Almost-dominant Strategy Implementation

James Schummer*

MEDS Department, Kellogg Graduate School of Management

Northwestern University, Evanston, IL 60208–2009

November 8, 1999

Abstract

Though some economic environments provide allocation rules that are implementable in dominant strategies (strategy-proof), a significant number of environments yield impossibility results. On the other hand, while there are quite general possibility results regarding implementation in Nash or Bayesian equilibrium, these equilibrium concepts make strong assumptions about the knowledge that players possess, or about the way they deal with uncertainty. As a compromise between these two notions, we propose a solution concept built on one premise: Players who do not have much to gain by manipulating an allocation rule will not bother to manipulate it.

We search for efficient allocation rules for 2-agent exchange economies that never provide players with large gains from cheating. Though we show that such rules are very inequitable, we also show that some such rules are significantly more flexible than those that satisfy the stronger condition of strategy-proofness.

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

JEL Classification Numbers: C70, D70.

*Email: `schummer@nwu.edu`.

1 Introduction

In the field of mechanism design, one of the most desirable incentives properties for a choice rule to possess is that of dominant strategy implementability. For a planner attempting to implement a rule with this property, certain issues—such as what information he 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—are irrelevant. These issues are also irrelevant for a participating agent calculating an appropriate (best) action to take. In fact, even the assumption that his fellow players are rational need not be made by the player concerned with his own best interests. Furthermore, calculating the appropriate action need not be more complex for a player than the act of determining his own preferences over outcomes.

Given the extreme desirability of this incentives property, it is an important question to determine the situations for which rules exhibiting the property exist. Indeed this question has been—and continues to be—answered for an increasingly diverse class of situations. Interestingly, though, the nature of the result depends strongly on the situation being described.

For example, the seminal works of Gibbard (1973) and Satterthwaite (1975) provided an early negative result for the situation of voting: no (non-dictatorial) voting rule satisfies even the slightly weaker condition of *strategy-proofness*, requiring truth-telling to be *a* weakly dominant strategy in the direct revelation mechanism. For the situation of choosing public alternatives and taxation levels, the Vickrey–Clarke–Groves mechanisms have been shown (by Green and Laffont (1977), Holmström (1979)) to be the only ones satisfying *strategy-proofness* that choose efficient public alternatives. Since these mechanisms are typically not budget balancing, this has been seen as a negative result.[1]

In contrast, in certain "simpler" situations, positive results have prevailed. Moulin (1980) describes the class of *strategy-proof*, onto voting rules for the situation in which agents have "single-peaked" preferences over an ordering

[1]Interestingly, this class of mechanisms has a much better reputation in private goods environments!

of public alternatives, generalizing the classic median-voter rule (also, see Ching (1997). For two classes of 2-sided matching problems (known as marriage markets and college admissions problems), Alcaldé and Barberà (1994) provide a domain of preferences for which certain stable matching rules are *strategy-proof*. In such problems, stability is arguably the most important property for a rule to possess. For 1-sided matching problems—in particular, Shapley and Scarf's (1974) "housing market"—Roth (1982) shows that the allocation rule central to the previous analysis of this domain—the Top Trading Cycles algorithm—is *strategy-proof*. For situations in which agents have single-peaked preferences over consumption of a single divisible private good, Sprumont (1991) shows the *strategy-proofness* of the Uniform Rule, which has subsequently been characterized in terms of many other desirable properties; see Ching (1992,1994), Schummer and Thomson (1997), and Thomson (1994a,b,1995).

As the literature on *strategy-proofness* grows, our picture of the dividing line between possibility and impossibility becomes clearer.[2] This leaves us with the need to address those situations in which no reasonable rules are implementable in dominant strategies. There are various ways to do this.

One approach is to require a weaker form of implementation. This is the approach taken in the large literature on *Nash implementation* (and its refinements), in which mechanisms have the property that their equilibrium outcomes are ones that would have been chosen by some given choice rule. For example, see Moore (1996). The results here tend to be more positive than those in the *strategy-proofness* literature. However, these results come with a price: Strong assumptions are made concerning the structure of information that agents possess (e.g., that players have common knowledge of each other's preferences, or that they have common priors).

A second 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. One application of this approach can be seen in the literature on *virtual implementation* (Abreu and Matsushima (1992), Duggan (1997)), in which the goal is to find an imple-

[2] For a more detailed survey of the *strategy-proofness* literature, including more positive results, see Thomson (1998).

mentable mechanism whose equilibrium outcomes approximate the desired outcomes.

Another application of the "approximation approach" is to measure, in some way, the manipulability of a mechanism. In fact, there are various ways of performing *this* analysis. Roberts and Postlewaite (1976) observe that as the number of agents becomes large, the Walrasian allocation rule is asymptotically *strategy-proof* (also see Córdoba and Hammond (1998) and Ehlers et al. (1999)). A similar analysis in an auction setting is performed by Rustichini, Satterthwaite, and Williams (1994), who not only show an analogous asymptotic result for double auctions, but also argue that convergence happens quickly as the number of agents increases.

Alternatively, Beviá and Corchón (2) show that in a public goods setting, any efficient and individually rational mechanism is manipulable on a dense set of preference profiles. Kelly (1993) and Smith (1999) suggest ways of counting manipulable situations in a discrete voting environment.

Our Approach

The approach taken in this paper can be seen as a different type of contribution to this approximation approach, involving an approximation to the notion of a dominant strategy. The motivation behind our approximation lies with a simple assumption about the strategic behavior of agents. Specifically, we will approach the problem with the premise that if a player does not have much to gain by manipulating an allocation rule, then he will not bother to manipulate it.

This assumption can be interpreted (or applied) in various ways. For example, it applies when agents have a relatively high cost of gathering information about each other. If such information is necessary for the player to compute a profitable way to manipulate the choice rule, it may not be worth the expense to gather it. Similarly, another application of this idea is to situations in which computation itself is costly to a player. Thirdly, for situations in which agents value morality (or honesty) to some degree, the small gains from cheating may not outweigh the losses ("guilt") incurred.

An important observation to make is that we make no assumption on the structure of information that agents possess.

The hypothesis of our premise is that "a player does not have *much* to gain by manipulating." The critical detail of our work is to precisely define *much*. One approach is to use a utility-based approach to preferences. Using this approach, a player would be assumed not to manipulate a rule unless his utility gain would exceed a predetermined amount. This approach, however, would depend heavily on the interpretation of utility and would imply certain interpersonal comparisons.

To avoid this difficulty, we use a different approach, which applies to situations in which there exists a transferable, divisible, private good (e.g., money). With respect to such a good (which we call the numeraire good), our behavioral assumption is that the only situations in which an agent will manipulate a choice rule are when his gains are equivalent to receiving (at least) a prespecified, additional amount (say, ϵ) of the numeraire good. If no such situation exists, we say that the choice rule provides truth-telling as ϵ-dominant.

In this paper, we search for such rules on a simple class of economies: 2-agent exchange economies with two goods. Furthermore, we restrict attention to the domain of linearly additively separable preferences. Aside from the fact that this is the first work on this project, There is another reason we restrict attention to such a simple class. It turns out that it is very straightforward to measure the gains in possibility for this domain as the truth-telling condition is relaxed (i.e., as ϵ is increased), as we discuss below.

One of the earlier works on mechanism design is a paper by Hurwicz (1972), concerning 2-agent exchange economies with a more general domain of preferences. He shows that it is impossible to construct a *strategy-proof*, efficient rule that provides allocations which both agents prefer to their original endowment. Zhou (1991) improves upon this result by showing that if a rule is *strategy-proof* and efficient, then it is dictatorial: it must always give all of the goods to a prespecified agent. Finally, Schummer (1997) strengthens these results by showing them to hold even on "small" domains of preferences, including the linear preferences we use here.

Our results show that when *strategy-proofness* is weakened to the condition that truth-telling be ϵ-dominant, a larger class of rules becomes admissible. The result, however, has both a negative and a positive flavor. First, it must be the case that such a rule always allocates almost all of the numeraire good to a prespecified agent. On the other hand, allocation of the non-numeraire good may range from giving it all to one agent to giving it all to the other agent. This flexibility is in strong contrast to the negative results with respect to *strategy-proofness*.

As a second positive interpretation of our results, we show that as the truth-telling condition approaches *strategy-proofness* (i.e., as ϵ converges to zero), the set of ranges of the set of admissible rules does *not* converge to the range of the only *strategy-proof*, efficient rules. That is, even if *strategy-proofness* is relaxed an arbitrarily small amount, there is a (relatively) large increase in the flexibility of admissible rules.

The paper is organized as follows. Section 2 formalizes the model. Section 2.1 defines the truth-telling condition. Section 3 provides the first result, demonstrating a bound on the range of a rule. Section 4 describes the "least dictatorial" rule satisfying our conditions. Section 8 concludes.

2 Model

The set of two agents is $N = \{1, 2\}$.[3] 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 $x_1 + x_2 = \Omega$. Subscripts refer to agents and superscripts refer to goods. The set of allocations is denoted A. The vector inequalities are $<$, $\geq$, and $\geqq$.

Each agent has a strictly monotonic, linear preference relation, R_i, over his consumption space $\mathbb{R}^2_+$.[4] Denote the set of such preference relations as $\mathcal{R}$. The strict (antisymmetric) preference relation and indifference (symmetric) relation associated with R_i are denoted P_i and I_i.

An *allocation rule* is a function $\varphi\colon \mathcal{R}^2 \to A$ mapping the set of preference

[3]Notation primarily follows Schummer (1997).

[4]Such preference relations are the ones representable by a utility function of the form $u(x_i) = ax_i^1 + bx_i^2$, $a, b > 0$.

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 relation 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 $E(R)$. 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, $E(R) = E^{\ulcorner} \equiv \{x \in A : x_1^1 = 0 \text{ or } x_2^2 = 0\}$.

2.1 A Definition of Nonmanipulability

Our goal is to find allocation rules that never afford agents the opportunity to gain much. The difficulty in formalizing this notion is to define *much*, for all possible preference relations.

Our approach will be to restrict attention to measures on the consumption space.[5] There are many ways to construct such measures of gains (and we will address more of them in the Conclusion). A simple one is to consider one of the two goods as a numeraire good, in terms of which gains can be measured. Such a measure is especially appropriate when one of the goods is to be interpreted as a medium of exchange (e.g., money). For the remainder, we interpret good 1 as such a numeraire good.

Consider a situation in which an allocation rule φ prescribes for a given preference profile, $R \in \mathcal{R}^2$, an allocation $x = \varphi(R)$. If agent i would not manipulate the rule for small gains, then there exists some $\epsilon > 0$ such that

[5] Alternatively, one could measure gains from manipulation in terms of some utility measure. However, we wish to avoid the implicit assumptions imposed by such modeling.

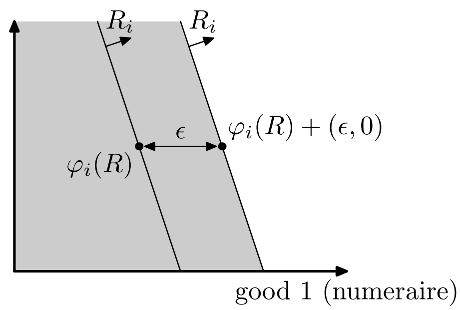

Figure 1: If agent i has no ϵ-improvement, then he can only obtain bundles in the shaded area.

if for some $R'_i \in \mathcal{R}$, $\varphi_i(R'_i, R_{-i}) = \varphi_i(R) + (\epsilon, 0)$, then agent i would not manipulate the rule with that particular misrepresentation R'_i. With similar reasoning, if for some $R''_i \in \mathcal{R}$, it is *not* the case that $\varphi_i(R''_i, R_{-i})\ P_i\ \varphi_i(R) + (\epsilon, 0)$, then agent i would not misreport R''_i. In Figure 1, agent 1 cannot gain *much* at the profile R if, no matter how he misreports his preferences, the bundle he receives lies in the shaded area.[6]

Formally, under an allocation rule φ, agent i has an *ϵ-improvement at* $R \in \mathcal{R}^2$ if there exists $R'_i \in \mathcal{R}$ such that $\varphi(R'_i, R_{-i})\ P_i\ \varphi_i(R) + (\epsilon, 0)$. We say that φ *provides truth-telling as ϵ-dominant* if there exist no $i \in N$ and $R \in \mathcal{R}^2$ such that agent i has an ϵ-improvement at R.

Of course as a special case, *strategy-proofness* is equivalent to providing truth-telling as ϵ-dominant when $\epsilon = 0$.

3 A Bound on the Range

As shown by Schummer (1997) for this class of exchange economies with linear preferences, the only rules that are both *efficient* and *strategy-proof* are those that always allocate all of the endowment to a prespecified agent (i.e., dictatorial rules). By relaxing the *strategy-proofness* condition to the requirement that truth-telling be only ϵ-dominant, the class of admissible allocation rules is enlarged, as we show further below in Example 1. Our first

[6] In the language of Barberà and Peleg (1), his *option set* should lie within the shaded area.

result shows, however, that a prespecified agent always must receive nearly all of the endowment of the numeraire good. On the other hand, allocation of the second good may vary from being given entirely to the prespecified agent to being given entirely to the other agent, as in Example 1.

THEOREM 1 *Let* φ *be an* efficient *rule that provides truth-telling as* ϵ*-dominant, where* $\epsilon < \Omega^1/5$*. There exists an agent* $i \in N$ *that always receives almost all of the numeraire good: for all* $R \in \mathcal{R}^2$, $\varphi_i^1(R) \geq \Omega^1 - 2\epsilon$.

To prove the result, we first provide the following lemma, which essentially states that for all preference profiles with the same set of efficient allocations, the chosen allocations are not much different in terms of the numeraire good.

LEMMA 1 *Let* φ *be* efficient *and provide truth-telling as* ϵ*-dominant. For all* $R, R' \in \mathcal{R}^2$*, if either* $E(R) = E(R') = E^{\ulcorner}$ *or* $E(R) = E(R') = E^{\lrcorner}$*, then* $|\varphi(R)_1^1 - \varphi(R')_1^1| \leq 2\epsilon$.

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

By *efficiency*, $\varphi(R_1, R_2') \in E^{\ulcorner}$. Since truth-telling is ϵ-dominant and $\varphi(R_1', R_2') \in E^{\ulcorner}$, we have $\varphi_1^1(R_1, R_2') - \varphi_1^1(R_1', R_2') \leq \epsilon$. Similarly, $\varphi_1^1(R_1', R_2') - \varphi_1^1(R_1, R_2') \leq \epsilon$, so

$$|\varphi_1^1(R_1, R_2') - \varphi_1^1(R_1', R_2')| \leq \epsilon \tag{1}$$

By the same argument, we have

$$|\varphi_2^1(R_1', R_2) - \varphi_2^1(R_1', R_2')| \leq \epsilon$$

implying

$$|\varphi_1^1(R_1', R_2) - \varphi_1^1(R_1', R_2')| \leq \epsilon \tag{2}$$

Therefore by the triangle inequality,

$$|\varphi_1^1(R_1, R_2) - \varphi_1^1(R_1', R_2')| \leq 2\epsilon$$

proving the result. $\square$

Now we can prove the theorem.

Proof of Theorem 1: Let φ be *efficient* and provide truth-telling as ϵ-dominant. There are three possible cases.

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

Step 1a: ($E^{\lrcorner}$) In this case, for all $\delta > 0$, there exists $R \in \mathcal{R}^2$ such that $E(R) = E^{\lrcorner}$ and $\varphi_1(R) \geqq (\Omega^1, \Omega^2 - \delta)$. To see this, let R_1 satisfy $(0, \Omega^2)$ P_1 $(\Omega^1 + \epsilon, \Omega^2 - \delta)$, let R_2 be such that $E(R) = E^{\lrcorner}$, and let R_1' be such that $E(R_1', R_2) = E^{\ulcorner}$. Since truth-telling is ϵ-dominant and $E(R_1', R_2) = E^{\ulcorner}$,

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

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

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

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

$$\begin{aligned}
y_1\ &I_1\ \varphi_1(R) + (\epsilon + \delta/3, 0) \\
y_1'\ &I_1\ \varphi_1(R) + (\epsilon + 2\delta/3, 0) \\
y_1''\ &I_1\ \varphi_1(R) + (2\epsilon + 2\delta/3, 0)
\end{aligned}$$

Let R_2' be such that $y_2\ I_2'\ \varphi_2(R) - (\epsilon, 0)$. Since $\varphi(R_1, R_2') \in E^{\lrcorner}$, the truth-telling condition implies $\varphi_2(R_1, R_2') \geqq y_2$. Let R_2'' be sufficiently flat so that both $y_2''\ P_2''\ (\Omega^1 + \epsilon, 0)$ and $y_2' + (\epsilon, 0)\ P_2''\ y_2$. The truth-telling condition implies $\varphi_2(R_1, R_2'') + (\epsilon, 0)\ R_2''\ \varphi_2(R_1, R_2')$, so $\varphi_2(R_1, R_2'') \geq y_2'$.

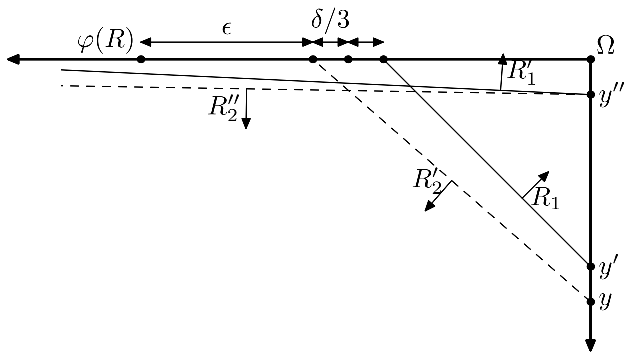

Figure 2: Proof of Theorem 1. The figure represents the upper-right corner of the Edgeworth Box.

Let R_1' satisfy $(0, \Omega^2)\ I_1'\ y_1'' + (\epsilon, 0)$. Then $E(R_1', R_2'') = E^{\lrcorner}$. Note that by construction, $y_1' + (\epsilon, 0)\ I_1\ y_1''$. The truth-telling condition implies $\varphi_1(R_1, R_2'') + (\epsilon, 0)\ R_1\ \varphi_1(R_1', R_2'')$. Therefore $\varphi_1(R_1', R_2'') \leq y_1''$.

By the hypothesis of Case 1, for all R_1'' such that $E(R'') = E^{\ulcorner}$, we have $\varphi_1(R'') \geqq (0, \Omega^2)$. But then for any such R_1'', we have $\varphi_1(R'')\ P_1'\ \varphi_1(R_1', R_2'') + (\epsilon, 0)$, which contradicts the truth-telling condition.

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

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

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

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

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

$$\varphi_1^1(R') - \varphi_1^1(R) > \epsilon \tag{3}$$

Let R_1 be such that $(2\epsilon, \Omega^2)\ P_1\ (\Omega^1 - 3\epsilon, 0)$. Let R_2, R_1' be such that $E(R) = E^{\ulcorner}$ and $E(R_1', R_2) = E^{\lrcorner}$. Then eqn. (3) implies $\varphi_1(R_1', R_2)\ P_1\ \varphi_1(R) + (\epsilon, 0)$, which contradicts the truth-telling condition. Therefore this case cannot hold. □

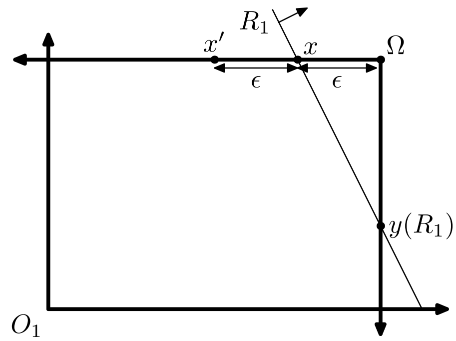

Figure 3: Constructing a non-dictatorial, *efficient* rule that provides truth-telling as ϵ-dominant.

4 An Important Rule

Theorem 1 states that under an *efficient* rule that provides truth-telling as ϵ-dominant, one agent always must receive at least $\Omega^1 - 2\epsilon$ of the numeraire good. The rule described below in Example 1 shows that (i) this bound is tight, and (ii) there is no such bound corresponding to the other (non-numeraire) good. That is, Agent 1 receives (i) from as little as $\Omega^1 - 2\epsilon$ of the numeraire good to as much as all of it, and (ii) from as little as none of the other good to as much as all of it.

Furthermore, and most importantly, this rule is unambiguously the "least dictatorial" (or most equitable) of all *efficient* rules providing truth-telling as ϵ-dominant. This statement is made more precise below as Theorem 2.

Example 1 Fix the allocations $x = ((\Omega^1 - \epsilon, \Omega^2), (\epsilon, 0))$, which gives agent 1 the entire endowment except for ϵ units of good 1, and $x' = ((\Omega^1 - 2\epsilon, \Omega^2), (2\epsilon, 0))$. For all $R_1 \in \mathcal{R}$, let $y(R_1) \in E^{\lrcorner}$ be the unique allocation in $E^{\lrcorner}$ that agent 1 considers indifferently to x (as in Figure 3), i.e., $x_1 \; I_1 \; y_1(R_1)$. Then for all $R \in \mathcal{R}^2$, let

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

We leave it to the reader to check that $\tilde{\varphi}$ is *efficient* and provides truth-telling as ϵ-dominant.

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 discussion of welfare appears in Section 5. The statement of Theorem 1 does not, by itself, rule out more equitable rules. As the next theorem shows, however, $\tilde{\varphi}$ is the most equitable *efficient* rule that provides truth-telling as ϵ-dominant. Under any other such rule, say φ, one of the two agents—specifically, the agent that does not always receive almost all of the numeraire good—would prefer the bundle Agent 2' receives under $\tilde{\varphi}$ to what this agent receives under φ, *regardless of the profile of preferences.*

THEOREM 2 *The rule $\tilde{\varphi}$ is the most equitable* efficient *rule that provides truth-telling as ϵ-dominant. Specifically, let φ be an* efficient *rule that provides truth-telling as ϵ-dominant, where $\epsilon < \Omega^1/5$. There exists an agent $i \in N$ such that for all $R \in \mathcal{R}^2$, $\tilde{\varphi}_2(R)\ R_i\ \varphi_i(R)$.*

Proof: Let φ be an *efficient* rule that provides truth-telling as ϵ-dominant. Suppose without loss of generality that Agent 1 always receives at least $\Omega^1 - 2\epsilon$ of the numeraire good. We show that for all $R \in \mathcal{R}^2$, $\tilde{\varphi}_2(R)\ R_2\ \varphi_2(R)$.

If $E(R) = E^{\ulcorner}$, the conclusion follows from Theorem 1, since $\varphi_2(R) \leqq (2\epsilon, 0) = \tilde{\varphi}_2(R)$.

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

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

Letting $y = \varphi(R)$ and $R_2' = R_2$, and defining y', y'', R_1', and R_1'' as in Case 1, Step 1b in the Proof of Theorem 1, leads to a contradiction as in that proof.

□

5 A Measure of Welfare

Theorem 2 provides a lower bound on the welfare of an agent under an *efficient* rule providing truth-telling as ϵ-dominant. In order to have a better understanding of exactly how well-off Agent 2 is when using 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

$$u(x_i) = \lambda_i x_i^1 + (1 - \lambda_1) x_i^2$$

We consider the case in which $\Omega = (1, 1)$. Therefore an agent's utility is always equal to one when receiving the entire endowment, and is equal to zero when receiving nothing. In particular, a utility level can be interpreted as a proportion of the entire endowment, that is, $u(\delta\Omega^1, \delta\Omega^2) = \delta$.

Under the rule $\tilde{\varphi}$, Agent 2's utility is a function of λ_1, λ_2, and ϵ. It is a straightforward geometric exercise to derive Agent 2's utility under $\tilde{\varphi}$:[7]

$$u_2(\lambda_1, \lambda_2, \epsilon, \omega^2) = \begin{cases} 2\lambda_2\epsilon & \text{if } \lambda_2 \geq \lambda_1 \\ (1 - \lambda_2)(\epsilon\lambda_1)/(1 - \lambda_1) & \text{if } \lambda_2 < \lambda_1 \leq \omega^2/(\omega_2 + \epsilon) \\ \lambda_2\epsilon + \omega_2(1 - (\lambda_2/\lambda_1)) & \text{otherwise} \end{cases}$$

Figure 4 shows the utility of Agent 2 when $\epsilon = 0.1$. We see that Agent 2 receives a non-negligible amount of utility at many profiles. The average utility that Agent 2 receives over the entire range of values for (λ_1, λ_2) is approximately 0.18.[8] This is significantly higher than the average utility Agent 2 would receive from a constant $\epsilon = 0.1$ units of the numeraire good, which would be 0.05. By the previously mentioned result of Schummer (1997), if *strategy-proofness* were required, one agent 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.

[7]Proof available upon request.

[8]Values were calculated with Microsoft Excel for values of λ_1 and λ_2 ranging from 0.01 to 0.99 in increments of .01.

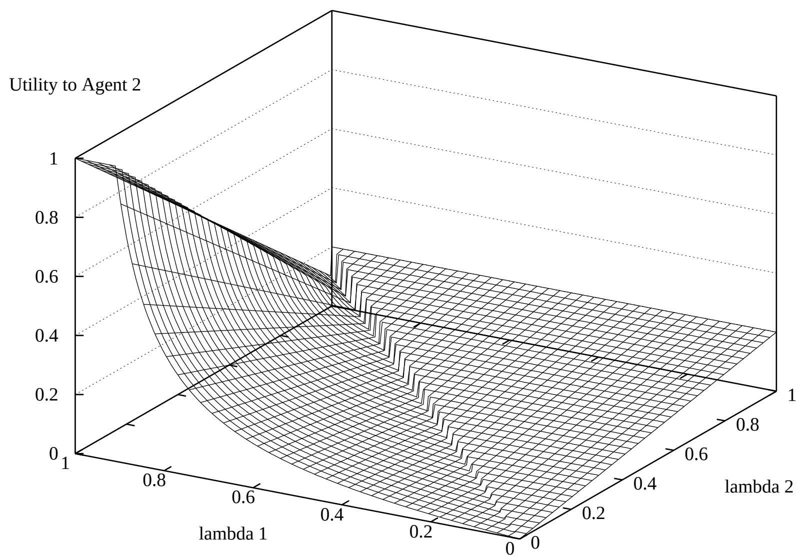

Figure 4: Utility to Agent 2 from the rule $\tilde{\varphi}$, when $\epsilon = 0.1$. Lambda i increases with relative preference toward the numeraire good. (Note: to facilitate viewing, the x-axis of this figure is reversed.)

6 The Truth-telling Condition as a Perturbation of Strategy-proofness

To further emphasize the idea that a small relaxation in *strategy-proofness* leads to a large increase in the flexibility of rules, consider the implications of the truth-telling condition (with *efficiency*) as ϵ approaches zero. The rule $\tilde{\varphi}$ was defined in Example 3 with respect to a fixed value of ϵ. The range of this rule, for a given ϵ, is

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

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

Therefore, as the *provision of truth-telling as ϵ-dominant* converges to *strategy-proofness*, the range of admissible rules does *not* converge to the class of *strategy-proof* and *efficient* rules (i.e., dictatorial rules) characterized in Schummer (1997).[9] 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 number of admissible rules. On domains for which impossibility results regarding *strategy-proofness* have been established, relaxing the condition even in a small way may significantly enlarge the class of admissible allocation rules.

7 Other Rules and Other Domains

The rule $\tilde{\varphi}$ can be generalized to yield a larger class of *efficient* rules providing truth-telling as ϵ-dominant. These rules, described in Example 2, are not particularly elegant; they are essentially the same as the rule $\tilde{\varphi}$, except for arbitrary ϵ-perturbations in a direction favoring Agent 1. Recall that by Theorem 2, such perturbations cannot favor Agent 2.

We provide these rules not to suggest their use, but to suggest that a

[9]Formally, this sequence of examples shows that the ranges of the admissible rules is a correspondence that is not upper-semi-continuous at $\epsilon = 0$. It is clearly lower-semi-continuous, because dictatorial rules provide truth-telling as ϵ-dominant for any ϵ.

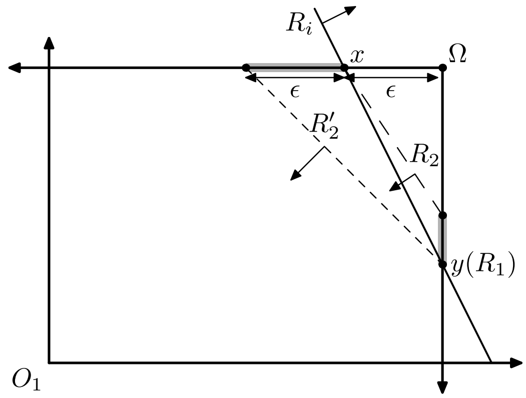

Figure 5: A more complicated, non-dictatorial, *efficient* rule that provides truth-telling as ϵ-dominant. Given R_1, if agent 2's preferences are steeper than R_1, choose a point in the horizontal shaded area. If agent 2's preferences are flatter than R_2', choose $y(R_1)$. Otherwise, choose a point $w(R)$ in the vertical shaded area.

characterization of rules satisfying the truth-telling condition (which we do not provide) may be best attempted using an approximation approach.

EXAMPLE 2 Fix the allocation $x = ((\Omega^1 - \epsilon, \Omega^2), (\epsilon, 0))$, and let $y\colon \mathcal{R} \to E^{\lrcorner}$ be defined as in Example 1. Let $\lambda\colon \mathcal{R}^2 \to [0, \epsilon]$ be an arbitrary function. For preference profiles that have the efficient set $E^{\lrcorner}$, choose an arbitrary function

$$w\colon \{R^2 : E(R) = E^{\lrcorner}\} \to \{x = \delta y(R_1) + (1-\delta)\Omega, \delta \in [0,1]\}$$

that always satisfies $w_2(R)\ R_2\ x_2$. For all $R \in \mathcal{R}^2$, let

$$\varphi(R) = \begin{cases} x - \lambda(R) & \text{if } x \text{ is efficient for } R \\ y(R_1) & \text{if } y_2(R_1)\ R_2\ x + (\epsilon, 0) \\ w(R) & \text{otherwise} \end{cases}$$

If our class of economies is generalized to include three or more goods, our rule $\tilde{\varphi}$ can be generalized in the natural way. Formally, alter the interpretation of the notation of Example 1 so that Ω^2 refers instead to the (vector of) total endowment of all non-numeraire goods, and for all $R \in \mathcal{R}$,

$y(R_1, R_2) \in E(R)$ is any allocation (efficient for R) that Agent 1 considers indifferently to x. Then the redefinition of $\tilde{\varphi}$ is well-defined, efficient, and provides truth-telling as ϵ-dominant.

We conjecture that results analogous to Theorems 1 and 2 can be obtained for the case of multiple goods. With an investment in additional notation, such results should be obtainable in the same way Schummer (1997) extends the results for 2-good economies to multiple-good economies.

Our rule $\tilde{\varphi}$ can also be generalized to the domain of economies in which the two agents may have any (possibly non-linear) convex preference relation.[10] Again using the notation from Example 1, the rule should let Agent 2 decide the final allocation by choosing his favorite from among the allocations Agent 1 considers indifferently to x. In the case that Agent 2 would choose x, however, the rule should allocate x'.

8 Conclusion

We have introduced a relaxation of the notion of dominant strategy implementation by requiring only that truth-telling be an ϵ-dominant strategy. This concept was formalized by defining an ϵ-improvement to be a misrepresentation that provides a gain equivalent to receiving an additional ϵ units of a prespecified numeraire good (which could be thought of as money).

On the simple class of 2-agent exchange economies with two goods, in which agents have linear preference relations, we have provided (in Theorem 1) a bound on the range of any *efficient* rule that provides truth-telling as ϵ-dominant: A prespecified agent must always receive almost all of the numeraire good; the second agent always receives at most 2ϵ units of the good. However, we provide a rule (in Example 1) which varies the allocation of the second good between the two agents to the degree that in some situations, one agent receives all of it, while in some other situations, the other agent receives all of it.

The flexibility of this rule is in stark contrast to the conclusions derived when truth-telling is required to be a dominant strategy, i.e., that the only

[10] We omit a formalization of this common domain.

strategy-proof, *efficient* rule in this context always allocates *all* of the goods to a prespecified agent (Schummer (1997)). Admittedly, the rule provided in Example 1 is not extremely flexible. However, we show (in Theorem 2) that this rule is actually the most equitable of all *efficient* rules that provide truth-telling as ϵ-dominant.

To summarize, the negative interpretation of the results is that the requirement that truth-telling be ϵ-dominant does restrict our choice of rules, at least for this economic environment. This is not surprising given the previous results concerning *strategy-proofness*. The positive interpretation of the results concerns the fact that even a small relaxation of *strategy-proofness* leads to a relatively large increase in the flexibility of rules. This is not only good news for domains with previously established impossibility results regarding *strategy-proofness*, but also for domains in which additional requirements, such as efficiency, may restrict our choice of reasonable allocation rules. The results extend to broader classes of 2-agent exchange economies, as described in Section 7.

There is an additional point that gives these results even more positive flavor. In models with additional agents, the rules satisfying the truth-telling condition may be even more flexible. The 2ϵ-bound of Theorem 1 crucially depends on the fact that there are only two agents. Roughly speaking, two unilateral changes in preferences change the welfare of agents by an amount comparable to at most 2ϵ units of the numeraire good (as in the proof of Theorem 1). 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.[11] This provides hope that for other economic environments for which impossibility results have been obtained regarding *strategy-proofness*, there is reason to consider our weaker truth-telling condition.

[11] This could be seen as a dual result to the asymptotic results of Roberts and Postlewaite (1976).

References

Abreu, D. and H. Matsushima (1992). "Virtual Implementability in Iteratively Undominated Strategies: Complete Information," *Econometrica*, **60**, 993–1008.

Alcaldé, J. and S. Barberà (1994). "Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems," *Economic Theory*, **4**, 417–35.

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., On the Generic Impossibility of Truthful Behavior: A Simple Approach, *Econ Theory* **6**, 365–71 (1995).

Ching, S. (1994). "An Alternative Characterization of the Uniform Rule," *Social Choice and Welfare*, **11**, 131-36.

Ching, S. (1992). "A Simple Characterization of the Uniform Rule," *Economics Letters*, **40**, 57–60

Ching, S. (1997). "Strategy-Proofness and 'Median Voters'," *International Journal of Game Theory*, **26**, 473–90.

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

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.

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

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

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

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

HURWICZ, L. (1972). "On Informationally Decentralized Systems," in *Decision and Organization*, ed. by C. B. McGuire and R. Radner. Amsterdam: North Holland, 297–336.

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.

MOULIN, H. (1980). "On Strategy-proofness and Single-peakedness," *Public Choice*, **35**, 437–55.

ROBERTS, D. AND A. POSTLEWAITE (1976). "The Incentives for Price-taking Behavior in Large Exchange Economies," *Econometrica*, **44**, 113–27.

ROTH, A. (1982). "Incentive Compatibility in a Market with Indivisible Goods," *Economics Letters*, **9**, 127–32.

RUSTICHINI, A., M. SATTERTHWAITE, AND S. WILLIAMS "Convergence to Efficiency in a Simple Market with Incomplete Information," *Econometrica*, **62**, 1041–63.

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

SCHUMMER, J. (1997). "Strategy-proofness versus Efficiency on Restricted Domains of Exchange Economies," *Social Choice and Welfare*, **14**, 47–56.

SCHUMMER, J., AND W. THOMSON (1997). "Two Derivations of the Uniform Rule and an Application to Bankruptcy," *Economics Letters*, **55**, 333-7.

SHAPLEY, L., AND H. SCARF (1974). "On Cores and Indivisibility," *Journal of Mathematical Economics*, **1**, 23–37.

SMITH, D. (1999). "Manipulability Measures of Common Social Choice Functions," *Social Choice and Welfare*, **16**, 639–61.

SPRUMONT, Y. (1991). "The Division Problem with Single-Peaked Preferences: A Characterization of the Uniform Allocation Rule," *Econometrica*, **59**, 509–19.

THOMSON, W. (1994a). "Consistent Solutions to the Problem of Fair Division when Preferences are Single-Peaked," *Journal of Economic Theory*, **63**, 219–45.

THOMSON, W. (1994b). "Resource-monotonic Solutions to the Problem of Fair Division when Preferences are Single-Peaked," *Social Choice and Welfare*, **11**, 205–23.

THOMSON, W. (1995). "Population-monotonic Solutions to the Problem of Fair Division when Preferences are Single-Peaked," *Economic Theory*, **5**, 229–46.

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.