

Almost-dominant Strategy Implementation: Exchange Economies

James Schummer*

*MEDS Department, Kellogg School of Management
Northwestern University, Evanston, IL 60208-2009
Email: schummer@kellogg.northwestern.edu*

*I thank Ehud Kalai and two anonymous referees for helpful comments.

Running Title: Almost-dominant Strategy Implementation

Send Proofs To: James Schummer
MEDS, Kellogg GSM
Northwestern University
Evanston, IL 60208-2009 USA
schummer@northwestern.edu
FAX: 847-467-1220

Abstract

We relax strategy-proofness (a form of dominant strategy implementation) by allowing “small” gains from manipulation. In 2-agent exchange economies, this relaxation is shown to have a discontinuous effect on the range of efficient rules, demonstrating a type of non-robustness in previous impossibility results. When gains are measured with respect to a single good and preferences are linear, we characterize a particular rule as being the most equitable among all efficient rules satisfying the relaxation.

Keywords: strategy-proof, almost dominant strategy, ϵ -dominant.

JEL Classification Numbers: C70, D70.

1. INTRODUCTION

Dominant strategy implementability is clearly a desirable incentives property of choice rules. It makes certain issues—such as what information the planner has about the agents, what information agents have about each other, and what information is revealed during intermediate stages of the execution of the mechanism—basically irrelevant. Even the assumption that his fellow players are rational need not be made by a player concerned with his own best interests. Furthermore, calculating a player’s best action cannot be more complex than determining his own preferences over outcomes.

Given the desirability of this incentives property, it is important to determine which rules satisfy it in various situations. Indeed this question has been—and continues to be—answered for an increasingly diverse class of situations.¹ The nature of such results depends on the environment being examined. Roughly speaking, possibility results can be obtained in “simpler” environments, while impossibility results are often obtained in richer ones.

As the literature on *strategy-proofness* clarifies the line between possibility and impossibility, we are left with the question as to what incentives properties are obtainable in those richer environments where no reasonable rules are *strategy-proof*. In this paper, we address this question by considering a weaker version of *strategy-proofness* which does not rule out “small” gains from manipulation.

Other ways of addressing this question have appeared in the literature, though many have unappealing modelling assumptions. For instance, while results under Nash (or Bayesian, etc.) implementation (e.g. see Moore (1996)) tend to be more positive, they come with a price: Strong assumptions are made concerning the agents’ and planner’s information (e.g. common knowledge or common prior beliefs about each other’s preferences). For many mechanism design environments, such an assumption is not realistic.

Another approach applies to situations in which the planner is satisfied with approximations; he may find it sufficient to implement a rule that is “close” to some other desirable choice rule. This literature on *virtual implementation* (Abreu and Matsushima (1992), Duggan (1997)), achieves very

¹See Barberà (2001) and Thomson (1998).

positive results, though under the same type of informational assumptions listed above.

Asymptotic non-manipulability is considered by others, such as Roberts and Postlewaite (1976), Córdoba and Hammond (1998), Ehlers et al. (1999), Rustichini, Satterthwaite, and Williams (1994), Satterthwaite (2001), Swinkels (2001), and Kalai (2002).

Finally, yet another way of addressing incentives where *strategy-proofness* is impossible is to measure the frequency of opportunities agents have to manipulate. This idea is considered by Beviá and Corchón (1995), Kelly (1993), Saari (1995), Smith (1999), and, in experiments, Harrison and McDaniel (2002).

Our Approach

The method used in this paper can be seen as a different type of approximation approach, involving an approximation to the notion of dominance. The motivation behind our notion lies with a simple assumption about the strategic behavior of agents. Specifically, we approach the problem with the premise that if a player does not have much to gain by lying, then he will not bother to do so. Under this modelling assumption, we search for rules in which gains are limited by some upper bound.

This assumption can be interpreted or applied in various ways. For example, it applies when gathering information (about other agents) is costly. If such costly information is necessary for a player to compute a profitable way to manipulate the choice rule, it would never be worth the expense to gather it if the potential gains were bounded above by this cost. Another application of this idea is to situations in which computation itself is costly.

A third application of our assumption is to situations in which agents value morality (or honesty) in some real, fixed terms. In such settings, small gains from cheating do not outweigh the losses (or “guilt”) incurred.

An important observation here is that we make no assumption on the structure of information that agents possess. Some of the work cited previously considers a rule to be almost non-manipulable in a Bayesian setting even if there is a small probability of a very large gain. Such a definition implicitly assumes that players not only have beliefs consistent with those assumed by the planner, but that the players *cannot* have *more* information

than that.

The critical detail of our work is to precisely define what it means to gain *much*. One approach that may come to mind is to use a utility-based notion of preferences, where a player would be assumed not to manipulate a rule unless his utility gain would exceed some bound. This approach, however, would depend heavily on the interpretation (and/or the parameterization) of utility functions.

To avoid this difficulty, we define our condition in terms of real commodities. In our exchange economy model, our behavioral assumption is that an agent will manipulate a choice rule only if his gains are perceived to be better than receiving a prespecified, additional amount of goods.

Overview and Interpretation of Results

We restrict attention to the domain of linear (additively separable) preferences in 2-agent exchange economies with two goods, i.e. an Edgeworth Box. We begin with this simple class of preferences for various reasons. First, the analysis is more tractable. Second, here we are able to obtain tight welfare bounds imposed by our relaxed truth-telling condition, allowing us to quantify the effect of relaxing *strategy-proofness*. It seems difficult to obtain results of similar strength on other standard preference domains.

Our results have both a positive and negative flavor. In order to interpret them, we recall a result concerning fully *strategy-proof* rules for this domain. Extending results on the classical domain by Hurwicz (1972) and Zhou (1991),² Schummer (1997) shows that even when preferences are restricted to the linear preference domain, a *strategy-proof*, efficient rule must be dictatorial, i.e. always give the entire endowment to a prespecified agent.

Our sharpest results are obtained when (Section 4) we measure gains only with respect to a single numeraire good. The first result (Theorem 2) is in the negative direction: Any efficient rule that satisfies such a relaxation of *strategy-proofness* must allocate most of the numeraire good to a prespecified agent; This bound is shown to be tight with the rule in Example 1.

²They show that for 2-agent exchange economies on the classical domain of preferences, a *strategy-proof*, efficient rule sometimes gives agents bundles worse than some original endowment (Hurwicz), and in fact must always give one agent everything (Zhou). Also, see Barberà and Jackson (1995), and Serizawa and Weymark (2002).

While Theorem 2 shows an asymmetry between the agents' consumption of the numeraire good, Theorem 3 states that the rule described in Example 1 is always (weakly) more equitable than any other efficient rule satisfying this relaxed version of *strategy-proofness*. Specifically, *regardless* of the preference profile reported by the agents, the allocation prescribed by this rule weakly Lorenz-dominates the allocation prescribed by any other such rule.

A negative interpretation of Theorem 3 (as an upper bound on equity) is clear. However, a positive interpretation comes from the observation (made in Section 4.2 that this rule demonstrates a type of discontinuity with respect to the impossibility result of Schummer (1997) described above. As allowable gains from manipulation are made arbitrarily small, the range of this (parameterized) rule does not converge to the range of the only *strategy-proof*, efficient rule, i.e. of a dictatorial rule. Once *strategy-proofness* is relaxed an arbitrarily small amount, there is a (discontinuous) increase in the flexibility of admissible rules.

In Section 5, we allow measures of gains to be made with respect to any good. Under this weaker version of almost-dominance, even more rules are admissible. Furthermore, another discontinuity occurs similar to the one discussed above.

We conduct a simple welfare analysis in Section 6, and quantify the effects of relaxing *strategy-proofness*. Under *strategy-proofness*, one agent must always consume nothing. Under the rules we discuss, one of the agents consumes relatively less than the other agent, but for some preference profiles consumes a bundle that he considers to be *almost as good as the entire endowment*.

To summarize, the paper is organized as follows. In Section 2 we formalize the exchange economy model, while we provide our main definition in Section 3. In Sections 4 and 5, we provide our results for exchange economies. Using these results, we quantify the consequences of relaxing *strategy-proofness* in this model in Section 6. In Section 7, we provide a brief discussion on extending the analysis to more general models of exchange.

2. EXCHANGE ECONOMY MODEL

The set of two agents is $N = \{1, 2\}$. There is a positive endowment of two infinitely divisible goods $\Omega = (\Omega^1, \Omega^2) \in \mathbb{R}_{++}^2$. Each agent $i \in N$ is to consume a bundle $x_i \in \mathbb{R}_+^2$. An *allocation* is a pair of bundles $x = (x_1, x_2) = ((x_1^1, x_1^2), (x_2^1, x_2^2)) \in \mathbb{R}_+^4$ such that $x_1 + x_2 = \Omega$; the set of allocations is denoted A . Subscripts refer to agents, superscripts refer to goods, and the vector inequalities are $>$, \geq , and $\underline{\geq}$.

Each agent has a strictly monotonic, linear preference relation, R_i , over his consumption space \mathbb{R}_+^2 . Precisely, such preference relations are the ones representable by a utility function of the form $u(x_i) = \lambda x_i^1 + (1 - \lambda)x_i^2$, $\lambda \in (0, 1)$. Denote the set of such preference relations as \mathcal{R} . The strict (antisymmetric) and indifference (symmetric) preference relations associated with R_i are denoted P_i and I_i .

An *allocation rule* is a function, $\varphi: \mathcal{R}^2 \rightarrow A$, mapping the set of preference profiles into the set of allocations. To simplify notation, when $\varphi(R) = x$, we denote $\varphi_i(R) = x_i$ for any agent $i \in N$. Furthermore, we write $-i$ to refer to the agent not equal to i . For example, if $i = 1$, then $x_{-i} = x_2$, and (R'_i, R_{-i}) is the same as (R'_1, R_2) .

We are interested in finding allocation rules that satisfy desirable properties not only in terms of incentives, but also in terms of efficiency. An allocation $x \in A$ is *efficient* with respect to a preference profile $R \in \mathcal{R}^2$ if there exists no $y \in A$ such that for some $i \in N$, $y_i P_i x_i$ and $y_{-i} R_{-i} x_{-i}$. We also call an allocation rule *efficient* if it assigns to every preference profile an allocation that is efficient with respect to that preference relation.

For any profile $R \in \mathcal{R}^2$, denote the set of efficient allocations for R as $E(R)$. On our domain of linear preferences, if both agents have the same preference relation ($R_1 = R_2$), then the set of efficient allocations is the entire set: $E(R) = A$. If R is such that agent 1 values good 1 relatively more than agent 2 does, then the set of efficient allocations is $E(R) = E^\lrcorner \equiv \{x \in A : x_1^2 = 0 \text{ or } x_2^1 = 0\}$. In the opposite, remaining case, $E(R) = E^\ulcorner \equiv \{x \in A : x_1^1 = 0 \text{ or } x_2^2 = 0\}$.

3. A DEFINITION OF NONMANIPULABILITY

A simple way to measure manipulability is to measure gains relative to either of the two goods. To be precise, consider a situation in which an allocation rule φ prescribes, for $R \in \mathcal{R}^2$, an allocation $x = \varphi(R)$. If we postulate that agent i would not falsely report his preferences for small gains, then there exists some number $\epsilon_1 \geq 0$ such that if for some $R'_i \in \mathcal{R}$, we have $\varphi_i(R'_i, R_{-i}) \leq \varphi_i(R) + (\epsilon_1, 0)$, then agent i would not manipulate the rule with that particular misrepresentation R'_i . That is, if agent i can gain only ϵ_1 (or fewer) units of good 1, then the gain is too small to be considered.

Similarly, for some $\epsilon_2 \geq 0$, we say that an agent does not manipulate φ if he simply gains ϵ_2 (or fewer) units of good 2.³

Finally, consider a situation in which a false report of preferences, R'_i , gives agent i the bundle $x_i = \varphi(R'_i, R_i)$, such that $\varphi_i(R) + (\epsilon_1, 0) R_i x_i$. Since the agent would not manipulate the rule in order to obtain the bundle $\varphi_i(R) + (\epsilon_1, 0)$, we conclude that he would not manipulate the rule in order to obtain the (worse) bundle x_i . A similar reasoning is to be applied with respect to good 2 and ϵ_2 .

Our formal definition of this reasoning is as follows.

(ϵ_1, ϵ_2) -strategy-proofness: For any $\epsilon \in \mathbb{R}_+^2$, a rule is (ϵ_1, ϵ_2) -strategy-proof if for all $R \in \mathcal{R}^2$, all $i \in \{1, 2\}$, and all $R'_i \in \mathcal{R}$, we have either

- (i) $\varphi_i(R) + (\epsilon_1, 0) R_i \varphi_i(R'_i, R_{-i})$, or
- (ii) $\varphi_i(R) + (0, \epsilon_2) R_i \varphi_i(R'_i, R_{-i})$.

That is, by misreporting his preferences, an agent cannot procure a gain that he considers, simultaneously, to be (i) better than simply acquiring an additional ϵ_1 units of good 1 and (ii) better than simply acquiring an additional ϵ_2 units of good 2. See Figure 1; in Figure 1a, part [i] of the definition is redundant, while in the case of Figure 1b, part [ii] is. In the language of Barberà and Peleg (1990), the agent's *option set* should be a subset of the shaded area.

³A more general definition makes the value of ϵ_j dependent on the identity of the agent in question, or, even more generally, his preference relation R_i . For simplicity, we do not go to this level of generality.

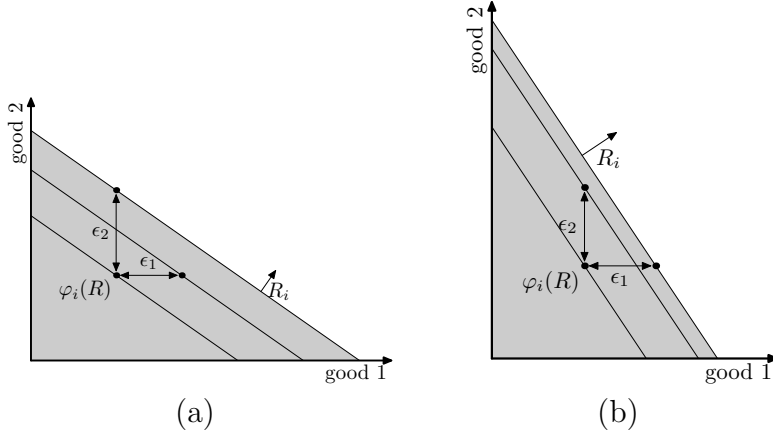


Figure 1: If φ is (ϵ_1, ϵ_2) -strategy-proof, then any false report by agent i results in a bundle somewhere within the shaded area.

It should be clear that (ϵ_1, ϵ_2) -strategy-proofness is a stronger condition than $(\epsilon'_1, \epsilon'_2)$ -strategy-proofness whenever $\epsilon \leq \epsilon'$, and that $(0, 0)$ -strategy-proofness is equivalent to the standard definition of strategy-proofness.

We close this section by observing a result previously established for the case $\epsilon = (0, 0)$. With a result related to that of Zhou (1991), Schummer (1997) shows that on this class of problems (with linear preferences), the only efficient rules that are $(0, 0)$ -strategy-proof are those that assign the entire endowment to a prespecified agent.

THEOREM 1 (SCHUMMER (1997)) *Let φ be an efficient rule that is $(0, 0)$ -strategy-proof. There exists an agent $i \in N$ that always receives the entire endowment: for all $R \in \mathcal{R}^2$, $\varphi_i(R) = (\Omega^1, \Omega^2)$ (and $\varphi_{-i}(R) = (0, 0)$).*

4. RESULTS FOR $(\epsilon_1, \mathbf{0})$ -strategy-proofness

We first examine the implications of (ϵ_1, ϵ_2) -strategy-proofness when $\epsilon_2 = 0$ (Figure 2). In this case, we are able to obtain tight bounds on the flexibility of efficient rules that satisfy this condition (Sections 4.1 and 4.2). Furthermore, even though this case yields a stronger condition than when $\epsilon_2 > 0$, there is a discontinuous increase in the range of such rules at $\epsilon_1 = 0$. This demonstrates a type of nonrobustness to the impossibility result of Theorem 1.

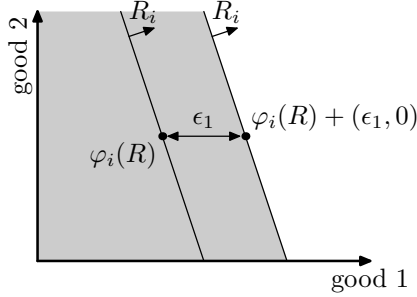


Figure 2: The special case of (ϵ_1, ϵ_2) -strategy-proofness when $\epsilon_2 = 0$.

4.1. A Bound on the Range

Our first result is that, under $(\epsilon_1, 0)$ -strategy-proofness, a rule always allocates nearly all of the endowment of good 1 to a prespecified agent.

THEOREM 2. *Let φ be an efficient rule that is $(\epsilon_1, 0)$ -strategy-proof, where $\epsilon_1 < \Omega^1/5$. There exists an agent $i \in N$ that always receives almost all of good 1: for all $R \in \mathcal{R}^2$, $\varphi_i^1(R) \geq \Omega^1 - 2\epsilon_1$.*

To prove the result, we first provide the following lemma, which states that for any pair of preference profiles with the same set of efficient allocations, the allocation of good 1 differs at those profiles by at most $2\epsilon_1$.

LEMMA 1. *Let φ be efficient and $(\epsilon_1, 0)$ -strategy-proof. For all $R, R' \in \mathcal{R}^2$, if either $E(R) = E(R') = E^\Gamma$ or $E(R) = E(R') = E^\Delta$, then $|\varphi_1^1(R) - \varphi_1^1(R')| \leq 2\epsilon_1$.*

Proof. Let $R, R' \in \mathcal{R}^2$ be such that $E(R) = E(R') = E^\Gamma$. It is either the case that $E(R_1, R'_2) = E^\Gamma$, or $E(R'_1, R_2) = E^\Gamma$. Without loss of generality, suppose $E(R_1, R'_2) = E^\Gamma$ (which is true, for example, if the indifference curves of R_1 are “flatter” than those of R'_1).

By efficiency, $\varphi(R_1, R'_2) \in E^\Gamma$. Since φ is $(\epsilon_1, 0)$ -strategy-proof and $\varphi(R'_1, R'_2) \in E^\Gamma$, we have $\varphi_1^1(R_1, R'_2) - \varphi_1^1(R'_1, R'_2) \leq \epsilon_1$. Similarly, $\varphi_1^1(R'_1, R'_2) - \varphi_1^1(R_1, R'_2) \leq \epsilon_1$, so $|\varphi_1^1(R_1, R'_2) - \varphi_1^1(R'_1, R'_2)| \leq \epsilon_1$.

By the same type of argument, we have $|\varphi_2^1(R_1, R_2) - \varphi_2^1(R_1, R'_2)| \leq \epsilon_1$, implying $|\varphi_1^1(R_1, R_2) - \varphi_1^1(R_1, R'_2)| \leq \epsilon_1$.

Therefore, by the triangle inequality, $|\varphi_1^1(R_1, R_2) - \varphi_1^1(R'_1, R'_2)| \leq 2\epsilon_1$, proving the result. \square

Now we can prove the theorem.

Proof. Let φ be efficient and $(\epsilon_1, 0)$ -strategy-proof. There are three possible cases.

Case 1: For all $R \in \mathcal{R}^2$, if $E(R) = E^\Gamma$, then $\varphi_1^2(R) = \Omega^2$.

Step 1a: (E^\perp) In this case, for all $\delta > 0$, there exists $R \in \mathcal{R}^2$ such that $E(R) = E^\perp$ and $\varphi_1(R) \geq (\Omega^1, \Omega^2 - \delta)$. To see this, let R_1 satisfy $(0, \Omega^2) P_1 (\Omega^1 + \epsilon_1, \Omega^2 - \delta)$, let R_2 be such that $E(R) = E^\perp$, and let R'_1 be such that $E(R'_1, R_2) = E^\Gamma$. Since φ is $(\epsilon_1, 0)$ -strategy-proof and $E(R'_1, R_2) = E^\Gamma$,

$$\varphi_1(R) + (\epsilon_1, 0) R_1 \varphi_1(R'_1, R_2) R_1 (0, \Omega^2)$$

by the hypothesis of Case 1. Therefore $\varphi_1(R) P_1 (\Omega^1, \Omega^2 - \delta)$. Since $\varphi_1(R) \in E^\perp$, we have $\varphi_1(R) \geq (\Omega^1, \Omega^2 - \delta)$.

Therefore by Lemma 1, for all $R \in \mathcal{R}^2$, if $E(R) = E^\perp$, then $\varphi_1^1(R) \geq \Omega^1 - 2\epsilon_1$.

Step 1b: (E^Γ and A) Let $R \in \mathcal{R}^2$ be such that $E(R) \in \{E^\Gamma, A\}$, and suppose in contradiction to the theorem that $\Omega^1 - \varphi_1^1(R) - 2\epsilon_1 = \delta > 0$. Let $y, y', y'' \in E^\perp$ satisfy (see Figure 3):

$$\begin{aligned} y_1 I_1 \varphi_1(R) + (\epsilon_1 + \frac{1}{3}\delta/3, 0) \\ y'_1 I_1 \varphi_1(R) + (\epsilon_1 + \frac{2}{3}\delta, 0) \\ y''_1 I_1 \varphi_1(R) + (2\epsilon_1 + \frac{2}{3}\delta, 0) &= (\Omega^1 - \frac{1}{3}\delta, 0) \end{aligned}$$

Let R'_2 be such that $y_2 I'_2 \varphi_2(R) - (\epsilon_1, 0)$. Since $\varphi(R_1, R'_2) \in E^\perp$, the truth-telling condition implies $\varphi_2(R_1, R'_2) \geq y_2$. Let R''_2 be sufficiently flat so that both $y''_2 P''_2 (\Omega^1 + \epsilon_1, 0)$ and $y_2 P''_2 y'_2 + (\epsilon_1, 0)$. The truth-telling condition implies $\varphi_2(R_1, R''_2) + (\epsilon_1, 0) R''_2 \varphi_2(R_1, R'_2)$, so $\varphi_2(R_1, R''_2) \geq y'_2$.

Let R'_1 satisfy $(0, \Omega^2) I'_1 y''_1 + (\epsilon_1, 0)$. Then $E(R'_1, R''_2) = E^\perp$. Note that by construction, $y'_1 + (\epsilon_1, 0) I_1 y''_1$. The truth-telling condition implies $\varphi_1(R_1, R''_2) + (\epsilon_1, 0) R_1 \varphi_1(R'_1, R''_2)$. Therefore $\varphi_1(R'_1, R''_2) \leq y''_1$.

By the hypothesis of Case 1, for all R''_1 such that $E(R'') = E^\Gamma$, we have $\varphi_1(R'') \geq (0, \Omega^2)$. But then for any such R''_1 , we have $\varphi_1(R'') P'_1 \varphi_1(R'_1, R''_2) + (\epsilon_1, 0)$, which contradicts the truth-telling condition.

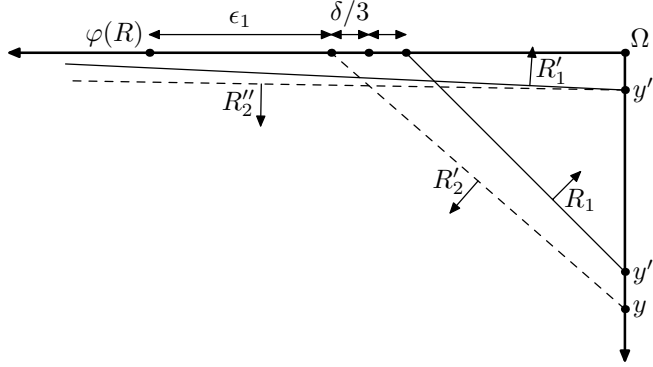


Figure 3: Proof of Theorem 2. The figure represents the upper-right corner of the Edgeworth Box.

Therefore, if Case 1 holds, we have derived the conclusion of the theorem.

Case 2: For all $R \in \mathcal{R}^2$, if $E(R) = E^\perp$, then $\varphi_2^2(R) = \Omega^2$.

This case is symmetric to Case 1. In this case, for all $R \in \mathcal{R}^2$, $\varphi_2^1(R) \geq \Omega^1 - 2\epsilon_1$.

Case 3: Neither Case 1 nor Case 2 holds, i.e., there exist $R, R' \in \mathcal{R}^2$ such that $E(R) = E^\top$, $E(R') = E^\perp$, $\varphi_1^2(R) < \Omega^2$, and $\varphi_2^2(R') < \Omega^2$.

In this case, by Lemma 1, for all $R, R' \in \mathcal{R}^2$, $E(R) = R^\top$ implies $\varphi_1^1(R) \leq 2\epsilon_1$, and $E(R') = R^\perp$ implies $\varphi_2^1(R') \leq 2\epsilon_1$. Since $\epsilon_1 < \Omega^1/5$, this implies that for all such R, R' ,

$$\varphi_1^1(R') - \varphi_1^1(R) > \epsilon_1 \quad (1)$$

Let R_1 be such that $(2\epsilon_1, \Omega^2) \succ P_1(\Omega^1 - 3\epsilon_1, 0)$. Let R_2, R_1' be such that $E(R) = E^\top$ and $E(R_1', R_2) = E^\perp$. Then eqn. (1) implies $\varphi_1(R_1', R_2) \succ P_1(\varphi_1(R) + (\epsilon_1, 0))$, which contradicts the truth-telling condition. Therefore this case cannot hold. \square

4.2. A Most-Equitable Rule

Theorem 2 states that under an efficient rule that is $(\epsilon_1, 0)$ -strategy-proof, one agent always must receive at least $\Omega^1 - 2\epsilon_1$ of good 1. The rule—described below in Example 1—simultaneously shows that (i) this bound is tight, and (ii) there is no such bound corresponding to good 2. In other words, agent 1 receives (i) from as little as $\Omega^1 - 2\epsilon_1$ of good 1 to as much as all of it, and

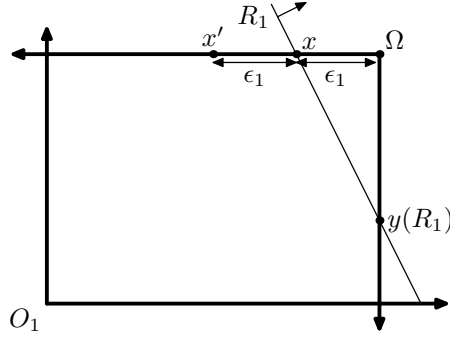


Figure 4: An efficient rule that is $(\epsilon_1, 0)$ -strategy-proof.

(ii) from as little as none of good 2 to as much as all of it.

Furthermore, and *most importantly*, we provide Theorem 3, showing that this rule is, unambiguously, the “least dictatorial” (or most equitable) of all efficient rules that are $(\epsilon_1, 0)$ -strategy-proof. This result does not, by itself, make the rule appealing. Instead, the rule is used to show another discontinuity when we relax (ϵ_1, ϵ_2) -strategy-proofness from $\epsilon_2 = 0$ to $\epsilon_2 > 0$.

EXAMPLE 1. Fix the allocations $x = ((\Omega^1 - \epsilon_1, \Omega^2), (\epsilon_1, 0))$, which gives agent 1 the entire endowment except for ϵ_1 units of good 1, and $x' = ((\Omega^1 - 2\epsilon_1, \Omega^2), (2\epsilon_1, 0))$. For all $R_1 \in \mathcal{R}$, let $y(R_1) \in E^\perp$ be the unique allocation in E^\perp that agent 1 considers indifferently to x (as in Figure 4), i.e., that $x_1 I_1 y_1(R_1)$. Define $\tilde{\varphi}$ so that for all $R \in \mathcal{R}^2$,

$$\tilde{\varphi}(R) = \begin{cases} x' & \text{if } x' \text{ is efficient for } R \\ y(R_1) & \text{otherwise} \end{cases}$$

We leave it to the reader to check that $\tilde{\varphi}$ is efficient and $(\epsilon_1, 0)$ -strategy-proof.⁴

This rule is clearly not symmetric. In fact, for most profiles of preferences, both agents would prefer agent 1’s consumption bundle to agent 2’s. A more formal welfare analysis appears in Section 6. The statement of Theorem 2 does not, by itself, rule out more equitable rules. As the next theorem shows, however, no efficient, $(\epsilon_1, 0)$ -strategy-proof rule can be more equitable than $\tilde{\varphi}$ (or the rule obtained from $\tilde{\varphi}$ by switching the roles of the agents). This

⁴Clearly, a mirror image to this rule exists in which the labels of the two agents are switched, and that rule also satisfies the two properties.

statement is true in a very strong sense: Under any other such rule, say φ , one of the two agents would, *under any profile of preferences*, prefer the bundle that $\tilde{\varphi}$ prescribes to agent 2 to the one he receives under φ .

THEOREM 3. *No efficient, $(\epsilon_1, 0)$ -strategy-proof rule is more equitable than $\tilde{\varphi}$.⁵ Specifically, let φ be an efficient rule that is $(\epsilon_1, 0)$ -strategy-proof, where $\epsilon_1 < \Omega^1/5$. Then one of the agents must (weakly) prefer playing the role of agent 2 under $\tilde{\varphi}$ to playing his own role under φ , i.e., one of the following is true.*

- (1) for all $R \in \mathcal{R}^2$, $\tilde{\varphi}_2(R) R_2 \varphi_2(R)$ (agent 2 prefers $\tilde{\varphi}$ to φ), or,
- (2) for all $R \in \mathcal{R}^2$, $\tilde{\varphi}_2(R') R_1 \varphi_1(R)$, where $R'_1 = R_2$ and $R'_2 = R_1$ (agent 1 prefers agent 2's consumption under $\tilde{\varphi}$ to his under φ).

Proof. Let φ be an efficient rule that is $(\epsilon_1, 0)$ -strategy-proof. Suppose (by Theorem 2) that agent 1 is the agent who always receives at least $\Omega^1 - 2\epsilon_1$ of the numeraire good under φ . In this case, we show statement (1) of the Theorem: for all $R \in \mathcal{R}^2$, $\tilde{\varphi}_2(R) R_2 \varphi_2(R)$. (Supposing the opposite leads to statement (2).)

If $E(R) = E^\top$, the conclusion follows from Theorem 2, since $\varphi_2(R) \leq (2\epsilon_1, 0) = \tilde{\varphi}_2(R)$.

If either $E(R) = E^\perp$ or $E(R) = A$, suppose in contradiction to the theorem that $\varphi_2(R) P_2 \tilde{\varphi}_2(R)$. Then there exists $\delta > 0$ such that

$$\varphi_1(R) I_1 (\Omega^1 - \epsilon_1 - \frac{2}{3}\delta, \Omega^2)$$

(otherwise the proof is trivial). Letting $y = \varphi(R)$ and $R'_2 = R_2$, and defining y' , y'' , R'_1 , and R''_1 as in Figure 3 (proof of Theorem 2), leads to a contradiction as it did in Step 1b of that proof. \square

In light of this result, a full characterization of efficient, $(\epsilon_1, 0)$ -strategy-proof rules does not appear to be interesting. For example, the rule $\tilde{\varphi}$ can be perturbed in many uninteresting ways (e.g., by giving slightly more of the goods to agent 1) while remaining $(\epsilon_1, 0)$ -strategy-proof.

⁵The same result obviously applies to the rule obtained from $\tilde{\varphi}$ by reversing the (asymmetric) roles of the two agents.

To emphasize the idea that a small relaxation in *strategy-proofness* leads to a large increase in the flexibility of rules, consider the implications of $(\epsilon_1, 0)$ -*strategy-proofness* as ϵ_1 approaches zero. The rule $\tilde{\varphi}$ was defined in Example 1 with respect to a given value of ϵ_1 . The range of this rule for a given ϵ_1 is

$$\{x \in E^1 : x_1^1 > \Omega^1 - \epsilon_1, x_1^2 < \Omega^2\} \cup \{((\Omega^1 - 2\epsilon_1, \Omega^2), (2\epsilon_1, 0))\}$$

As ϵ_1 converges to zero, this set converges to the right-hand border of the Edgeworth Box, i.e., to $\{x \in A : x_1^1 = \Omega^1\}$.

Therefore, as $(\epsilon_1, 0)$ -*strategy-proofness* converges to *strategy-proofness*, the range of an admissible rule does *not* converge to the support of the ranges of *strategy-proof* and efficient rules (i.e. dictatorial rules) characterized in Schummer (1997) for this domain.⁶ This discontinuity is important to observe because it reinforces the notion that a small relaxation of *strategy-proofness* leads to a relatively large increase in the range of admissible rules. On domains for which impossibility results regarding *strategy-proofness* have been established, relaxing the condition even in a small way may allow for significantly more flexible allocation rules.

There is an additional point that gives these results even more positive flavor. In models with additional agents, rules satisfying the truth-telling condition may be even more flexible. The $2\epsilon_1$ -bound of Theorem 1 is derived from the fact that there are only two agents. Roughly speaking, two unilateral changes in preferences can change the welfare of agents by an amount comparable to at most $2\epsilon_1$ units of good 1 (as in the proof of Theorem 2). With more agents, there is reason to believe that changes in preferences by more agents will lead to even greater flexibility in rules satisfying our condition. Owing to the difficulty of this model with more than two agents⁷ we leave this topic to future research.

⁶Formally, this sequence of examples shows that the ranges of the admissible rules is a correspondence that is not upper-semi-continuous at $\epsilon_1 = 0$, fixing $\epsilon_2 = 0$. It is clearly lower-semi-continuous: dictatorial rules are $(\epsilon_1, 0)$ -*strategy-proof* for any ϵ_1 .

⁷See Serizawa and Weymark (2002) for the latest development on the consequences of efficiency and full *strategy-proofness* for $n > 2$.

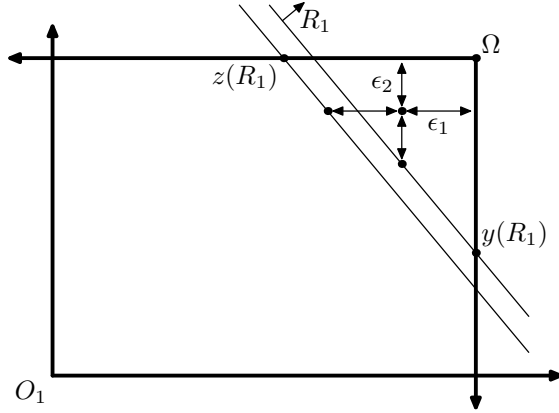


Figure 5: Defining an efficient rule that is (ϵ_1, ϵ_2) -strategy-proof.

5. RULES FOR (ϵ_1, ϵ_2) -strategy-proofness

We now turn our attention to the weaker condition of (ϵ_1, ϵ_2) -strategy-proofness when $\epsilon_2 > 0$. The rule $\tilde{\varphi}$ can be generalized in various ways, not all of which are obvious. For example, one obvious generalization could be obtained by redefining x (in Example 1) to be $(\Omega_1 - \epsilon_1, \Omega_2 - \epsilon_2)$, and generalizing the rule in the obvious way.

In this section, we focus our attention on a generalization that is slightly less asymmetric (and perhaps less obvious) than that one. Though we are unable to obtain results analogous to Theorems 2 and 3, the purpose of this example is to show that relaxing $(\epsilon_1, 0)$ -strategy-proofness to (ϵ_1, ϵ_2) -strategy-proofness results in another discontinuity (at $\epsilon_2 = 0$) in the ranges of admissible rules.

EXAMPLE 2. For all $R_1 \in \mathcal{R}$, let $y(R_1) \in E^\perp$ and $z(R_1) \in E^\Gamma$ be the unique allocations (as in Figure 5) such that $y_1(R_1) \geq \Omega_1 - \epsilon_1$, $\Omega_2 - 2\epsilon_2$ and $z_1(R_1) \geq \Omega_1 - 2\epsilon_1$, $\Omega_2 - \epsilon_2$. Define $\hat{\varphi}$ so that for all $R \in \mathcal{R}^2$,

$$\hat{\varphi}(R) = \begin{cases} z(R_1) & \text{if } z(R_1) \text{ is efficient for } R \\ y(R_1) & \text{otherwise} \end{cases}$$

We leave it to the reader to check that $\hat{\varphi}$ is efficient and (ϵ_1, ϵ_2) -strategy-proof.

In Section 4.2 we showed that relaxing strategy-proofness to $(\epsilon_1, 0)$ -strategy-proofness results in a discontinuous enlargement in the possible range

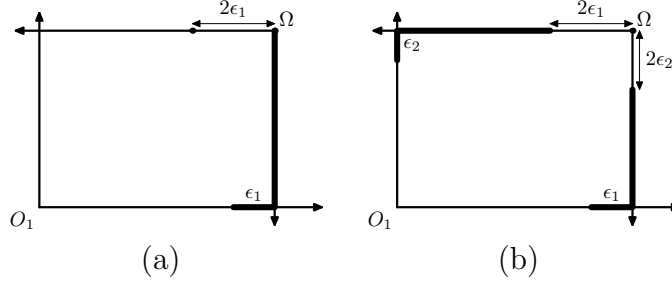


Figure 6: The ranges of (a) $\tilde{\varphi}$ and (b) $\hat{\varphi}$.

of admissible, efficient rules. The rule $\hat{\varphi}$ shows that further relaxing to (ϵ_1, ϵ_2) -*strategy-proofness* ($\epsilon_2 > 0$) results in another striking discontinuity: Fixing $\epsilon_1 \geq 0$, the range of $\hat{\varphi}$ —defined with respect to (ϵ_1, ϵ_2) —is discontinuous at $\epsilon_2 = 0$.

To see this, observe that the range of $\hat{\varphi}$ (for a given (ϵ_1, ϵ_2)) is

$$\{x \in E^d : x_1^1 > \Omega^1 - \epsilon_1, x_1^2 < \Omega^2 - 2\epsilon_2\} \cup \{x \in E^r : x_1^2 > \Omega^2 - \epsilon_2, x_1^1 < \Omega^1 - 2\epsilon_1\}$$

which is highlighted in Figure 6b. As (ϵ_1, ϵ_2) converges to $(0, 0)$ from above, this range converges to the upper and right-hand borders of the Edgeworth Box. However, by Theorem 2, the range of any efficient, $(\epsilon_1, 0)$ -*strategy-proof* rule is contained in the set $\{x \in A : x_1^1 \geq \Omega^1 - 2\epsilon_1\}$, which converges to the right-hand side of the Edgeworth Box as ϵ_1 converges to zero.

6. MEASURES OF WELFARE

First consider the stronger condition of $(\epsilon_1, 0)$ -*strategy-proofness*. Theorem 3 provides an upper bound on the welfare of the “unfavored” agent under an efficient, $(\epsilon_1, 0)$ -*strategy-proof* rule. In order to have a better understanding of how well-off agent 2 is under the rule $\tilde{\varphi}$, it is useful to consider a class of normalized utility functions. We parameterize each preference relation $R_i \in \mathcal{R}$ with $\lambda_i \in]0, 1[$ such that the preference relation is represented by the utility function

$$u(x_i) = \lambda_i x_i^1 + (1 - \lambda_i) x_i^2$$

Below we consider the case in which $\Omega = (1, 1)$. In this case, an agent’s utility is always equal to one when he receives the entire endowment, and

is always equal to zero when he receives nothing. In particular, a utility level can be interpreted as a proportion of the entire endowment, that is, $u(\delta, \delta) = \delta$.

Under the rule $\tilde{\varphi}$ (defined with respect to a given ϵ_1), agent 2's utility is a function of λ_1 , λ_2 , ϵ_1 , and Ω . It is a fairly straightforward geometric exercise to derive agent 2's utility under $\tilde{\varphi}$:⁸

$$u_2(\tilde{\varphi}; \lambda, \epsilon_1, \Omega) = \begin{cases} 2\lambda_2\epsilon_1 & \text{if } \lambda_2 \geq \lambda_1; \\ (1 - \lambda_2)\epsilon_1\lambda_1/(1 - \lambda_1) & \text{if } \lambda_2 < \lambda_1 \leq \Omega^2/(\Omega^2 + \epsilon_1); \\ \lambda_2\epsilon_1 + \Omega^2(\lambda_1 - \lambda_2)/\lambda_1 & \text{otherwise.} \end{cases}$$

Figure 7 graphs $u_2(\tilde{\varphi}; \cdot)$ when $\epsilon_1 = 0.1$ and $\Omega = (1, 1)$. We see that agent 2 receives a non-negligible amount of utility at most profiles. The average utility that agent 2 receives over this entire range of values for (λ_1, λ_2) is approximately 0.18 (under a uniform distribution).⁹ This is significantly higher than the average utility agent 2 would receive by consuming a constant $\epsilon_1 = 0.1$ units of good 1, which would be 0.05 units of utility.

By the previously mentioned result of Schummer (1997), if *strategy-proofness* were required, one of the agents would receive a constant utility of zero. These numbers encourage the idea that a “small” relaxation of *strategy-proofness* leads in some sense to a “larger” relaxation of dictatorship.

When the condition is weakened to (ϵ_1, ϵ_2) -*strategy-proofness* with $\epsilon_2 > 0$, the rule $\hat{\varphi}$ is admissible. Under that rule (defined with respect to a given ϵ), agent 2's utility is

$$u_2(\hat{\varphi}; \lambda, \epsilon_1, \epsilon_2, \Omega) = \begin{cases} \lambda_2\Omega^1 + (1 - \lambda_2)(\epsilon_2 + (2\epsilon_1 - \Omega^1)\lambda_1/(1 - \lambda_1)) & \text{if } \lambda_2 \geq \lambda_1 \text{ and } 2\epsilon_1 + \epsilon_2(1 - \lambda_1)/\lambda_1 > \Omega^1; \\ \lambda_2(2\epsilon_1 + \epsilon_2(1 - \lambda_1)/\lambda_1) & \text{if } \lambda_2 \geq \lambda_1 \text{ and } 2\epsilon_1 + \epsilon_2(1 - \lambda_1)/\lambda_1 < \Omega^1; \\ (1 - \lambda_2)\Omega^2 + \lambda_2(\epsilon_1 + (2\epsilon_2 - \Omega^2)(1 - \lambda_1)/\lambda_1) & \text{if } \lambda_2 < \lambda_1 \text{ and } 2\epsilon_2 + \epsilon_1\lambda_1/(1 - \lambda_1) > \Omega^2; \\ (1 - \lambda_2)(2\epsilon_2 + \epsilon_1\lambda_1/(1 - \lambda_1)) & \text{if } \lambda_2 < \lambda_1 \text{ and } 2\epsilon_2 + \epsilon_1\lambda_1/(1 - \lambda_1) < \Omega^2. \end{cases}$$

⁸A proof is available upon request.

⁹Upon request, an Excel file is available to compute utility values in this section.

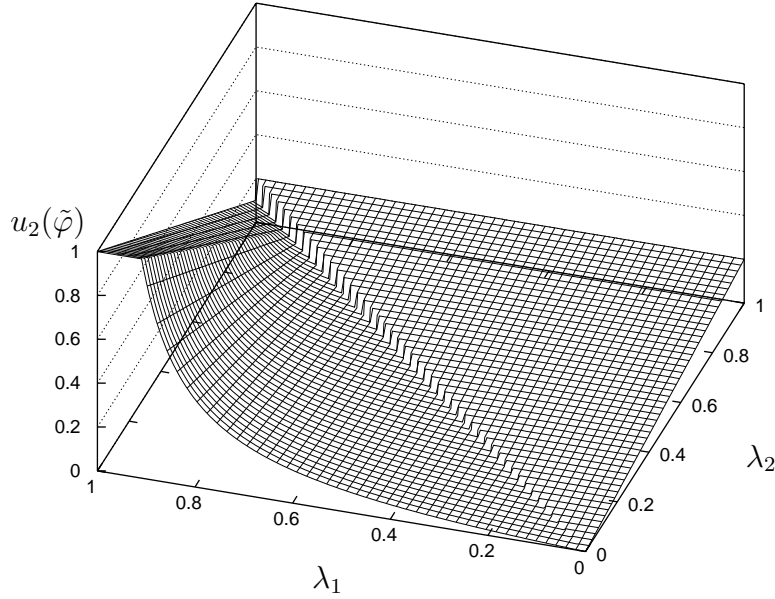


Figure 7: Utility to agent 2 from the rule $\tilde{\varphi}$, when $\epsilon_1 = 0.1$. Higher λ_i indicates higher relative preference toward good 1.

Figure 8 graphs agent 2's utility under this rule when $\epsilon_1 = \epsilon_2 = .1$ and $\Omega = (1, 1)$. The average value is approximately .34.

7. OTHER DOMAINS

So far we have restricted attention to exchange economies with only two goods. The definition given in Section 3 can be generalized in the obvious way when there are $k > 2$ goods, with respect to some $\epsilon \in \mathbb{R}_+^k$. Furthermore, the definition can be applied to any domain of preferences, instead of just the domain of linear preferences, \mathcal{R} . In this section, we briefly mention some implications of such generalizations.

7.1. More goods

For economies with k goods (and linear preferences), the natural definition of our condition is as follows.

ϵ -strategy-proofness: For any $\epsilon \in \mathbb{R}_+^k$, a rule is ϵ -strategy-proof if for all

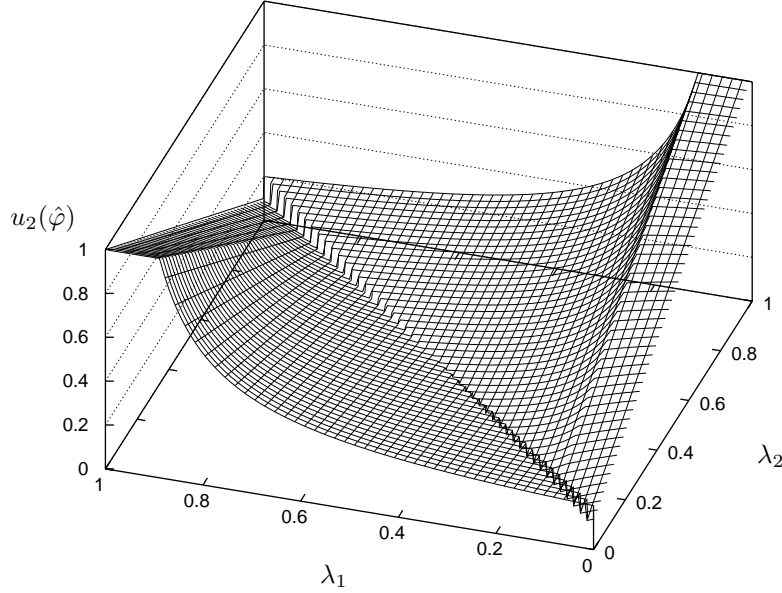


Figure 8: Utility to agent 2 from the rule $\hat{\varphi}$, when $\epsilon_1 = \epsilon_2 = 0.1$.

$R \in \mathcal{R}^2$, all $i \in \{1, 2\}$, and all $R'_i \in \mathcal{R}$, we have for some j , $1 \leq j \leq k$,

$$\varphi_i(R) + (0, \dots, 0, \epsilon_j, 0, \dots, 0) R_i \varphi_i(R'_i, R_{-i}).$$

Considering the case of $\epsilon = (\epsilon_1, 0, \dots, 0)$, the rule $\tilde{\varphi}$ described in Section 4.2 can be generalized as follows. Re-define x, x' to be the allocations such that $x_1 = \Omega - (\epsilon_1, 0, \dots, 0)$ and $x'_1 = \Omega - (2\epsilon_1, 0, \dots, 0)$; redefine $y(R_1, R_2)$ to be the set of efficient allocations that agent 1 considers indifferently to x . Then, $\tilde{\varphi}$ (defined as before with respect to x, x') is efficient and $(\epsilon_1, 0, \dots, 0)$ -strategy-proof.

Furthermore, it is clear that for this case, results analogous to Theorems 2 and 3 can be obtained, showing this generalization of $\tilde{\varphi}$ to be a “most equitable” such rule on this domain. With an investment in additional notation, these analogous results can be obtained from the original ones in the same way Schummer (1997) extends the results for 2-agent/2-good economies to multiple-good economies. For brevity, we omit this notationally tedious task.

7.2. Other preferences

When agents' preferences may be other than linear, some of the previous results can be easily extended. For instance, suppose agents may have any quasi-linear preference relation over, say, two goods, i.e., preferences represented by a utility function of the form $u(x^1, x^2) = x^1 + v(x^2)$ for some concave $v(\cdot)$. In this case, $\tilde{\varphi}$ can be generalized by letting $y(R_1, R_2)$ be any efficient allocation that agent 1 considers indifferently to x (with ties broken arbitrarily). This generalization chooses the efficient point that agent 1 considers indifferently to x , unless that point would be x itself, in which case x' is chosen.

Obviously the generalization is efficient. To see that it is $(\epsilon_1, 0)$ -*strategy-proof*, note that if $y(R_1, R_2)$ is chosen, then agent 1 cannot possibly gain enough to violate the condition; agent 2 weakly prefers $y_2(R_1, R_2)$ to x_2 , so cannot prefer x'_2 to $y_2(R_1, R_2) + (\epsilon_1, 0)$. If x' is chosen, similar arguments apply.

Finally, consider a “standard economic” domain of all convex, strictly monotonic, continuous preferences.¹⁰ A standard result in the *strategy-proofness* literature is that an increase in the domain of preferences can only make implementability more difficult. Indeed, it is trivial to show that if an efficient, $(\epsilon_1, 0)$ -*strategy-proof* rule is defined on any domain larger than \mathcal{R} , then *whenever both agents' preferences are linear*, a prespecified agent must receive almost all of the endowment of good 1, as in Theorem 2.

It turns out to be more difficult to extend this bound to the entire domain of convex, monotonic preferences. While we are unable to show that it is tight, we provide the following bound over this larger domain.

THEOREM 4. *Suppose φ is an efficient, $(\epsilon_1, 0)$ -strategy-proof rule defined over the domain of convex, strictly monotonic, continuous preferences. Then for some $i \in N$, for any profile (R_1, R_2) in that domain, we have $\varphi_i^1(R_1, R_2) \geq \Omega^1 - 3\epsilon_1$.*

Proof. Denote any superdomain of the linear preference domain by $\mathcal{R}^* \supseteq \mathcal{R}$, and let φ be an efficient, $(\epsilon_1, 0)$ -*strategy-proof* rule for that domain.

Note that the restriction of φ to the subdomain \mathcal{R} defines an efficient, $(\epsilon_1, 0)$ -*strategy-proof* rule for that domain. Therefore, Theorems 2 and 3

¹⁰A formal definition of this type of domain is omitted.

apply to that restriction. Therefore, and without loss of generality, suppose that whenever, $(R_1, R_2) \in \mathcal{R}^2$, we have $\varphi_1^1(R_1, R_2) \geq \Omega^1 - 2\epsilon_1$. We show that for all $(R_1, R_2) \in (\mathcal{R}^*)^2$, $\varphi_1^1(R_1, R_2) \geq \Omega^1 - 3\epsilon_1$.

Step 1: for all $R_2 \in \mathcal{R}^*$ and all $R_1 \in \mathcal{R}$, $\varphi_1(R_1, R_2) R_1 x'_1$.

To see this, suppose by contradiction that $x'_1 P_1 \varphi_1(R_1, R_2)$. Then there exists a linear preference ordering $R'_2 \in \mathcal{R}$ (sufficiently “close” to R_1) such that both $E(R_1, R'_2) = E^\perp$ and $y_2(R_1, R_2) + (\epsilon_1, 0) P'_2 \varphi_2(R_1, R_2)$ (where $y(\cdot)$ is defined as in Example 1). However, Theorem 3 implies $y_2(R_1) R'_2 \varphi_2(R_1, R'_2)$, contradicting the fact that φ is $(\epsilon_1, 0)$ -strategy-proof.

This step implies that by reporting a sufficiently “flat” preference relation, agent 1 can obtain a bundle that “almost” vector-dominates x'_1 . Intuitively, this leads to the conclusion that he can never be more than “ ϵ_1 -worse off” than x' , as we formalize in Step 2.

Step 2: For all $R_1, R_2 \in \mathcal{R}^*$, $\varphi_1(R_1, R_2) + (\epsilon_1, 0) R_1 x'_1$.

To see this, suppose by contradiction the opposite. Then there exists (sufficiently flat) $R'_1 \in \mathcal{R}$ such that for any allocation z , we have

$$[z_1 R'_1 x'_1] \implies [z_1 P_1 \varphi_1(R_1, R_2) + (\epsilon_1, 0)].$$

By Step 1, this is a contradiction: agent 1 can successfully manipulate by declaring R'_1 .

The conclusion of the Theorem follows from Step 2 and monotonicity. \square

We leave it as an open question whether a tighter bound exists over that entire domain, and whether the discontinuity result of Section 4.2 extends.

REFERENCES

- ABREU, D. AND H. MATSUSHIMA (1992). “Virtual Implementability in Iteratively Undominated Strategies: Complete Information,” *Econometrica*, **60**, 993–1008.
- BARBERÀ, S. (2001). “An Introduction to Strategy-proof Social Choice Functions,” *Social Choice and Welfare*, **18**, 619–653.
- BARBERÀ, S. AND M. JACKSON (1995). “Strategy-proof Exchange,” *Econometrica*, **63**, 51–87.

- BARBERÀ, S. AND B. PELEG (1990). “Strategy-proof Voting Schemes with Continuous Preferences,” *Soc Choice Welfare*, **7**, 31–8.
- BEVIÁ, C. AND CORCHÓN, L. (1995). “On the Generic Impossibility of Truthful Behavior: A Simple Approach,” *Econ Theory* **6**, 365–71.
- CÓRDOBA, J. AND P. HAMMOND (1998). “Asymptotically Strategy-Proof Walrasian Exchange,” *Mathematical Social Sciences*, **36**, 185–212.
- DUGGAN, J. (1997). “Virtual Bayesian Implementation,” *Econometrica*, **65**, 1175–99.
- EHLERS, L., H. PETERS, AND T. STORCKEN (1999). “Threshold Strategy-proofness,” mimeo, University of Maastricht.
- HARRISON, G. AND T. MCDANIEL (2002). “Voting Games and Computational Complexity,” mimeo, Univ. of South Carolina.
- HURWICZ, L. (1972). “On Informationally Decentralized Systems,” in *Decision and Organization*, ed. by C. B. McGuire and R. Radner. Amsterdam: North Holland, 297–336.
- KALAI, E. (2002). “Large Robust Games,” CMS-EMS working paper 1350, Northwestern University.
- KELLY, J. (1993). “Almost all Social Choice Rules are Highly Manipulable, but a Few Aren’t,” *Social Choice and Welfare*, **10**, 161–75.
- MOORE, J. (1996). “Implementation, Contracts, and Renegotiation in Environments with Complete Information,” in *Advances in Economic Theory*, ed. by J.-J. Laffont. Cambridge: Cambridge University Press, 182–282.
- ROBERTS, D. AND A. POSTLEWAITE (1976). “The Incentives for Price-taking Behavior in Large Exchange Economies,” *Econometrica*, **44**, 113–27.
- RUSTICHINI, A., M. SATTERTHWAITTE, AND S. WILLIAMS (1994). “Convergence to Efficiency in a Simple Market with Incomplete Information,” *Econometrica*, **62**, 1041–63.
- SAARI, D. (1995). *Basic Geometry of Voting*, Springer, Berlin.
- SATTERTHWAITTE, M. (2001). “Strategy-Proofness and Markets,” *Social Choice and Welfare*, **18**, 37–58.
- SCHUMMER, J. (1997). “Strategy-proofness versus Efficiency on Restricted Domains of Exchange Economies,” *Social Choice and Welfare*, **14**, 47–56.

- SERIZAWA, S. AND J. WEYMARK (2002). "Efficient Strategy-Proof Exchange and Minimum Consumption Guarantees," *Journal of Economic Theory*, **109**, 246–263.
- SMITH, D. (1999). "Manipulability Measures of Common Social Choice Functions," *Social Choice and Welfare*, **16**, 639–61.
- SWINKELS, J. (2001). "Efficiency of Large Private Value Auctions," *Econometrica*, **69**, 37–68.
- THOMSON, W. (1998). "Strategy-proof Allocation Rules," mimeo, University of Rochester.
- ZHOU, L. (1991). "Inefficiency of Strategy Proof Allocation Mechanisms in Pure Exchange Economies," *Social Choice and Welfare*, **8**, 247–54.