by jointly misrepresenting their preferences. As above, we show that the range of a solution satisfying this property must contain at most one allocation per assignment of the indivisible objects to the agents.

Third, we introduce what appears to be a new condition. We require that no agent should be able to bribe another agent to change his preferences, making both agents better off after the bribe. Surprisingly, this condition has the strongest of consequences — only constant solutions satisfy it.

All of these results extend to the model in which there may be more than one copy of a single object (*e.g.*, identical offices) between which agents do not differentiate in terms of preferences. One application of such a model is to a situation in which a number of agents may share offices, but do not care about with whom they share it.

Finally, we return to the combination of *strategy-proofness* and *efficiency*. Given that agents have quasi-linear preferences, one may make the prediction that Groves-type mechanisms[5] are the only *strategy-proof* rules that always assign the indivisible goods efficiently, corresponding to what Green and Laffont (1977) showed for a model with public goods. Even though the Green-Laffont result has no direct implications for our model, Holmström (1979) shows this pre-

[5] See Clarke (1971), Groves (1973), Green and Laffont (1977).

diction to be correct: He generalized the Green-Laffont characterization to any "connected" domain of quasi-linear preferences over an arbitrary set of public goods, which *does* have implications for our model with private goods. The result is that any *strategy-proof* and *efficient* solution must be equivalent to a Groves mechanism. However, as is the case with public goods models, we establish that for our model, a Groves mechanism does not always allocate the entire endowment of the divisible good, which we require. Therefore no *strategy-proof* solution is fully efficient.

Inherent in this result, as in the Groves mechanism literature, is the assumption that agents may consume arbitrary (negative) levels of the divisible good. That is why, for example, a dictatorial mechanism does not define a *strategy-proof* and *efficient* solution — a dictator wants to consume an infinite quantity of the divisible good! Disallowing negative consumption leads to a natural alternate formulation of efficiency, requiring only that an allocation not be Pareto-dominated by any other allocation *prescribing non-negative consumption*, instead of the stronger requirement of always choosing an efficient assignment of the indivisible goods. For the 2-agent case, we show that dictatorial solutions are in fact the *only strategy-proof* solutions satisfying this Pareto-optimality condition.

A similar model in which there is no money, often called a housing market, has been considered by Shapley and Scarf (1974), who show the non-emptiness of the core when each agent is endowed with an object. Roth (1982) then shows the *strategy-proofness* of solutions that make a selection from the core. Disallowing indifference between different objects, Ma (1994) then shows that the core is characterized by the properties of *strategy-proofness*, *efficiency*, and *individual rationality*.

Our model with money has appeared in the fairness literature. Requiring consumption of the divisible good to be non-negative, Svensson (1983) gives a set of conditions on preferences guaranteeing, for a given allocation, the existence of prices and an income distribution such that the allocation is part of a competitive equilibrium. He also gives conditions for the existence of efficient, envy-free allocations, and of efficient, egalitarian-equivalent allocations. Maskin (1987) identifies another set of conditions on preferences and on the amount of the divisible good to be divided that guarantee the existence of efficient, envy-free allocations. Alkan *et al.* (1991) and Tadenuma and Thomson (1995a) consider the problem of selecting an efficient, envy-free allocation, as there typically exists a continuum of them. In particular they show (for the 2-agent and n-agent cases, respectively) that when there is exactly one indivisible object, no *strategy-proof* solution always chooses an efficient, envy-free allocation. Aragones (1995) connects the work of

Alkan *et al.* to the existence problem on the domain of quasi-linear preferences. The reader is referred to these works and to Tadenuma and Thomson (1991,1995b) for more detailed references.

1.3 General Notation

In this section we introduce some general notation that will apply to all three of the following chapters. There is a set of agents, $N = \{1, 2, \ldots, n\}$. Each agent is to consume a bundle from his own consumption space, say A_i. For example, in an exchange economy setting with ℓ goods, an agent's consumption space is $\mathbb{R}^{\ell}_{+}$. Each agent $i \in N$ has a preference relation on A_i, say R_i, that is an element of a set of admissible preference relations, say $\mathcal{R}_i$.[6] This set will depend on what economic situation we are modeling. The notation $\arg\max_S R_i$ refers to the set of maximal elements of a set $S \subset A_i$ according to R_i.[7] Finally, there is a set of feasible allocations, say A. In economies with just private goods, we have $A \subset A_1 \times A_2 \times \cdots \times A_n$, and in economies with just pure public goods, we have $A = A_i$ for all $i \in N$.

A **class of economies** is therefore a specification of these sets of agents, allocations, and admissible preferences, say $(N, A, \mathcal{R}_1, \ldots, \mathcal{R}_n)$. The **domain** of preferences for a given class is the cross product of the agents' sets of admissible preferences, $\mathcal{R} = \mathcal{R}_1 \times \mathcal{R}_2 \times \cdots \times \mathcal{R}_n$. An element of the domain, $R = (R_1, R_2, \ldots, R_n) \in \mathcal{R}$, is a **profile** of preferences. An allocation rule, or **solution (for $\boldsymbol{\mathcal{R}}$)**, associates an allocation with any profile preferences, hence it is a mapping $\varphi\colon \mathcal{R} \to A$. A typical solution will be denoted $\boldsymbol{\varphi}$, throughout. For the remainder of this section, a class of economies is assumed to be given.

For any profile $R \in \mathcal{R}$ and agent $i \in N$, the notation R_{-i} refers to the list of the preferences of each agent except i. Similar notation will be used replacing i with a *set* of agents, $C \subset N$.

In a private goods setting, such as Chapters 2 and 4, the following notation will be useful. For any profile $R \in \mathcal{R}$ and agent $i \in N$, $\varphi_i(R) \in A_i$ is the bundle in the allocation $\varphi(R)$ that agent i consumes.

Our primary goal is to determine what type of solution, φ, satisfies the following condition.

***Strategy-proofness*:** for all $(R_1, \ldots, R_n) \in \mathcal{R}$, all $i \in N$, and all $R'_i \in \mathcal{R}_i$, we

[6] This is a set of weak orders. The implied strict preference relation is denoted P_i and indifference relation is denoted I_i.

[7] There will be a non-empty set of maximal elements whenever we use this piece of notation.

have

$$\varphi_i(R)\ R_i\ \varphi_i(R_i', R_{-i}) \quad \text{(private good economy notation)}$$
$$\varphi(R)\ R_i\ \varphi(R_i', R_{-i}) \quad \text{(public good economy notation)}$$

A useful notion throughout all three of the following chapters is one that is closely related to what is known as *strong positive association* in the social choice literature, and *Maskin-monotonicity* in the implementation literature. For any agent $i \in N$, let $R_i, R_i' \in \mathcal{R}_i$ and $x \in A_i$. Suppose that for all $y \in A_i$ such that $y \neq x$, we have $x\ R_i\ y$ implies $x\ P_i'\ y$. In that case, we will say that R_i' is a **strict monotonic transformation of $\boldsymbol{R_i}$ at $\boldsymbol{x}$**. That is, define the set of strict monotonic transformations of $R_i \in \mathcal{R}_i$ at $x \in A_i$ to be

$$\boldsymbol{SMT}(\boldsymbol{R_i}, (\boldsymbol{a}, \boldsymbol{m_i})) = \{R_i' \in \mathcal{R}_i : \forall y \in A_i \setminus \{x\},\ x\ R_i\ y \text{ implies } x\ P_i'\ y\}$$

The following lemma appears throughout the literature on *strategy-proofness*, and is used throughout this work.

Lemma 1.1 *Let φ be a* strategy-proof *solution on a domain $\mathcal{R}$. Let $i \in N$, $R \in \mathcal{R}$, and $R_i' \in SMT(R_i, \varphi_i(R))$. Then $\varphi_i(R_i', R_{-i}) = \varphi_i(R)$.*

Proof: Let φ, i, R, and R_i' be defined as above. *Strategy-proofness* implies $\varphi_i(R)\ R_i\ \varphi_i(R_i', R_{-i})$ and $\varphi_i(R_i', R_{-i})\ R_i'\ \varphi_i(R)$. By definition of R_i', we have $\varphi_i(R) = \varphi_i(R_i', R_{-i})$. ■

Note that in a public goods setting, the conclusion of the lemma is that $\varphi(R_i', R_{-i}) = \varphi(R)$, since agents share a common consumption space.

Chapter 2

Exchange Economies

In this chapter[1] we consider 2-agent exchange economies. First we show that if agents are assumed to have only homothetic, strictly convex preferences over their consumption space, then any *strategy-proof* and *efficient* solution is *dictatorial*. Second, we show that the same conclusion holds when agents have only linear preferences. Finally, these results extend to any larger domain of monotonic preferences for the agents.

2.1 The Model

We examine 2-agent exchange economies with a positive endowment $\Omega \in \mathbb{R}^{\ell}_{++}$ of $\ell \geq 2$ infinitely divisible goods.[2] So, the set of agents is $N = \{1, 2\}$ and each agent's consumption space is $\mathbb{R}^{\ell}_{+}$. An **allocation** is a vector[3] $(z_1, z_2) \in \mathbb{R}^{2\ell}_{+}$ such that $z_1 + z_2 = \Omega$. Here, z_i represents agent i's consumption bundle. Let $\boldsymbol{Z}$ be the set of allocations.

Each agent has a continuous preference relation, R_i, on $\mathbb{R}^{\ell}_{+}$. We will make assumptions about preferences from the following list.

Monotonic: $x > y$ implies $x \; P_i \; y$.

Strictly monotonic: $x \geq y$ implies $x \; P_i \; y$.

Strictly convex (in the interior): $x R_i y$, $x \neq y > 0$, and $\lambda \in (0, 1)$ imply $(\lambda x + (1-\lambda) y) \; P_i \; y_i$.

[1] This chapter is a version of Schummer (1997a).

[2] See Section 1.3 for preliminary notation.

[3] Vector inequalities are denoted as follows: $x \geqq y$ means $x_k \geq y_k$ for each k; $x \geq y$ means $x \geqq y$ and $x \neq y$; $x > y$ means $x_k > y_k$ for each k. In addition, 0 represents a list $(0, 0, \ldots, 0)$ whose length can be inferred from the context.

Homothetic: $x \; R_i \; y$ implies $\lambda x \; R_i \; \lambda y$ for all $\lambda \in \mathbb{R}_+$.

Linear: There exists $\lambda \in \mathbb{R}^\ell_+$ such that $x \; R_i \; y$ if and only if $\sum \lambda_k x^k \geq \sum \lambda_k y^k$, where superscripts denote goods.

Let $\mathcal{R}_H$ be the set of continuous preference relations over $\mathbb{R}^\ell_+$ that are monotonic, strictly convex, and homothetic. Let $\mathcal{R}_L$ be the class of preference relations over $\mathbb{R}^\ell_+$ that are strictly monotonic and linear.

Note that $SMT(R_i, x)$ is non-empty if either (1) $R_i \in \mathcal{R}_H$ and $x_i > 0$ or (2) $R_i \in \mathcal{R}_L$, $\ell = 2$, and x is on the boundary of Z.

In Sections 2.2 and 2.3, we will consider the domains $\mathcal{R}^2_H$ and $\mathcal{R}^2_L$, respectively. To that end, the notation in the rest of this section is given with respect to an arbitrary domain of preferences, $\mathcal{R} = \mathcal{R}_1 \times \mathcal{R}_2$.

An allocation is in the **Pareto set** of an economy, $\boldsymbol{P(R)}$, if there exists no other allocation that both agents weakly prefer and one agent strictly prefers:

$$\forall R \in \mathcal{R},\; P(R) = \{z \in Z : \nexists z' \in Z \text{ and } i \in N \text{ such that } z'_i \; P_i \; z_i \text{ and } z'_{-i} \; R_{-i} \; z_{-i}\}$$

A solution $\varphi \colon \mathcal{R} \to Z$ is ***efficient*** **(for $\mathcal{R}$)** if for all $R \in \mathcal{R}$, $\varphi(R) \in P(R)$. It is ***dictatorial*** **(for $\mathcal{R}$)** if for some $i \in N$, for all $R \in \mathcal{R}$, and all $x \in Z$, we have $\varphi_i(R) \; R_i \; x_i$. Note that *dictatorial* solutions defined on $\mathcal{R}^2_H$ or $\mathcal{R}^2_L$ always give the entire endowment Ω to the dictator.

2.2 Homothetic Preferences

We turn our attention to the first domain. The next result states that when agents have identical preferences, the Pareto set is the diagonal of the (ℓ-dimensional) Edgeworth box.

Lemma 2.1 *For all $R \in \mathcal{R}^2_H$, if $R_1 = R_2$ then $P(R) = D \equiv \{(x_1, x_2) \in Z : x_1 = \lambda\Omega, \lambda \in [0,1]\}$.*

Proof: Let $R_1 = R_2 \in \mathcal{R}_H$. Clearly $(0, \Omega) \in P(R)$ and $(\Omega, 0) \in P(R)$. Let $x = (\lambda\Omega, (1-\lambda)\Omega)$ for some $\lambda \in (0,1)$. By homotheticity, there is a hyperplane supporting agent 1's indifference surface at x_1 that is parallel to one supporting his indifference surface at x_2. Since $R_2 = R_1$, $x \in P(R)$.

Let $y \notin D$. Let H_1 be a hyperplane that (1) passes through y_1 and (2) is parallel to a hyperplane supporting the indifference surface of $R_1 = R_2$ passing through Ω. Let H_2 be parallel to H_1 and pass through y_2. Then there exist $\delta \in (0,1)$ and $z \in Z$ such that $z = (\delta\Omega, (1-\delta)\Omega)$, z_1 is in H_1, and z_2 is in H_2.

By homotheticity, for all $i \in N$, H_i supports the indifference surface of R_i passing through z_i. By strict convexity, for all $i \in N$, $z_i \; P_i \; y_i$. Hence $y \notin P(R)$. ■

A lemma in Zhou (1991b) states that if an allocation is in the range of a *strategy-proof* solution, then no other allocation in which one agent receives more of all goods is also in the range. Due to the homotheticity of preferences, we can at this point only show the following limited version of that result.

Lemma 2.2 *Let* $\varphi\colon \mathcal{R}_H^2 \to Z$ *be* strategy-proof. *For all* $R, R' \in \mathcal{R}_H^2$ *and all* $i \in N$*, if* $\varphi_i(R) \neq \varphi_i(R') > 0$*, then for all* $\lambda \in [0, \infty)$*,* $\varphi_i(R) \neq \lambda \varphi_i(R')$*.*

Proof: Let $x = \varphi(R)$, $y = \varphi(R')$, and suppose, by contradiction and without loss of generality, that $\lambda \in [0,1) \cup (1, \infty)$, $y_1 > 0$, and $x_1 = \lambda y_1$.
Case 1: $\lambda > 0$.

Let $R_1'' \in \mathcal{R}_H$ be such that $R_1'' \in SMT(R_1, x) \cap SMT(R_1', y)$. (Such preferences can be found by approximating Leontieff preferences.) By Lemma 1.1, $x = \varphi(R_1'', R_2)$ and $y = \varphi(R_1'', R_2')$. However, monotonicity and strict convexity imply (i) $x_2 \; P_2' \; y_2$ if $\lambda < 1$, and (ii) $y_2 \; P_2 \; x_2$ if $\lambda > 1$. Either case contradicts *strategy-proofness*.
Case 2: $\lambda = 0$.

We have $x_1 = 0$ and $y_2 < x_2 = \Omega$. Let $z = \varphi(R_1, R_2')$. By *strategy-proofness*, $z_2 \; R_2' \; x_2$. Hence by monotonicity and strict convexity, $z_2 = \Omega$ and $z_1 = 0$. By *strategy-proofness*, $z_1 \; R_1 \; y_1$, contradicting monotonicity. ■

Lemmas 2.1 and 2.2 imply the following:

Corollary 2.1 *Let* $\varphi\colon \mathcal{R}_H^2 \to Z$ *be* efficient *and* strategy-proof*. There exists* $d \in Z$ *such that for all* $R \in \mathcal{R}_H^2$*, if* $R_1 = R_2$ *then* $\varphi(R) = d$*.*

Proof: Let $R_1 = R_2 \in \mathcal{R}_H$ and $R_1' = R_2' \in \mathcal{R}_H$. By Lemma 2.1, $\varphi(R) \in D$ and $\varphi(R') \in D$. Therefore Lemma 2.2 implies $\varphi(R) = \varphi(R')$. ■

The corollary implies that by "matching" the other agent's preferences, an agent can guarantee himself the bundle he receives at d. Thus we have the following.

Corollary 2.2 *Let* $\varphi\colon \mathcal{R}_H^2 \to Z$ *be* efficient *and* strategy-proof*, and let* d *be given as in Corollary 2.1. Then there is no* $R \in \mathcal{R}_H^2$ *and* $i \in N$ *such that* $d_i \; P_i \; \varphi_i(R)$*.*

We have already done most of the work in proving the first main result.

Theorem 2.1 *If* $\varphi\colon \mathcal{R}_H^2 \to Z$ *is* efficient *and* strategy-proof*, then it is* dictatorial.

Proof: Suppose $\varphi\colon \mathcal{R}_H^2 \to Z$ is *efficient* and *strategy-proof.* Let d be given as in Corollary 2.1. According to Corollary 2.2, if $d_i = \Omega$ for some $i \in N$, then for all $R \in \mathcal{R}_H^2$, $\varphi_i(R) = \Omega$, implying φ is *dictatorial.* Therefore suppose by contradiction that $d > 0$.

Let $R' \in \mathcal{R}_H^2$ be such that (1) $P(R') \cap D = \{(\Omega, 0), (0, \Omega)\}$, and (2) for all $z \in P(R') \backslash \{(\Omega, 0), (0, \Omega)\}$, $z > 0$. For example, R' may be any profile of two different Cobb-Douglas preference relations. Let $\varphi(R') = x$. By *efficiency*, we have $x \neq d$. We know by Corollary 2.2 and the monotonicity and strict convexity of preferences that $x_1 \not\lneq d_1$ and $x_2 \not\lneq d_2$. Hence by (2), we also have $x > 0$. Let $\overline{P} = \{z \in Z : z_1 = \lambda x_1 \text{ or } z_2 = \lambda x_2, \text{ for } \lambda \in [0,1]\}$.

We can construct a preference profile $R \in \mathcal{R}_H^2$ such that $P(R) = \overline{P}$ and $d_1\ P_1\ x_1$. By *efficiency*, and Lemma 2.2, $\varphi(R) = x$. However $d_1\ P_1\ x_1$, contradicting Corollary 2.2. Hence $d \not> 0$, and φ is *dictatorial.*

We conclude the proof by describing a construction of R for the case $\ell = 2$. Construction for $\ell > 2$ follows the same lines of reasoning. We assume without loss of generality, that $x_1^1 > d_1^1$, where the superscript denotes the good.

Because R_1 and R_2 are homothetic and $d > 0$, it suffices to define the indifference curve of R_i through d_i for each $i \in N$. In particular, we only need to specify the slopes of lines supporting this indifference curve at points where it intersects rays from agent i's origin through bundles in $\overline{P}$. (Slopes are in the form of "change in good 2" over "change in good 1".) The rest of the preference maps can then be arbitrarily completed, within the bounds of admissibility.

Let $r = |(d_1^2 - x_1^2)/(d_1^1 - x_1^1)|$ be the (absolute value of the) slope of the line segment from d_1 to x_1. Let $f_\epsilon\colon \overline{P} \to (0, r)$ be a continuous function such that $z_1 \geq z_1'$ implies $f_\epsilon(z) > f_\epsilon(z')$, and $f((\Omega, 0)) - f((0, \Omega)) = \epsilon > 0$. This function evaluated at an allocation $z \in \overline{P}$ provides the absolute value of the slope of lines supporting the agents' indifference curves of R through their respective bundles at z.

Note that (i) along a ray from agent 1's origin through x_1, the indifference curves of R_1 have a kink, and are supported by lines with slopes ranging from $f_\epsilon((0, \Omega))$ through $f_\epsilon(x)$; (ii) along a ray from agent 2's origin through x_2, the indifference curves of R_2 have a kink, and are supported by lines with slopes ranging from $f_\epsilon(x)$ through $f_\epsilon((\Omega, 0))$; and (iii) if ϵ is sufficiently small, R is a strictly convex approximation of linear preferences. Therefore by choosing ϵ sufficiently small, $d_1\ P_1\ x_1$. ■

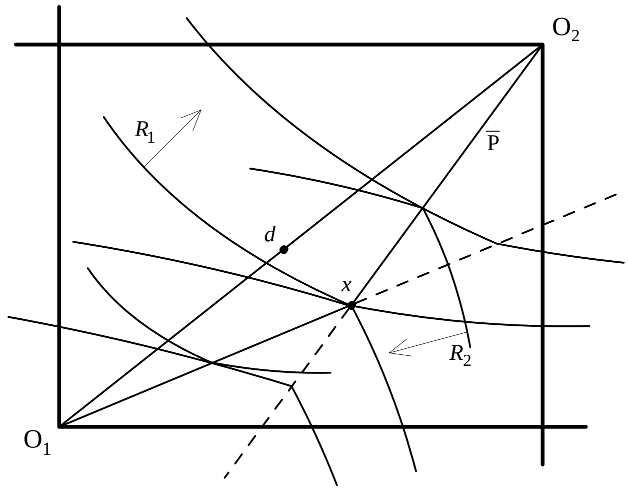


Figure 2.1: Constructing R_1 and R_2 in the proof of Theorem 2.1.

Remark 2.1 Let $\mathcal{R}_{HS} \subset \mathcal{R}_H$ be the set of preference relations in $\mathcal{R}_H$ that are strictly convex.[4] It is clear that a version of Theorem 2.1 corresponding to $\mathcal{R}_{HS}$ can be obtained. First, Lemma 2.2 can be modified by requiring φ to be *efficient* but allowing $\varphi_i(R) \not> 0$. Second, the preference relation R, constructed in the proof of Theorem 2.1, can be made to be strictly convex everywhere. Note also that $\mathcal{R}_{HS}$ is a subdomain of the one used in Zhou (1991b). These facts are of interest in light of the Corollary to Theorem 2.1, stated in the next section (see Remark 2.4).

2.3 Linear Preferences

We now examine the second domain. We first consider economies with only two goods. The proof of the result for this case gives the intuition behind the proof of the result for the general case.

Note that for linear preferences over two goods, the Pareto set is one of three sets. The first set corresponds to the case in which agent 1 values the second good relatively more than agent 2 does: $P^1 = \{x \in Z : x_1^1 = 0 \text{ or } x_2^2 = 0\}$. The second set corresponds to the case in which agent 1 values the first good relatively more than agent 2 does: $P^2 = \{x \in Z : x_1^2 = 0 \text{ or } x_2^1 = 0\}$. The third set is, of course, Z, corresponding to cases in which the agents have identical preferences.

[4]That is, strictly convex everywhere, including the boundary. It follows that each preference relation in $\mathcal{R}_{HS}$ is strictly monotonic.

Theorem 2.2 *If $\ell = 2$ and $\varphi\colon \mathcal{R}_L^2 \to Z$ is* efficient *and* strategy-proof*, then it is* dictatorial.

Proof: Let $\varphi\colon \mathcal{R}_L^2 \to Z$ be *efficient* and *strategy-proof.*
Step 1: There exist $x, x' \in Z$ such that $P(R) = P^1$ implies $\varphi(R) = x$, and $P(R) = P^2$ implies $\varphi(R) = x'$.

Let $R, R' \in \mathcal{R}_L^2$ be such that $P(R) = P(R') = P^1$. Let $x = \varphi(R)$ and $y = \varphi(R')$. We must show $y = x$. Note that since preferences are linear and x is on the boundary of Z, we have for all $i \in N$ either (1) for all $z \in P^1$, $R'_i \in SMT(R_i, z_i)$, (2) for all $z \in P^1$, $R_i \in SMT(R'_i, z_i)$, or (3) $R_i = R'_i$. For all $i \in N$, let $R''_i = R'_i$ if (1) holds, and let $R''_i = R_i$ otherwise. By Lemma 1.1, $x = \varphi(R_1, R''_2) = \varphi(R'')$. Similarly $y = \varphi(R'_1, R''_2) = \varphi(R'') = x$. Therefore for all $R \in \mathcal{R}_L^2$, $P(R) = P^1$ implies $\varphi(R) = x$.

Likewise we can show the existence of x'.
Step 2: For all $i \in N$ and all $R_i \in \mathcal{R}_L$, $x_i \; I_i \; x'_i$.

Let $R_1 \in \mathcal{R}_L$. Let $R_2 = R_1$ and $z = \varphi(R)$. Since preferences are strictly monotonic, for all $k \in \{1, 2\}$ there exists $R_1^k \in \mathcal{R}_L$ such that $P(R_1^k, R_2) = P^k$. Therefore *strategy-proofness* implies $z_1 \; R_1 \; x_1$ and $z_1 \; R_1 \; x'_1$. Similarly, $z_2 \; R_2 \; x_2$ and $z_2 \; R_2 \; x'_2$. However $\{x, x'\} \subset P(R) = Z$, so $x_i \; I_i \; z_i \; I_i \; x'_i$ for all $i \in N$.

Since $R_1 \in \mathcal{R}_L$ was arbitrary, either for all $R_1 = R_2 \in \mathcal{R}_L$, we have $\varphi(R) = x = x' = (\Omega, 0)$, or for all $R_1 = R_2 \in \mathcal{R}_L$, we have $\varphi(R) = x = x' = (0, \Omega)$. Hence φ is *dictatorial.* ■

Remark 2.2 An even more striking fact is that this result holds on a domain containing only *four* (or more) linear preferences. Step 2 of the proof must be modified so that its claim holds only for each preference relation that has neither the steepest nor the flattest indifference curves in the domain. For a domain consisting of no more than three linear preference relations, *efficient* and *strategy-proof* solutions that are not *dictatorial* do exist.[5]

Two difficulties arise in applying this proof to the case of more than two goods. First, the number of possible Pareto sets increases quickly as the number of goods increases. Second, many of these possible Pareto sets contain Pareto-indifferent allocations. At such an allocation, there may be no strict monotonic transformation of a preference relation. For these reasons, in the proof for $\ell > 2$ we first consider a subset of the admissible preferences that have a common rate of substitution between any of the last $\ell - 1$ goods. For all pairs of preferences from

[5] Suppose $\Omega = (1, 1)$, and the three linear preference relations have indifference curves of slopes $-1/2$, -1, and -2. One *strategy-proof* and *efficient* solution yields the allocation $((1, 0), (0, 1))$ whenever it is in the Pareto set, and yields $((0, 1), (1, 0))$ otherwise.

that subset, the Pareto set is one of three sets. In the 2-good case, this subset was the set of *all* admissible preferences. After showing that for this subset of preferences, an *efficient* and *strategy-proof* solution always gives one of the agents the entire endowment, we then use the strict monotonicity of preferences to show that when we allow for *any* linear preference profile, the solution still gives everything to that agent.

To do this, we temporarily transform the last $\ell - 1$ goods into a composite commodity. Let $\tilde{\mathcal{R}} \subset \mathcal{R}$ be a maximal set of preferences with respect to the following condition:[6] for all $R_i, R_i' \in \tilde{\mathcal{R}}$, all $x \in \mathbb{R}_+$, and all $y, y' \in \mathbb{R}_+^{\ell-1}$, we have $(x, y)\ I_i\ (x, y')$ if and only if $(x, y)\ I_i'\ (x, y')$.

Note that for $R \in \tilde{\mathcal{R}}^2$, $P(R)$ may be one of three sets. As we did for the case of two goods, we will denote two of these sets as follows. Let $P^1 = \{x \in Z : x_1^1 = 0 \text{ or } (x_2^2, \ldots, x_2^\ell) = 0\}$ and $P^2 = \{x \in Z : (x_1^2, \ldots, x_1^\ell) = 0 \text{ or } x_2^1 = 0\}$. Again, the third possible Pareto set is Z.

Theorem 2.3 *If* $\varphi \colon \mathcal{R}_L^2 \to Z$ *is* efficient *and* strategy-proof*, then it is* dictatorial*.*

Proof: Let $\varphi \colon \mathcal{R}_L^2 \to Z$ be *efficient* and *strategy-proof*.
Step 1a: There exists $X \subset P^1$ such that for all $R \in \tilde{\mathcal{R}}^2$, (1) $P(R) = P^1$ implies $\varphi(R) \in X$, and (2) for all $x, y \in X$ and all $i \in N$, $x_i\ I_i\ y_i$.

Let $R, R' \in \tilde{\mathcal{R}}^2$ be such that $P(R) = P(R') = P^1$. Let $x = \varphi(R)$ and $y = \varphi(R')$. We must show that for all $i \in N$, $y_i\ I_i\ x_i$.

For all $i \in N$, let $R_i'' \in \{R_i, R_i'\}$ be such that for all $z \in Z$ and all $z' \in P^1$, $z_i\ R_i''\ z_i'$ if and only if $z_i\ R_i\ z_i'$ and $z_i\ R_i'\ z_i'$. By *strategy-proofness*, $\varphi_1(R_1'', R_2)\ R_1''\ x_1$ and $x_1\ R_1\ \varphi_1(R_1'', R_2)$. Then by definition of R_1'', $\varphi_1(R_1'', R_2)\ R_1\ x_1$, so $\varphi_1(R_1'', R_2)\ I_1\ x_1$.

Note that $P(R_1'', R_2) = P^1$. Hence by *efficiency*, $\varphi_2(R_1'', R_2)\ I_2\ x_2$. By *strategy-proofness*, $\varphi_2(R_1'', R_2)\ R_2\ \varphi_2(R'')$ and $\varphi_2(R'')\ R_2''\ \varphi_2(R_1'', R_2)$. Then by definition of R_2'', $\varphi_2(R'')\ R_2\ \varphi_2(R_1'', R_2)$, so $\varphi_2(R'')\ I_2\ \varphi_2(R_1'', R_2)$, hence $\varphi_2(R'')\ I_2\ x_2$.

Noting that $\varphi(R'') \in P(R'')$, $x \in P^1$, and $P(R'') = P^1$, we then have $\varphi_1(R'')\ I_1\ x_1$.

Similarly, we can show that for all $i \in N$, $\varphi_i(R'')\ I_i'\ y_i$. Hence by the construction of $\tilde{\mathcal{R}}$, for all $i \in N$, $x_i\ I_i\ y_i\ I_i'\ x_i$. This demonstrates the existence of X as described above. In the 2-good case, X is a singleton. In the 3-good case, X is a subset of a line segment.
Step 1b: There exists $X' \subset P^2$ such that for all $R \in \tilde{\mathcal{R}}^2$, (1) $P(R) = P^2$ implies $\varphi(R) \in X'$, and (2) for all $x, y \in X'$ and all $i \in N$, $x_i\ I_i\ y_i$.

[6] A set is maximal with respect to a condition if no superset of the set also satisfies the condition.

This follows from the argument in Step 1a.

Step 2: If $i \in N$, $R_i \in \tilde{\mathcal{R}}$, $x \in X$, and $x' \in X'$, then $x_i \; I_i \; x'_i$.

Let $R_1 \in \tilde{\mathcal{R}}$. Let $R_2 = R_1$ and $z = \varphi(R)$. As in the proof of Theorem 2.2, note that for all $k \in \{1, 2\}$ and all $i \in N$, there exists $R_i^k \in \tilde{\mathcal{R}}$ such that $P(R_i^k, R_{-i}) = P^k$. Therefore by *strategy-proofness*, for all $i \in N$ and all $x \in X \cup X'$, $z_i \; R_i \; x_i$. Since $X \cup X' \subset P(R) = Z$, we have for all $i \in N$ and all $x \in X \cup Y$, $z_i \; I_i \; x_i$. Since $R_1 \in \tilde{\mathcal{R}}$ was arbitrary, either for all $R_1 = R_2 \in \tilde{\mathcal{R}}$, we have $\varphi(R) \in X = X' = \{(\Omega, 0)\}$, or for all $R_1 = R_2 \in \tilde{\mathcal{R}}$, we have $\varphi(R) \in X = X' = \{(0, \Omega)\}$.

Step 3: For all $R \in \mathcal{R}_L^2$, $\{\varphi(R)\} = X$.

Without loss of generality, suppose that $X = \{(\Omega, 0)\}$, *i.e.* that for all $R \in \tilde{\mathcal{R}}^2$, $\varphi(R) = (\Omega, 0)$. Let $R_1 \in \mathcal{R}_L$ and $R_2 \in \tilde{\mathcal{R}}$. *Strategy-proofness* implies $\varphi_1(R) \; R_1 \; \Omega$. Since R_1 is strictly monotonic, $\varphi(R) = (\Omega, 0)$. Let $R'_2 \in \mathcal{R}_L$. *Strategy-proofness* implies $\varphi_2(R) \; R_2 \; \varphi_2(R_1, R'_2)$. Strict monotonicity then implies $\varphi(R_1, R'_2) = (\Omega, 0)$. We have shown that φ is *dictatorial*: for all $R \in \mathcal{R}_L^2$, $\varphi(R) = (\Omega, 0)$. ■

Remark 2.3 As with Theorem 2.2, this result holds on any smaller domain for which we can find a set $\tilde{\mathcal{R}}$, as defined above, that contains at least four elements.

2.4 Extensions

Many properties of solutions, such as *strategy-proofness* or *efficiency*, are stronger on larger domains; it is simply the case that the definitions of such properties imply more conditions on larger domains. One way to formalize such a statement is by showing that the results above apply to any domain of monotonic preferences containing either $\mathcal{R}_H$ or $\mathcal{R}_L$.[7]

Corollary to Theorem 2.1 *Let $\tilde{\mathcal{R}} \supset \mathcal{R}_H$ be a class of preference relations over $\mathbb{R}_+^\ell$ that are monotonic and strictly monotonic on the interior. If $\varphi \in \Phi(\tilde{\mathcal{R}})$ is* strategy-proof *and* efficient*, then it is* dictatorial.

Proof: For $\tilde{\mathcal{R}}$ as defined above, let $\tilde{\varphi} \in \Phi(\tilde{\mathcal{R}})$ be *strategy-proof* and *efficient*. Let $\varphi \colon \mathcal{R}_H^2 \to Z$ be such that for all $R \in \mathcal{R}_H$, $\varphi(R) = \tilde{\varphi}(R)$. Then φ is *strategy-proof* and *efficient*, and hence by Theorem 2.1 is *dictatorial*. Without loss of generality, let agent 1 be that dictator.

Let $R_1 \in \mathcal{R}_H$ and $R_2 \in \tilde{\mathcal{R}}$. Let $R'_2 \in \mathcal{R}_H$ satisfy strict monotonicity. We've shown that $\tilde{\varphi}_2(R_1, R'_2) = \varphi_2(R_1, R'_2) = 0$. By *strategy-proofness* we have $\tilde{\varphi}_2(R_1, R'_2) \; R'_2 \; \tilde{\varphi}_2(R_1, R_2)$, hence $\tilde{\varphi}(R_1, R_2) = (\Omega, 0)$.

---

[7] I thank Lin Zhou for noting this, with respect to the domain in his work.

Let $R_1 \in \tilde{\mathcal{R}}$, $R_2 \in \tilde{\mathcal{R}}$, and $R'_1 \in \mathcal{R}_H$. We've shown that $\tilde{\varphi}(R'_1, R_2) = (\Omega, 0)$. By *strategy-proofness* we have, $\tilde{\varphi}_1(R_1, R_2)\ R_1\ \tilde{\varphi}_1(R'_1, R_2)$. Therefore agent 1 is a dictator on $\tilde{\mathcal{R}}^2$. ■

Remark 2.4 As noted in Remark 2.1, $\mathcal{R}_{HS}$ is a subdomain of Zhou's (1991b) domain, and if $\varphi \in \Phi(\mathcal{R}_{HS})$ is *efficient* and *strategy-proof*, then it is *dictatorial*. By an argument identical to the one proving the Corollary to Theorem 2.1, Zhou's result is a corollary to that Remark.

We also have the following.

Corollary to Theorem 2.3 *Let $\tilde{\mathcal{R}} \supset \mathcal{R}_L$ be a class of preference relations over $\mathbb{R}^\ell_+$ that are monotonic. If $\varphi \in \Phi(\tilde{\mathcal{R}})$ is* strategy-proof *and* efficient*, then it is* dictatorial*.*

2.5 Summary

We may have hoped that *strategy-proofness* would not be as strong on certain "small" economic domains as it has been shown to be on more general ones. Unfortunately we have shown that when combined with *efficiency*, it forces us to use a *dictatorial* solution on these domains, at least in the 2-person case. That the negative result for exchange economies holds on any class of monotonic preferences containing the class of linear preferences is particularly discouraging, as this domain is already substantially smaller than the usual domains of exchange economies. It appears that no reasonable domain will allow *strategy-proof* and *efficient* solutions that are not *dictatorial*.

Of interest, however, is whether similar results hold for economies with more than two agents. Even though *strategy-proof* and *efficient* solutions for n-agent exchange economies may not be *dictatorial* when $n > 2$,[8] Zhou (1991b) conjectures that they must be *inversely dictatorial*: some agent always receives nothing. The methods used here may facilitate resolving this issue.

[8] For example, in a 3-agent exchange economy, define a solution to give the entire endowment either to agent 1 or agent 2, depending only on the preferences of agent 3 (Satterthwaite and Sonnenschein, 1981).

Chapter 3

Economies with Public Goods

In this chapter, we consider economies with public goods. Our first, and strongest, result for this type of environment is on our *smallest* domain of preferences. We show that on the class of 2-agent economies in which agents have linear preferences over one private good and one public good (produced according to a constant-returns production function), a *strategy-proof* and *efficient* solution must be *dictatorial*. This result is then extended to a class of economies with many public goods. In turn, the result is finally extended to the class of economies with an arbitrary number of agents with linear preferences over just public goods, coinciding with the domain of von Neumann-Morgenstern preferences over lotteries. We discuss the connection of this work to the literature on the Clarke-Groves mechanisms in the Summary.

3.1 The Model

There is a set of agents, $N = \{1, 2, \ldots, n\}$.[1] In Sections 3.2 and 3.3.1, we consider the case where $n = 2$, and in Section 3.3.2, the case where $n \geq 2$. The set of allocations is a k-dimensional simplex ($k \geq 3$), denoted by $\triangle \equiv \{\delta \in \mathbb{R}^k_+ : \sum \delta_j = 1\}$. The points in this simplex can represent, for example, (1) allocations of an aggregate endowment of a divisible good among two agents and the linear production of $k-2$ public goods, or (2) lotteries over k public alternatives.

An **extreme point** of $\triangle$ is an allocation $\delta \in \triangle$ such that for some $j \in \{1, \ldots, k\}$, we have $\delta_j = 1$. Denote these points as $\boldsymbol{E} = \{\boldsymbol{e^1}, \boldsymbol{e^2}, \ldots, \boldsymbol{e^k}\} \equiv \{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$. Typical elements of E are denoted $a, a', b, c, e^j, \ldots$ etc.

[1] See Section 1.3 for preliminary notation.

The set of admissible preferences on $\triangle$ depends on our interpretation of $\triangle$. Considering interpretation (2) of $\triangle$ as a set of lotteries, to say that agents have von Neumann-Morgenstern preferences over lotteries is the same as allowing to be admissible only those preferences representable by a utility function of the form $u(\delta) = \sum \lambda_j \delta_j$ for some $\lambda \in \mathbb{R}^k$. Let $\boldsymbol{\mathcal{R}_L}$ denote that class of "linear" preference relations.

Alternatively, consider interpretation (1) of $\triangle$, above. For a given agent, there should be an extreme point, interpreted as the allocation at which no public goods are produced and the agent receives none of the private good, that should be ranked as the worst allocation in $\triangle$ by any admissible preference relation for that agent. In addition, we may choose to impose the usual monotonicity and convexity assumptions. Given an extreme point $e^j \in E$, a preference relation R_i is **monotonic with respect to $\boldsymbol{e^j}$** if for all $\delta, \delta' \in \triangle$ and all $\lambda \in (0,1)$, $\delta = \lambda a + (1-\lambda)\delta'$ implies $\delta' \; P_i \; \delta$ (where P_i is the asymmetric part of R_i). Let $\mathcal{R}^{e^j} \subset \mathcal{R}_L$ be the class of preference relations that are both linear and monotonic with respect to e^j. Finally, let $\boldsymbol{\mathcal{R}_C^{e^j}}$ be the class of preference relations that are only convex[2] and monotonic with respect to e^j. For any of these preference relations, R_i, the notation $\arg\max R_i$ refers to the set of maximal elements of $\triangle$ under R_i.

For the remainder of this section, let $\mathcal{R}$ be an arbitrary domain. As in Chapter 2, the **Pareto set** for a preference profile is defined as follows.

$$\forall R \in \mathcal{R},\ P(R) = \{\delta \in \triangle : \nexists \delta' \in,\ i \in N \text{ such that } z' \; P_i \; z \text{ and } z' \; R_j \; z, \forall j \neq i\}$$

Elements of the Pareto set of R are **Pareto-optimal** allocations for R.

A solution $\varphi \colon \mathcal{R} \to \triangle$ is ***efficient*** if for all $R \in \mathcal{R}$, we have $\varphi(R) \in P(R)$. A solution is ***dictatorial*** if there exists $i \in N$ such that for all $R \in \mathcal{R}$, $\varphi(R) \in \arg\max_{\triangle} R_i$.

3.2 Two Agent Mixed Economies

In this section, we interpret an allocation as the division of an aggregate endowment of one unit of a divisible good among k uses: direct consumption of some of the good by each of the two agents, and production of $k-2$ public goods according to linear production functions.[3] For an allocation $\delta = (x_1, x_2, y_1, \ldots, y_{k-2}) \in \triangle$, the first component, x_1, is interpreted as agent 1's consumption of the private

[2] Convex: for all $\delta, \delta' \in \triangle$, if $\delta' \; R_i \; \delta$ and $\lambda \in [0,1]$, then $\lambda\delta' + (1-\lambda)\delta \; R_i \; \delta$.

[3] More accurately, the production is actually according to *identity* functions, due to our definition of $\triangle$. However we could have specified $\triangle$ to be any weighted simplex (*i.e.* $\triangle^\lambda = \{\delta \in \mathbb{R}^k_+ : \sum \lambda_j \delta_j = 1\}$) and all results in this paper would continue to hold.

good, the second component, x_2, as agent 2's, and the remaining components, $(y_1, \ldots, y_{k-2})$, as levels of the public goods.[4]

Therefore, the extreme point e^2 is interpreted as the allocation at which agent 2 consumes all of the endowment as a private good. This allocation is always considered to be the unique worst allocation according to agent 1's preferences. Similarly, e^1 is interpreted as the allocation at which agent 1 receives all of the endowment, which is always worst according to agent 2's preferences. We will first consider the domain $\mathcal{R}^{e^2} \times \mathcal{R}^{e^1}$.

3.2.1 Linear Preferences

In Section 3.2.1, we simplify notation by denoting the domain $\mathcal{R} \equiv \mathcal{R}^{e^2} \times \mathcal{R}^{e^1}$.

One Public Good

Given a preference relation $R_i \in \mathcal{R}^{e^j}$ we denote by r_i the amount of private good an agent with preferences R_i is willing to give up in exchange for an additional unit of the public good. For example, if $r_i = 1$, then agent i is indifferent between receiving one additional unit of the public good and receiving one additional unit of the private good. If $r_i > 1$, then agent i strictly prefers one additional unit of the public good to one additional unit of the private good.

Note that the Pareto set for a preference profile $R \in \mathcal{R}$ is one of six possible sets (see Figure 3.1).

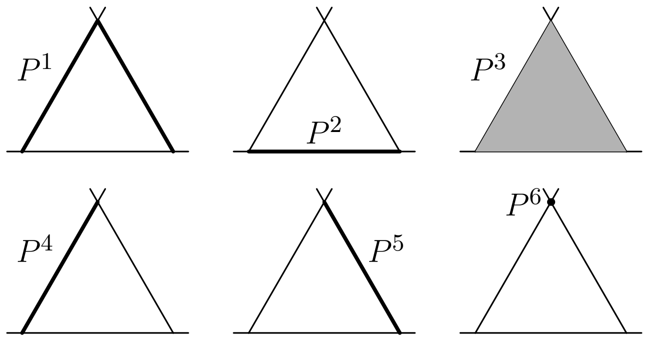
Theorem 3.1 (One Public Good, One Private Good) *Suppose* $k = 3$. *If* $\varphi\colon \mathcal{R} \to \triangle$ *is* strategy-proof *and* efficient*, then it is* dictatorial.

Proof: Let φ be *strategy-proof* and *efficient*. First we show that there exists $\delta^1 \in P^1$ such that for all $R \in \mathcal{R}$, $P(R) = P^1$ implies $\varphi(R) = \delta^1$.

Let $R, R' \in \mathcal{R}$ be such that $P(R) = P(R') = P^1$. By *efficiency*, $\varphi(R) \in P^1$ and $\varphi(R') \in P^1$. For all $i \in N$, if $r_i \geq r'_i$ let $R''_i = R_i$, and $R''_i = R'_i$ otherwise. By Lemma 1.1 and the construction of R'', $\varphi(R) = \varphi(R''_1, R_2) = \varphi(R'')$ and $\varphi(R') = \varphi(R''_1, R'_2) = \varphi(R'')$. Hence $\varphi(R) = \varphi(R')$.

Similarly for all $j \in \{2, 4, 5, 6\}$, there exists $\delta^j \in P^j$ such that for all $R \in \mathcal{R}$, $P(R) = P^j$ implies $\varphi(R) = \delta^j$.

[4] S.-C. Kolm introduces a simple representation of the (non-wasteful) allocations in an economy of two agents, one private good, and one public good produced with a linear production function and fixed endowment; these are the points in an equilateral triangle, where each corner represents an allocation at which all of the endowment is devoted to either one agent's consumption, or production of the public good. See Schlesinger (1989) for a brief history of, and justification for, Kolm's representation.



If (R_1, R_2) satisfies ...	then $P(R)$ is ...
$r_1 < 1$, $r_2 < 1$, $r_1 + r_2 > 1$	$P^1 \equiv \{(x_1, x_2, y) \in \triangle : x_1 = 0 \text{ or } x_2 = 0\}$
$r_1 < 1$, $r_2 < 1$, $r_1 + r_2 < 1$	$P^2 \equiv \{(x_1, x_2, y) \in \triangle : y = 0\}$
$r_1 < 1$, $r_2 < 1$, $r_1 + r_2 = 1$	$P^3 \equiv \triangle$
$r_1 \geq 1$, $r_2 < 1$	$P^4 \equiv \{(x_1, x_2, y) \in \triangle : x_1 = 0\}$
$r_1 < 1$, $r_2 \geq 1$	$P^5 \equiv \{(x_1, x_2, y) \in \triangle : x_2 = 0\}$
$r_1 \geq 1$, $r_2 \geq 1$	$P^6 \equiv \{(0, 0, 1)\}$

Figure 3.1: Pareto sets of the Kolm triangle for linear preferences.

It is simple to check that for all $R \in \mathcal{R}$ such that $P(R) = P^1$, there exists $R' \in \mathcal{R}$ such that for all $i \in N$, $P(R'_i, R_{-i}) = P^2$. *Strategy-proofness* then implies that for all $R \in \mathcal{R}$ such that $P(R) = P^1$ and all $i \in N$, $\delta^1 \; R_i \; \delta^2$.

Similarly for all $R \in \mathcal{R}$ such that $P(R) = P^2$, there exists $R'_i \in \mathcal{R}_L^{e^i}$ such that for all $i \in N$, $P(R'_i, R_{-i}) = P^1$. *Strategy-proofness* then implies that for all $R \in \mathcal{R}$ such that $P(R) = P^2$ and for all $i \in N$, $\delta^2 \; R_i \; \delta^1$.

It follows that $\delta^1 = \delta^2$. Hence $\delta^1 \in \{e^1, e^2\}$. Without loss of generality, we will assume $\delta^1 = e^1$.

For all $R \in \mathcal{R}$, the following hold:

$$\begin{aligned}
P(R) = P^4 &\implies \exists R'_1 \in \mathcal{R}^{e^2} \text{ such that } P(R'_1, R_2) = P^1 \\
P(R) = P^3 &\implies \exists R'_1, R''_1 \in \mathcal{R}^{e^2} \text{ such that } P(R'_1, R_2) = P^1,\ P(R''_1, R_2) = P^4 \\
P(R) = P^1 &\implies \exists R'_2 \in \mathcal{R}^{e^1} \text{ such that } P(R_1, R'_2) = P^5
\end{aligned}$$

Hence, *strategy-proofness* implies that for all $R \in \mathcal{R}$,

$$P(R) = P^4 \implies \delta^4 \; R_1 \; \delta^1 \tag{3.1}$$

$$P(R) = P^3 \implies \varphi(R) \; R_1 \; \delta^1 \text{ and } \varphi(R) \; R_1 \; \delta^4 \tag{3.2}$$

$$P(R) = P^1 \implies \delta^1 \; R_2 \; \delta^5 \tag{3.3}$$

Since there exists $R \in \mathcal{R}$ such that $P(R) = P^4$ and $r_1 = 1$, (3.1) implies $\delta^4 = e^3$. Since preferences are linear, (3.2) implies that for all $R \in \mathcal{R}$, we have $P(R) = P^3$ implies $\varphi(R) \in \arg\max R_1$. Finally, (3.3) implies $\delta^5 = e^1$.

We have shown that for all $R \in \mathcal{R}$, we have $\varphi(R) \in \arg\max R_1$, that is, φ is *dictatorial*. ■

Multiple Public Goods

The result of the previous section can be extended to the case of more than one public good in an intuitively simple way. We consider a subclass of the linear preferences for which the public goods are "equivalent". That is, for any two allocations $(x_1, x_2, y_1, \ldots, y_{k-2})$ and $(x'_1, x'_2, y'_1, \ldots, y'_{k-2})$, if $\sum y_j = \sum y'_j$, then the agents are indifferent between the two allocations. This subdomain is essentially the same as the domain of section 3.2.1. Hence there is a dictator on that subdomain. It is then simple to show that this agent must be a dictator on the entire domain.

Corollary 3.1 (Many Public Goods, One Private Good) *If* $\varphi\colon \mathcal{R}^2 \to \triangle$ *is* strategy-proof *and* efficient*, then it is* dictatorial*.*

Proof: Let φ be *strategy-proof* and *efficient*. Define the subdomain of preference relations for which agent one is indifferent among all the public goods as follows:

$$\overline{\mathcal{R}}^{e^2} = \{R_1 \in \mathcal{R}^{e^2} : \forall (x_1, x_2, y_1, \ldots, y_{k-2}), (x'_1, x'_2, y'_1, \ldots, y'_{k-2}) \in \triangle, \text{ if } x_1 = x'_1 \text{ and } x_2 = x'_2 \text{ then } (x_1, x_2, y_1, \ldots, y_{k-2}) \; I_1 \; (x'_1, x'_2, y'_1, \ldots, y'_{k-2})\}$$

Define $\overline{\mathcal{R}}^{e^1}$ similarly. From Theorem 3.1, we know that the restriction of φ to the subdomain $\overline{\mathcal{R}}^{e^1} \times \overline{\mathcal{R}}^{e^2}$ is *dictatorial* on that subdomain.[5] Without loss of generality, suppose φ chooses an allocation giving agent 1 one of his most preferred bundles on that subdomain. We must show that one of agent 1's most preferred allocations is chosen for each profile in $\mathcal{R}^2$.

[5] We do not show this explicitly, though it can be done by allowing the δ^j's in the proof of Theorem 3.1 to be *sets* of allocations, among which both agents are always indifferent when preferences are in the subdomain, instead of requiring each δ^j to be a single allocation.

Let $R \in \mathcal{R}$ be such that $R_1 \in \overline{\mathcal{R}}^{e^2}$. If $\arg\max R_1 = \{e^1\}$, then by supposition, for all $R_2' \in \overline{\mathcal{R}}^{e^1}$, $\varphi(R_1, R_2') = e^1$. *Strategy-proofness* implies $\varphi(R)\ R_2\ \varphi(R_1, R_2')$, so $\varphi(R) = e^1$. If, on the other hand, $e^1 \notin \arg\max R_1$, then note that $e^3 \in \arg\max R_1$. By *efficiency*, for all $R_2' \in \overline{\mathcal{R}}^{e^1}$, $\varphi(R_1, R_2')$ is of the form $(0, 0, y_1, \ldots, y_{k-2})$. Hence by *strategy-proofness*, for all $R_2' \in \overline{\mathcal{R}}^{e^1}$, $(0, 0, y_1, \ldots, y_{k-2})\ R_2'\ \varphi(R)$. It follows that $\varphi(R)\ I_1\ e^3$.

Now let $R \in \mathcal{R}$. The previous arguments imply that there exists $R_1' \in \overline{\mathcal{R}}^{e^2}$ such that $\varphi(R_1', R_2) = e^2$. Therefore *strategy-proofness* implies $\varphi(R)\ R_1\ e^2$. Let $a \in E\backslash\{e^1\}$. Since preferences are linear, we can conclude that φ is *dictatorial* by showing that $\varphi(R)\ R_1\ a$. Let $R_1'' \in \mathcal{R}^{e^2}$ satisfy $a\ I_1''\ e^1\ P_1''\ b$ for each $b \in E\backslash\{\delta, e^1\}$. Again by *strategy-proofness*, $\varphi(R_1'', R_2)\ R_1''\ e^2$. Hence *efficiency* implies $\varphi(R_1'', R_2) = a$. Finally, *strategy-proofness* implies $\varphi(R)\ R_1\ a$. ■

3.2.2 Convex Preferences

The negative results above may be extended to broader domains of preferences. We show how to extend Theorem 3.1, the case of one public good, to the domain of convex, monotonic preferences. The domain in this section is $\mathcal{R}_C \equiv \mathcal{R}_C^{e^2} \times \mathcal{R}_C^{e^1}$.

Corollary 3.2 (One Public Good, One Private Good) *Suppose* $k = 3$. *If* $\varphi\colon \mathcal{R}_C \to \triangle$ *is* strategy-proof *and* efficient*, then it is* dictatorial.

Proof: Let φ be *strategy-proof* and *efficient*. Then the restriction of φ to the subdomain $\mathcal{R}^{e^2} \times \mathcal{R}^{e^1}$ is *strategy-proof* and *efficient* on that subdomain. By Theorem 3.1, then, there exists an agent $i \in N$ such that for all $R \in \mathcal{R}^{e^2} \times \mathcal{R}^{e^1}$ and all $\delta \in \triangle$, we have $\varphi(R)\ R_i\ \delta$. Assume without loss of generality that this is agent 1.

Let $R \in \mathcal{R}_C$. Let $\delta \in (\arg\max R_1) \cap P(R)$ (see Figure 3.2). Note that of agent 1's favorite allocations in $\triangle$, δ is the one providing the highest level of the public good. We need to show that $\varphi(R) = \delta$.

Let $R_2' \in \mathcal{R}^{e^1}$, and let $R_1' \in \mathcal{R}_C^{e^2}$ and $R_1'' \in \mathcal{R}^{e^2}$ satisfy the following (see Figure 3.2):

$$\arg\max R_1' = \{\delta' \in \triangle : \text{for some } \lambda \in [0,1],\ \delta' = \lambda\delta + (1-\lambda)e^1\}$$
$$\arg\max R_1'' = \{e^1\}$$

Since agent 1 is the dictator on the subdomain of linear preferences, we have $\varphi(R_1'', R_2') = e^1$. *Strategy-proofness* implies $\varphi(R')\ R_1'\ \varphi(R_1'', R_2') = e^1$, so by *efficiency*, $\varphi(R') = \delta$. *Strategy-proofness* implies $\varphi(R_1, R_2')\ R_1\ \delta$, so by *efficiency*,

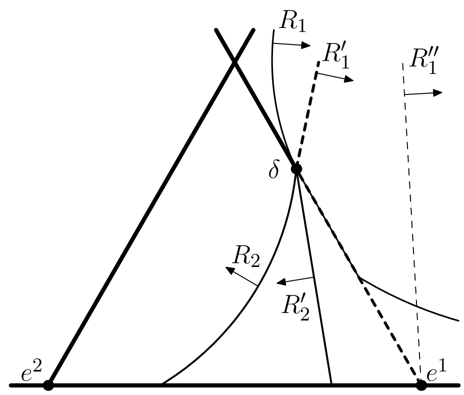


Figure 3.2: (Proof of Corollary 3.2) The dictator on the linear subdomain must be a dictator on the convex domain.

$\varphi(R_1, R_2') = \delta$. *Strategy-proofness* implies $\varphi(R_1, R_2')\ R_2'\ \varphi(R)$, *i.e.* $\delta\ R_2'\ \varphi(R)$. However R_2' was chosen as an arbitrary element of $\mathcal{R}^{e^1}$, so for some $\lambda \in [0,1]$, $\varphi(R) = \lambda\delta + (1-\lambda)e^1$. *Efficiency* then implies $\varphi(R) = \delta$. ■

3.3 Pure Public Good Economies and Lotteries over Alternatives

3.3.1 Two Agents

We now extend the first result to the larger class of 2-agent economies in which agents may have any linear preferences over $\triangle$; *i.e.* the domain $\mathcal{R}_L^2$. The most obvious interpretation of this domain is as the situations in which agents have von Neumann-Morgenstern preferences over lotteries. We then extend the result to the case of more than two agents.

Let $a \neq b \in E$ be two distinct extreme points. Consider the class of economies for which preference profiles are elements of $\mathcal{R}^a \times \mathcal{R}^b \subset \mathcal{R}_L^2$. Such a class is equivalent to the one described in the previous section. Hence an *efficient* and *strategy-proof* solution on $\mathcal{R}_L^2$, which is *efficient* and *strategy-proof* on the subdomain $\mathcal{R}^a \times \mathcal{R}^b$, must choose one of a given agent's favorite points in the simplex for any preference profile in $\mathcal{R}^a \times \mathcal{R}^b$. This fact is exploited in the proof of the following result.

Corollary 3.3 (Many Public Goods, Two Agents) *If* $k \geq 3$*,* $n = 2$*, and* $\varphi\colon \mathcal{R}_L^2 \to \triangle$ *is* strategy-proof *and* efficient*, then* φ *is* dictatorial.

Proof: Let φ be *strategy-proof* and *efficient*. Let $a, b \in E$ be such that $a \neq b$. Then the restriction of φ to the subdomain $\mathcal{R}^a \times \mathcal{R}^b$ is a *strategy-proof* and *efficient* solution for the domain of preferences in Section 3.2.1. Hence by Theorem 3.1 (or Corollary 3.1), there must exist a "subdomain-dictator" for that subdomain, *i.e.* there is an agent for whom φ always chooses one of his favorite allocations when the preference profile is in that subdomain.

This is true for any subdomain defined in terms of two different extreme points $a, b \in E$. We can show that each such subdomain-dictator must in fact be the same agent; this is Step 1 of the proof. In Step 2 we show that this agent is a dictator over the domain of profiles at which agents have unique (but possibly identical) least favorite extreme points. In Step 3 we show that the agent is a dictator for the entire domain.

Step 1: There is one agent who is the dictator of each subdomain $\mathcal{R}^a \times \mathcal{R}^b$ $(a \neq b)$.

Let $a, b, c, d \in E$ be such that $a \neq b$ and $c \neq d$. Without loss of generality, assume that for any profile $R \in \mathcal{R}^a \times \mathcal{R}^b$, $\varphi(R)$ is one of agent 1's favorite allocations according to R_1. We want to show that this is also the case for any profile $R' \in \mathcal{R}^c \times \mathcal{R}^d$.

Let $a' \in E$ be such that: if $a \neq d$ then $a' = a$; if $a = d$ and $b = c$ then $a \neq a' \neq b$; if $a = d$ and $b \neq c$ then $a' = c$. Since $a' \neq b$, note that the subdomain $\mathcal{R}^{a'} \times \mathcal{R}^b$ is one for which we have shown that there exists a dictator. If $a' = a$, then agent 1 is, trivially, the dictator for that subdomain, by our assumption. Otherwise, let $R_1 \in \mathcal{R}^a$ and $R'_1 \in \mathcal{R}^{a'}$ be such that $\arg\max R_1 = \arg\max R'_1 = \{b\}$. Let $R_2 \in \mathcal{R}^b$. By assumption, $\varphi(R_1, R_2) = b$. *Strategy-proofness* implies that $\varphi(R'_1, R_2)\ R'_1\ \varphi(R) = b$. Hence $\varphi(R'_1, R_2) = b$ is not one of agent 2's favorite allocations, so agent 1 must be the dictator on the subdomain $\mathcal{R}^{a'} \times \mathcal{R}^b$.

Since $a' \neq d$, note that the subdomain $\mathcal{R}^{a'} \times \mathcal{R}^d$ is one for which we have shown that there exists a dictator. Again, if $b = d$, agent 1 is, trivially, the dictator for the subdomain. Otherwise, let $R'_2 \in \mathcal{R}^b$ and $R''_2 \in \mathcal{R}^d$ be such that $\arg\max R'_2 = \arg\max R''_2 = \{a'\}$ (since $a' \neq b$). *Strategy-proofness* implies that $\varphi(R'_1, R'_2)\ R'_2\ \varphi(R'_1, R''_2)$, and we have shown that $\varphi(R'_1, R'_2) = b$. Therefore $\varphi(R'_1, R''_2) = b$, implying that agent 1 is the dictator on the subdomain $\mathcal{R}^{a'} \times \mathcal{R}^d$.

If $a' = c$ then we are done with this step. Otherwise, since $a' \neq d$, let $R''_1 \in \mathcal{R}^{a'}$ and $\mathcal{R}'''_1 \in \mathcal{R}^c$ be such that $\arg\max R''_1 = \arg\max R'''_1 = \{d\}$. *Strategy-proofness* implies that $\varphi(R'''_1, R''_2)\ R'''_1\ \varphi(R''_1, R''_2) = d$. Hence $\varphi(R'''_1, R''_2) = d$ is not one of agent 2's favorite allocations, so agent 1 must be the dictator on the subdomain $\mathcal{R}^c \times \mathcal{R}^d$.

Step 2: The subdomain-dictator is a dictator when agents have a unique worst alternative.

In the previous step, we showed that for any two extreme points $a \neq a'$, and any profile $R \in \mathcal{R}^a \times \mathcal{R}^{a'}$, $\varphi(R)$ is one of agent 1's favorite allocations according to R_1. Consider $R \in \mathcal{R}^a \times \mathcal{R}^a$. Let $b, c \in E$ be such that $b \in \arg\max R_1$ and $b \neq c$. Note that there exist $R_1' \in \mathcal{R}^c$ such that $\arg\max R_1' = \{b\}$. *Strategy-proofness* implies $\varphi(R_1, R_2)\ R_1\ \varphi(R_1', R_2) = b$, that is, $\varphi(R)$ is one of agent 1's favorite allocations according to R_1.

Similarly, for any extreme point $d \in E$ and any profile $R' \in \mathcal{R}^d \times \mathcal{R}^d$, $\varphi(R')$ is one of agent 1's favorite allocations according to R_1'. Let $\overline{\mathcal{R}} = \cup_{a,b \in E} \mathcal{R}^a \cup \mathcal{R}^b$. So far we have shown that for any profile $R'' \in \overline{\mathcal{R}}^2$, $\varphi(R'')$ is one of agent 1's favorite allocations according to R_1''.

Step 3: The subdomain-dictator is a dictator on the whole domain.

Let $R_1, R_1' \in \overline{\mathcal{R}}$, $R_2 \in \mathcal{R}_L \backslash \overline{\mathcal{R}}$, and $a \in E$ be such that for all $b \in E \setminus \{a\}$, we have $a\ P_1\ b$, $a\ R_1'\ b$, and $b\ R_2\ a$. Then for all $R_2' \in \mathcal{R}^a$, $\varphi(R_1, R_2')\ R_2'\ a$. Also, *strategy-proofness* implies $\varphi(R_1, R_2')\ R_2'\ \varphi(R_1, R_2)$. Therefore, $\varphi(R) = a$. *Strategy-proofness* implies $\varphi(R_1', R_2)\ R_1'\ \varphi(R)$, hence $\varphi(R_1', R_2)$ is one of agent 1's favorite allocations according to R_1'.

Similarly, for any profile $R \in \overline{\mathcal{R}} \times (\mathcal{R}_L \setminus \overline{\mathcal{R}})$, $\varphi(R)$ is one of agent 1's favorite allocations according to R_1.

Finally, let $R \in \mathcal{R}_L \times \mathcal{R}_L$, $R_1' \in \overline{\mathcal{R}}$, and $a \in E$ be such that $a \in \arg\max R_1$ and $\{a\} = \arg\max R_1'$. We have shown that $\varphi(R_1', R_2) = a$. Hence *strategy-proofness* implies $\varphi(R_1, R_2)\ R_1\ a$. Since R was arbitrary, φ is *dictatorial*. ■

As stated in the Introduction, this 2-agent result is implied by Hylland (1980). However, our geometric proof is more intuitive than the one in that work, and, by using Theorem 3.1 in the proof, demonstrates that the essential incompatibility between *strategy-proofness* and *efficiency* is deeper than Corollary 3.3 suggests.

3.3.2 Many Agents

We now use Corollary 3.3 to show the corresponding result for more than two agents. The proof is by induction, and may remind the reader of proofs, such as in Kalai and Muller (1977) and Aswad and Sen (1996), which show that results like the Arrow Impossibility Theorem and the Gibbard-Satterthwaite Theorem hold on certain 2-agent domains if and only if they hold on the corresponding n-agent domains. Those results have no implications for our smaller domain of preferences.

Corollary 3.4 (Many Public Goods, Many Agents) *If $k \geq 3$, $n \geq 3$, and $\overline{\varphi} \colon \mathcal{R}^n \to \triangle$ is* strategy-proof *and* efficient*, then $\overline{\varphi}$ is* dictatorial.

Proof: Let $\overline{\varphi}$ be *strategy-proof* and *efficient*. Let $M = \{n-1, n\}$. Define an $(n-1)$-agent solution $\varphi\colon \mathcal{R}_L^{n-1} \to \triangle$ as follows: for all $(R_1, R_2, \ldots, R_{n-2}) \in \mathcal{R}_L^{n-2}$ and all $R_m \in \mathcal{R}_L$, where $R'_1 = R_1, R'_2 = R_2, \ldots, R'_{n-2} = R_{n-2}$, and $R'_{n-1} = R'_n = R_m$, let $\varphi(R_1, \ldots, R_{n-2}, R_m) = \overline{\varphi}(R')$. When agents $(n-1)$ and n have the same preferences, we are joining them into one "agent", whom we call m.

Step 1: φ is a *strategy-proof* and *efficient* $(n-1)$-agent solution.

Efficiency is obvious, and since $\overline{\varphi}$ is *strategy-proof*, agents 1 through $n-2$ cannot manipulate φ. Let $R_m, R'_m \in \mathcal{R}_L$. Let $R, R' \in \mathcal{R}_L^n$ be such that both $R_{n-1} = R_n = R_m$ and $R'_{n-1} = R'_n = R'_m$, and denote $R_{-M} = (R_1, \ldots, R_{n-2})$. *Strategy-proofness* implies that $\overline{\varphi}(R_{-M}, R_{n-1}, R_n)\ R_{n-1}\ \overline{\varphi}(R_{-M}, R'_{n-1}, R_n)$ and $\overline{\varphi}(R_{-M}, R'_{n-1}, R_n)\ R_n\ \overline{\varphi}(R_{-M}, R'_{n-1}, R'_n)$. Hence $\varphi(R_{-M}, R_m)\ R_m\ \varphi(R_{-M}, R'_m,)$, so φ is *strategy-proof*.

Step 2: If φ is *dictatorial* and agent $i \neq m$ is the dictator for φ, then $\overline{\varphi}$ is *dictatorial*, and agent i is the dictator for $\overline{\varphi}$.

Suppose that φ is *dictatorial* and, without loss of generality, that agent 1 is the dictator. Let $R \in \mathcal{R}_L^n$. Let $R'_2, \ldots, R'_{n-1}, R_m$ be such that $R'_2 = \cdots = R'_{n-1} = R_m = R_n$. *Strategy-proofness* implies the following:

$$
\begin{aligned}
&\overline{\varphi}(R_1, \ldots, R_{n-3}, R_{n-2}, R'_{n-1}, R_n)\ R'_{n-1}\ \overline{\varphi}(R_1, \ldots, R_{n-1}, R_n) \\
&\overline{\varphi}(R_1, \ldots, R_{n-3}, R'_{n-2}, R'_{n-1}, R_n)\ R'_{n-2}\ \overline{\varphi}(R_1, \ldots, R_{n-2}, R'_{n-1}, R_n) \\
&\qquad\vdots \\
&\overline{\varphi}(R_1, R'_2, \ldots, R'_{n-1}, R_n)\quad R'_2\quad \overline{\varphi}(R_1, R_2, R'_3, \ldots, R'_{n-1}, R_n)
\end{aligned}
$$

Letting $\delta^n = \overline{\varphi}(R_1, R'_2, \ldots, R'_{n-1}, R_n)$, we have $\delta^n\ R_n\ \overline{\varphi}(R)$. By construction, $\delta^n = \varphi(R_1, R'_2, \ldots, R'_{n-2}, R_m) \in \arg\max R_1$.

Similarly, for all $j \neq 1$, there exists $\delta^j \in \arg\max R_1$ such that $\delta^j\ R_j\ \overline{\varphi}(R)$. Suppose that δ^n is the only element of $\arg\max R_1$, that is, for all $j \neq 1$, $\delta^j = \delta^n$. Then for all $j \neq 1$, $\delta^n\ R_j\ \overline{\varphi}(R)$, hence *efficiency* implies $\overline{\varphi}(R)\ R_1\ \delta^n$, *i.e.* $\overline{\varphi}(R) = \delta^n$.

Suppose that δ^n is not the only element of $\arg\max R_1$. Then using the above argument, for any $R'_1 \in \mathcal{R}_L$ such that $\{\delta^n\} = \arg\max R'_1$, $\overline{\varphi}(R'_1, R_{-1}) = \delta^n$. *Strategy-proofness* implies $\overline{\varphi}(R)\ R_1\ \overline{\varphi}(R'_1, R_{-1})$. Hence $\overline{\varphi}$ is *dictatorial*, and agent 1 is the dictator.

Step 3: Define similar $(n-1)$-agent solutions with respect to other pairs of agents.

Let $M \subset N$ be such that $|M| = 2$. Define $\varphi^M\colon \mathcal{R}_L^{n-1} \to \triangle$ as follows: for all $R_{-M} \in \mathcal{R}_L^{n-2}$ and all $R_m \in \mathcal{R}_L$, $\varphi^M(R_{-M}, R_m) = \overline{\varphi}(R')$, where for all $i \notin M$ we have $R'_i = R_i$, and for all $j \in M$ we have $R'_j = R_m$. Again, we are joining the agents in M into one "agent", m.

As above, if for any $M \subset N$ with $|M| = 2$, there is an agent $i \notin M$ who is a dictator for φ^M, then $\overline{\varphi}$ is *dictatorial*. Hence either (1) $\overline{\varphi}$ is *dictatorial*, or (2) for all $M \subset N$ such that $|M| = 2$, there is no agent $i \notin M$ that is a dictator for φ^M. The proof concludes with an induction argument on the number of agents, showing that (2) cannot hold.

Step 4a: For $n = 3$, the existence of a certain Condorcet triple implies there must be a dictator.

Suppose $n = 3$. According to Corollary 3.3, a 2-agent solution that is *strategy-proof* and *efficient* must be *dictatorial*.

Therefore, for all $M \subset N$ with $|M| = 2$, φ^M is *dictatorial*. Furthermore, if (2) holds, "agent m" must be the dictator for φ^M.

Let $M = \{2, 3\}$, $M' = \{1, 3\}$, and $M'' = \{1, 2\}$. Let $a, b, c \in E$ be three distinct extreme points. Without loss of generality, let $a = (1, 0, \dots, 0)$, $b = (0, 1, 0 \dots, 0)$, and $c = (0, 0, 1, 0, \dots, 0)$. Let $R \in \mathcal{R}_L^3$ be such that for all $d \in E \setminus \{a, b, c\}$,

$$
\begin{aligned}
&a\ P_1\ b\ P_1\ c\ P_1\ d \quad \text{and} \quad b\ I_1\ (2/3, 0, 1/3, 0, \dots, 0)\\
&c\ P_2\ a\ P_2\ b\ P_2\ d \quad \text{and} \quad a\ I_2\ (0, 1/3, 2/3, 0, \dots, 0)\\
&b\ P_3\ c\ P_3\ a\ P_3\ d \quad \text{and} \quad c\ I_3\ (1/3, 2/3, 0, \dots, 0)
\end{aligned}
$$

Let $R'_1 = R_{m'} = R_3$, $R'_2 = R_{m''} = R_1$, and $R'_3 = R_m = R_2$. Supposing that m, m', and m'' are the respective dictators for φ^M, $\varphi^{M'}$, and $\varphi^{M''}$, *strategy-proofness* implies:

$$
\begin{aligned}
\overline{\varphi}(R)\ R_1\ \overline{\varphi}(R'_1, R_2, R_3) &= \varphi^{M'}(R_2, R_{m'}) = b\\
\overline{\varphi}(R)\ R_2\ \overline{\varphi}(R_1, R'_2, R_3) &= \varphi^{M''}(R_3, R_{m''}) = a\\
\overline{\varphi}(R)\ R_3\ \overline{\varphi}(R_1, R_2, R'_3) &= \varphi^{M}(R_1, R_m) = c
\end{aligned}
$$

It is simple to check (see Figure 3.3) that there exists no such $\overline{\varphi}(R)$. Hence there must be a dictator for $\overline{\varphi}$.

Step 4b: For $n \geq 4$, a certain profile puts two pairs of agents in opposition.

Suppose $n \geq 4$. Furthermore, suppose that for all $n' < n$, any *strategy-proof* and *efficient* n'-agent solution must be *dictatorial*. Therefore if (2) holds above, it must be the case that for each φ^M, "agent m" is the dictator.

Let $R \in \mathcal{R}_L^n$ and $a, f \in E$ satisfy $a \neq b$, $R_1 = R_2$, $R_3 = R_4 = \cdots = R_n$, $\arg\max R_1 = \{a\}$, and $\arg\max R_3 = \{b\}$. Let $M = \{1, 2\}$ and $M' = \{3, 4\}$, and let $R_m = R_1$ and $R_{m'} = R_3$.

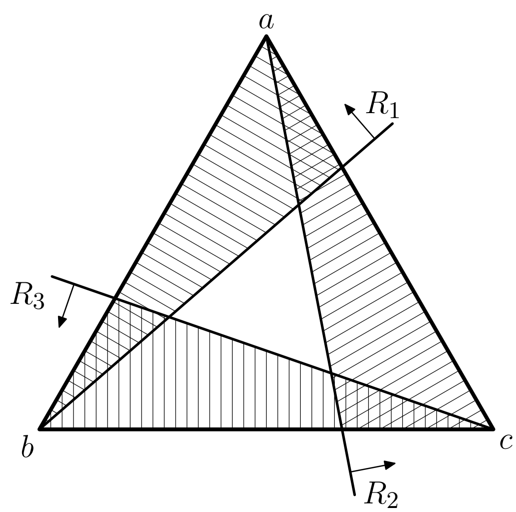


Figure 3.3: (Proof of Corollary 3.4) If $n = 3$, the existence of a Condorcet triple implies that there is a dictator.

If (2) holds, then $\varphi^M(R_m, R_{-M}) = a$ and $\varphi^{M'}(R_{m'}, R_{-M'}) = b$. However by definition, $\varphi^M(R_m, R_{-M}) = \overline{\varphi}(R) = \varphi^{M'}(R_{m'}, R_{-M'})$, which is a contradiction. ■

3.4 Summary

We have shown for various 2-agent economic models that a severe restriction of the class of admissible preferences does not weaken *strategy-proofness* in the presence of *efficiency*. In fact the incompatibility shown for the model of Hylland (1980) stems from the incompatibility on a much smaller domain of preferences over a simplex — those corresponding to mixed economies with one public and one private good, and two agents with linear preferences.

This result may seem to be tied to the literature on Clarke-Groves mechanisms (see Clarke (1971) and Groves (1973)) for domains of quasi-linear preferences. It is not, for the following reasons. Inherent in that literature is the assumption that agents can transfer arbitrary amounts of the divisible good among themselves, hence *efficiency* implies that a chosen level of the public good must maximize the sum of the agents' valuations. This notion of *efficiency* is stronger than our concept of Pareto-optimality, which was subject to our assumption that agents consume *non-negative* quantities of the divisible good. For the same reasons, the results of Walker (1980) and Hurwicz and Walker (1990), which show the "failure" of Clarke-Groves mechanisms at almost all preference profiles in certain domains, have no connection with ours.

This is not the first work to consider a restriction to the domain of linear preferences. Bossert and Weymark (1993) consider linear monotonic preferences over the non-negative 2-dimensional orthant. There, they search for social welfare functions satisfying *anonymity* and Arrow's *independence of irrelevant alternatives*, in turn characterizing median-voter-type solutions; the social preferences are the same as those of the agent whose indifference curves have the m^{th} steepest slope, for some pre-specified m.

In a very recent work, Duggan (1996) provides a "geometric proof" of another result of Gibbard (1977), stating that a *strategy-proof* and *sovereign*[6] solution that depends only on the agents' preferences over the extreme points must be a random-dictator solution.

Given the negative results mentioned in the Introduction and contained herein, and given the appeal of the *efficiency* condition, it may be fruitful to search for *efficient* solutions that "almost" satisfy *strategy-proofness*. That is, while it is desirable to have allocation mechanisms for which it is *always* in each agent's best interest to truthfully reveal his preferences, a reasonable compromise would be a solution for which this is *usually* the case, particularly in situations where agents have little information regarding the preferences of the other agents; without making formal specifications about uncertainty, it is of interest to find *efficient* solutions for which only a "few" preference profiles are manipulable. Unfortunately, on the domain of quasi-linear preferences for the model of Section 3.2, Beviá and Corchón (1995) have shown that *efficient* and *individually-rational* solutions are almost always manipulable. Perhaps this is not the case, however, on a smaller domain of preferences.

[6]Sovereign: each extreme point is chosen for some preference profile.

Chapter 4

Economies with Indivisibilities

In this chapter, we consider the problem of allocating a set of indivisible objects and a fixed amount of a divisible good (*e.g.*, money) to a set of agents with the restriction that each agent consumes precisely one object.[1]

For the case of two agents and two objects, we characterize the entire class of *strategy-proof* solutions. For the case of more than two agents, we show that if a solution is *strategy-proof* and *non-bossy*, or is *coalitionally strategy-proof*, then despite the existence of a continuum of allocations, its range contains at most one allocation per assignment of the indivisible objects to the agents. We also show that if a solution is not manipulable by agents who bribe each other to misrepresent their preferences, then it is constant.

These results extend to the model in which there may be more than one copy of a single object (*e.g.*, identical offices) between which agents do not differentiate in terms of preferences.

Finally we show a general incompatibility between *strategy-proofness* and *efficiency.*

4.1 The Model

There is a set of agents, $N = \{1, 2, \ldots, n\}$.[2] In Section 4.2 we consider the case where $n = 2$, and in Section 4.3, the case where $n \geq 2$. Each agent consumes one indivisible object from a set $\boldsymbol{\Omega} = \{\alpha, \beta, \ldots, \omega\}$, plus an amount of a divisible good. Arbitrary objects are denoted by $a, b \in \Omega$; arbitrary amounts of the divisible good are denoted by $m_i, m'_j \in \mathbb{R}$.

[1]See the Introduction for more information on this model.
[2]See Section 1.3 for preliminary notation.

An allocation has two components: (1) an **assignment** of objects to the agents, $\bar{\sigma} \in \Omega^n$, such that $\bar{\sigma}_i = \bar{\sigma}_j$ implies $i = j$ (agents consume different objects), and (2) a list of amounts of the divisible good, $\bar{m} \in \mathbb{R}^n$, such that $\sum \bar{m}_i = M$, where M is interpreted as the total endowment of the divisible good. Let $\boldsymbol{A}$ denote the set of allocations. Arbitrary allocations are denoted $(\bar{\sigma}, \bar{m})$, (σ^a, m^a), (σ^b, m^b), etc.

Note that for the sake of generality, we allow negative consumption of the divisible good. However, it is important to note that none of the results in Sections 4.2 and 4.3 would be changed by requiring non-negative consumption of the divisible good, and the effect of such a requirement on one result of Section 4.4, though trivial, will be given. Denote the set of allocations giving each agent a non-negative amount of the divisible good by $\boldsymbol{A}_+ = \{(\bar{\sigma}, \bar{m}) \in A : \bar{m} \in \mathbb{R}^n_+\}$.

For notational convenience, we will assume that $|\Omega| \geq n$. To consider the case in which $|\Omega| < n$, we may create enough "null" objects to reverse the inequality, and all of the results in this paper continue to hold — a more formal discussion of this and more general situations is given in Section 4.3.4.[3]

Each agent has a preference relation on $\Omega \times \mathbb{R}$, R_i, that satisfies the following assumptions:

Strict Monotonicity: for all $a \in \Omega$ and all $m_i, m_i' \in \mathbb{R}$, $m_i > m_i'$ implies $(a, m_i)\ P_i\ (a, m_i')$.

Compensability: for all $a, b \in \Omega$ and all $m_i \in \mathbb{R}$, there exists $m_i' \in \mathbb{R}$ such that $(a, m_i)\ I_i\ (b, m_i')$.

Quasi-linearity: for all $a, b \in \Omega$ and all $m_i, m_i', x \in \mathbb{R}$, $(a, m_i)\ I_i\ (b, m_i')$ implies $(a, m_i + x)\ I_i\ (b, m_i' + x)$.

Let $\boldsymbol{\mathcal{Q}}$ denote the set of preferences satisfying these three conditions. An arbitrary subset of these preferences is denoted $\mathcal{R} \subseteq \mathcal{Q}$, and a **domain** is a set of profiles $\mathcal{R}^n$.[4] For a convenient graphical representation of an agent's consumption space and preferences, see Figure 4.1.

For an arbitrary domain $\mathcal{R}^n$, a solution $\varphi\colon \mathcal{R}^n \to A$ is defined by two functions, (1) $\sigma\colon \mathcal{R}^n \to \Omega^n$ and (2) $m\colon \mathcal{R}^n \to \mathbb{R}^n$, such that for all $R \in \mathcal{R}^n$, $(\sigma(R), m(R)) \in A$. At times it will be convenient to denote a solution by $\varphi(\cdot) \equiv (\sigma(\cdot), m(\cdot))$. In such a case, for any $i \in N$, let $\varphi_i(R) = (\sigma_i(R), m_i(R))$. The range of φ over a (sub)domain $\mathcal{D} \subseteq \mathcal{R}^n$ is denoted $\boldsymbol{\varphi(\mathcal{D})}$.

---

[3]See also Remark 4.5.

[4]Note that the notation in this chapter differs slightly from the previous chapter in that the symbol $\mathcal{R}$ now represents a set of preferences for *one* agent instead of representing a domain.

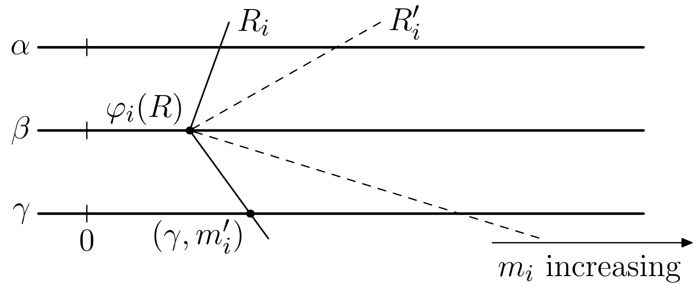


Figure 4.1: A consumption bundle is a point on one of the horizontal lines. The line determines the object; horizontal distance measures the amount of the divisible good. An "indifference curve" connects bundles between which a preference relation expresses indifference, *e.g.*, $\varphi_i(R)\ I_i\ (\gamma, m'_i)$ and $\varphi_i(R)\ P'_i\ (\gamma, m'_i)$. Here, $R'_i \in SMT(R_i, \varphi_i(R))$.

4.2 Two Agents, Two Objects

We begin with the simple case where $n = |\Omega| = 2$, which contains some intuition for the more general cases. Note for this case that if preferences change and one agent's bundle stays constant, then the other agent's bundle must also remain constant by feasibility. We can use this fact with Lemma 1.1 to show that in the particular 2-agent, 2-object case, *strategy-proofness* restricts the range of a solution to contain at most two allocations.

Lemma 4.1 *Suppose $n = |\Omega| = 2$. Let $\varphi = (\sigma, m)$ be a* strategy-proof *solution on any domain $\mathcal{R}^n$, and let $R, R' \in \mathcal{R}^n$. Then $\sigma(R) = \sigma(R')$ implies $m(R) = m(R')$. Hence $\varphi(\mathcal{R}^n)$ contains at most two allocations.*

Proof: Let $R, R' \in \mathcal{R}^n$ and let $\varphi = (\sigma, m)$ be *strategy-proof*. Suppose $\sigma(R) = \sigma(R')$.

Since preferences are quasi-linear and $|\Omega| = 2$, there are three possibilities for all $i \in N$:

$$\begin{aligned} &R_i \in SMT(R'_i, \varphi_i(R')) \qquad (4.1)\\ &R'_i \in SMT(R_i, \varphi_i(R))\\ &R_i = R'_i \end{aligned}$$

If (4.1) holds, then let $\tilde{R}_i = R_i$, otherwise let $\tilde{R}_i = R'_i$.

By Lemma 1.1, we have $\varphi(R) = \varphi(\tilde{R}_1, R_2) = \varphi(\tilde{R})$ and $\varphi(R') = \varphi(\tilde{R}_1, R'_2) = \varphi(\tilde{R})$. Hence $m(R) = m(R')$. ■

Lemma 4.1 makes it easy to describe *all* of the *strategy-proof* solutions for the simple case of two agents and two objects. The only solutions that are *strategy-proof* are (1) constant solutions, (2) solutions that let one agent choose his favorite of two pre-specified allocations, and (3) solutions that choose one allocation only if both agents prefer it to a second pre-specified allocation. Formally, we show that if $\varphi = (\sigma, m)$ is a *strategy-proof* solution, then it is one of these types:

Constant: $\sigma(\cdot)$ and $m(\cdot)$ are constant.

Dictatorial on *k* allocations: there exists $i \in N$ such that for all $R \in \mathcal{R}^n$ and all $(\bar{\sigma}, \bar{m}) \in \varphi(\mathcal{R}^n)$, we have $\varphi_i(R)\ R_i\ (\bar{\sigma}_i, \bar{m}_i)$; furthermore, $|\varphi(R^n)| = k$.

Status-quo-preserving: there exist $(\bar{\sigma}, \bar{m}), (\bar{\sigma}', \bar{m}') \in A$ such that $\varphi(\mathcal{R}^n) = \{(\bar{\sigma}, \bar{m}), (\bar{\sigma}', \bar{m}')\}$, and for all $i \in N$ and all $R_i \in \mathcal{R}$, if $(\bar{\sigma}_i, \bar{m}_i)\ P_i\ (\bar{\sigma}'_i, \bar{m}'_i)$, then $\varphi_i(R) = (\bar{\sigma}, \bar{m})$.

The following minor technical condition for 2-agent domains will complete our characterization. Essentially, it is *strategy-proofness* applied only to one agent when the other agent is indifferent between the allocations in the range of the solution.

Tie-breaking condition: For all $R \in \mathcal{R}^2$, all $i, j \in N$ with $i \neq j$, if for all $(\bar{\sigma}, \bar{m}), (\bar{\sigma}', \bar{m}') \in \varphi(\mathcal{R}^n)$, we have $(\bar{\sigma}_j, \bar{m}_j)\ I_j\ (\bar{\sigma}'_j, \bar{m}'_j)$, then $\varphi_i(R)\ R_i\ \varphi_i(R'_i, R_j)$ for all $R'_i \in \mathcal{R}$,

Now we have our characterization.

Theorem 4.1 *Suppose $n = |\Omega| = 2$, and let φ be a solution for any domain $\mathcal{R}^2$. Then the following are equivalent:*

(a) φ is strategy-proof*;*

(b) φ is either constant, dictatorial on two allocations, or status-quo-preserving; furthermore φ satisfies the tie-breaking condition.

Proof: We show that *(a)* implies *(b)*. Let φ be a *strategy-proof* solution on $\mathcal{R}^2$ that is not constant. Lemma 4.1 implies that $\varphi(\mathcal{R}^n)$ contains two allocations, $(\sigma^a, m^a), (\sigma^b, m^b) \in A$, such that $\sigma^a \neq \sigma^b$. For all $i \in N = \{1, 2\}$, partition $\mathcal{R}$ as follows:

$$\mathcal{R}_i^a \equiv \{R_i \in \mathcal{R} : (\sigma_i^a, m_i^a)\ P_i\ (\sigma_i^b, m_i^b)\} \tag{4.2}$$

$$\mathcal{R}_i^b \equiv \{R_i \in \mathcal{R} : (\sigma_i^b, m_i^b)\ P_i\ (\sigma_i^a, m_i^a)\} \tag{4.3}$$

$$\{\overline{R}_i\} \equiv \{R_i \in \mathcal{R} : (\sigma_i^a, m_i^a)\ I_i\ (\sigma_i^b, m_i^b)\}$$

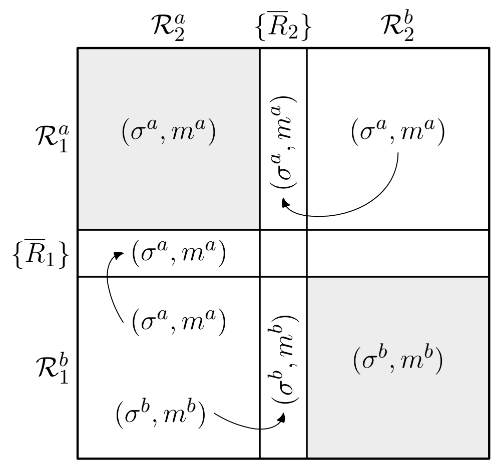


Figure 4.2: (Proof of Theorem 4.1) The arrows show directions of implication, *e.g.*, if $\varphi(\mathcal{R}_1^b \times \mathcal{R}_2^a) = \{(\sigma^a, m^a)\}$, then $\varphi(\{\overline{R}_1\} \times \mathcal{R}_2^a) = \{(\sigma^a, m^a)\}$, implying that φ is status-quo-preserving.

Some of these sets may be empty, depending on the domain $\mathcal{R}$. It is without loss of generality that we assume not.[5]

We first show that φ is constant on any cross-product of the sets defined in (4.2) and (4.3). Let $R \in \mathcal{R}_1^a \times \mathcal{R}_2^a$. Let $R' \in \mathcal{R}^2$ be such that $\varphi(R') = (\sigma^a, m^a)$. *Strategy-proofness* implies that $\varphi_1(R)\ R_1\ \varphi_1(R'_1, R_2)$ and $\varphi_2(R'_1, R_2)\ R_2\ \varphi(R')$. Hence $\varphi(R) = (\sigma^a, m^a)$. Similarly, for all $R \in \mathcal{R}_1^b \times \mathcal{R}_2^b$, we have $\varphi(R) = (\sigma^b, m^b)$. See the shaded regions of Figure 4.2.

We now show that φ is also constant on the subdomain $\mathcal{R}_1^a \times \mathcal{R}_2^b$. Let $R, R' \in \mathcal{R}_1^a \times \mathcal{R}_2^b$ and suppose without loss of generality that $\varphi(R) = (\sigma^a, m^a)$. *Strategy-proofness* implies $\varphi_1(R'_1, R_2)\ R'_1\ \varphi_1(R)$, therefore $\varphi(R'_1, R_2) = (\sigma^a, m^a)$. *Strategy-proofness* also implies $\varphi_2(R'_1, R_2)\ R_2\ \varphi_2(R')$, therefore $\varphi(R') = (\sigma^a, m^a)$. Similarly, φ is constant on $\mathcal{R}_1^b \times \mathcal{R}_2^a$.

Assuming for the remainder that $\varphi(\mathcal{R}_1^a \times \mathcal{R}_2^b) = \{(\sigma^a, m^a)\}$, we now handle the cases when one agent's preference relation is $\overline{R}_i$.

Let $R \in \mathcal{R}_1^a \times \mathcal{R}_2^b$. *Strategy-proofness* implies that $\varphi_2(R)\ R_2\ \varphi(R_1, \overline{R}_2)$. Therefore $\varphi(\mathcal{R}_1^a \times \mathcal{R}) = \{(\sigma^a, m^a)\}$. What we have just shown is demonstrated by the arrow in the upper-right corner of Figure 4.2.

Similarly we can show that if $\varphi(\mathcal{R}_1^b \times \mathcal{R}_2^a) = \{(\sigma^b, m^b)\}$, then $\varphi(\mathcal{R} \times \mathcal{R}_2^a) = \{(\sigma^b, m^b)\}$, implying that φ is dictatorial (see Figure 4.2).

If, instead, $\varphi(\mathcal{R}_1^b \times \mathcal{R}_2^a) = \{(\sigma^a, m^a)\}$, then *strategy-proofness* implies that for all $R'_2 \in \mathcal{R}_2^a$, we have $\varphi(\overline{R}_1, R'_2) = (\sigma^a, m^a)$, in which case φ is status-quo-

[5]This becomes apparent after one reads the proof.

preserving.

It is obvious that if φ does not satisfy the tie-breaking condition, then it is not *strategy-proof*.

Showing that *(b)* implies *(a)* is left to the reader. ■

4.3 Many agents

A characterization like the one of Theorem 4.1 does not hold for the case of more than two agents. As is the case for many models with private goods, there exist *strategy-proof* solutions that give one agent a fixed bundle, and vary the other agents' bundles arbitrarily according only to the fixed agent's preferences.[6] In fact, we can give an even more interesting example.

Example 4.1 Let $N = \{1, 2, 3, 4\}$, $\Omega = \{\alpha, \beta, \gamma, \delta\}$, and $M = 0$. Let $\mathcal{R}^n$ be an arbitrary domain. For all $i \in N$, let $f^i\colon \mathcal{R} \to \mathbb{R}$ be an arbitrary function. Let $g^1\colon \mathcal{R} \to \{\gamma, \delta\}$ and $g^3\colon \mathcal{R} \to \{\alpha, \beta\}$ also be chosen arbitrarily. Define the solution φ to satisfy for all $R \in \mathcal{R}^n$,

$$\begin{aligned}
\varphi_1(R) &= (\qquad g^3(R_3) \qquad, f^2(R_2) - f^3(R_3))\\
\varphi_2(R) &= (\{\alpha, \beta\} \setminus \{g^3(R_3)\}, f^3(R_3) - f^4(R_4))\\
\varphi_3(R) &= (\qquad g^1(R_1) \qquad, f^4(R_4) - f^1(R_1))\\
\varphi_4(R) &= (\{\gamma, \delta\} \setminus \{g^1(R_1)\}, f^1(R_1) - f^2(R_2))
\end{aligned}$$

Such a solution is obviously *strategy-proof* since no agent can ever vary his own bundle. However it is clear that solutions like these use the information of agents' preferences in an extremely arbitrary way.

4.3.1 Non-bossiness

To be able to describe the more interesting *strategy-proof* solutions, it will be helpful to introduce a condition eliminating the solutions as in Example 4.1. We begin this avenue of search by using the concept of *non-bossiness*, introduced in Satterthwaite and Sonnenschein (1981).

***Non-bossiness*:** for all $R \in \mathcal{R}^n$, all $i \in N$, and all $R'_i \in \mathcal{R}$, if $\varphi_i(R'_i, R_{-i}) = \varphi_i(R)$ then $\varphi(R'_i, R_{-i}) = \varphi(R)$.

[6] See, for example, Satterthwaite and Sonnenschein (1981).

The first question that arises is whether or not the addition of *non-bossiness* narrows the class of *strategy-proof* solutions as severely as *strategy-proofness* did by itself for the case of two agents. The answer is yes, as the next result shows by generalizing Lemma 4.1.

Theorem 4.2 *Let $\varphi = (\sigma, m)$ be a* strategy-proof *and* non-bossy *solution on the domain $\mathcal{Q}^n$, and let $R, R' \in \mathcal{R}^n$. Then $\sigma(R) = \sigma(R')$ implies $m(R) = m(R')$. Hence $\varphi(\mathcal{Q}^n)$ contains at most $n!/(n-k)!$ allocations, where $k = |\Omega|$.*

Proof: Let $R, R' \in \mathcal{Q}^n$ and let φ be as above. Suppose $\sigma(R) = \sigma(R')$.

Note that for all $i \in N$, there exists $R''_i \in SMT(R_i, \varphi_i(R)) \cap SMT(R'_i, \varphi_i(R'))$. By Lemma 1.1, $\varphi_1(R''_1, R_{-1}) = \varphi_1(R)$. Hence *non-bossiness* implies $\varphi(R''_1, R_{-1}) = \varphi(R)$. Similarly, $\varphi(R''_1, R''_2, R_3, \ldots, R_n) = \varphi(R)$. Repeating this argument for each agent, $\varphi(R'') = \varphi(R)$. Similarly, $\varphi(R'') = \varphi(R')$, so $m(R) = m(R')$. ■

It is intuitive that if a solution is *strategy-proof* and *non-bossy*, then on the subdomain of preferences for which agents have strict preferences over the (bundles in the) allocations in the range of the solution, the solution should only use information concerning the agents' preferences over the (bundles in the) allocations in the range, and no other preference information. This is a type of independence condition, which suggests the usual notion that *strategy-proof* solutions are never "too sensitive" to preference information.

Lemma 4.2 *Let φ be a* strategy-proof *and* non-bossy *solution on $\mathcal{Q}^n$, and denote its range by $\varphi(\mathcal{Q}^n) \subsetneq A$. Let $R, R' \in \mathcal{Q}^n$ be such that for all distinct $(\bar{\sigma}^a, \bar{m}^a), (\bar{\sigma}^b, \bar{m}^b) \in \varphi(\mathcal{Q}^n)$ and all $i \in N$, the following two conditions hold:*

$$(\bar{\sigma}^a_i, \bar{m}^a_i)\ P_i\ (\bar{\sigma}^b_i, \bar{m}^b_i) \Leftrightarrow (\bar{\sigma}^a_i, \bar{m}^a_i)\ P'_i\ (\bar{\sigma}^b_i, \bar{m}^b_i)$$
$$(\bar{\sigma}^a_i, \bar{m}^a_i)\ P_i\ (\bar{\sigma}^b_i, \bar{m}^b_i) \text{ or } (\bar{\sigma}^b_i, \bar{m}^b_i)\ P_i\ (\bar{\sigma}^a_i, \bar{m}^a_i)$$

Then $\varphi(R) = \varphi(R')$.

Proof: Let φ, R and R' be as above. By *strategy-proofness* $\varphi_1(R'_1, R_{-1})\ R'_1\ \varphi_1(R)$ and $\varphi_1(R)\ R_1\ \varphi_1(R'_1, R_{-1})$. By our assumptions on R and R', we have $\varphi(R'_1, R_{-1}) = \varphi(R)$.

Similarly for $j = 2, 3, \ldots, n$,

$$\varphi(R'_1, \ldots, R'_j, R_{j+1}, \ldots, R_n) = \varphi(R'_1, \ldots, R'_{j-1}, R_j, \ldots, R_n) = \cdots = \varphi(R)$$

Hence $\varphi(R') = \varphi(R)$. ■

This result brings us back around to the Gibbard-Satterthwaite Theorem[7] in the following way. Suppose the range of a *strategy-proof* and *non-bossy* solution contains between three and n allocations, and no agent receives the same object at any two of those allocations. These allocations represent the alternatives of the Gibbard-Satterthwaite domain. Since all quasi-linear preferences are admissible and no agent receives the same object at two of those allocations, no preference ordering over the *allocations* is ruled out *a priori*. Lemma 4.2 implies that essentially, only the agents preferences over the allocations matter, so these preferences correspond to those of the Gibbard-Satterthwaite domain. Hence the solution must allow one agent to always choose his best allocation, *i.e.* it must be a dictatorial solution.

Proposition 4.1 *Let* $\varphi = (\sigma, m)$ *be a* strategy-proof *and* non-bossy *solution on* $\mathcal{Q}^n$*. Suppose the range of* φ *contains at least three allocations. In addition, suppose that for all* $i \in N$ *and all* $R, R' \in \mathcal{Q}^n$ *we have* $\sigma_i(R) = \sigma_i(R')$ *implies* $\varphi(R) = \varphi(R')$*. Then* φ *is dictatorial.*

Sketch of Proof: For all $i \in N$, and all $R, R' \in \mathcal{Q}^n$, define a partial *preference ordering over allocations*, $\succ_{R_i}$, as follows: $\varphi(R) \succ_{R_i} \varphi(R')$ if and only if $\varphi_i(R)\ R_i\ \varphi_i(R')$. Note by the hypotheses of the Proposition that for all $i \in N$, the set $\mathcal{P}_i \equiv \bigcup_{R_i \in \mathcal{Q}} \succ_{R_i}$ is the set of all possible preference orderings over the allocations in the range of φ.

Define the function $f\colon \mathcal{P}_1 \times \cdots \times \mathcal{P}_n \to A$ as follows: $f(\succ_{R_1}, \ldots, \succ_{R_n}) = (\bar{\sigma}, \bar{m})$ if and only if $\varphi(R_1, \ldots, R_n) = (\bar{\sigma}, \bar{m})$. It is simple to check that since φ is *strategy-proof*, f is a *strategy-proof* decision rule in the Gibbard-Satterthwaite framework (see also Lemma 4.2). By our hypothesis of there being at least three allocations in the (finite) range of φ, f must be dictatorial, at least on the subdomain of preferences that are strict over those allocations. It then follows that φ is dictatorial on that subdomain, and it can be shown that φ is in fact dictatorial (on the whole domain). ■

If, on the other hand, two different allocations in the range of a *strategy-proof* and *non-bossy* solution give an agent the same object, we can not derive unrestricted preference orderings over the allocations. Hence we can not just invoke the Gibbard-Satterthwaite result. We can show, however, that no *strategy-proof* and *non-bossy* solution always treats agents with the same preferences equally.

***Equal treatment of equals*:** for all $R \in \mathcal{R}^n$ and all $i, j \in N$, if $R_i = R_j$ then we have $\varphi_i(R)\ I_i\ \varphi_j(R)$.

[7]See Gibbard (1973), Satterthwaite (1975).

Proposition 4.2 *If φ is a* strategy-proof *and* non-bossy *solution on $\mathcal{Q}^n$, then φ does not satisfy* equal treatment of equals.

Proof: Suppose by contradiction that φ does satisfy the three properties. For all $\epsilon \in \mathbb{R}$ and all $i \in N$, let $R_i^\epsilon \in \mathcal{Q}^n$ satisfy for all $m \in \mathbb{R}$ and all $a \in \Omega \setminus \{\alpha\}$,

$$(\alpha, m)\ I_i^\epsilon\ (a, m + \epsilon) \tag{4.4}$$

Equal treatment of equals implies that for all $i, j \in N$ and all $\epsilon \in \mathbb{R}$,

$$\varphi_i(R_1^\epsilon, \ldots, R_n^\epsilon)\ I_i^\epsilon\ \varphi_j(R^\epsilon, \ldots, R_n^\epsilon) \tag{4.5}$$

Let $(\epsilon_k)_{k=1}^{n!+1}$ be a strictly increasing list of numbers. By equations (4.4) and (4.5), the list $(\varphi(R^{\epsilon_k}))_{k=1}^{n!+1}$ consists of $n!+1$ different allocations, which contradicts Theorem 4.2. ■

Remark 4.1 Some authors consider a technically weaker version of *equal treatment of equals* for which a solution must treat agents equally only when *all* agents have the same preferences. Note that the impossibility in Proposition 4.2 still holds if we so weaken our condition.

Finally, on a more positive note, one method for constructing a non-trivial *strategy-proof* and *non-bossy* solution is to transform the problem into one resembling the housing market of Shapley and Scarf (1974). As we mentioned earlier, Roth (1982) constructs ways to choose core allocations of the housing market (with respect to an endowment) that are *strategy-proof*. We can use that result here by fixing one allocation as our "endowment", and considering the bundles of that allocation (an indivisible object plus an amount of money) to be the "houses". A solution choosing "core" allocations of this resulting housing market is *strategy-proof* as Roth showed. It can also be shown to be *non-bossy*. Of course it should be noted that the resulting allocation is not any type of core allocation for our model due to the presence of the (transferable) divisible good, which is absent in the housing market model. However this method does define a somewhat rich *strategy-proof* solution.

4.3.2 Coalitional Strategy-proofness

The property of *non-bossiness* is a controversial one. It has what is in some circumstances a desirable normative implication in that an agent cannot influence the other agents' consumption without affecting the physical bundle he receives (though his well-being may be unaffected.) On the other hand, the following requirement makes the condition look more technical and less prescriptive: when a change in one agent's preferences does not affect his bundle, the other agents' bundles must not change, *even if such a change would keep the other agents' welfare levels constant.*

To loosen this technical condition, but still maintain the flavor of the requirement that an agent should not be able to inflict a loss (or gain) for any other agent(s) at no cost to himself, we will consider instead the condition of *coalitional strategy-proofness.*

***Coalitional strategy-proofness*:** for all $R \in \mathcal{R}^n$, and all $C \subseteq N$, there exists no $R'_C \in \mathcal{R}^{|C|}$ such that for all $i \in C$, $\varphi_i(R'_C, R_{-C})\ R_i\ \varphi_i(R)$, and for some $j \in C$, $\varphi_j(R'_C, R_{-C})\ P_i\ \varphi_j(R)$.

It turns out that the conclusion of Theorem 4.2 remains valid after we make this change.

Theorem 4.3 *Let* $\varphi = (\sigma, m)$ *be a* coalitionally strategy-proof *solution on the domain* $\mathcal{Q}^n$*, and let* $R, R' \in \mathcal{Q}^n$*. Then* $\sigma(R) = \sigma(R')$ *implies* $m(R) = m(R')$*. Hence* $\varphi(\mathcal{Q}^n)$ *contains at most* $n!/(n-k)!$ *allocations, where* $k = |\Omega|$*.*

Proof: Let φ satisfy *coalitional strategy-proofness*. The bulk of the proof involves showing the following claim to be true for any coalition $C \subseteq N$.

Claim: For all $R \in \mathcal{Q}^n$, and all $R'_C \in \mathcal{Q}^{|C|}$, if for all $i \in C$, we have $R'_i \in SMT(R_i, \varphi_i(R))$, then $\varphi_C(R'_C, R_{-C}) = \varphi_C(R)$.

The proof of this claim is by induction on $|C|$. If C is a singleton, the claim is proven by Lemma 1.1. Let $C \subseteq N$, and suppose that the claim is true for any coalition $B \subset N$ such that $|B| < |C|$. Let R and R'_C be given as above. We must show that for all $i \in C$, $\varphi_i(R'_C, R_{-C}) = \varphi_i(R)$.

Let a coalition $B \subset N$ be such that $|B| < |C|$. By our induction hypothesis, $\varphi_B(R'_B, R_{-B}) = \varphi_B(R)$. Therefore, for all $i \notin B$, *coalitional strategy-proofness* implies $\varphi_i(R'_B, R_{-B})\ R_i\ \varphi_i(R)$ (via the coalition $B \cup \{i\}$). Similarly, for all $i \notin B$, we have $\varphi_i(R)\ R_i\ \varphi_i(R'_B, R_{-B})$. Hence, for any such coalition B,

$$\text{for all } i \in N,\ \varphi_i(R)\ I_i\ \varphi_i(R'_B, R_{-B}) \tag{4.6}$$

Let $i \in C$. By *strategy-proofness*, $\varphi_i(R_i, R'_{C\setminus i}, R_{-C})\ R_i\ \varphi_i(R'_C, R_{-C})$. Letting $C \setminus i$ play the role of B in (4.6), this implies the following (see Figure 4.3):

$$\varphi_i(R)\ R_i\ \varphi_i(R'_C, R_{-C}), \text{ implying} \tag{4.7}$$

$$\varphi_i(R)\ R'_i\ \varphi_i(R'_C, R_{-C}) \tag{4.8}$$

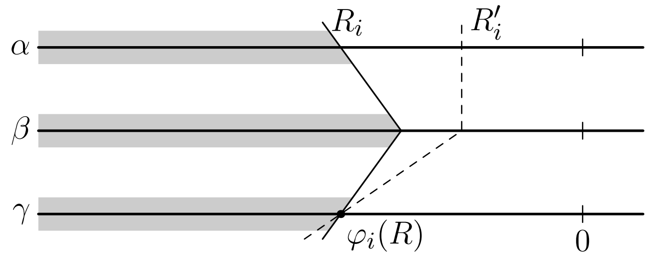


Figure 4.3: (Proof of Theorem 4.3) By equation (4.7), $\varphi_i(R'_C, R_{-C})$ must lie in one of the shaded areas.

Since (4.8) holds for all members of C, *coalitional strategy-proofness* implies that there exists no $j \in C$ such that $\varphi_j(R)\ P'_j\ \varphi_j(R'_C, R_{-C})$, *i.e.* for all $j \in C$, we have $\varphi_j(R'_C, R_{-C})\ R'_j\ \varphi_j(R)$. With (4.7) and the definition of R'_C, this implies $\varphi_C(R'_C, R_{-C}) = \varphi_C(R)$, proving the claim.

To finish the proof of the theorem, let $R, R' \in \mathcal{Q}^n$ and suppose $\sigma(R) = \sigma(R')$. Then for each $i \in N$, there exists $R''_i \in SMT(R_i, \varphi_i(R)) \cap SMT(R'_i, \varphi_i(R'))$. Two applications of the claim to the coalition $C = N$ imply both $\varphi(R) = \varphi(R'')$ and $\varphi(R') = \varphi(R'')$. Hence $m(R) = m(R')$. ■

Remark 4.2 Note that this result holds on any domain, $\mathcal{R}$, of quasi-linear preferences that satisfies the following richness condition: for any allocation $(\bar{\sigma}, \bar{m}) \in A$, any agent $i \in N$, and any two preference relations, $R_i, R'_i \in \mathcal{R}$, there exists $R''_i \in \mathcal{R}$ such that $R''_i \in SMT(R_i, (\bar{\sigma}_i, \bar{m}_i)) \cap SMT(R'_i, (\bar{\sigma}_i, \bar{m}_i))$. An interesting such domain (see Tadenuma and Thomson, 1991) is one for which there exists a fixed ordering on the set of objects such that for any admissible preference relation, objects that come earlier in the order are "better" than those that come later in the order, in the following sense: two bundles are considered indifferent only if there is a greater amount of the divisible good with the object that comes later in the order.

Remark 4.3 With the same reasoning we used with Proposition 4.1, we can see that many *coalitionally strategy-proof* rules with "very small" ranges are dictatorial.

Finally, note that with more than two agents, even dictatorial rules need not satisfy our strong version of *coalitional strategy-proofness*. To see this, consider a solution choosing among two allocations, depending on which of the two makes agent 1 better off. Suppose preferences are such that agent 1 is indifferent between the bundles he receives at the two allocations, agent 2 strictly prefers the bundle he receives from the first allocation, and agent 3 strictly prefers the bundle he receives from the second allocation. Depending on which allocation is chosen at that profile, agent 1 can either make agent 2 better off, or make agent 3 better off, at no loss to himself.

This problem, arising when agents may not care which allocation in the range of a solution is chosen, also has implications for the adaptation of Roth's method discussed in Section 4.3.1. However, non-trivial examples of *coalitionally strategy-proof* solutions exist. Consider a status-quo-preserving solution that chooses one of two allocations only when all agents strictly prefer it to the "status-quo" allocation. This solution is *coalitionally strategy-proof*. More generally, consider dividing the agents, objects, and divisible good into two "sub-economies", and operating two different such status-quo-preserving solutions on each. This also defines a *coalitionally strategy-proof* solution for the original economy.

4.3.3 Bribe-proofness

While *coalitional strategy-proofness* is a desirable property, it may be unreasonable in some situations to suspect that large coalitions can coordinate enough to jointly misrepresent their preferences in order to gain.

It may also be unreasonable to suspect that an agent will misrepresent his preferences in order for another agent to gain, at no gain to himself. This second criticism may be countered by the argument that an agent that *does* gain with his misrepresentation could offer a small "bribe" to the agent who neither gains nor loses from his his misrepresentation. If this is the case, however, it is not unreasonable to suspect that one agent may bribe another agent to misrepresent his preferences even in situations where the bribed agent would be strictly worse off without the bribe, but better off with it.

We address these two points simultaneously by introducing the condition of *bribe-proofness*. A solution is *bribe-proof* if it is never the case that one agent can bribe another to misrepresent his preferences, making both agents strictly better off. Formally, letting $\varphi = (\sigma, m)$ be a solution,

Bribe-proofness: there exist no $R \in \mathcal{R}^n$, $i, j \in N$, R'_i, and $\tau \in \mathbb{R}_+$ such that

$$(\sigma_i(R'_i, R_{-i}), m(R'_i, R_{-i}) + \tau) \ P_i \ \varphi_i(R), \text{ and}$$
$$(\sigma_j(R'_i, R_{-i}), m(R'_i, R_{-i}) - \tau) \ P_j \ \varphi_j(R).$$

If the definition were violated, agent j would bribe agent i to misrepresent his preferences with τ units of the divisible good. Note that *bribe-proofness* implies *strategy-proofness* (let $\tau = 0$ and $i = j$).

It may appear at first that this condition neither implies nor is implied by *coalitional strategy-proofness*. It turns out, however, that *bribe-proofness* is an extremely strong property; the only solutions to satisfy it are essentially constant solutions. The proof is completed by two lemmata. First, it will be convenient to introduce some extra notation as in Barberà and Peleg (1990).

Agent i's **option set**, or attainable set, is defined as follows: for all $R_{-i} \in \mathcal{R}^{n-1}$,

$$O_i(R_{-i}) = \{(a, m_i) \in \Omega \times \mathbb{R} : \exists R_i \in \mathcal{R} \text{ such that } \varphi_i(R_i, R_{-i}) = (a, m_i)\}$$

The first lemma already shows the strength of *bribe-proofness*. It says that an agent's well-being cannot be affected when other agents change their preferences.

Lemma 4.3 *Let $\varphi = (\sigma, m)$ be a* bribe-proof *solution on $\mathcal{Q}^n$. For all $i \in N$, all $R_i \in \mathcal{R}$, and all $R_{-i}, R'_{-i} \in \mathcal{R}^{n-1}$, we have $\varphi_i(R_i, R_{-i})\ I_i\ \varphi_i(R_i, R'_{-i})$.*

Proof: Let φ be *bribe-proof* and let $R_j, R'_j \in \mathcal{Q}$ and $i, j \in N$ be such that $j \neq i$. It will suffice to show that $\varphi_i(R_j, R_{-j})\ I_i\ \varphi_i(R'_j, R_{-j})$, because repeated application of this statement proves the lemma.

Suppose this is not the case, that is, suppose without loss of generality that $\varphi_i(R_j, R_{-j})\ P_i\ \varphi_i(R'_j, R_{-j})$. Let $\delta > 0$ be such that

$$(\sigma_i(R_j, R_{-j}), m(R_j, R_{-j}) - 3\delta)\ I_i\ \varphi_i(R'_j, R_{-j}) \tag{4.9}$$

That is, when R_j changes to R'_j, agent i is made "3δ worse off." *Bribe-proofness* implies $\varphi_j(R'_j, R_{-j})\ P'_j\ \varphi_j(R_j, R_{-j})$. Let $(a, m_j) = \varphi_j(R_j, R_{-j})$ and $(a', m'_j) = \varphi_j(R'_j, R_{-j})$.

Note that *strategy-proofness* implies that for all $(b, \tilde{m}_j) \in O_j(R_{-j})$, if $b = a$ then $\tilde{m}_j = m_j$. Hence $a \neq a'$. Therefore, there exists $\overline{R}_j \in \mathcal{R}$ such that $\arg\max_{O_j(R_{-j})} \overline{R}_j = \{(a', m'_j)\}$ and $(a, m_j + \delta)\ \overline{I}_j\ (a', m'_j)$ (see Figure 4.4).

Strategy-proofness clearly implies $\varphi_j(\overline{R}_j, R_{-j}) = (a', m'_j)$. Therefore *bribe-proofness* implies $\varphi_i(\overline{R}_j, R_{-j})\ R_i\ \varphi_i(R'_j, R_{-j})$ and $\varphi_i(R'_j, R_{-j})\ R_i\ \varphi_i(\overline{R}_j, R_{-j})$. Hence $\varphi_i(\overline{R}_j, R_{-j})\ I_i\ \varphi_i(R'_j, R_{-j})$. With (4.9), this implies

$$\varphi_i(\overline{R}_j, R_{-j})\ I_i\ (\sigma_i(R_j, R_{-j}), m_i(R_j, R_{-j}) - 3\delta)$$

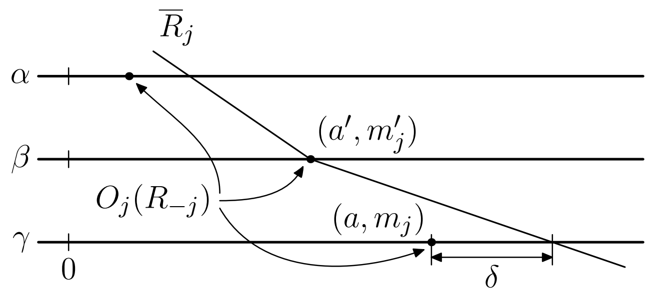


Figure 4.4: (Proof of Lemma 4.3) The construction of $\overline{R}_j$.

But then

$$(\sigma_j(R_j, R_{-j}), m_j(R_j, R_{-j}) + 2\delta)\ \overline{P}_j\ \varphi_j(\overline{R}_j, R_{-j}), \text{ and}$$
$$(\sigma_i(R_j, R_{-j}), m_i(R_j, R_{-j}) - 2\delta)\ P_i\ \varphi_i(\overline{R}_j, R_{-j})$$

violating *bribe-proofness*. ■

The next theorem follows intuitively from Lemma 4.3.

Theorem 4.4 *Let φ be a* bribe-proof *solution on $\mathcal{Q}^n$. Then for all $i \in N$, $O_i(\cdot)$ is constant.*

Proof: Let $R_{-i}, R'_{-i} \in \mathcal{R}^{n-1}$ and $(a, m_i) \in O_i(R_{-i})$. We must show that $(a, m_i) \in O_i(R'_{-i})$.

If there exists $m'_i \in \mathbb{R}$ such that $(a, m'_i) \in O_i(R'_{-i})$, assume without loss of generality that $m'_i \leq m_i$. Recall that if $(a, m'_i) \in O_i(R_{-i})$ then $m_i = m'_i$. Therefore there exists $R_i \in \mathcal{R}$ such that $\arg\max_{O_i(R_{-i}) \cup O_i(R'_{-i})} R_i = \{(a, m_i)\}$.

Strategy-proofness therefore implies $\varphi(R_i, R_{-i}) = (a, m_i)$. Lemma 4.3 then implies $\varphi(R_i, R'_{-i})\ I_i\ \varphi(R_i, R_{-i})$. Therefore by choice of R_i, $\varphi(R_i, R'_{-i}) = (a, m_i)$, hence $(a, m_i) \in O_i(R'_{-i})$. ■

Corollary 4.1 *If $\varphi = (\sigma, m)$ is a* bribe-proof *solution on $\mathcal{Q}^n$, then an agent always receives the same amount of the divisible good, and receives his favorite object from a pre-specified set. That is, m is constant, and for all $R, R' \in \mathcal{Q}^n$ and all $i, j \in N$, $\sigma_i(R) \neq \sigma_j(R')$. Hence if $|\Omega| = n$, φ is constant.*

Proof: Letting $\varphi = (\sigma, m)$ be *bribe-proof*, Theorem 4.4 implies m is constant. Suppose by contradiction that there exist $i, j \in N$, $a \in \Omega$, and $R, R' \in \mathcal{Q}^n$ such that $\sigma_i(R) = a = \sigma_j(R')$.

Let $\overline{R}_i$ and $\overline{R}_j$ be such that $\arg\max_{O_i} \overline{R}_i = \{\varphi_i(R)\}$ and $\arg\max_{O_j} \overline{R}_j = \{\varphi_j(R')\}$. By Theorem 4.4 and *strategy-proofness*, agents i and j receive object a, at the allocation $\varphi(\overline{R}_i, \overline{R}_j, R_{-\{i,j\}})$, which is a contradiction. ■

So if there are more objects than agents and a solution is *bribe-proof*, then the set of objects is partitioned into $n+1$ subsets, so that each agent receives his favorite object from his own subset in the partition, and objects in the $(n+1)^{\text{th}}$ subset are never assigned. When $|\Omega| = n$, this implies that a *bribe-proof* solution is constant.

4.3.4 Restricted Domains of Preferences

Consider a situation in which the set of indivisible objects to be assigned to the agents contains objects that are identical. For instance, some offices in a department may be so similar as to be indistinguishable in terms of preferences. Such a situation is described by our model above with one added requirement: for any domain of preferences, $\mathcal{R}$, and any preference relation $R_i \in \mathcal{R}$, if α and α' are to be considered identical objects, then for all $m_i, m_i' \in \mathbb{R}$, we must have $(\alpha, m_i)\ I_i\ (\alpha', m_i')$ if and only if $m_i = m_i'$. That is, we must impose a **restriction on the domain of preferences** so that *copies* of the same *type* of object are always considered equivalent in terms of preferences.

One application of this domain restriction is to the case in which there are fewer "real" objects than agents. Agents that do not receive a real object instead consume a "null" object, and all null objects are equivalent in terms of preferences.

A second application of this domain restriction is to the case in which a number of agents (but perhaps not all) may share an office, or other indivisible public good, as long as they are indifferent about sharing it (*i.e.* there are no externalities). If an office can hold no more than ℓ occupants, then ℓ copies of this office belong to the set of objects to be assigned.

Theorems 4.2, 4.3, and 4.4 hold on such domains of preferences. However, the following should be noted:

- The *non-bossiness* condition should require that if a change in an agent's preferences affects neither his consumption of the divisible good nor the *type* of object he receives, then it affects neither the consumption of the divisible good of another agent nor the *type* of object another agent consumes, *i.e.* in the example above, consuming α is the same as consuming α'.

- The wording of Theorems 4.2 and 4.3 should be slightly strengthened, stating that if at two allocations any agent receives the same *type* of object, then at the two allocations, each agent receives the same amount of the divisible good.

- The wording of Theorem 4.4 does not need to be adjusted, but Corollary 4.1 does not apply. However a similar corollary can be obtained.

Finally, note that to obtain proofs of these results, the definition of a strict monotonic transformation must be slightly modified. Given a domain $\mathcal{R}$, restricted as above, for any preference relation $R_i \in \mathcal{R}$ and any bundle (a, m_i), a preference relation $R'_i \in \mathcal{R}$ is a **strict monotonic transformation of $\boldsymbol{R_i}$ at $\boldsymbol{(a, m_i)}$ with respect to the domain $\boldsymbol{\mathcal{R}}$** when for any other bundle (a', m'_i), if both (1) $(a, m_i)\ R_i\ (a', m'_i)$ and (2) a and a' are *not the same type of object according to our domain restriction*, then $(a, m_i)\ P'_i\ (a', m'_i)$.

4.4 Efficiency

Much of the work on other economic models (*e.g.*, Hurwicz and Walker (1990), Schummer (1996,1997a), and Zhou (1991a,b)) has shown a general incompatibility between *strategy-proofness* and efficiency, *e.g.*, a solution satisfying those properties must allow a given agent to choose his favorite feasible allocation. This model provides no exception. There are two ways of defining efficiency in this model, depending on whether agents may consume negative quantities of the divisible good. The choice results either in the conclusion that no *strategy-proof* and efficient solution exists, or in the conclusion that only certain dictatorial solutions are *strategy-proof* and efficient — a trivial difference.

First we address the case in which agents' consumption of the divisible good is not bounded. The corresponding definition of efficiency is the usual one used in the context of quasi-linear preferences (*e.g.*, the literature on the Groves mechanism).

***Assignment-Efficiency*:** for all $R \in \mathcal{R}^n$, there exists no $(\bar{\sigma}, \bar{m}) \in A$ such that for all $i \in N$, $(\bar{\sigma}_i, \bar{m}_i)\ P_i\ \varphi_i(R)$.

One may check that this condition only requires that the surplus-maximizing assignment of objects be chosen, where surplus is measured in terms of the divisible good; the condition is independent of $m(\cdot)$, but only because we require $\sum m_i(R) = M$.

4.4.1 Two Agents, Two Objects

Beginning again with the 2-agent, 2-object case, we show that even on domains of extremely small size, *assignment-efficiency* is completely incompatible with

	R_2^w	R_2^x	R_2^y	R_2^z
R_1^w		(σ^b, m^b)	(σ^b, m^b)	(σ^b, m^b)
R_1^x	(σ^a, m^a)	$\varphi(R_1^x, R_2^x)$	(σ^b, m^b)	(σ^b, m^b)
R_1^y	(σ^a, m^a)	(σ^a, m^a)		(σ^b, m^b)
R_1^z	(σ^a, m^a)	(σ^a, m^a)	(σ^a, m^a)	

Figure 4.5: (Proof of Theorem 4.5) When agent 1 likes object α relatively more than agent 2 does, *efficiency* implies that φ gives α to agent 1.

strategy-proofness. Before showing that result, we introduce the following notation to be used only for the case of two agents and two objects. When $n = |\Omega| = 2$, for any number $x \in \mathbb{R}$, define the preference relation R^x as follows:

$$\text{for all } m_i \in \mathbb{R},\ (\alpha, m_i)\ I^x\ (\beta, m_i + x) \tag{4.10}$$

A higher value of x implies a relatively greater worth of the object β in comparison to α.

Theorem 4.5 *Suppose $n = |\Omega| = 2$. Further, suppose that $|\mathcal{R}| \geq 4$, that is, $\mathcal{R}$ contains at least four preference relations. Let φ be a* strategy-proof *solution on $\mathcal{R}^2$. Then φ violates* assignment-efficiency*.*

Proof: Let $\Omega = \{\alpha, \beta\}$. Let $\{R^w, R^x, R^y, R^z\} \subset \mathcal{R}$ be defined as in (4.10), where $w < x < y < z$. Hence an agent with the preference relation R^y cares more for good α (compared to good β) than does an agent with either the preference relation R^w or R^x.

Suppose by contradiction that there exists a *strategy-proof* and *assignment-efficient* solution, $\varphi = (\sigma, m)$. Note that *assignment-efficiency* implies both $\sigma_1(R_1^x, R_2^w)) = \alpha$ and $\sigma_1(R_1^w, R_2^x) = \beta$. Therefore denote $(\sigma^a, m^a) = \varphi(R_1^x, R_2^w))$ and $(\sigma^b, m^b) = \varphi(R_1^w, R_2^x)$.

Lemma 4.1 and *assignment-efficiency* imply (see Figure 4.5):

$$\varphi(R_1^x, R_2^w) = \varphi(R_1^y, R_2^w) = \varphi(R_1^z, R_2^w) = \varphi(R_1^y, R_2^x) = \varphi(R_1^z, R_2^x) = \varphi(R_1^z, R_2^y) = (\sigma^a, m^a)$$

$$\varphi(R_1^w, R_2^x) = \varphi(R_1^w, R_2^y) = \varphi(R_1^w, R_2^z) = \varphi(R_1^x, R_2^y) = \varphi(R_1^x, R_2^z) = \varphi(R_1^y, R_2^z) = (\sigma^b, m^b)$$

Therefore *strategy-proofness* implies that for all $i \in N$, $\varphi_i(R_1^x, R_2^x)\ R_i^x\ (\sigma_i^a, m_i^a)$ and $\varphi_i(R_1^x, R_2^x)\ R_i^x\ (\sigma_i^b, m_i^b)$. Note that both assignments are efficient for the profile (R_1^x, R_2^x). Therefore $(\sigma_i^a, m_i^a)\ I_i^x\ \varphi_i(R_1^x, R_2^x)\ I_i^x\ (\sigma_i^b, m_i^b)$.

Similarly, we can show that $(\sigma_i^a, m_i^a)\ I_i^y\ (\sigma_i^b, m_i^b)$. However this is a contradiction since $R_1^x \neq R_1^y$, and $\sigma_1^a \neq \sigma_1^b$. ■

Remark 4.4 A similar result holds for the case $|\Omega| > n = 2$. The proof uses the existence of four admissible preference relations and two objects such that: (1) the restrictions of those preferences to those two objects ($\times\mathbb{R}$) correspond to R^w through R^z in the proof of Theorem 4.5, and (2) for any profile made up of only those preference relations, *efficiency* implies that each agent receives one of those two objects.

Notice that in our model, there is no favorite feasible allocation for an agent because the consumption of the divisible good is unbounded. Therefore a solution which gives to one agent M units of the divisible good plus his favorite object is not *efficient*; if the other agent likes his favorite object relatively more than this dictator, both agents can be made better off by switching objects and making a transfer of money. It is this sort of reason that causes the impossibility of Theorem 4.5. However, the transfer in this example results in negative consumption of the divisible good for one of the agents, which may not be reasonable. At least, unbounded consumption (from below) is not reasonable. If we require non-negative consumption, though, it makes no sense to consider the transfer in our above example as a way for the agents to jointly improve upon the allocation prescribed by the dictatorial solution. More precisely, we should only require that an allocation (with non-negative consumption of the divisible good) not be Pareto-dominated by any allocation that *also* prescribes non-negative consumption of the divisible good, hence weakening our previous notion of *efficiency* in an appropriate way. Formally, consider the following conditions:

***Non-negativity*:** for all $R \in \mathcal{R}^n$, $\varphi(R) \in A_+$.

***Pareto-optimality*:** for all $R \in \mathcal{R}^n$, there exists no $(\bar{\sigma}, \bar{m}) \in A_+$, such that for all $i \in N$, $(\bar{\sigma}_i, \bar{m}_i)\ P_i\ \varphi_i(R)$.

While Theorem 4.5 tells us that none of the solutions characterized in Theorem 4.1 is *assignment-efficient*, it is simple to check that the particular dictatorial solution just described above satisfies both *non-negativity* and *Pareto-optimality*. It turns out that for the domain $\mathcal{Q}^2$, such dictatorial solutions are the *only* ones that are *strategy-proof*, *non-negative*, and *Pareto-optimal*. We again use the notation defined with (4.10).

Theorem 4.6 *Suppose $n = |\Omega| = 2$ and $M > 0$. Let φ be a solution on the domain $\mathcal{Q}^2$. Then the following are equivalent:*

(a) φ is strategy-proof, non-negative, *and* Pareto-optimal*;*

(b) φ is dictatorial on two allocations that give all of the divisible good to the dictator; furthermore φ satisfies the tie-breaking condition.

Proof: Let φ be a *strategy-proof*, *non-negative*, and *Pareto-optimal* solution. Clearly by *Pareto-optimality*, φ cannot be constant, so by Lemma 4.1, it is either dictatorial on two allocations or status-quo-preserving. Denote the two allocations in its range by (σ^a, m^a) and (σ^b, m^b), where $\sigma_1^a = \alpha$ and, therefore, $\sigma_1^b = \beta$.
Step 1: $m^a = m^b$.

Let $R_1^x, R_1^y \in \mathcal{Q}$ be such that $x > 0$ and $y < 0$ (see the notation of (4.10)). By *Pareto-optimality*, $\varphi(R_1^x, R_2^0) = (\sigma^a, m^a)$ and $\varphi(R_1^y, R_2^0) = (\sigma^b, m^b)$. By *strategy-proofness*, $(\sigma_1^a, m_1^a)\ R_1^x\ (\sigma_1^b, m_1^b)$ and $(\sigma_1^b, m_1^b)\ R_1^y\ (\sigma_1^a, m_1^a)$. Since x and y were arbitrary positive and negative numbers, respectively, it follows that $m_1^a = m_1^b$, hence $m^a = m^b$.
Step 2: Either $m_1^a = 0$ or $m_1^a = M$.

Let $\mathcal{R}_i^a$ and $\mathcal{R}_i^b$ be defined as in the proof of Theorem 4.1 and let $R \in \mathcal{R}_1^a \times \mathcal{R}_2^b$. Suppose that $\varphi(R) = (\sigma^a, m^a)$. Let $\epsilon = m_2^a$, and suppose $\epsilon > 0$. Since φ is either dictatorial or status-quo-preserving, $\varphi(R_1^{\epsilon/2}, R_2^{-2\epsilon}) = \varphi(R)$. However the allocation at which agent 1 receives the bundle (β, M) Pareto-dominates (σ^a, m^a), violating *Pareto-optimality*. Therefore $\epsilon = 0$, *i.e.* $m^a = m^b = (M, 0)$. Similarly, $\varphi(R) = (\sigma^b, m^b)$ implies $m^a = m^b = (0, M)$.

Let $R' \in \mathcal{R}_1^b \times \mathcal{R}_2^a$. As above, $\varphi(R') = (\sigma^b, m^b)$ implies $m^a = m^b = (M, 0)$, and $\varphi(R') = (\sigma^b, m^b)$ implies $m^a = m^b = (0, M)$.
Step 3: φ is dictatorial.

Since either $m_1^a = 0$ or $m_1^a = M$ but not both, it follows that either (1) $\varphi(R) = (\sigma^a, m^a)$ and $\varphi(R') = (\sigma^b, m^b)$, or (2) $\varphi(R) = (\sigma^b, m^b)$ and $\varphi(R') = (\sigma^a, m^a)$. Therefore, φ is dictatorial and the dictator always receives M units of the divisible good.

Strategy-proofness clearly implies the tie-breaking condition. Showing that *(b)* implies *(a)* is left to the reader. ■

4.4.2 Many Agents

The issue of efficiency for agents with quasi-linear preferences over *public* goods and a divisible good has been thoroughly addressed in the literature. For such a domain, the only *strategy-proof* solutions that satisfy *assignment-efficiency* (appropriately defined for that domain) are the Groves mechanisms[8] as shown by Green and Laffont (1977). However these mechanisms fail to always allocate the exact endowment of the divisible good.

While this particular result has no direct implication on our model, one may ask the question of whether such a characterization also holds here. If it does hold, the second question would concern whether any such mechanisms always completely allocate all of the divisible good ("budget balance"). If none do, we would conclude that no *strategy-proof* solutions satisfy *assignment-efficiency* (since our definition requires budget balance).

The first of these two questions is answered by a strong, clean result of Holmström (1979). He shows that the Green-Laffont characterization holds not only on the domain of quasi-linear preferences over a set of public decisions plus a divisible good, but that it holds on any convex subdomain of such preferences.[9] To give this result, we will need the following notation.

Fix an arbitrary assignment σ^v (*e.g.*, $\sigma_1^v = \alpha$, $\sigma_2^v = \beta$, $\sigma_3^v = \gamma$, etc.). For any agent $i \in N$ and preference relation $R_i \in \mathcal{Q}$, define the **valuation function for $\boldsymbol{R_i}$**, $v_i(\,\cdot\,; R_i)$, which maps assignments into real numbers, to satisfy for any assignment σ^a, we have $v_i(\sigma^a; R_i) = m$ if and only if $(\sigma_i^a, 0)\ I_i\ (\sigma_i^v, m)$. It is the "value" of the assignment σ^a to agent i with preferences R_i, normalized with respect to σ^v. This normalization shall remain fixed throughout.

One may check that *assignment-efficiency* is merely the requirement that for any profile $R \in \mathcal{Q}^n$, the chosen assignment maximizes $\sum v_i(\,\cdot\,, R_i)$.

Lemma 4.4 *Let $\varphi = (\sigma, m)$ be a* strategy-proof *and* assignment-efficient *solution defined on $\mathcal{Q}^n$. Then for all $i \in N$, there exists $h_i\colon \mathcal{Q}^{n-1} \to \mathbb{R}$ such that for all $R \in \mathcal{Q}^n$,*

$$m_i(R) = \sum_{j \neq i} v_j(\sigma(R); R_j) - h_i(R_{-i})$$

Proof: Follows directly from Holmström (1979). ■

As we mentioned above, solutions of this form do not always balance for the domain of quasi-linear preferences over *public goods*. Though that result implies nothing for our model, the same conclusion does in fact hold for our model.

Theorem 4.7 *Let φ be a* strategy-proof *solution defined on $\mathcal{Q}^n$. Then φ violates* assignment-efficiency.

[8]See Clarke (1971), Groves (1973), and Green and Laffont (1977).

[9]In fact his result is even stronger — the term *convex* can be replaced with the weaker term *connected*. See Holmström (1979) for a more precise statement.

Proof:[10] Suppose $\varphi = (\sigma, m)$ satisfies both properties. We will show that for some $R \in \mathcal{Q}^n$, $\sum m_i(R) \neq M$, violating our definition of a solution (balancedness). We will assume without loss of generality that $M = 0$. Clearly if no *strategy-proof* and *assignment-efficient* solution exists for this case, then none exists for the general case.

Recall the assignment σ^v, from the definition of the valuation functions. Let the assignment σ^a satisfy $\sigma^a_n = \sigma^v_1$, and for all $i \neq n$, $\sigma^a_i = \sigma^v_{i+1}$. Now define $R, R' \in \mathcal{Q}^n$ as follows:

$$\begin{array}{rl} \forall i \in N, \forall a \in \Omega \setminus \{\sigma^v_i\}, & (\sigma^v_i, 0)\ I_i\ (a, n) \\ \forall i \in N, \forall a \in \Omega \setminus \{\sigma^v_i, \sigma^a_i\}, & (\sigma^v_i, 0)\ I'_i\ (\sigma^a_i, -1)\ I'_i\ (a, n) \end{array}$$

Hence for all $i \in N$, we have $v_i(\sigma^v, R_i) = 0$ and $v_i(\sigma^a, R'_i) = 1$.

Finally, for any (possibly empty) coalition $C \subseteq N$, define $R^C_i \in \mathcal{Q}^n$ as follows: for all $i \in N$, if $i \in C$ then $R^C_i = R'_i$; $R^C_i = R_i$ otherwise. Starting from the profile R, C is the coalition of agents whose preferences have changed to R'_C. Therefore $R^{\emptyset} = R$ and $R^N = R'$.

It is easy to verify that *assignment-efficiency* implies that $\sigma(R) = \sigma^v$ and $\sigma(R') = \sigma^a$. Furthermore, for any subcoalition $C \subsetneq N$,

$$\sigma(R^C) = \sigma^v. \tag{4.11}$$

Since $M = 0$, the definition of a solution requires that for any $C \subseteq N$, we have $\sum m_i(R^C) = 0$. With Lemma 4.4, this implies the existence of n functions $h_i \colon \mathcal{Q}^{n-1} \to \mathbb{R}$ such that for each coalition $C \subseteq N$,

$$\sum_{i \in N} h_i(R^C_{-i}) = \sum_{i \in N} \sum_{j \neq i} v_j(\sigma(R^C); R^C_j) \tag{4.12}$$

Let $i \in N$, and let $B \subset N$ be such that $i \notin B$. Then $R^{B \cup i}_{-i} = R^B_{-i}$ follows by definition. Hence we can write for each $i \in N$,

$$\begin{aligned} 0 &= \sum_{C \not\ni i} (-1)^{|C|+1} h_i(R^{C \cup i}_{-i}) + \sum_{C \not\ni i} (-1)^{|C|} h_i(R^C_{-i}) \\ &= \sum_{C \ni i} (-1)^{|C|} h_i(R^C_{-i}) + \sum_{C \not\ni i} (-1)^{|C|} h_i(R^C_{-i}) \\ &= \sum_{C \subseteq N} (-1)^{|C|} h_i(R^C_{-i}) \end{aligned}$$

[10] This proof is based on one in Moulin (1988) for the case of unrestricted quasi-linear preferences over public goods.

Therefore

$$0 = \sum_{i \in N} \sum_{C \subseteq N} (-1)^{|C|} h_i(R^C_{-i}) = \sum_{C \subseteq N} (-1)^{|C|} \sum_{i \in N} h_i(R^C_{-i})$$

With (4.12), we have

$$\sum_{C \subseteq N} (-1)^{|C|} \sum_{i \in N} \sum_{j \neq i} v_j(\sigma(R^C); R^C_j) = 0 \tag{4.13}$$

By definition of the valuation functions and our earlier remarks (4.11) about *assignment-efficiency*,

$$\begin{aligned} \sum_{C \subseteq N} (-1)^{|C|} \sum_{i \in N} \sum_{j \neq i} v_j(\sigma(R^C); R^C_j) &= \left(\sum_{C \subsetneq N} (-1)^{|C|} \sum_{i \in N} \sum_{j \neq i} 0 \right) + (-1)^n \sum_{i \in N} \sum_{j \neq i} 1 \\ &= (-1)^n n(n-1) \neq 0 \end{aligned}$$

which contradicts (4.13). ■

Remark 4.5 It should be noted that Theorem 4.7 also holds for the case in which there are fewer objects than agents. The necessary modifications to the above proof are given in Section 4.6. Note that in such a case, if a solution always chooses an envy-free allocation (*i.e.*, no agent should prefer another agent's bundle), then the solution is *assignment-efficient* (see Svensson, 1983). Therefore this result implies a result of Tadenuma and Thomson (1995a), stating that there is no *strategy-proof* solution that always chooses an envy-free allocation, when $|\Omega| = 1$.

Following this general negative result but still desiring *strategy-proofness*, it is natural to weaken *assignment-efficiency* to *Pareto-optimality* in an attempt to find possibility.[11] At this point, we provide no general theorem but make an observation for the special case $M = 0$. Requiring *non-negativity* when $M = 0$, we obtain the model of Roth (1982).[12] One can see that by applying his solutions to our model (with respect to some arbitrarily chosen "endowments" — see Section 4.3.1) we obtain a class of *strategy-proof*, *Pareto-optimal*, *non-negative*

[11]It is interesting to note that in a model (of quasi-linear preferences) for which (1) arbitrary consumption of the divisible good is permitted, and (2) the definition of a solution allows disposal of some of the divisible good, the concept of *Pareto-optimality* is actually *stronger* than the concept of *assignment-efficiency*.

[12]We also have cardinal preference information over Roth's "houses"; however, this type of preference information is, irrelevant in the presence of *strategy-proofness* for our model, as it is in most models.

solutions. Whether possibility other than dictatorship obtains when $M > 0$ is an open question at this time.

4.5 Summary

For a model with $n \geq 4$ agents and n indivisible goods of which agents consume exactly one (and no divisible good), Bordes and Le Breton (1990) prove an Arrovian impossibility result. In contrast, Roth (1982) shows that reasonable *strategy-proof* solutions exist for that same model. We extend the search to a similar model — one with a transferable divisible good.

For the simple case of two agents and two objects, the *strategy-proof* solutions are quite basic. However with the addition of agents and objects, many *strategy-proof* solutions exist that use preference information in an arbitrary way. In an effort to eliminate such solutions from our search, we considered in turn three additions to the condition of *strategy-proofness*. The first two of these additions (*non-bossiness* and *coalitional strategy-proofness*) restrict the range of a *strategy-proof* solution to a very small size — at most one allocation per assignment of the indivisibles to the agents. The last addition (*bribe-proofness*) requires that a solution be, essentially, constant.

Finally we addressed the issue of *efficiency*. Adapting arguments from work on public goods models, we showed that no *strategy-proof* solution always chooses the *efficient* assignment of the objects to the agents. Weakening this requirement to *Pareto-optimality* while requiring non-negative consumption of the divisible good gives us a previously known possibility when $M = 0$, *i.e.* when there is no divisible good; the answer for the general case ($M \geq 0$) is unknown at this point.

4.6 Supplement: Efficiency with Few Objects

We provide an outline of the proof of Theorem 4.7 for the case of fewer objects than agents by providing the substitutions that must be made in to the proof for the original case.

Suppose there are m "real" objects, and therefore $n-m$ null objects. Without loss of generality, suppose that σ^v is the assignment giving the "real" objects to the agents in the set $M \equiv \{1, \ldots, m\}$.

Let $N' \equiv \{m+1, \ldots, n\}$. Define R, R' so that:

$$\begin{aligned}
\forall i \in N, \forall a \neq \sigma^v_i, &\quad (\sigma^v, 0)\ I_i\ (a, n) \\
\forall i \in M, \forall a \neq \sigma^v_i, \sigma^v_{i+1}, &\quad (\sigma^v, 0)\ I'_i\ (\sigma^v_{i+1}, -1) I'_i (a, n) \\
\forall i \in N', \forall a \neq \sigma^v_i, \alpha &\quad (\sigma^v, 0)\ I'_i\ (\alpha, -1)\ I'_i\ (a, n)
\end{aligned}$$

Define R^C as in the original proof.

Assignment-efficiency implies that if either $C \not\supset M$ or $C \cap N' = \emptyset$, then the assignment chosen is σ^v, or is equivalent to σ^v in terms of the valuation functions. Hence $\sum v_j(\sigma(R^C); R^C_j) = 0$.

Assignment-efficiency also implies that if either $C \supset M$ or $C \cap N' \neq \emptyset$, then each agent $i \in M$ receives object σ^v_{i+1}, and one agent $j \in N' \cap C$ receives α. In this case, $\sum v_j(\sigma(R^C); R^C_j) = m + 1$.

Now, the equation immediately following equation (4.13) becomes:

$$\begin{aligned}
& \sum_{C \subseteq N} (-1)^{|C|} \sum_{i \in N} \sum_{j \neq i} v_j(\sigma(R^C); R^C_j) \\
= \; & \sum_{C \cap N' = \emptyset} (-1)^{|C|} \cdot 0 + \sum_{C \not\supset M} (-1)^{|C|} \cdot 0 + \sum_{C \supset M, C \cap N' \neq \emptyset} (-1)^{|C|} (m+1)(n-1) \\
= \; & \sum_{k=1}^{n'} (-1)^{k+m} \binom{n'}{k} (m+1)(n-1) = (-1)^{m+1}(m+1)(n-1) \neq 0
\end{aligned}$$

where $n' = |N'|$, $m = |M|$, and the last equality follows from that fact that for any integer $\bar{n}$,

$$\sum_{k=1}^{\bar{n}} (-1^k) \binom{\bar{n}}{k} = -1$$

concluding the proof.

Bibliography

[1] Alkan, A., G. Demange, and D. Gale (1991) Fair Allocation of Indivisible Objects and Criteria of Justice, *Econometrica* **59**, 1023–39.

[2] Aragones, E. (1995) A Derivation of the Money Rawlsian Solution, *Soc Choice Welfare* **12**, 267–76.

[3] Aswad, N., and A. Sen (1996) Restricted Domains and the Gibbard-Satterthwaite Theorem, University of Minnesota mimeo.

[4] Barberà, S. (1977) The Manipulation of Social Choice Mechanisms that Do Not Leave "Too Much" to Chance, *Econometrica* **45**, 1573–88.

[5] Barberà, S., and M. Jackson (1994) A Characterization of Strategy-proof Social Choice Functions for Economies with Pure Public Goods, *Soc Choice Welfare* **11**, 241–52.

[6] Barberà, S. and M. Jackson (1995) Strategy-proof Exchange, *Econometrica* **63**, 51–87.

[7] Barberà, S. and B. Peleg (1990) Strategy-proof Voting Schemes with Continuous Preferences, *Soc Choice Welfare* **7**, 31–8.

[8] Beviá, C., and Corchón, L., On the Generic Impossibility of Truthful Behavior: A Simple Approach, *Econ Theory* **6**, 365–71 (1995).

[9] Bordes, G. and M. Le Breton (1990) Arrovian Theorems for Economic Domains: Assignments, Matchings, and Pairings, *Soc Choice Welfare* **7**, 193–208.

[10] Bossert, W., and J. Weymark (1993) Generalized Median Social Welfare Functions, *Soc Choice Welfare* **10**, 17–33.

[11] Chipman, J. (1974) Homothetic Preferences and Aggregation. *J Econ Theory* **8**, 26–38.

[12] Clarke, E. (1971) Multipart Pricing of Public Goods, *Public Choice* **8**, 19–33.

[13] Duggan, J. (1996) A Geometric Proof of Gibbard's Random Dictatorship Theorem, *Econ Theory* **7**, 365–9.

[14] Gibbard, A. (1973) Manipulation of Voting Schemes: A General Result, *Econometrica* **41**, 587–602.

[15] Gibbard, A. (1977) Manipulation of Schemes that Mix Voting with Chance, *Econometrica* **45**, 665–81.

[16] Green, J. and J.-J. Laffont (1977) Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods, *Econometrica* **45**, 427–38.

[17] Groves, T. (1973) Incentives in Teams, *Econometrica* **41**, 617–31.

[18] Holmström, B. (1979) Groves' Scheme on Restricted Domains, *Econometrica* **47**, 1137–44.

[19] Hurwicz, L. (1972) On Informationally Decentralized Systems, in Decision and Organization, (C. B. McGuire and R. Radner, eds.), North Holland, Amsterdam, 297–36.

[20] Hurwicz, L. and M. Walker (1990) On the Generic Non-optimality of Dominant Strategy Allocation Mechanisms, *Econometrica* **58**, 683–704.

[21] Hylland, A. (1980) Strategy Proofness of Voting Procedures with Lotteries as Outcomes and Infinite Sets of Strategies, Harvard University mimeo.

[22] Kalai, E., and E. Muller (1977) Characterization of Domains Admitting Nondictatorial Social Welfare Functions and Nonmanipulable Voting Procedures, *J Econ Theory*, **16**, 457–69.

[23] Ma, J. (1994) Strategy-Proofness and the Strict Core in a Market with Indivisibilities, *Intl J Game Theory* **23**, 75–83.

[24] Maskin, E. (1987) On the Fair Allocation of Indivisible Goods, in "Arrow and the Foundations of the Theory of Economic Policy," G. Feiwel ed., MacMillan.

[25] Moulin, H. (1988) "Axioms of Cooperative Decision Making," Econometric Society Monograph, Cambridge University Press, Cambridge.

[26] Nagahisa, R.-I. (1995) A Topological Characterization of Domains Admitting Nondictatorial, Efficient, and Strategy-proof Rules to Operate in Pure Exchange Economies, Kansai University mimeo.

[27] Roth, A. (1982) Incentive Compatibility in a Market with Indivisible Goods, *Econ Letters* **9**, 127–32.

[28] Satterthwaite, M. (1975) Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions, *J Econ Theory* **10**, 187–217.

[29] Satterthwaite, M. and H. Sonnenschein (1981) Strategy Proof Allocation Mechanisms at Differentiable Points, *Rev Econ Stud* **48**, 587–97.

[30] Schlesinger, H. (1989) On the Analytics of Pure Public Good Provision, *Public Finance* **44**, 102–9.

[31] Schummer, J. (1996) Strategy-proofness versus Efficiency for Small Domains of Preferences over Public Goods, University of Rochester mimeo.

[32] Schummer, J. (1997a) Strategy-proofness versus Efficiency on Restricted Domains of Exchange Economies, *Soc Choice Welfare* **14**, 47–56.

[33] Schummer, J. (1997b) Eliciting Preferences to Assign Positions and Compensation, University of Rochester mimeo.

[34] Serizawa, S. (1996) Strategy-proof and Individually Rational Social Choice Functions for Public Good Economies, *Econ Theory* **7**, 501–12.

[35] Shapley, L. and H. Scarf (1974) On Cores and Indivisibility, *J Math Econ* **1**, 23–37.

[36] Sprumont, Y. (1995) A Note on Strategyproofness in Edgeworth-Box Economies. *Econ Letters* **49**, 45–50.

[37] Svensson, L. (1983) Large Indivisibles: An Analysis with Respect to Price Equilibrium and Fairness, *Econometrica* **51**, 939–54.

[38] Tadenuma, K. and W. Thomson (1991) No-envy and Consistency in Economies with Indivisible Goods, *Econometrica* **59**, 1755–67.

[39] Tadenuma, K. and W. Thomson (1995a) Games of Fair Divison, *Games and Economic Behavior* **9**, 191–204.

[40] Tadenuma, K. and W. Thomson (1995b) Refinements of the No-envy Solution in Economies with Indivisible Goods, *Theory and Decision* **39**, 189–206.

[41] Walker, M. (1980) On the Nonexistence of a Dominant-Strategy Mechanism for Making Optimal Public Decisions, *Econometrica* **48**, 1521–40.

[42] Zhou, L. (1991a) Impossibility of Strategy-proof Mechanisms in Economies with Pure Public Goods, *Rev Econ Stud* **58**, 107–19.

[43] Zhou, L. (1991b) Inefficiency of Strategy Proof Allocation Mechanisms in Pure Exchange Economies, *Soc Choice Welfare* **8**, 247–54.