IIM AHMEDABAD

Working Paper

COALITIONAL FAIRNESS AND DISTORTION OF UTILITIES

By

Somdeb Lahiri

W P No. 781
January 1989

The main objective of the working paper series of the IIMA is to help faculty members to test out their research findings at the pre-publication stage.

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380 056
INDIA

ABSTRACT

Given a finite number of agents with utilities who wish to divide a finite number of commodities, consider the non-cooperative game with strategies consisting of concave, increasing utility functions and whose outcomes are coalitionally fair solutions to the underlying equity problem determined by the strategies used. It is shown that for such a game any equal-income competitive equilibrium allocation for the true utilities is a Nash equilibrium outcome for the non-cooperative game.

ACKNOWLEDGEMENT

I have had the benefit of enlightening conversations on this paper and related topics with Prakash Chander, Arunava Sen and Gopal Naik. To them I would like to put on record my very deep appreciation.

1. Introduction:- The concept of coalitional fairness for an equity problem (as defined in Pazner and Schmeidler (1978)) originated in the work of Schmeidler and Vind (1972). It was subsequently refined and developed upon by Varian (1974), Jaskold-Gabszewicz (1975) and Svenslon (1983), in possibly different directions. The central idea is the allocation of a fixed supply of goods fairly amongst a population of fixed size.

In many economic or game - theoretical models predictions are based on information that is not observable. For example, Nash's (1950) theory of bargaining determines an outcome that depends on the bargainers' von Neumann-Morgenstern utility functions. Kurz (1977,1980) introduced a technique for analysing such models that yields predictions about the outcome of a game without relying on unobserved information. The technique of Kurz has been adopted by Crawford and Varian (1979) and Sobel (1981) to analyse the outcomes of the Nash Bargaining solutions over the division of commodities. Thomson (1979a and b) studies the Nash equilibria for the distortion game derived from a class of performance correspondences that yield individually rational and Pareto-efficient outcomes. Kalai and Rosenthal (1978) consider an arbitration model under ignorance, where a cooperative two-player bi-matrix game is transformed into a non-cooperative game by an arbitrator. Players are asked to report a mixed strategy (threat) and two payoff matrices. The arbitrator then determines an outcome using a procedure that generalizes Nash's (1953) extended bargaining solution if the players report the same payoff matrices. If the players report different payoff matrices then they receive the threat outcomes. Assuming that the players know the underlying cooperative game and that the arbitrator knows only the dimensions of its payoff matrices, Kalai and Rosenthal show that reporting the true payoff matrices and appropriate mixed strategies forms a Nash equilibrium for the arbitration games. Moreover, the equilibrium outcome is Pareto efficient and individually rational.

Lahiri (1988) considers a distortion game with fixed initial supply of resources where an impartial arbitrator allocates fair allocations amongst agents [redacted]. There it is shown that if a normalized price vector is an equal-income competitive equilibrium with respect to true preferences, then it is a Nash equilibrium for the utility distortion game.

The problem considered in this paper may be viewed as an arbitration problem under ignorance. An arbitrator is assigned the task of determining a 'coalitionally fair' outcome to an equity problem. A possible technique for the arbitrator would be to ask the players to report their utility functions and then determine an outcome to the resulting equity problem according to the 'coalition fairness' criterion. If the arbitrator has no knowledge about the players' true utilities except that they belong to a certain prespecified set, then the agents will be playing the distortion game we study here. We, however, assume that agents have perfect information about the game situation they are facing.

In section 2 of this paper we define the distortion game. In section 3 we establish the main results.

2. The Concept of Coalitional Fairness And The Distortion Game:-

We consider an economy with l commodities indexed by k, $k = 1,\dots,l$ and n traders, indexed by i, $i = 1,\dots,n$. The commodity space is IR^l_+. For each trader i, we assume the existence of a true utility function $u_i: IR^l_+ \longrightarrow IR$, which is continuous, concave and monotonic (i.e. $x \geq y$, $x \neq y \in IR^l_+$ implies $u_i(x) > u_i(y)$). Let $\underline{1} = (1,\dots,1) \in IR^l_+$ be the initial aggregate endowment of the commodities. We adopt the convention to denote the vector $(a,\dots,a) \in IR^l_+$ by $\underline{a}$. An allocation is an n-tuple $x = (x_1,\dots,x_n)$ of vectors in IR^l_+ verifying $\sum_{i=1}^{n} x_i = \underline{1}$. The equity problem is the assignment of an allocation to every economy specified by $\left[u_i\right]_{i=1}^{n}$. For the sake of convenience we will agree to denote the set of all allocations by

$$\pi = \left\{(x_1,\dots,x_n) \in (IR^l_+)^n \,/\, \sum_{i=1}^{n} x_i = \underline{1}\right\}.$$

The n players report utilities that are restricted to lie in the class $\mathcal{U}$, where $\mathcal{U}$ consists of those functions $U: T \to [0,1]$ such that

(a) U is continuous, strictly increasing and concave in T;

(b) U is normalized so that $U(\underline{0}) = 0$ and $U(\underline{1}) = 1$.

The class of admissible utilities should include those functions that are credible representations of the true preferences of the players. Condition (a) is a regularity assumption on the range of potential players. Condition (b) though inessential, greatly facilitates the presentation.

The distortion game for the equity problem is played by each agent revealing a utility function in $\mathcal{U}$. Given these reports, U_i for player i, a set of outcomes $G(U_1,\ldots,U_n)$ is selected. $G(U_1,\ldots,U_n)$ is the set of coalitionally fair outcomes (to be defined below) for the reported economy $[U_i]_{i=1}^{n}$.

Let $[U_i]_{i=1}^{n}$ be a set of reports. An allocation x is coalitionally fair (and thereby belongs to $G(U_1,\ldots,U_n)$) if and only if there exists no $S_1, S_2 \subseteq \{1,\ldots,n\}$ with $|S_1| \geq |S_2|$ and $y \in T$, such that

$$U_i(y_i) > U_i(x_i) \text{ for } i \in S_1$$

and

$$\sum_{i \in S_1} y_i = \sum_{i \in S_2} x_i$$

Intuitively, an allocation x would not be coalitionally fair if and only if there would exist two coalitions S_1 and S_2 with S_1 at least as large as S_2, such that S_1 could benefit from achieving the aggregate allocation of S_2 under x. Setting $S_1 = S_2 = \{1, \ldots, n\}$, it can easily be seen that any coalitionally fair allocation is Pareto-efficient, (i.e. $x \in T$ is Pareto efficient if and only if $y \in T$ and $\forall\, i \in \{1, \ldots, n\}$, $u_i(y_i) \geq u_i(x_i)$ implies $u_i(y_i) = u_i(x_i)$ $\forall\, i \in \{1, \ldots, n\}$). The continuity of preferences is crucial in establishing this result.

In what follows, we shall require the following definition of an <u>equal-income competitive equilibrium</u> (EICE).

<u>Definition</u> :- An <u>equal-income competitive equilibrium</u> (EICE) is an ordered (n+1)- tuple, $(p^*; x_1^*, \ldots, x^*_n)$ where

a) $p^* \in IR^l_+$, $p^* \neq 0$, $\sum_{k=1}^{l} p^*_k = 1$

b) $x^* = (x_1^*, \ldots, x^*_n) \in T$

c) x^*_i solves

$$\max u_i(x_i) \text{ subject to } p^*.x \leq \frac{1}{n} p^*.\underline{1} = \frac{1}{n}, \quad x \in IR^l_+.$$

The vector x^* will be referred to as the competitive allocation. On occasion, a vector $p = (p_1, \ldots, p_l)$ will be used to refer to the linear function from IR^l_+ to IR, where

$$p(x) = p.x = \sum_{i=1}^{l} p_i x_i$$

No confusion should arise.

<u>Proportion 1</u>:- Let $(p^*; x^*)$ be an EICE for $[u_i]_{i=1}^{n}$. Then $x^* \in G(u_1, \ldots, u_n)$.

<u>Proof</u>: Assume that $(p^*; x^*)$ is an EICE and $x^* \notin G(u_1, \ldots, u_n)$. Then there exists S_1 and S_2, $|S_2| \leq |S_1|$, and $y \in T$ such that

$u_i(y_i) > u_i(x^*_i)$, $i \in S_1$, and

$$\sum_{i \in S_1} y_i = \sum_{i \in S_2} x_i \, .$$

Now $u_i(y_i) > u_i(x_i^*)$ for all $i \in S_1 \Rightarrow p \cdot y_i > \frac{1}{n}$ for all $i \in S_1$

So $p \cdot \sum_{i \in S_1} y_i > \frac{|S_1|}{n} \geq \frac{|S_2|}{n} = p \cdot \sum_{i \in S_2} x_i$ (the last equality being valid by montonicity of preferences).

This contradicts that $\sum_{i \in S_1} y_i = \sum_{i \in S_2} x_i$.

Hence $x^* \in G(u_1, \ldots, u_n)$.

Q.E.D.

In the above proof the only property of preferences which we effectively used was its montonicity.

Finally, we introduce the notion of a Nash equilibrium for the distortion game defined above.

Definition :- The strategies $(U^*_1, \ldots, U^*_n) \in \mathcal{U}^n$ constitute a Nash equilibrium for the distortion game determined by G if and only if there exists $\bar{x} \in G(U^*_1, \ldots, U^*_n)$ so that for all $i \in \{1, \ldots, n\}$, $\bar{x}_i$ solves

$$\max \; u_i(x_i)$$

$$\text{subject to } x \in G(U^*_1, \ldots, U^*_{i-1}, U_i, U^*_{i+1}, \ldots, U^*_n),\; U_i \in \mathcal{U}$$

and $x = (x_1, \ldots, x_i, \ldots, xn)$.

The desirability of such an equilibrium strategy follows from the self enforceability of the Nash non-cooperative solution concept.

3. Main Results :- The main theorem can now be stated.

Theorem 1 :- If $(p^*; x^*_1, \ldots, x^*_n)$ is an EICE for true preferences, then $(p^*, \ldots, p^*)$ is a Nash equilibrium for the distortion game.

Proof :- The proof proceeds by a sequence of lemmas.

Lemma 1:- If $x \in G(p^*, \ldots, p^*, U, p^*, \ldots, p^*)$, then $p^*.x_i = p^*.x_j$ $\forall\, i, j \in \{1, \ldots, m-1, m+1, \ldots, n\}$ and $p^*.\, x_m \leq p^*.x_i$ for all $i \in \{1, \ldots, m-1, m+1, \ldots, n\}$, where U is the strategy of player [redacted] 'm'.

Further, $p^*.\, x_m \leq \frac{1}{n}$.

<u>Proof</u> : Suppose, $p^*.\ x_m > p^*.x_i$ for some $i \in \{1,\ldots,m-1,m+1,\ldots,n\}$.

Define, $y \in T$ as follows:

$$y_j = x_j \quad \text{for } j \neq i,\ m$$

$$y_i = x_m$$

$$y_m = x_i$$

$S_1 = \{m\}$, $S_2 = \{i\}$.

Clearly, $p^*.y_i > p^*.x_i$ and $\sum_{j \in S_1} y_j = \sum_{j \in S_2} x_j$

Hence $x^* \notin G(p^*,\ldots,p^*,\ U,\ p^*,\ldots,p^*)$ a contradiction.

So, $p^*.\ x_m \leqslant p^*.x_i$ for all $i \in \{1,\ldots,\ m-1,\ m+1,\ldots,n\}$.

Also, $p^*.x_i = p^*.x_j$ for all $i,\ j \in \{1,\ldots,\ m-1,\ m+1,\ldots,n\}$, follows by a symmetric application of the above argument.

Since, $p^*.\ x_m \leqslant p^*.x_i$ for all $i \in \{1,\ldots,\ m-1,\ m+1,\ldots,\ n\}$ and since $x \in T$, $p^*.\sum_{j=1}^{n} x_j = 1$,

$$\therefore\ p^*.\ x_m + p^*.\sum_{j \neq m} x_j = 1 \geqslant n\ (p^*.x_m).$$

$$\therefore\ p^*.\ x_m \leqslant \frac{1}{n}.$$

This proves the lemma.

Q.E.D.

<u>Lemma 2</u> : $x^* \in G(p^*,p^*,\ldots,p^*) \Leftrightarrow p^*.x^*_i = \frac{1}{n}$, $i \in \{1,\ldots,n\}$.

<u>Proof</u> : Let $x^* \in G(p^*,p^*,\ldots, p^*)$. Then by Lemma 1, $p^*.x_i = p^*. x_j = \frac{1}{n}$ for all $i, j \in \{1,\ldots, n\}$.

Now suppose, $x^* \in T$ and $p.^* x_i = \frac{1}{n}$, $i \in \{1,\ldots, n\}$.

Assume towards a contradiction, that $x^* \notin G(p^*, p^*,\ldots, p^*)$.

Then, there exists $S_1, S_2 \subseteq \{1,\ldots, n\}$, $|S_2| \leq |S_1|$ and $y \in T$ such that

$$p^*.y_i > p^*.x^*_i \quad \forall \; i \in S_1.$$

$$\sum_{i \in S_1} y_i = \sum_{i \in S_2} x^*_i.$$

$$\therefore \; p^*. \sum_{i \in S_1} y_i > p^*. \sum_{i \in S_1} x^*_i \geq p^*. \sum_{i \in S_2} x^*_i$$

This contradicts that $\sum_{i \in S_1} y_i = \sum_{i \in S_2} x^*_i$.

Q.E.D.

<u>Proof of Theorem 1</u> : Without loss of generality suppose,

$$x \in G\,(U, p^*, \ldots, p^*), \quad U \in \mathcal{U}$$

By Proposition 1, if $(p^*; x^*_1,\ldots, x^*_n)$ is an EICE then

$$x^* = (x^*_1,\ldots, x^*_n) \in G(p^*,\ldots,p^*) =$$

$$\left\{ x \in T / p^*.x_i = \frac{1}{n}, \; i \in \{1,\ldots, n\} \right\} \quad \text{(by Lemma 2)}$$

Further by the definition of an EICE, x^*1 solves,

max $u_1 \; (x_1)$

subject to $p^*. \; x_1 \leq 1/n$.

$\therefore$ x^*_1 solves,

$\max u_1(x_1)$

subject to $x \in G(U, p^*, \ldots, p^*)$, $U \in \mathcal{U}$.

Repeating the same argument for $i \in \{1, \ldots, n\}$, we get that $(p^*, \ldots, p^*)$ is a Nash equilibrium for the distortion game.

Q.E.D.

The EICE is attained as follows: each player reports linear preferences with indifference surfaces parallel to the supporting prices. The set of solutions to the equity problem then consists of an entire hyperplane. However, since the hyperplane supports the EICE, both agents can agree on a most preferred point (with respect to their true preferences). This point is a competitive allocation.

Theorem 1 has a partial converse.

<u>Theorem 2</u> :- If $(p^1, \ldots, p^n)$ in a Nash equilibrium for the distortion game with each p^i satisfying condition (a) of the definition of an EICE and if an interior allocation is Pareto-efficient with respect to reported utilities $(p^1, \ldots, p^n)$, then there exists $x^* \in G(p^1, \ldots, p^n)$ such that x^* is an EICE allocation with respect to true preferences.

<u>Proof</u> : Choose $x^* \in G(p^1, \ldots, p^n)$ such that x^* is the Nash equilibrium outcome. Clearly x^*_i solves

$$\max u_i(x_i)$$

$$\text{subject to } p^i . x_i \leq p^i . x^*_i$$

for all $i \in \{1, \ldots, n\}$. It remains to show that

(i) $p^i = p^j$, $i, j \in \{1, \ldots, n\}$. Then by montonicity of preferences we get that x^* is a competitive equilibrium allocation.

(ii) $p^i . x^*_i = \frac{1}{n} \quad \forall \; i \in \{1, \ldots, n\}$. Then we get that x^* is an EICE allocation.

Observe that $p^i . x^*_i \geq \frac{1}{n} \quad \forall \; i \in \{1, \ldots, n\}$

Suppose not i.e. $p^i x^*_i < \frac{1}{n}$.

Then since $\sum_{j=1}^{n} x^*_j = \underline{1}$ and $p^i . \sum_{j=1}^{n} x^*_j = 1$, there must exist some j for which

$$p^i . x^*_j > \frac{1}{n}$$

Let $S_1 = \{i\}$, $S_2 = \{j\}$ and $y \in T$ with

$$y_h = x^*_h \text{ if } h \neq i, j$$

$$y_i = x^*_j$$

$$y_j = x^*_i$$

Clearly, $p^i . y_i > p^i . x^*_i$

and $\sum_{h \in S_1} y_h = \sum_{h \in S_2} x^*_h$.

This contradicts that $x^* \in G(p^1, p^2, \ldots, p^n)$.

First let us show that $p^i = p^j \;\forall\; i, j \in \{1, \ldots, n\}$.

Since an interior allocation is Pareto efficient with respect to reported utilities $p^1, \ldots, p^n$, we must have $p^i = p^j \;\forall\; i, j \in \{1, \ldots, n\}$

This proves (i).

Now since $p^i \cdot x^*_i \geq \frac{1}{n} \;\forall\; i \in \{1, \ldots, n\}$ and equal division is Pareto efficient, we must have $p^i \cdot x^*_i = \frac{1}{n} \;\forall\; i \in \{1, \ldots, n\}$.
This proves (ii) and the theorem.

4. Conclusion : In this paper we show that coalitionally fair allocations are implementable in Nash strategies. Similar results for other performance correspondences are available in Sobel (1981) and Lahiri (1988). However, the underlying game considered here is different from the ones considered earlier.

In Hurwicz (1972), it is shown that for the same class of environments, the individually rational and Pareto optimal correspondence (in the context of a private ownership economy) is non-implementable. The same example (example D) in his paper would show that the Pareto correspondence for the class of economies considered by us is non-implementable. What we have shown however is that by suitably restricting the performance correspondence, we are able to implement our desired social goals. The reason behind this is simple. The Pareto correspondence being too large, any player given the Nash strategy of the other player can deviate to a strategy which yields a better non-Pareto optimal allocation with regard to true preferences and yet remains Pareto-optimal with regard to stated preferences. By reducing the size of the performance correspondence, we are not eliminating better allocations for the players. We only guarantee that these better allocations do not belong to the image of the performance correspondence obtained by an unilateral derivation from the Nash strategies.

REFERENCES:

1. CRAWFORD, V.P. and Varian H.R. (1979): Distortion of Preferences and the Nash theory of bargaining, Economics Letters 3, 203-6.

2. HURWICZ, L. (1972): On informationally decentralized systems, In 'Decision And Organization', a volume in honour of Jacob Marschak, edited by C.B. McGuire and Roy Radmer, North-Holland.

3. Jaskold-Gabszewics, J. (1975): Coalitional Fairness of Allocations in Pure Exchange Economics, Econometrica, 43, 661-669.

4. Kalai, E and Rosenthal, R.W. (1978): Arbitration of two-party disputes under ignorance. International Journal of Game Theory, 7, 65-72.

5. Kurz, M. (1977): Distortion of Preferences, Income distribution and the case for a linear income tax. Journal of Economic Theory, 14, 291-8.

6. Kurz, M. (1980): Income distribution and distortion of preferences: The 1-commodity case. Journal of Economic Theory, 22, 99-106.

7. Lahiri, S. (1988): Fair Allocations and Distortion of Utilities: A Note, Indian Institute of Management (Mimeo).

8. Nash, J. (1950): The bargaining problem, Econometrica, 18. 155-162.

9. Nash, J. (1953): Two-person cooperative games, Econometrica. 21, 128-40.

10. Pazner, E.A. and Schmeidler, D. (1978): Egalitarian Equivalent Allocations: A New Concept of Economic Equity, Quarterly Journal of Economics, 92-671-687.

11. Schmeidler, D. and K. Vind (1972): Fair Net Trades, Econometrica 40, 637-642.

12. Sobel, J (1981): Distortion of Utilities and the Bargaining Problem, Econometrica 19, 597-617.

13 Svensson, L.G. (1983): Large Indivisibilities : An Analysis With Respect To Price Equilibrium And Fairness, Econometrica 51, 939-954.

14. Thomson, W (1979a): The manipulability of the Shapley-value. University of Minnessota. Discussion Paper No. 79-115 (reprinted in International Journal of Game Theory (1988)).

15. Thomson, W (1979b): On the manipulability of rescurces allocation mechanisms designed to achieve individually rational and Pareto-optimal outcomes. University of Minnesota Discussion Paper No. 79-116.

16. Varian, H.R. (1974): Equity, Envy and Efficiency. Journal of Economic Theory. Journal of Economic Theory, 9, 63-91.