ON THE MANIPULABILITY OF RESOURCE ALLOCATION MECHANISMS DESIGNED TO ACHIEVE INDIVIDUALLY-RATIONAL AND PARETO-OPTIMAL OUTCOMES

by

William Thomson

Discussion Paper No. 79-116, September 1979

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

On the Manipulability of Resource Allocation Mechanisms Designed to Achieve Individually-Rational and Pareto-Optimal Outcomes*

by

William Thomson

September 1979

Given a class of economies E , a performance correspondence Φ defined on E associates to every element e of E a set $\Phi(e)$ of allocations feasible for e . Φ embodies the aspirations of the social planner. The determination of Φ-optimal allocations typically requires information that is initially dispersed throughout the economy whence the need for implementation mechanisms: an implementation mechanism for Φ specifies rules to generate, exchange and process information so as to reach these Φ-optimal allocations.

Let then a performance correspondence Φ be given as well as an implementation mechanism for Φ . By not following the rules of behavior assigned to him by the mechanism, an agent may often influence the final outcomes in his favor. If all agents engage in such manipulation, the allocations eventually reached may be quite different from the ones that the mechanism was designed to achieve. In order to characterize the set of allocations that result from joint manipulation, it is natural to associate to the mechanism a manipulation game, to specify an equilibrium

---

*Presented at the IMSSS Seminar, Stanford, August 1979.

**I am indebted to L. Hurwicz for formulating the conjecture that the results of Thomson [28] concerning the manipulation of the Shapley-value were susceptible of generalization, and for several helpful conversations. Typing assistance was provided by National Science Foundation grant No. S)C-7825734.

concept, and to determine its equilibria. The vulnerability of the mechanism to manipulative behavior can then be evaluated by how different or "far" from the set of Φ-optimal allocations the set of equilibrium allocations of its associated manipulation game is. In this paper, we will consider direct implementation mechanisms in which the information communicated by each agent corresponds in a direct way to his potential characteristics. The manipulation game associated to the direct implementation mechanism associated with a performance correspondence Φ will be referred to as the "Φ-manipulation game" (or "Φ-manipulative" quasi-game for technical reasons elaborated on later).

The descriptive study of the manipulability of various mechanisms has been the object of several papers; in [9], Hurwicz studied the Walras manipulation game in pure-exchange, 2-person, 2-good economies, and characterized its set of equilibrium allocations as the lens-shaped area determined by the agent's true offer curves. This result was extended by Otani and Sicilian [19] to 2-person and n-good economies, and a counterpart of it for the Lindahl mechanism in public good economies was established by Thomson [29]. For both the Walras and the Lindhal mechanisms, most of the equilibrium allocations are not Pareto-optimal. Assuming preferences to be known, manipulation of endowments is shown in Thomson [29] to also lead the Walras and Lindhal mechanism to non-optimal allocations. However, the Walras mechanism enjoys a form of asymptotic cheat-proofness, as established in Roberts and Postlewaite [23].

Kurz [14], [15] showed the existence of dominant strategies in the manipulation game associated to an economy with taxes and an infinite number of players while Crawford and Varian [2] for the one-commodity case found dominant strategies in the manipulation game associated with the Nash bargaining problem; these dominant strategies are linear functions in [2], [14] and homogeneous functions in [15]. Similarly, Kannai [12] discovered that announcing the least concave representation of one's preferences would have dominance property in certain games and Kihlstrom, Roth, and Schmeidler [13] showed that, for a class of bargaining solutions that include the Nash [18] and Raiffa [20], Kalai, and Smorodinsky [11] solutions, agents will report linear utilities; Sobel [27] established that constrained equal-income competitive allocations for the true utilities are always equilibria of the misrepresentation game associated to bargaining solutions in a class containing the two just mentioned. Finally, Thomson [28] explores the manipulation game associated with the Shapley-value and proves that the true constrained Walrasian allocations or constrained Lindahl allocations in private and public good economies respectively will be the only outcomes under a certain smoothness condition on the reported preferences. (The constrained Walrasian [resp. Lindhal] correspondence is an extension of the Walrasian [resp. Lindhal] correspondence which coincides with it in the interior of the feasible set and always yield Pareto-optimal outcomes.)

A second branch of this literature concerns dynamic procedures. Such procedures involve iterated exchanges ofmessages between consumers and producers and a planning authority. In the procedures for the provision of public goods devised by Malinvaud [16] and Drèze and de la Vallée Poussin [3] agents are assumed to announce their true rates of substitution. Roberts [21], [22] studied the effect of manipulative behavior on these procedures by defining a local (or myopic) manipulation game, in which agents announce false rates of substitution in an attempt to maximize the instantaneous rate of increase of their utilities, and showed that the path generated in this way will lead the economy to optimal allocation. This result is extended by Henry [4], [5]. Schoumaker [25] showed that a game similarly defined for the discrete-line procedure devised by Champsaur, Drèze, and Henry [1] would converge to optimality. A particular specification of a planning procedure for economies with private goods only, due to Tulkens and Zamir [30] is also shown by Schoumaker [26] to generate a manipulation game whose local Nash equilibrium converges to optimality.

These results indicate that manipulative behavior does not always have bad consequences, although one should not in general expect to get what was originally desired, as follows from the impossibility theorem of Hurwicz's [8]. In addition, one may sometimes obtain as equilibria of a Φ-manipulation game the outcomes of some other performance correspondence Φ' whose Φ'-manipulation games yield something else still. How general is this phenomena will be studied in this paper. Its main conclusion is that the equilibrium sets of the Φ-manipulation games associated to Φ's that select individually rational and Pareto-optimal outcomes all bear a close relationship to the equilibria of the Walrasian manipulation game, and that there is a large class of Φ's whose Φ-manipulation games yield all and only the constrained Walrasian allocations. The Walrasian performance correspondence is then central to the general study of the manipulability of resource allocation mechanisms.

1. Definitions and Notations: E is the class of transferable utility economies with n agents and m privately appropriable commodities, defined as follows: each agent is indexed by the subscript i in $I = \{1,\ldots,n\}$ and characterized by a list $(X_i,\omega_i,\succsim_i)$ where $X_i = R \times R_+^{m-1}$ is his consumption space, ω_i his initial endowment, a point of X_i and $\succsim_i$ his preference relation defined over X_i. It is assumed that for each i in I, $\succsim_i$ can be represented by a function of the form $x_i + v_i(y_i)$ where $z_i = (x_i,y_i)$, with x_i in R and y_i in R_+^{m-1} is agent i's consumption. The first commodity, called money, allows for transfer of utility among the agents. Given e in E, the set of feasible allocations of e, $F(e)$, is the set of lists $z = (z_1,\ldots,z_n)$ in $(R \times R_+^{m-1})^n$ with $\sum_{i\in I} z_i \leqq \sum_{i\in I} \omega_i$. We will assume endowments to be known and dependent on ω will not usually be explicitly stated.

A performance correspondence Φ: $E \to X = X_1 \times \ldots \times X_n$ associates to every e in E a non-empty subset of $F(e)$. Given the special structure of the preferences, for an economy to be completely specified, it is sufficient that one indifference surface be known for each agent. Preference maps can then be completed by translating this indifference surface parallel to the money axis. Given a performance correspondence Φ, we will construct the direct revelation mechanism (M,Φ) by specifying (a) for each agent a message space M_i $(M = M_1 \times \ldots \times M_n)$, which is the space of possible indifference surfaces m_i containing ω_i from which can be derived by translation parallel to the money axis a preference map consistent with the traditional requirements of convexity and continuity; on occasion additional requirements, in particular a certain degree of smoothness, may be imposed, and (b) an <u>outcome correspondence</u> which is taken to be Φ itself. A list $m = (m_1, \ldots, m_n)$ in M specifies an economy in E whose Φ-optimal outcomes are $\Phi(m)$.

Truthful information will be denoted with a circle. The list of truthful messages $(m_1^o, \ldots, m_n^o)$ characterizes the true economy e . The direct revelation mechanism (M,Φ) implements Φ if the agents are truthful. To analyze manipulative behavior, one considers the pair (M,Φ) as a <u>quasi-game</u> where the M_i are now seen as strategy spaces, and Φ plays the role of an outcome function. It is because of the multivaluedness of Φ that we use the term quasi-game. Multiplicities make a definition of equilibrium more delicate; however, we will see that in many cases of interest, it will turn out that these quasi-games will in fact be analyzable as games.

Given m_i in M_i , $h_i(m_i)$ is the offer curve corresponding to the preference map generated from m_i . It contains ω_i . The aggregate net offer curve faced by agent i is denoted $h(m_{-i})$, where m_{-i} designates the list $(\ldots m_{i-1}, m_{i+1}, \ldots)$. The list $m = (m_1, \ldots, m_n)$ will also be denoted (m_i, m_{-i}) . Also, $\succsim_{m_i}$ designates the preference relation corresponding to the preference map generated from m_i . For z_i in X_i , $V(m_i, z_i) = \{z_i' \in X_i | z_i' \succ_{m_i} z_i\}$. $S^m = \{p \in R_+^m | \; \|p\| = 1\}$. For z_i in m_i , $H(m_i, z_i) = \{p \in S^m | \forall \; z_i' \in m_i \; , \; pz_i' \geq pz_i\}$.

We now come to the equilibrium notions. Given a quasi-game (M,Φ) and an economy e in E , we introduce the following definitions.

<u>Definition</u>. A pair $(m,z) \in M \times X$ is a <u>weak equilibrium</u> of (M,Φ) for e ,

written as $(m,z) \in \tilde{N}(\Phi,e)$ iff,

(a) $z \in \Phi(m)$

(b) $\forall\ i \in I\ ,\ \forall\ m_i' \in M_i\ ,\ [\exists\ z' \in \Phi(m_i',m_{-i})\ ,\ z' \succ_{m_i^o} z] \Rightarrow$

$\exists\ z'' \in \Phi(m_i',m_{-i})\ ,\ z \succ_{m_i^o} z''$.

If (m,z) is a weak equilibrium, z is a weak equilibrium allocation.

According to this definition, agents are rather pessimistic and would not switch to another strategy unless all of the resulting allocations were at least as good (according to their true preferences) as the allocation currently selected from $\Phi(m)$.

Definition. A pair $(m,z) \in M \times X$ is an equilibrium of (M,Φ) for e written as $(m,z) \in N(\Phi,e)$ iff,

(a) $z \in \Phi(m)$

(b) $\forall\ i \in I\ ,\ \forall\ m_i' \in M_i\ ,\ \forall\ z' \in \Phi(m_i',m_{-i}) \Rightarrow z \succsim_{m_i^o} z'$.

If (m,z) is an equilibrium, z is an equilibrium allocation.

Every equilibrium is a weak equilibrium, but the notion of equilibrium is much stronger.

Definition. A quasi-game (M,Φ) satisfies property A, if $\forall\ e \in E$, $\forall\ (m,z) \in M \times X\ ,\ [\exists\ i \in I\ ,\ \exists\ m_i' \in M_i\ ,\ \exists\ z' \in \Phi(m_i',m_{-i})$ with $z' \succ_{m_i^o} z] \Rightarrow$ $\exists\ m_i'' \in M_i\ ,\ \forall\ z'' \in \Phi(m_i'',m_{-i})\ ,\ z'' \succ_{m_i^o} z$.

This says that if an agent has available a strategy that could lead to a selection that he strictly prefers to z , then he also has available a strategy for which all possible selections would be strictly preferred to z .

Clearly, if (M,Φ) satisfies property A, the concept of weak equilibrium is not the interesting one. Unfortunately, property A seems to be a strong requirement, and the concept of equilibrium appears to be very demanding. We will see, however, that there exists a large class of quasi-games (M, Φ) satisfying property A and having a non-empty set of equilibria.

The performance correspondences associating to every economy its set of individually rational and Pareto-optimal allocations are denoted I and P respectively, and their intersection IP .

We will consider here the class ψ of performance correspondences Φ satisfying

(a) $\forall\ m \in M\ ,\ \Phi(m) \subset IP(m)$.

(b) $\forall\ m \in M\ ,\ \omega \in P(m) \Rightarrow \Phi(m) = I(m)$.

ψ contains the Walrasian correspondence, the core correspondence, the Shapley-value correspondence, the individually-rational and Pareto-optimal correspondence and many others that will be used as examples to illustrate the following results.

<u>Lemma 1</u>: $\forall\ \Phi \in \psi\ ,\ \forall\ (m,z) \in M \times X\ ,\ (m,z) \in \tilde{N}(\Phi,e) \Rightarrow \forall\ i \in I\ ,\ h(m_{-i}) \cap V(m_i^o, z_i) = \emptyset$.

If (m,z) is a weak equilibrium, the net offer curve that each agent i faces, as determined according to the announced preferences of the other agents, does not intersect his strict upper contour set at z_i , as determined according to his true preferences.

<u>Proof</u>: Suppose by way of contradiction that,

$$\exists\ \Phi \in \psi\ ,\ \exists (m,z) \in M \times X\ ,\ z \in \Phi(m)\ ,\ \exists\ i \in I\ ,\ \exists\ z_i' \in X_i \text{ with } z_i' \in h(m_{-i}) \cap V(m_i^o, z_i)\ .$$

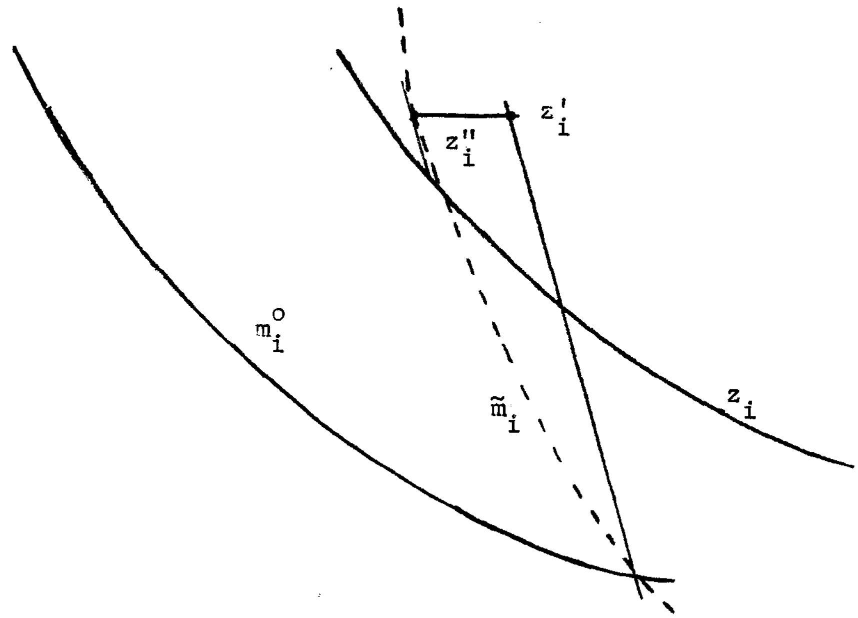


Figure 1

This means that there exists z_j' , $j \neq i$ and $p' \in S^m$ such that z_j' maximizes $\succsim_{m_j}$ at prices p' and $\sum_{j \neq i} (z_j' - \omega_j) = -(z_i' - \omega_i)$. Since $z_i' \succ_{m_i^o} z_i$, and $\succsim_{m_i^o}$ is strictly monotonic in x , there exists $z_i'' \in X_i$ with $y_i'' = y_i'$, $p'z_i'' < p'z_i'$ and $z_i'' \;{}_{m_i^o}\; z_i$. Then, let $\tilde{m}_i \in M_i$ be such that $z_i'' \in \tilde{m}_i$, $p' \in H(\tilde{m}_i, z_i'')$ and $\tilde{m}_i$ be strictly convex at z_i'' . It follows that if $\tilde{z} \in P(\tilde{m}_i, m_{-i})$, $\tilde{y}_i = y_i''$, and if in addition $\tilde{z} \in I(\tilde{m}_i, m_{-i})$, then $\tilde{x}_i = x_i'' + \varepsilon$ for $\varepsilon > 0$. Monotonicity in x of $\succsim_{m_i^o}$ implies that for all $\tilde{z} \in \Phi(\tilde{m}_i, m_{-i}) \subset IP(\tilde{m}_i, m_{-i})$, $\tilde{z}_i \not\succsim_{m_i^o} z_i$, which in turn implies that $(m,z) \notin \tilde{N}(\Phi, e)$ and proves the claim.

<u>Lemma 2</u>: $\forall\ \Phi \in \psi$, $\forall\ (m,z) \in M \times X$, $\forall\ z' \in W(m)$; $(m,z) \in \tilde{N}(\Phi,e)$, $y' = y \Rightarrow z' = z$.

If (m,z) is a weak equilibrium, there does not exist an allocation z' different from z but with the same distribution of the last m-1 commodities, which is Walrasian for m.

Proof: Suppose, by way of contradiction, that,

$$\exists\ \Phi \in \psi\ ,\ \exists\ (m,z) \in M \times X\ ,\ z \in \Phi(m)\ ,\ \exists\ z' \in W(m) \text{ with } y' = y \text{ and } z' \neq z\ .$$

For all $i \in I$, since $z' \in W(m)$, $z_i' \in h(m_{-i})$. Since $y' = y$ but $z' \neq z$, for at least one i, say $i = 1$, $x_1 < x_1'$. Then by strict monotoncity of $\succsim_{m_1^o}$ in x, $z_1' \succ_{m_1^o} z_1$. By Lemma 1, we conclude that $(m,z) \notin \tilde{N}(\Phi,e)$, which proves the claim.

Lemma 3: $\forall\ \Phi \in \psi\ ,\ \forall\ (m,z) \in M \times X\ ;\ (m,z) \in \tilde{N}(\Phi,e) \Rightarrow z \in W(m)$.

If (m,z) is a weak equilibrium allocation, it is a Walrasian allocation with respect to the announced preferences.

Proof: Let $\Phi \in \psi$ be given as well as $(m,z) \in \tilde{N}(\Phi,e)$. Since $z \in \Phi(m) \subset P(m)$, there exists $p \in S^m$ such that

$$\forall\ i \in I\ ,\ \forall\ z_i' \in X_i\ ,\ z_i' \succsim_{m_i} z_i \Rightarrow pz_i' \geqq pz_i\ .$$

For every $i \in I$, let $\tilde{z}_i'$ be such that $p\tilde{z}_i' = p\omega_i$ and $\tilde{y}_i = y_i$. By Walras' Law, $\Sigma\tilde{x}_i = 0$, and therefore $\tilde{z} \in W(m)$. By Lemma 2, $z = \tilde{z}$, which proves the claim.

Lemma 4: The quasi-game (M,W) satisfies property A. In addition,

$$\forall\ (m,z) \in M \times X\ ,\ \forall\ i \in I\ ,\ h(m_{-i}) \cap V(m_i^o,z_i) = \emptyset \Leftrightarrow (m,z) \in N(W,e)\ .$$

This says that one need not investigate the set of weak equilibria of the Walrasian quasi-game.

Proof: To establish the first statement, let (m,z) in $M \times X$ be such that $z \in W(m)$ and for one agent, say agent 1, there is $\tilde{m}_1 \in M_1$

and $\tilde{z} \in W(m_1', m_{-1})$ with $\tilde{z} \;_{m_1^o}\; z$. This means that $h(m_{-1}) \cap V(m_1^o, z) \neq \emptyset$. Then, by an argument similar to the one appearing in Lemma 1, one could show that agent 1 would have a strategy m_1'' such that for all $z'' \in W(m_1'', m_{-1})$, $z'' \succ_{m_1^o} z$.

To establish the second claim, note that, for all i, given m_{-i} in M_{-i}, for all m_i' in M_i, and for all z' in $W(m_i', m_{-i})$, $z_i \in h(m_{-i})$ and therefore $z \succsim_{m_i^o} z'$. This proves Lemma 4.

Note that "most" of the equilibrium allocations z of $N(W,e)$ can be obtained as the second component of equilibrium pairs (m,z) with $\{z_i\} = h(m_{-i}) \cap \{z' \in X_i | z_i' \succsim_{m_i^o} z_i\}$. Indeed, given e in E and $z \in N(W,e)$, a small perturbation of e consisting in increasing the curvature of the preferences of each agent around z_i yields an economy e' such that $z \in N(W,e')$ and for all i, $\{z_i\} = h(m_{-i}) \cap \{z_i' \in X_i | z_i' \succsim_{m_i'^o} z_i\}$. This observation will be used later on.

An explicit characterization of $N(W,e)$ was obtained by Hurwicz for $(m,n) = (2,2)$: denoting by $L(e)$ the (closed) lens-shaped areas delineated by the agents' true offer curves, we have $N(W,e) = L(e)$. (Hurwicz obtained in fact a variant of this result as he imposed somewhat different requirements on the strategy spaces.)

The analysis of the Shapley-value quasi-game carried out in Thomson [28] reveals that smoothness conditions on the strategy spaces may quite substantially restrict the set of equilibria. In fact this phenomenon is general, as revealed by the following lemmas, where M_i^s designates the subset of M_i of C^2 indifference surfaces. Before stating Lemma 5, we need one more definition.

The <u>constrained demand correspondence</u> of an agent is obtained by maximization of his preferences subject to the traditional budget constraint and the requirement that his consumption should not exceed the aggregate resources available in the economy as a whole. The <u>constrained Walrasian performance correspondence CW</u> is obtained by the usual procedure with the demand correspondences being replaced by the constrained demand correspondences. CW coincides with W in the interior of F(e) , and extends W on the boundary of F(e) . CW is a sub-correspondence of P , and is the smallest monotonic extension of W (the monotonicity of a performance correspondence is defined later on). This concept was introduced by Hurwicz, Maskin and Postlewaite in [ ]. In what follows, we will state the Lemmas for F(e) , and prove them for int F(e) . Taking care of the boundary of F(e) does not create any difficulties.

<u>Lemma 5</u>: $\forall\ \Phi \in \psi$, $\forall\ (m,z) \in M \times X$, $(m,z) \in \tilde{N}(\Phi,e)$, $\omega \in P(m) \Rightarrow z \in CW(e)$ and $(m,z) \in N(\Phi,e)$.

If $\omega \in P(m)$, there is no gain from trade according to the announced preferences. However, since $z \in P(m)$ also, this reveals that all agents have announced indifference surfaces that contains their respective $[\omega_i, z_i]$, and this means that the mechanism "almost looks like" the price mechanism.

<u>Proof</u>: By hypothesis, $z \in W(m)$, and for all i , $z_i \sim_{m_i} \omega_i$. Since $z \in P(m)$, there exists $p \in S^m$ such that every m_i has a hyperplane of support at z_i with normal p . The aggregate net offer curve faced by agent i, $h(m_{-i})$ contains z_i and admits there of a unique (by the smoothness assumption on m) hyperplane of support with normal p . We claim that

(p, z) is a competitive equilibrium for e. If not, for at least one agent, say agent 1, the set $\{z_1' \varepsilon X_i | z_1' \succsim_{m_i^o} z_1\}$ is not supported at z_1 by a hyperplane with normal p. But then $h(m_{-i})$ is transversal to agent 1's true indifference surface through z_1. By Lemma 1, $(m, z) \notin \tilde{N}(\Phi, e)$.

To prove the other part of the statement, we have that by individual rationality, given (m, z) as in the hypothesis

$$\forall i \in I, \forall m_i' \in M_i, \forall z_i' \in I(m_i', m_{-i}), [\forall j \neq i, pz_j' \geqq pz_j]$$

and consequently $pz_i' \leqq pz_i$. Since z_i' maximizes $\succsim_{m_i^o}$ on a subset of $\{z_i' \in X_i | pz_i' \leqq pz_i\}$, this proves that in fact (m, z) is an equilibrium.

Lemma 7: $\forall \Phi \in \Psi, \forall e \in E, \forall z \in F(e), z \in CW(e) \Rightarrow \exists m \in M^S$ s.t $(m, z) \in N(\Phi, e)$.

The true constrained Walrasian allocations can always be reached as equilibria.

Proof: Let Φ in Ψ, e in E, and z in W(e) be given. There exists p such that (p, z) is a competitive equilibrium for e. Then, let m_i in M_i^S be given with $z_i \in m_i$, and $p \in H(m_i, z_i)$. A list m obtained in this way is such that $(m, z) \in N(\Phi, e)$. This proves the lemma.

Definition: A quasi-game (M, Φ) implements a performance correspondence Φ' iff it satisfies property A and for any e in E, $N(\Phi, e) = \Phi'(e)$.

Theorem 1: $\forall \Phi \in \Psi, [\forall (m, z) \in M^S \times X, (m, z) \in \tilde{N}(\Phi, e) \Rightarrow \omega \in P(m)] \Rightarrow (M, \Phi)$ implements CW.

Proof: It follows easily from Lemmas 6 and 7.

To show the usefulness of this theorem, we now provide several examples that satisfy its hypotheses. The first two are well-known mechanisms. The other two, which are here defined only for two agents and two commodities, are very different in their informational requirements, and illustrate the richness of the class of mechanisms that implement the constrained Walrasian correspondence.

Example 1: Φ_1 is the Shapley-value performance correspondence. It is shown in [12] that it satisfies the hypothesis of Theorem 1.

Example 2: Φ_2^q with q in R_+^n satisfying $q_i \neq 1$ for all i, and $\sum q_i = 1$, is the performance correspondence distributing the social surplus generated from optimal trading, according to the distributional weights q. Formally, let

$$Y^* = \{y \in R_+^{n(m-1)} \mid \sum y_i \leqq \omega_y \text{ and } \sum v_i(y_i) \geqq \sum v_i(y_i') \quad \forall y' \in R_+^{n(m-1)} \text{ with } \sum y_i' \leqq \omega_y\}.$$

The difference $\sum v_i(y_i) - \sum v_i(\omega_{iy})$ where $y \in Y^*$ measures the gain from trade in terms of the first commodity. Φ_2^q is then defined by

$$\Phi_2^q(m) = \{z \in F(m) \mid x_i + v_i(y_i) = \omega_{ix} + v_i(\omega_{iy}) + q_i(\sum v_i(y_i) - \sum v_i(\omega_{iy})),\ y \in Y^*\}$$

Φ_2^q satisfies (a) and (b) and therefore belongs to Ψ.

Lemma 8: $\forall (m, z) \in M \times X$, $(m, z) \in \tilde{N}(\Phi_2^q, e) \Rightarrow \omega \in P(m)$.

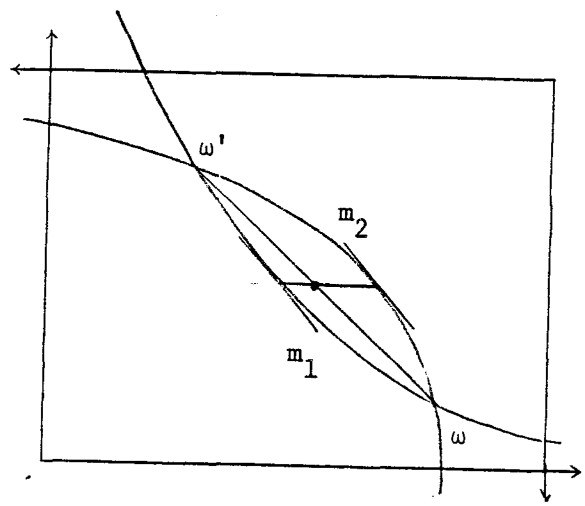
Proof: Let $(m, z) \in M \times X$ be given with $\omega \notin P(m)$. Since $z \in P(m)$, there is $p \in S^m$ such that

$$\forall i \in I,\ \forall z_i' \in X_i,\ z_i' \succ_{m_i} z_i \Rightarrow pz_i' \geqq pz_i.$$

Let $\tilde{z}_i \in X_i$ such that $\tilde{z}_i \sim_{m_i} \omega_i$ and $\tilde{y}_i = y_i$. Since $\omega \notin P(m)$, there is at least one agent, say agent 1, such that $p\tilde{z}_1 < p\omega_1$. Let $\tilde{\tilde{z}}_1 = (\tilde{x}_1 + \varepsilon, y_1)$ with ε such that $p\tilde{\tilde{z}}_1 < p\omega_1$, and $\tilde{m}_1 \in M_1$ such that $\tilde{\tilde{z}}_1 \in \tilde{m}_1$, $p \in H(\tilde{m}_1, \tilde{\tilde{z}}_1)$ and $\tilde{m}_1$ be strictly convex at $\tilde{\tilde{z}}_1$. The social surplus for the economy $(\tilde{m}_1, m_{-1})$ goes down by ε. Also, if $z' \in IP(\tilde{m}_1, m_{-1})$, then $z_1' = (x_1 + (1 - q_1)\varepsilon, y_1)$ and $z_1' \succ_{m_1^o} z_1$, since $q_1 \neq 1$. This shows that agent 1 gains by changing his strategy from m_1 to $\tilde{m}_1$, and therefore $(m, z) \notin N(\Phi_2^q, e)$ which proves the claim.

From the theorem, we can conclude that Φ_2^q implements CW.

<u>Example 3</u>: Let $(m, n) = (2, 2)$.



<u>Figure 2</u>

Also, to simplify the definition of Φ_3, we will assume that preferences are defined on $R \times R_{++}$ instead of $R \times R_+$. Then indifference curves are asymptotic to the x-axis. Now if $\omega \notin P(m)$, ther exists a unique $\omega' \in F(e)$ with $\omega' \neq \omega$ and $\omega_i \in m_i$ for $i = 1,2$. Then Φ_3 is defined by

$$\Phi_3(m) = \begin{cases} I(m) & \text{if } \omega \in P(m) \\ [\omega, \omega'] \cap P(m) & \text{if } \omega \notin P(m) \,. \end{cases}$$

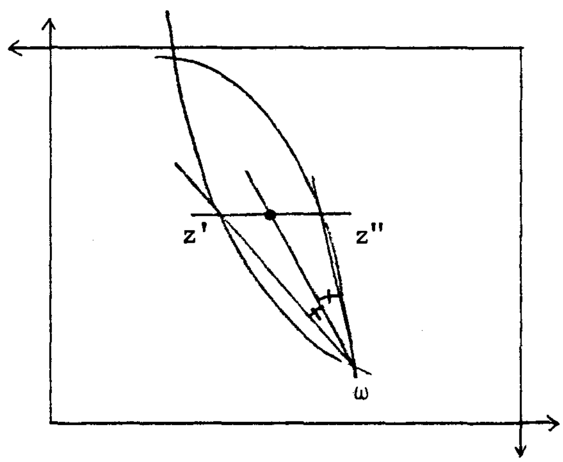
(This mechanism was used by Hurwicz in another context.) If P(m) has a non-empty interior, or if $\omega \varepsilon P(m)$ but $P(m) \neq \{\omega\}$, $\Phi_3(m)$ will not be a singleton. Then the various elements of $\Phi_2(m)$ may not be indifferent to each other for any one agent, even according to his announced performances, in contrast to what happens for Φ_1 and Φ_2^q in the event of such multiplicities. Note that $\Phi_3 \varepsilon \Psi$.

<u>Lemma 9</u>: $\forall (m, z) \varepsilon M \times X$, $(m, z) \varepsilon \tilde{N}(\Phi_3, e) \Rightarrow \omega \varepsilon P(m)$.

<u>Proof</u>: See appendix.

By Theorem 1, it follows that Φ_3 implements CW.

<u>Example 4</u>: Again, let $(m, n) = (2, 2)$.



<u>Figure 3</u>

Given $z \varepsilon P(m)$, let z' and z" $\varepsilon F(m)$ be such that $y_1' = y_1$, $y_2'' = y_2$. If z', z" $\neq \omega$, let B(z', ω, z") denote the bisector of the angle z'ωz". Then Φ_4 is defined by

$$\Phi_4(m) = \begin{cases} I(m) & \text{if } \omega \varepsilon P(m) \\ \bigcup_{z \varepsilon P(m)} B(z', \omega, z'') \cap [z', z''] & \text{if } \omega \notin P(m) . \end{cases}$$

Note that $\Phi_4 \in \Psi$.

Lemma 10: $\forall(m, z) \in M \times X$, $(m, z) \in \tilde{N}(\Phi_4, e) \Rightarrow \omega \in P(m)$.

Proof: See appendix.

By Theorem 1, it follows that Φ_4 implements CW.

We will now argue that whether or not a quasi-game (M, Φ) implements CW depends on "how often" $\Phi(m)$ contains $CW(m)$. To start with, we demonstrate the existence of a whole class of performance correspondences whose associated quasi-games do not implement CW. Paradoxically, these are precisely the performance correspondences that can be implemented by a game.

Definition: $\Phi : E \to X$ is monotonic iff

$$\forall m, m' \in E, \forall z \in \Phi(m), [\forall i \in I, \forall z' \in F(m)\ z \succsim_{m_i} z' \Rightarrow z \succsim_{m'_i} z'] \Rightarrow z' \in \Phi(m') .$$

Definition: $\Phi : E \to X$ is implementable by a game (S, h) with $S = (S_1, \ldots, S_n)$ as strategy spaces and $h : S \to X$ as outcome function iff, for all e in E

(a) s^* is a Nash-equilibrium of $(S, h) \Rightarrow h(s^*) \in \Phi(e)$.

(b) $z^* \in \Phi(e) \Rightarrow \exists\ s^*$, a Nash-equilibrium of (S, h) with $h(s^*) = z^*$.

Lemma 11: $\Phi : E \to X$ is implementable by a game only if it is monotonic

This lemma is proved by Maskin in [17].

Lemma 12: $\forall \Phi \in \Psi$, Φ is implementable by a game $\Rightarrow \forall e \in E$, $\Phi(e) \supset CW(e)$.

Proof: Let e in E be given, and let $z \in CW(e)$. There exists $m \in E$ such that (a) $\omega \in P(m)$, (b) $z \in IP(m)$ and (c) e is obtained from m by a "monotonic transformation," i.e., by changing preferences in such a way that z does not fall in any of the agents' preferences. Then, by property (b) of the elements of Ψ, $z \in \Phi(m)$. By monotonicity of Φ, $z \in \Phi(e)$.

This lemma appears in Hurwicz [6] in a somewhat different form. The role played here by the property (b) mentioned above, "$\omega \in P(m) \Rightarrow \Phi(m) = I(m)$," is played there by a requirement of upper semi-continuity on Φ. In fact, his continuity condition is used by Hurwicz only for sequences of economies converging to "flat" economies similar to m as appears in the proof of the above lemma.

Theorem 2: $\forall\Phi \in \Psi$, Φ is implementable by a game $\Rightarrow (M^s, \Phi)$ does not implement CW.

Proof: More precisely, we will prove that for such a performance correspondence, $\tilde{\tilde{N}}(\Phi, e) \supset N(W, e)$ where $\tilde{\tilde{N}}(\Phi, e)$ is defined by

$$(m, z) \in \tilde{\tilde{N}}(\Phi, e) \Leftrightarrow \forall i \in I, \forall m_i' \in M_i, [\exists z' \in \Phi(m_i', m_{-i}), z' \succ_{m_i^o} z] \Rightarrow$$
$$\exists z'' \in \Phi(m_i', m_{-i}), z \succsim_{m_i^o} z''.$$

The difference between $\tilde{N}(\Phi, e)$ and $\tilde{\tilde{N}}(\Phi, e)$ lies in the weakening of the last statement: $z \succsim_{m_i^o} z''$ instead of $z \succ_{m_i^o} z''$. $\tilde{\tilde{N}}(\Phi, e)$ contains $\tilde{N}(\Phi, e)$, but $\tilde{\tilde{N}}(\Phi, e) \setminus \tilde{N}(\Phi, e)$ is a "small set" since, starting from a pair (m, z) such that $h(m_{-i}) \cap V(m_i, z) = \phi$ for all i, small perturbations in $e = m^o$ guarantee that $h(m_{-i}) \cap \{z_i' \in X_i | z_i' \succsim_{m_i^o} z_i\} = \{z_i\}$.

Now let $(m, z) \in N(W, e)$ be given. We have $z \in W(m)$, and by Lemma 12, $z \in \Phi(m)$. Also, for all i, $h(m_{-i}) \cap V(m_i^o, z_i) = \phi$. Now let agent i consider changing his strategy for m_i to m_i'. Since $z' \in W(m_i', m_{-i}) \Rightarrow z_i \succsim_{m_i^o} z'$ and $z' \in \Phi(m_i', m_{-i})$, it follows that $(m, z) \in \tilde{\tilde{N}}(\Phi, e)$.

The usefulness of this theorem is illustrated by the following examples:

Example 5: Φ_5 is the constrained Walrasian performance correspondence.

As mentioned earlier, if (m, n) = (2, 2), N(W, e) = L(e), the lens-shaped area determined by the true offer curves.

Example 6: Φ_6 = IP. This is the largest of all the elements of Ψ, in the sense that

$$\forall \Phi \in \Psi, \forall e \in E, \Phi(e) \subset IP(e).$$

By Theorem 2, we can conclude that Φ_6 does not implement CW since Φ_6 is monotonic. To illustrate the theorem, we show that if (m, n) = (2, 2), $\tilde{\tilde{N}}(\Phi_6, e) \supset L^o(e)$.

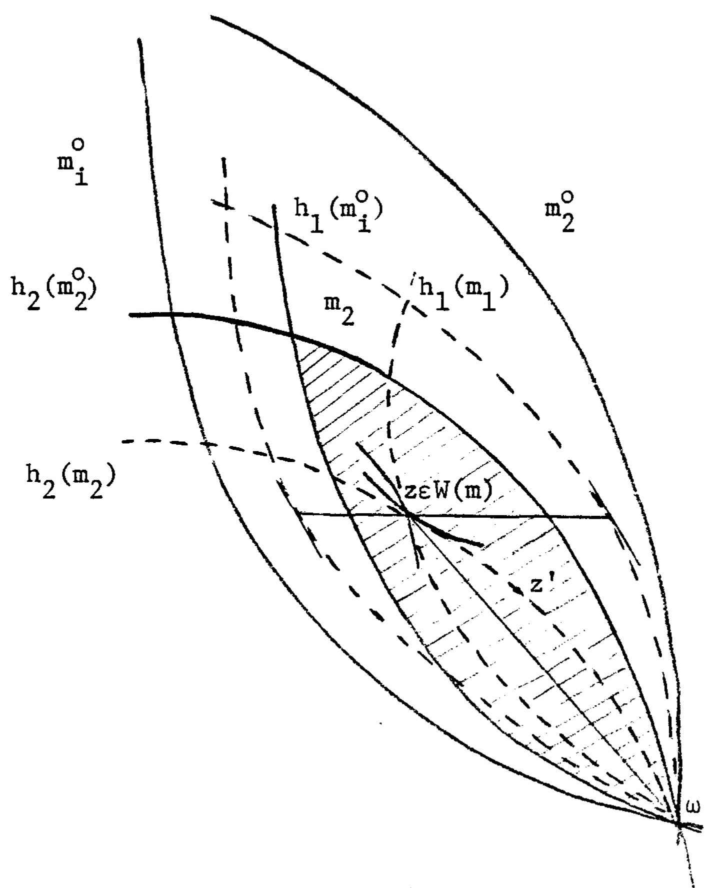


Figure 4

Let $z \varepsilon L^{o}(e)$, and let m_1, m_2 such that $z \varepsilon W(m)$, and for each i, $h(m_j) \cap V(m_i , z_i) = \phi$. Then, for every strategy choice of agent i, say, there would be a new Φ-optimal allocation z' to which z would be preferred, namely the new Walrasian allocation.

(Φ_5 and Φ_6 are indeed implementable, since in addition to being monotonic, they satisfy the "no-veto power", in conjunction with which monotonicity becomes sufficient for implementability; see Maskin [17] for details.)

If Φ_5 and Φ_6 are monotonic and their quasi-games (M^S, Φ) implement CW, note that Φ_1, Φ_2, Φ_3, Φ_4 are not monotonic and their quasi games (M^S, Φ) do implement CW. However, this does not mean that for any non-monotonic Φ, (M^S, Φ) will implement CW. Indeed, by modifying CW at one point e only so as to violate monotonicity, one can easily construct a non-monotonic performance correspondence W' such that $\tilde{N}(W', e)$ will clearly almost be the same as $N(W, e)$. (M^S, W') will therefore not implement CW.

Nevertheless, Theorem 2 suggests that monotonic performance correspondences do not implement CW because they always contain CW. Conversely, by constructing a performance correspondence that never contains CW unless $\omega \varepsilon P(m)$, one ensures implementation of CW. This is illustrated by the following example.

Example 7: Let $m \varepsilon M^S$ be given and $Z_i(m)$ be defined by $Z_i(m) = \{z' \epsilon P(m) | z_i' \epsilon m_i\}$. Given $z' \varepsilon P(m)$, let p' be the common normal to the unique (by smoothness) supporting hyperplane to m_i at z_i', for $i = 1,\ldots,n$.

$$\begin{cases} Z_1(m) & \text{if } (a_1)\ \exists z' \in P(m) \text{ with } p'z_1' < p'\omega_1 \\ Z_2(m) & \text{if } (a_1) \text{ is not true and } (a_2)\ \exists z' \in P(m) \text{ with } p'z_2' < p'\omega_2 \\ Z_n(m) & \text{if } (a_1,\ldots,(a_{n-1}) \text{ are not true and } \exists z' \in P(m) \text{ with } p'z_n' < p'\omega_n \\ I(m) & \text{if } \omega \in P(m)\ . \end{cases}$$

This mechanism is defined in such a way that $z \in \Phi_7(m)$ and $\omega \notin P(m) \Rightarrow z \notin W(m)$. Lemma 3 can never be satisfied, and equilibria occur only when $\omega \in P(m)$. Then but Theorem 1, $N(\Phi_7, e) = CW(e)$.

A performance correspondence with "large" images is more likely to give difficulties for implementation of CW because of the greater chance that it contains Walrasian allocations. However, if one is careful in eliminating these Walrasian allocations, one ensures implementation of CW, as illustrated by the next example:

Example 8:

$$\Phi_8(m) = \begin{cases} IP(m) \setminus CW(e) & \text{if } \omega \notin P(m) \\ I(m) & \text{if } \omega \in P(m)\ . \end{cases}$$

As before, one can show that Φ_8 implements CW.

Economies Without Transferable Utility

Some of the preceding results were proved for general strategy spaces M, while others were based on additional smoothness assumptions. Two interpretations can be provided. If one interprets the preceding quasi-games as direct revelation quasi-games, any requirement placed on the M_i should be thought as restricting the true economy also; in more abstract games, there can be a discrepancy between the spaces of true characteristics of the agents and their strategy spaces. The results established above are compatible with both interpretations. However, once direct games are abandoned, a still more radical interpretation is possible:

Theorem 3: Let E' be the class of pure-exchange economies with n agents and m commodities. Each agent is characterized by a list $(X_i, \omega_i, \succsim_i)$, where $X_i = R^m_+$, $\omega_i \in R^m_+$, and $\succsim_i$ is a convex, continuous preference relation strictly monotonic in the first commodity. Then let Φ be any performance correspondence in Ψ such that (M^s, Φ) implements CW in E, let e' in E' be given, and let the agents in e' play the quasi-game (M^s, Φ) as defined previously. Then, $N(\Phi, e') = CW(e')$.

In other words, the quasi-games defined in transferable utility economies can be used to implement the constrained Walrasian correspondence even in economies without transferable utility.

Proof: It is identical to the proof of implementability in E.

Remark 1: Property (b) placed on the elements of Ψ is a sort of continuity requirement, as argued previously. If this condition were not imposed, equilibria would in general fail to exist.

Remark 2: Is there a way of avoiding multiplicities in $\Phi(m)$? Some Φ in Ψ have always many outcomes (e.g., IP), but several that we have considered have multiple outcomes only exceptionally (e.g., Φ_1, Φ_2, Φ_5). One could restrict strategy lists m by the condition that $\Phi(m)$ be a singleton. Then, the strategies available to an agent would depend on the strategies chosen by the other agents, and one should speak of a "generalized" game. Unfortunately, equilibria would not in general exist in these conditions. By considering ε-equilibria instead of equilibria, where each agent is content to maximize his utility up to an

ε approximation, as measured in terms of the first commodity, existence would be guaranteed. However, the set of ε-equilibria would not shrink under smoothness conditions, and implementation of CW would not be achieved, even approximately.

Concluding Comments: The significance of the results presented here is twofold. First, from a descriptive viewpoint, they provide a general characterization of the equilibrium allocations that result from the manipulation of many economic mechanisms of interest, and indicate under what conditions Pareto-optimality is compatible with manipulative behavior. The price mechanism appeared central to the analysis in several ways: the equilibrium sets of the manipulation games we considered all bear a simple relationship to the equilibrium set of the manipulation game associated with the price mechanisms, and their characterization involves the true offer surfaces in an essential way. Also, under certain assumptions, the sets of equilibrium allocations precisely coincide with the set of constrained Walrasian allocations.

Second, from a normative viewpoint, a by-product of the present analysis is the identification of a whole class of mechanisms that do implement the Walrasian correspondence. Certain games have been devised by several authors to achieve this purpose (see Schmeidler [24], Hurwicz [7], Walker [31]). The mechanisms presented here have the advantage of enjoying better continuity properties, while they have the serious disadvantage of involving considerably more complex strategy spaces. Note, however, that equilibrium lists have a much simpler structure, and that less complicated games that also implement CW could probably be constructed using this fact.

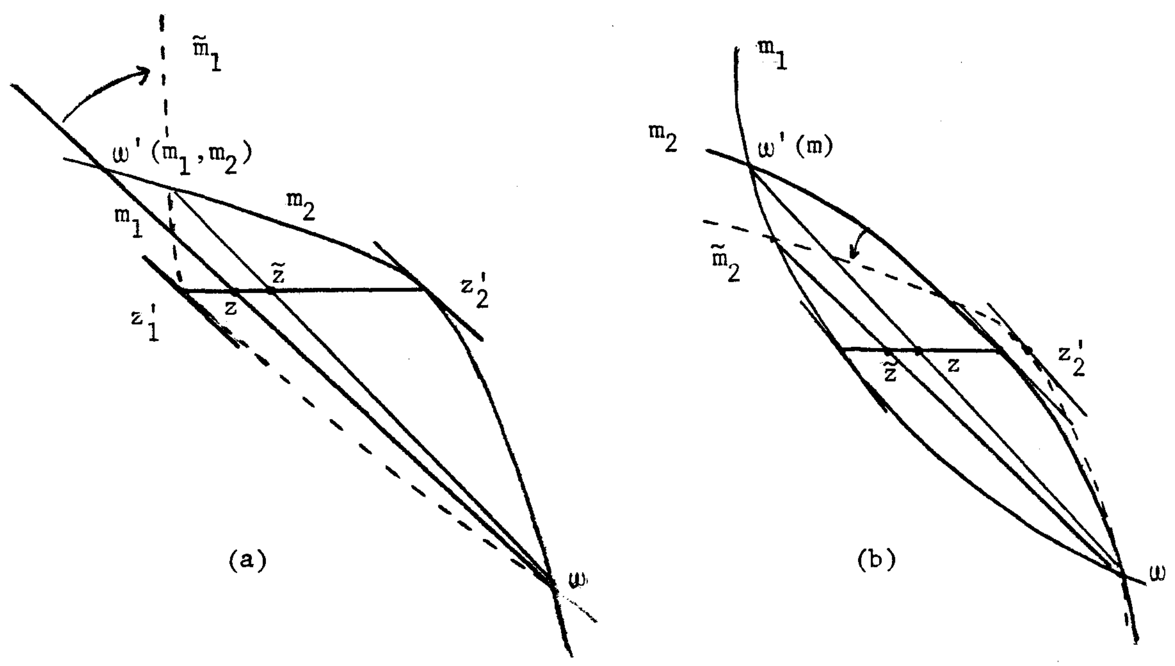
Appendix

We provide here the proofs of Lemmas 9 and 10.

Lemma 9: $\forall\ (m,z) \in M \times X\ ,\ (m,z) \in \tilde{N}(\Phi_3,e) \Rightarrow \omega \in P(m)$.

Proof: Let $(m,z) \in M \times X$ be given, and assume $\omega \notin P(m)$. Define $z_1' \in X_1$ and $z_2' \in X_2$ by $y_1' = y_1$, $z_1' \in m_1$ and $y_2'' = y_2$, $z_2' \in m_2$. Since $\omega \notin P(m)$ either $y_1 > \omega_{1y}$ or $y_2 > \omega_{2y}$. Suppose $y_1 > \omega_{1y}$. Also, either $m_1 \not\supset [\omega,\omega'(m)]$ or $m_2 \not\supset [\omega,\omega'(m)]$, where the dependence of ω' on the strategy pair m is explicitly indicated. Finally, let $p \in H(m_1,z_1) \cap H(m_2,z_2'')$.

Claim 1. $m_2 \not\supset [\omega,\omega'(m)]\ ,\ m_1 \supset [\omega,\omega'(m)] \Rightarrow (m,z) \notin \tilde{N}(\Phi_3,e)$.



(a) (b)

Figure 5

Figure 5 indicates that by rotating m_1 as indicated by the arrow in Figure (5a), agent 1 can ensure the existence of a Φ-optimal allocation $\tilde{z}$ such that $\tilde{y}_1 = y_1$ and $\tilde{x}_1 = x_1 + \varepsilon$ for $\varepsilon > 0$. Since there may be several Φ_3-optimal allocation for the pair $(\tilde{m}_1, m_2)$, agent 1 should in fact select z_1' with $y_1' = y_1$, and $x_1' = x_1 - \eta$ for $\eta > 0$ and $\tilde{m}_1$ so that (a) $z_1' \in \tilde{m}_1$, (b) $p \in H(\tilde{m}_1, z_1')$ and (c) $\tilde{m}_1$ is strictly convex at z_1'. It is clear that for η small enough. $\Phi_3(\tilde{m}_1, m_2)$ is a singleton to the right of z.

<u>Claim 2</u>. $m_1 \not\supset [\omega, \omega'(m)] \Rightarrow (m,z) \notin \tilde{N}(\Phi_3, e)$.

Indeed, if $m_1 \not\supset [\omega, \omega'(m)]$, then m_2 is not optimal against m_1: by rotating m_2 as indicated by the arrow in Figure 5b, agent 2 can make appear as Φ_3-optimal allocation, an allocation $\tilde{z}$ that he strictly prefers to z. <u>Enforcing</u> a preferred allocation can be done as explained in Claim 1. Lemma 9 follows from Claims 1 and 2.

<u>Proof of Lemma 10</u>

Let $(m,z) \in N(\Phi_4, e)$ and assume $\omega \notin P(m)$. Define $z_1' \in X_1$ and $z_2'' \in X_2$ by $y_1' = y_1$, $z_1' \in m_1$ and $y_2'' = y_2$, $z_2' \in m_2$. Let $p \in S^2$ be the non-negative normal to the parallel lines of support of m_1 and m_2 at z_1' and z_2'. We have $pz_1' \leqq p\omega_1$ and $pz_2' \leqq p\omega_2$. Since $\omega \notin P(m)$, the inequality is strict for at least one agent, say $pz_1' < p\omega_1$. Let then $\tilde{z}_1 = z_1' + \varepsilon(1,0)$ be such that $pz_1' < p\tilde{z}_1 < p\omega_1$ and let $\tilde{m}_1 \in M_1^s$ be a strategy for agent 1 such that (a) $\tilde{z}_1 \in \tilde{m}_1$, (b) $p \in H(\tilde{m}_1, \tilde{z}_1)$, (c) $\tilde{m}_1$ be strictly convex at $\tilde{z}_1$. Then, clearly, $z^* \in P(\tilde{m}_1, m_2) \Rightarrow y_1^* = y_1$, and $\Phi_4(\tilde{m}_1, m_2)$ is a singleton $\tilde{\tilde{z}}$ such that $\{\tilde{\tilde{z}}\} = B(\tilde{z}_1, \omega, z_2) \cap [\tilde{z}_1, z_2]$, and $\tilde{\tilde{x}}_1 > x_1$. This implies $\tilde{\tilde{z}} \succ_{m_1^o} z$, in contradiction with the statement that $(m,z) \in N(\Phi_4, e)$.

REFERENCES

[1] Champsaur, P., J. Drèze, and C. Henry, "Stability Theorems with Economic Applications," Econometrica 45 (1977), pp. 273-294.

[2] Crawford, V. P. and H. R. Varian, "Distortion of Preferences and the Nash Theory of Bargaining," Economic Letters (forthcoming).

[3] Drèze, J. and D. de la Vallée Poussin, "A Tâtonnement Process for Public Goods," Review of Economic Studies 38 (1971), pp. 133-150.

[4] Henry, C., "On the Free Rider Problem in the M.D.P. Procedure," Review of Economic Studies 46 (1979), pp. 293-303.

[5] Henry, C., "Comment on "Stategic Behavior in the M.D.P. Planning Procedure," Review of Economic Studies 46 (1979), pp. 363-369.

[6] Hurwicz, L., "On Allocations Attainable Through Nash-Equilibria," Journal of Economic Theory (forthcoming); also in, Laffont, J.-J. (ed.), Aggregation and Revelation of Preferences, North-Holland, 1979, pp. 397-419.

[7] Hurwicz, L., "Outcome Functions Yielding Walrasian and Lindhal Allocations at Nash Equilibrium Points," Review of Economic Studies, 46 (1979), pp. 217-225.

[8] Hurwicz, L., "On Informationally Decentralized Systems," Chapter 14 in Decision and Organization, C. B. McGuire and R. Radner (eds.), North-Holland, Amsterdam, 1972.

[9] Hurwicz, L., "On the Interaction Berween Information and Incentives in Organization," in Communication and Control in Society, K. Krippendorff (ed.), Scientific Publishers, Inc., New York, 1978.

[10] Hurwicz, E. Maskin, and A. Postlewaite, "Implementation with Unknown Endowments," mimeograph, June 1979.

[11] Kalai, E. and M. Smorodinsky, "Other Solutions to Nash's Bargaining Problem," Econometrica 43 (1979), pp. 513-518.

[12] Kannai, Y., "Concavifiability and Construction of Concave Utility Functions," Journal of Mathematical Economics 4 (1977), pp. 1-56.

[13] Kihlstrom, R.E., A. Roth, and D. Schmeidler, "Risk Aversion and Solutions to Nash's Bargaining Problem," mimeograph, March 1979.

[14] Kurz, M., "Distortion of Preferences, Income Distribution, and the Case for a Linear Income Tax," Journal of Economic Theory 14 (1977), pp. 291-298.

[15] Kurz, M., "Income Distribution and Distortion of Preferences: the ℓ-Commodity Case," Technical Report No. 266, IMSSS, Stanford University, July 1978.

[16] Malinvaud, E., "Prices for Individual Consumption, Quantity Indicators for Collective Consumption," Review of Economic Studies 39 (1972), pp. 385-405.

[17] Maskin, E.S., "Nash Equilibrium and Welfare Optimality," Mathematics of Operations Research (forthcoming).

[18] Nash, J.F., "The Bargaining Problem," Econometrica 18 (1950), pp. 155-162.

[19] Otani, Y. and J. Sicilian, "Equilibrium of Walras Preference Games," University of Kansas mimeograph, October 1978.

[20] Raiffa, H., "Arbitration Schemes for Generalized Two-Person Games," in Contributions to the Theory of Games II, H.W. Kuhn and A.W. Tucker (eds.), Princeton University Press, (1953).

[21] Roberts, J., "Strategic Behavior in the M.D.P. Planning Procedure," in Aggregation and Revelation of Preferences, J.-J. Laffont (ed.), 1979 North-Holland, pp. 353-362.

[22] Roberts, J., "Incentives in Planning Procedures for the Provision of Public Goods," Review of Economic Studies 46 (1979), pp. 283-292.

[23] Roberts, J. and A. Postlewaite, "The Incentives for Price-Taking in Large Economies," Econometrica 44 (1976), pp. 115-127.

[24] Schmeidler, D., "A Remark on Microeconomic Models of an Economy and on a Game Theoretic Interpretation of Walras Equilibria," mimeograph, Minneapolis, March 1976.

[25] Schoumaker, F., "Strategic Behavior in a Discrete-Time Procedure," in Aggregation and Revelation of Preferences, J.-J. Laffont (ed.), 1979, North-Holland, pp. 367-376.

[26] Schoumaker, F., "Incentives in Planning with Private Goods," Review of Economic Studies 46 (1979), pp. 315-318.

[27] Sobel, J., "Distortion of Utilities and the Bargaining Problem," University of California at San Diego Discussion Paper 79-29.

[28] Thomson, W., "On the Manipulability of the Shapley-Value," mimeograph, June 1979.

[29] Thomson, W., "The Equilibrium Allocations of Walras and Lindhal Manipulation Games," University of Minnesota Discussion Paper No. 111, July 1979.

[30] Tulkens, H. and S. Zamir, "Surplus-Sharing Local Games in Dynamic Exchange Economies," Review of Economic Studies 46 (1979), pp. 305-313.

[31] Walker, M., "An Informationally Efficient Auctioneerless Version of the Walras Mechanism," S.U.N.Y. discussion paper, September 1977.