GAMES AND ECONOMIC BEHAVIOR 4, 252-260 (1992)

Self-Optimality and Efficiency in Utility Distortion Games

HANS PETERS

Department of Mathematics, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands

Received December 4, 1989

In social choice problems where players may strategically misrepresent their preferences, we call a profile of preferences self-optimal if reporting them is a Nash equilibrium given that they are the true preferences. Self-optimality can be interpreted as a very weak honesty requirement. We apply the self-optimality concept to a utility distortion game in the context of bargaining and obtain a characterization of efficient Nash equilibria. *Journal of Economic Literature* Classification Numbers: 020,210,610. © 1992 Academic Press, Inc.

1. INTRODUCTION AND GENERAL FORMULATION

We consider the following *n*-person social choice problem. $N = \{1, 2, ..., n\}$ denotes the set of individuals, A is a nonempty set of alternatives, and, for each individual *i*, \mathcal{U}^i denotes a nonempty collection of utility functions $u^i: A \to \mathbb{R}$ representing the possible preferences of *i* over A. A solution is a function $\varphi: \mathcal{U} \to 2^A$, where $\mathcal{U} := \mathcal{U}^1 \times \cdots \times \mathcal{U}^n$, such that $u^i(a) = u^i(b)$ for all $i \in N$, $u = (u^1, \ldots, u^i, \ldots, u^n) \in \mathcal{U}$, $a, b \in \varphi(u)$. Note that a solution is a social choice correspondence; the converse, however, does not necessarily hold in view of the utility-equivalence constraint implicit in the definition of a solution.

Suppose a solution φ were single-valued, $|\varphi(u)| = 1$ for every $u \in \mathfrak{A}$. Then φ would be a *game form*, and, for each $\hat{u} \in \mathfrak{A}$, would give rise to a noncooperative game with N as the set of players, \mathfrak{A}^i as the strategy set of player *i*, and $\hat{u}(\varphi(u)) \in \mathbb{R}^n$ as the payoff vector resulting from a strategy *n*-tuple $u \in \mathfrak{A}$. Since, in general, we consider solutions φ that are not single-valued, we give the following definition of a Nash equilibrium. DEFINITION. We call an (n + 1)-tuple $(u, a) \in \mathcal{U} \times A$ a Nash equilibrium for φ and $v \in \mathcal{U}$ if the following two conditions are satisfied:

$$a \in \varphi(u) \tag{1}$$

$$\forall i \in N \;\forall \; \hat{u}^i \in \mathcal{U}^i \;\forall \; \hat{a} \in \varphi(u^{-i}, \; \hat{u}^i) \; [v^i(a) \ge v^i(\hat{a})]. \tag{2}$$

Here, we use the notation (u^{-i}, \hat{u}^i) for the vector obtained from u by replacing u^i by \hat{u}^i . In a Nash equilibrium, no player can possibly gain from unilaterally reporting a different utility function. This Nash equilibrium concept is equivalent to the equilibrium notion introduced by Thomson (1984, p. 451).

An appropriate context for this model is the following setting. There is a central planner who is going to use some solution φ to determine a final set of outcomes. However, he does not know the true utility functions of the individuals or players, and can only rely on the information given to him by these players. The players report (not necessarily true) utility functions to the central planner. We assume that the players report an *n*-tuple of utility functions leading to a Nash equilibrium for the given solution and the true utility functions. For this assumption to be reasonable, one might assume that the players known not only their own but also the other players' utility functions, and—especially in the case of multiple Nash equilibria—that there is some preplay communication between the players. Further the players might suggest an equilibrium selection *a* from $\varphi(u)$ as well.

We suppose that the central planner in this model wishes to use a solution that has appealing properties (such as the Nash bargaining solution discussed in the next section). Using such a solution, he will in general not elicit the players' true preferences in a Nash equilibrium, and, indeed, some of the solution's attractive properties, notably efficiency, may be lost ex post. The question we raise in this paper is: can one find restrictions on the allowed reports of the players, such that the set of possible Nash equilibria is narrowed down to the set of efficient Nash equilibria?

The restriction we impose in this paper is self-optimality:

DEFINITION. An *n*-tuple $u \in \mathfrak{A}$ is called *self-optimal for* (a solution) φ if $(u, a) \in \mathfrak{A} \times A$ is a Nash equilibrium for φ and u, for every $a \in \varphi(u)$.

(By the utility equivalence implied in the definition of a solution φ , it is of course sufficient for self-optimality that (u, a) be a Nash equilibrium for some $a \in \varphi(u)$.) Self-optimality of a vector of reports u means that these reports constitute a Nash equilibrium given that they are the true reports. In requiring the players' reports to be self-optimal, the central planner might reason as follows.

HANS PETERS

Suppose I receive reports that are not self-optimal. If these reports are the true utility functions of the players, then I know that at least one player could have deviated and thereby gained. So there must be one or more players lying, since I assume the players to be utility maximizers.

In this situation, the central planner could punish the collective of the players for being *provably* dishonest, e.g., by choosing a known bad alternative. Instead of modifying the game in this way, we will equivalently assume that the players are obliged always to come up with a self-optimal Nash equilibrium vector of reports. Thus, self-optimality can be viewed as a very mild honesty requirement.

There is a close relationship between self-optimality and strategy proofness (in the social choice literature—for instance, Moulin, 1983; Peleg, 1984) or incentive compatibility (in mechanism theory—for instance, Hurwicz, 1972; Myerson, 1979). In the present setting, these concepts (which are statements about a solution φ) would mean self-optimality of every $u \in \mathcal{U}$ for φ . Thus, self-optimality is much weaker, and our approach is more in line with Thomson (1984), and, for the specific context we study in the next section, with Sobel (1981). The next section studies bargaining over the division of a commodity bundle; we show, mainly, that self-optimality leads to a characterization of efficient Nash equilibria. The final Section 3 concludes with some discussion.

2. DISTORTION OF UTILITIES IN BARGAINING

Let there be two players who are to divide a bundle of *m* commodities. There is exactly one unit of each commodity. So the set of alternatives *A* can be described as $\{x \in \mathbb{R}^m: 0 \le x \le 1\}$ where 0 (1) denotes the vector with only zeros (ones). The interpretation of $x \in A$ is that player 1 receives x and player 2 receives 1 - x. Let \mathcal{U}^1 denote the collection of functions $u^1: A \to [0, 1]$ that satisfy:

(i) u^1 is concave and strictly increasing, i.e., $x \ge \hat{x}$ and $x \ne \hat{x} \Rightarrow u^1(x) > u^1(\hat{x})$;

(ii)
$$u^{1}(\mathbf{0}) = 0, u^{1}(\mathbf{1}) = 1;$$

(iii) u^1 is twice continuously differentiable on the interior of A.

Condition (iii), in particular the word "twice," is needed in order to be able to apply Lemma 2 in Sobel (1981), below.

We assume that player 1's set of utility functions or strategy set equals \mathfrak{A}^1 , and that player 2's strategy set is $\mathfrak{A}^2 = \{u^2 \colon A \to \mathbb{R}: \text{ there exists } u^1 \in \mathfrak{A}^1 \text{ with } u^2(x) = u^1(1-x) \text{ for all } x \in A\}$. Note that $u^2(x)$ denotes player 2's utility from receiving 1 - x. Further, we denote $\mathfrak{A} := \mathfrak{A}^1 \times \mathfrak{A}^2$. A solution assigns to each $u \in \mathfrak{A}$ a subset of A such that all alternatives in this subset are utility equivalent.

 $x \in A$ is called *efficient for* $u \in \mathcal{U}$ if there is no $\hat{x} \in A$ with $u^{1}(\hat{x}) \ge u^{1}(x)$, $u^{2}(\hat{x}) \ge u^{2}(x)$, and with at least one of these inequalities strict. A solution φ is called *efficient* if x is efficient for u for every $u \in \mathcal{U}$ and $x \in \varphi(u)$. A solution φ is called *symmetrically monotonic* if $u(x) \ge (\frac{1}{2}, \frac{1}{2})$ for every $u \in \mathcal{U}$ and $x \in \varphi(u)$. Symmetric monotonicity can be seen as a very weak symmetry or monotonicity property.

Suppose the players report a pair of utility functions $u \in \mathcal{U}$. Given the solution φ , the *attainable set* for player 1 is defined as

$$A^{1}(u^{2}, \varphi) := \{ x \in A : \exists v^{1} \in \mathcal{U}^{1}, y \in \varphi(v^{1}, u^{2}) | x \leq y] \},\$$

and the attainable set $A^2(u^1, \varphi)$ for player 2 is defined analogously. In what follows, we will need the requirement that such attainable sets have smooth boundaries, at least in the interior of A. Formally, a solution φ is called *smooth-regular* if for any attainable set $A^1(u^2, \varphi)$ there exists a function $F: A \to \mathbb{R}$, continuously differentiable on the interior of A and strictly increasing, such that

$$A^{\mathsf{I}}(u^2,\varphi) = \{x \in A \colon F(x) \le 0\},\$$

and for any attainable set $A^2(u^1, \varphi)$ there exists a function $G: A \to \mathbb{R}$, continuously differentiable on the interior of A and strictly decreasing, such that

$$A^{2}(u^{1}, \varphi) = \{x \in A : G(x) \leq 0\}.$$

Note that, in general, F and G will depend on u^2 and u^1 , respectively. The monotonicity conditions on F and G guarantee that the (preferred) boundaries of these attainable sets are given by F(x) = 0 and G(x) = 0, respectively.

A further requirement to be imposed later on is the following one. A solution φ is called *convex-regular* if all attainable sets are convex.

An example is the solution ν derived from the well-known Nash bargaining solution (Nash, 1950), as follows: to each pair $(u^1, u^2) \in \mathcal{U}, \nu$ assigns the subset of all $x \in A$ such that the product $u^1(x)u^2(x)$ is maximal on A. For simplicity, we call ν the Nash solution. This solution is efficient and symmetrically monotonic. Smooth-regularity and convex-regularity of the Nash solution are consequences of Lemma 2 in Sobel (1981, p. 612).

It is easy to see that ν is not "strategy-proof," that is, that not every $u \in \mathcal{U}$ is self-optimal for ν : for instance, for the case of one commodity, Crawford and Varian (1979) have already shown that, for each player, reporting the (unique) linear utility function is dominant. Also the follow-

ing example, taken from Sobel (1981, p. 617), can be used to this end. We include it, however, to show that a Nash equilibrium may lead to an alternative that is inefficient for the true utility functions.

EXAMPLE. Let m = 2, let $v^1(x) = x_1^{5/6}x_2^{1/6}$, $u^1(x) = (5x_1 + 3x_2)/8$, $v^2(x) = u^2(x) = \sqrt{(1 - x_1)(1 - x_2)}$. Then $(u^1, u^2, (\frac{3}{5}, \frac{1}{3}))$ is a Nash equilibrium for ν and (v^1, v^2) , as can be verified with the aid of Lemma 2 in Sobel (1981). Consider the allocation $(\frac{2}{3}, \frac{1}{5})$. Then $v^1(\frac{2}{3}, \frac{1}{5}) > v^1(\frac{3}{5}, \frac{1}{3})$ whereas $v^2(\frac{2}{3}, \frac{1}{5}) = v^2(\frac{3}{5}, \frac{1}{3})$. So the above Nash equilibrium allocation is inefficient for the true preferences (v^1, v^2) .

Thus, the Nash solution ν admits inefficient Nash equilibrium alternatives. Besides, there may be inefficient Nash equilibria not Pareto dominated by some efficient Nash equilibrium (Sobel, 1981, p. 617, same example): therefore, it may be plausible that the players actually come up with an inefficient equilibrium. How can a central planner avoid this, not knowing the true preferences and still using the Nash solution ν ? The following observation gives an answer to this question.

In the following, " ∇ " denotes "the gradient of."

THEOREM 2.1. Let φ be an efficient and smooth-regular solution. Let (u, \hat{x}) be a Nash equilibrium for φ and $v \in \mathfrak{A}$. Suppose \hat{x} is an interior point of A, and suppose u is self-optimal for φ . Then \hat{x} is efficient for v.

Proof. Let the functions F and G correspond to the attainable sets $A^{1}(u^{2}, \varphi)$ and $A^{2}(u^{1}, \varphi)$, repectively, as in the definition of smooth-regularity. Since (u, \hat{x}) is a Nash equilibrium for φ and v, \hat{x} maximizes v^{1} on $A^{1}(u^{2}, \varphi)$ and v^{2} on $A^{2}(u^{1}, \varphi)$. Since v^{1} and F are increasing, v^{2} and G decreasing, and \hat{x} is by assumption an interior point of A, $F(\hat{x}) = G(\hat{x}) = 0$ and there are numbers λ and λ' with $\nabla v^{1}(\hat{x}) = \lambda \nabla F(\hat{x})$, $\nabla v^{2}(\hat{x}) = \lambda' \nabla G(\hat{x})$. For analogous reasons and the self-optimality of u, there exist numbers μ and μ' with $\nabla u^{1}(\hat{x}) = \mu \nabla F(\hat{x})$, $\nabla u^{2}(\hat{x}) = \mu' \nabla G(\hat{x})$. By the efficiency of φ and hence of \hat{x} for u, there is a number κ with $\nabla u^{1}(\hat{x}) = \kappa \nabla u^{2}(\hat{x})$. Combining all these equalities, we find that $\nabla v^{1}(\hat{x})$ is a multiple of $\nabla v^{2}(\hat{x})$. Since v^{1} and v^{2} are concave, this implies efficiency of \hat{x} for v.

Thus, when using an efficient and smooth-regular solution, the central planner can achieve efficiency by requiring the reports to be self-optimal, that is, by requiring the reports to be not provably dishonest. There is also a converse to this theorem. We start with a definition.

DEFINITION. An equal income competitive equilibrium (EICE) for $v \in \mathcal{U}$ is a pair (p, \hat{x}) where

- (i) $p \in \mathbb{R}^m, p \ge 0, \hat{x} \in A$
- (ii) \hat{x} solves

max $v^{1}(x)$ subject to $p \cdot x \leq \frac{1}{2}p \cdot \mathbf{1}$ and $x \in A$

and

max
$$v^2(x)$$
 subject to $p \cdot (1 - x) \le \frac{1}{2}p \cdot 1$ and $x \in A$.

So an equal income competitive equilibrium is a competitive equilibrium starting from equal division of the goods. Note that a price vector p in an EICE must be positive, since the utility functions are strictly monotonic. Hence, such a price vector p gives rise to an element \overline{p}^1 of \mathfrak{A}^1 by $\overline{p}^1(x) := p \cdot x(\sum_{i=1}^m p_i)^{-1}$ and an element \overline{p}^2 by $\overline{p}^2(x) := p \cdot (1 - x)(\sum_{i=1}^m p_i)^{-1}$ for all $x \in A$.

LEMMA 2.1. Let (p, \hat{x}) be an EICE for v. Let φ be a symmetrically monotonic solution. Then $((\bar{p}^1, \bar{p}^2), \hat{x}) \in \mathfrak{A} \times A$ is a Nash equilibrium for φ and v.

Proof. By symmetric monotonicity of φ , for all $u \in {}^{0}\!\!u$, if $x \in \varphi(u^{1}, \overline{p}^{2})$, then $\overline{p}^{2}(x) \geq \frac{1}{2}$, which implies $\overline{p}^{1}(x) \leq \frac{1}{2}$. Similarly, $x \in \varphi(\overline{p}^{1}, u^{2})$ implies $\overline{p}^{2}(x) \leq \frac{1}{2}$. Further, since (p, \hat{x}) is an EICE for v and the utility functions are strictly monotonic, we have $p \cdot \hat{x} = \frac{1}{2}p \cdot \mathbf{1} = p \cdot (\mathbf{1} - \hat{x})$, which implies $\overline{p}^{1}(\hat{x}) = \overline{p}^{2}(\hat{x}) = \frac{1}{2}$. Therefore, \hat{x} is efficient for $(\overline{p}^{1}, \overline{p}^{2})$, and hence $\hat{x} \in \varphi(\overline{p}^{1}, \overline{p}^{2})$ by symmetric monotonicity. We conclude that $((\overline{p}^{1}, \overline{p}^{2}), \hat{x})$ is a Nash equilibrium for φ and v.

A consequence of Lemma 2.1 is the existence of a Nash equilibrium since, by standard arguments, an EICE always exists. Let 1/2 denote the vector in \mathbb{R}^m with all coordinates equal to $\frac{1}{2}$.

LEMMA 2.2. Let φ be a symmetrically monotonic and efficient solution. Let (\hat{x}, u) be a Nash equilibrium for φ and v. Then $v^{1}(\hat{x}) \ge v^{1}(1/2)$ and $v^{2}(\hat{x}) \ge v^{2}(1/2)$.

Proof. We prove only the first inequality. Suppose to the contrary that $v^{1}(\hat{x}) < v^{1}(1/2)$. Given u^{2} , player 1 can report some utility function \tilde{u}^{1} which is linear on the diagonal

$$D := \{(t, t, \ldots, t) \in \mathbb{R}^m : 0 \le t \le 1\},\$$

and such that *D* is exactly the set of alternatives that are efficient for $\tilde{u} := (\tilde{u}^1, u^2)$. By efficiency and symmetric monotonicity of φ , $\varphi(\tilde{u}) = \{(\tilde{t}, \ldots, \tilde{t})\}$ for some $\tilde{t} \ge 1/2$. So $v^1(\tilde{t}, \ldots, \tilde{t}) \ge v^1(1/2) > v^1(\hat{x})$, contradicting the assumption that (u, \hat{x}) is a Nash equilibrium.

LEMMA 2.3. Let φ be a convex-regular, efficient, and symmetrically monotonic solution. Let (u, \hat{x}) be a Nash equilibrium for φ and v with \hat{x} efficient for v. Then $(\nabla v^{1}(\hat{x}), \hat{x})$ is an EICE for v.

Proof. Since \hat{x} is efficient for v, and since the attainable sets $A^{1}(u^{2}, \varphi)$ and $A^{2}(u^{1}, \varphi)$ are convex, the hyperplane $\nabla v^{1}(\hat{x}) \cdot x = \nabla v^{1}(\hat{x}) \cdot \hat{x}$ separates

these sets at \hat{x} . Let $p := \nabla v^1(\hat{x})$; then it follows that (u, \hat{x}) is also a Nash equilibrium for (\bar{p}^1, \bar{p}^2) . Hence, by Lemma 2.2, $\bar{p}^1(\hat{x}) \ge \bar{p}^1(1/2)$ and $\bar{p}^2(\hat{x}) \ge \bar{p}^2(1/2)$. Combined, these inequalities imply $\nabla v^1(\hat{x}) \cdot \hat{x} = \nabla v^1(\hat{x}) \cdot 1/2 = 1/2\nabla v^1(\hat{x}) \cdot 1$. Hence, at \hat{x} the function v^1 is maximized subject to the constraint $\nabla v^1(\hat{x}) \cdot x \le 1/2\nabla v^1(\hat{x}) \cdot 1$, and v^2 is maximized subject to $\nabla v^1(\hat{x}) \cdot (1-x) \le 1/2\nabla v^1(\hat{x}) \cdot 1$. In other words, $(\nabla v^1(\hat{x}), \hat{x})$ is an EICE for v.

Lemma 2.3 is the only result in which convexity of the attainable sets is used. For the Nash solution, a direct proof of this result is given by Sobel (1981, Theorem 5).

LEMMA 2.4. Let φ be a symmetrically monotonic solution. Let (p, \hat{x}) be an EICE for $v \in \mathfrak{A}$. Then $(\overline{p}^1, \overline{p}^2)$ is self-optimal for φ .

Proof. From the definition of EICE and the strict monotonicity of the utility functions it follows that $p \cdot \hat{x} = p \cdot (1 - \hat{x}) = \frac{1}{2}p \cdot 1$. This implies $\hat{x} \in \varphi(\bar{p}^1, \bar{p}^2)$ by symmetric monotonicity of φ . Suppose there were a $u^1 \in \mathcal{U}^1$ and an $x \in \varphi(u^1, \bar{p}^2)$ with $\bar{p}^1(x) > \bar{p}^1(\hat{x}) = \frac{1}{2}$. Then $\bar{p}^2(x) < \frac{1}{2}$, which contradicts the symmetric monotonicity of φ . One similarly shows that player 2 cannot gain from unilaterally deviating. So $((\bar{p}^1, \bar{p}^2), \hat{x})$ is a Nash equilibrium for φ and (\bar{p}^1, \bar{p}^2) ; hence (\bar{p}^1, \bar{p}^2) is self-optimal for φ .

We can now prove:

THEOREM 2.2. Let φ be a convex-regular, efficient, and symmetrically monotonic solution. Let \hat{x} be a Nash equilibrium allocation for φ and $v \in$ ^Q that is efficient for v. Let $p := \nabla v^1(\hat{x})$. Then $((\vec{p}^1, \vec{p}^2), \hat{x})$ is a Nash equilibrium for φ and v with a self-optimal pair of reports.

Proof. First apply Lemma 2.3, then Lemma 2.1, and finally Lemma 2.4. ■

Summarizing, we note that for an efficient and smooth-regular solution, self-optimality leads to an allocation on the contract curve in the Edgeworth box associated with the division problem. Actually, the reported indifference curves must coincide locally with the true indifference curves (which supports our intuition of self-optimality as a very weak honesty requirement). This observation follows from the proof of Theorem 2.1, which is based mainly on the smoothness of the boundaries of the attainable sets. Further, Theorem 2.2 states that for a convex-regular, efficient, and symmetrically monotonic solution, any efficient Nash equilibrium allocation can be obtained by a self-optimal pair of reports. Requiring self-optimality does not narrow down the set of efficient Nash equilibria.

By applying Lemmas 2 and 3 in Sobel (1981), finally, it can be seen that our results hold for the Nash solution as well as for the Kalai–Smorodinsky solution.

3. DISCUSSION

We have introduced the concept of self-optimality for a general class of social choice problems. Application to a specific bargaining context has led to a characterization of efficient Nash equilibria for solutions satisfying a number of reasonable conditions. Efficiency and symmetric monotonicity are easily verifiable conditions. Convex- and smooth-regularity are properties stated only indirectly, in terms of attainable sets, and therefore are less readily verifiable. This last point may well be considered a drawback.

Another way to obtain efficiency of the final outcome is to allow only linear preferences. As Sobel (1981, Theorem 2) shows, if the reported preferences in a Nash equilibrium are linear, then (under certain conditions) they must support an EICE allocation, which is always efficient. Furthermore, the results above show that (under certain conditions again) all efficient Nash equilibrium allocations can be reached by linear preferences. Comparing the two approaches—self-optimality and linearity however, we think that the former has a number of advantages.

First, the self-optimality criterion is a more general principle than linearity. Indeed, it can be formulated even if linearity of preferences has no meaning: linearity comes out in the specific application discussed in this paper.

Second, in this specific application, self-optimal preferences do not have to be linear, as is shown by the following example. Although the difference with linear preferences in this example is not very essential, it remains true that self-optimality admits a larger class of preferences.

EXAMPLE. Let m = 1 (one commodity), let $u^1(x) := \frac{3}{2}x$ for all $0 \le x \le \frac{1}{2}$, $u^1(x) := \frac{1}{2}x + \frac{1}{2}$ for all $\frac{1}{2} \le x \le 1$, $u^2(x) := 1 - x$ for all $0 \le x \le 1$. Then $(u, \frac{1}{2})$ is a Nash equilibrium for u and (say) the Nash solution v. Note that u^1 is not linear—although it is linear on a "ray" connecting 0 and the solution alternative.

Third, and of interest by itself, self-optimality gives an alternative characterization of efficient Nash equilibria in the utility distortion game.

References

- CRAWFORD, V. P., AND VARIAN, H. R. (1979). "Distortion of Preferences and the Nash Theory of Bargaining," *Econ. Lett.* **3**, 203–206.
- HURWICZ, L. (1972). "On Informationally Decentralized Systems," in *Decision and Organization* (R. Radner and B. McGuire, Eds.), pp. 297–336. Amsterdam: North-Holland.
- MOULIN, H. (1983). The Strategy of Social Choice. Amsterdam: North-Holland.

HANS PETERS

- MYERSON, R. (1979). "Incentive Compatibility and the Bargaining Problem." *Econometrica* **47**, 61–73.
- NASH, J. F. (1950). "The Bargaining Problem," Econometrica 18, 155-162.
- PELEG, B. (1984). Game Theoretic Analysis of Voting in Committees. New York: Cambridge Univ. Press.
- SOBEL, J. (1981). "Distortion of Utilities and the Bargaining Problem," *Econometrica* 49, 597-620.
- THOMSON, W. (1984). "The Manipulability of Resource Allocation Mechanisms," *Rev. Econ. Stud.* 51, 447-460.