Economics Letters 5 (1980) 119–123

IGNORANCE AND MANIPULATION

Lucian A. BEBCHUK *
Harvard University, Cambridge, MA 02138, USA

Received 11 August 1980

This note proves that, if individuals are under ignorance regarding other's preferences, every neutral and positively responsive social choice function is non-manipulable.

1. Introduction

Gibbard (1973) and Satterthwaite (1975) proved that every social choice function which is non-imposed, non-dictatorial and has a range of at least three alternatives is manipulable. Since then much work has been done to identify different conditions and schemes that will rule out strategic manipulation [e.g., Barbera (1977), Gibbard (1978), Kelly (1977), Pasner and Wesley (1977)].

This note deals with the question of manipulability in the informational setting of ignorance. Under ignorance individuals have no information about the preferences the others will express and therefore perceive all the possible profiles of the others' preferences as equally likely. A theorem is proved that in a situation of ignorance every social choice function which is neutral and positively responsive is non-manipulable. This theorem demonstrates the role of information in manipulation.

2. Definitions and notation

Let $N = \{1, 2, ..., n\}$ be the set of individuals. A is the finite set of m alternatives. Each individual $i \in N$ has prefęrence R_i which is a complete, transitive, asymmetric ordering over A. Let $R = (R_1, ..., R_n)$ be the preference profile. Let Σ be the set of all possible preferences and let $D = \Sigma^n$, the n-fold Cartesian product of Σ be the set of all possible preference profiles. A social choice function (SCF) is a function $f : \Sigma^n \to A$, choosing for every profile of preferences an alternative in A.

Some functions that will be used in the proof of the theorem are defined below for every $x, y \in A, j \in N$,

$T^1_{xy} : \Sigma \to \Sigma$.

* I wish to thank Professor Kenneth Arrow for commenting on a draft of this note.

$\forall k \in \Sigma$, $T^1_{xy}(k)$ is the preference k with the only modification that the places of x and y were exchanged. If we represent the elements of Σ by the appropriate permutations then T^1_{xy} is multiplying by the transposition (xy).

$T^2_{jxy}, T^2_{xy} : D \to D$,

$$\forall R \in D, \qquad \begin{aligned} T^2_{jxy}(R)_i &= R_i\,, \quad \forall i \neq j\,, \\ T^2_{jxy}(R)_j &= T^1_{xy}(R_j)\,, \end{aligned}$$

$$\forall R \in D, \qquad T^2_{xy}(R)_i = T^1_{xy}(R_i)\,, \quad \forall i\,.$$

T^2_{jxy}, T^2_{xy} modify a profile R by exchanging the places of x and y in the preference of individual j, or all the preferences in the profile, respectively.

$T^3_{jxy}, T^3_{xy} : 2^D \to 2^D$,

$$\forall D' \subseteq D, \qquad \begin{aligned} T^3_{jxy}(D') &= \{R \in D | R = T^2_{jxy}(R'),\, R' \in D'\}\,, \\ T^3_{xy}(D') &= \{R \in D | R = T^2_{xy}(R'),\, R' \in D'\}\,. \end{aligned}$$

The theorem requires that the SCF be neutral and positively responsive.

A SCF f is *neutral* if for every $R \in D$, $x, y \in A$,

$$f(R) = x \quad \Leftrightarrow \quad f[T^2_{xy}(R)] = y\,,$$

$$f(R) = y \quad \Leftrightarrow \quad f[T^2_{xy}(R)] = x\,,$$

$$f(R) = z,\ z \neq x, y \quad \Leftrightarrow \quad f[T^2_{xy}(R)] = z\,,$$

A SCF f is *positively responsive* if for every $R \in D$, $x, y \in A$, $j \in N$,

$$\text{if } yR_jx \text{ and } f(R) = x \quad \text{then} \quad f[T^2_{jxy}(R)] = x\,.$$

It can be easily shown that a SCF which is both neutral and positively responsive must choose a Pareto superior alternative (if such exists), and cannot choose a Pareto inferior alternative.

3. Individual's ignorance and decision

It is assumed that individuals are in ignorance regarding the preferences that the others will express [1] and therefore perceive every possible profile of the others' preferences equally likely to be expressed. If an individual $i \in N$ expresses prefer-

[1] The ignorance required is regarding the others' expressed preferences (which may differ from their true ones due to manipulation). However, if the SCF is neutral, ignorance regarding the true preferences will lead to ignorance also regarding the expressed preferences.

ence $k \in \Sigma$, the set of possible profiles he faces is $D_k = \{R \in D | R_i = k\}$ and he assigns equal probability to every $R \in D_k$.

The individual has to decide which preference k, from the set Σ, to express. For every $x \in A$, $k \in \Sigma$, let $D_{kx} = \{R \in D_k | f(R) = x\}$. The choice of k determines D_k and the subsets D_{kx} for every $x \in A$ and thus assigns to each alternative a probability of being adopted.

Let P_{kx} be the probability of alternative x being adopted, given that the expressed preference is k. [2] Let P_k be the vector of the probabilities P_{kx} for all the alternatives in A. The choice of k is therefore a choice between lotteries and depends on the individual's ordering over these lotteries. This ordering depends on the individual's ordering over the alternatives in A and the method he employs for decisions under uncertainty. Assuming that the method for decision under uncertainty is rational it follows that for every two lotteries P', P'' such that $P'_x \geqslant P''_x$, $P'_y \leqslant P''_y$, $P'_z = P''_z$, $\forall z \neq x, y$, for some $x, y \in A$, if alternative x is preferred to alternative y by the individual, then the lottery P'' will not be preferred by him to the lottery P'. This is the only assumption about the ordering of the individual over the lotteries used in the proof of the main result. An individual will manipulate if as a result he will face a lottery which he prefers to the one he faces expressing his true preference.

4. The theorem and its proof

Theorem. *If a SCF f is neutral and positively responsive and if individuals know f but are ignorant about others' expressed preferences, then f is non-manipulable.*

Proof. Let k_0 be the true preference of an individual $i \in N$.

Lemma. *Let $k', k'' \in \Sigma$, $k'' = T^1_{xy}(k')$, for some $x, y \in A$. If $xk'y$ xk_0y then $P_{k''}$ is not preferred to $P_{k'}$. The meaning of this lemma is that an individual cannot benefit from modifying his expressed preference by exchanging, in constrast with their order in his true preference, the places of two alternatives.*

Proof. From the construction of k'',

$$D_{k''} = T^3_{ixy}(D_{k'}) \,, \tag{1}$$

$$D_{k''} = T^3_{xy}(D_{k'}) \,; \tag{2}$$

from (1), $xk'y$ and the positive responsiveness of f,

$$P_{k'y} \leqslant P_{k''y} \,; \tag{3}$$

[2] This probability is equal to the ratio $|D_{kx}|/|D_k|$ of the number of profiles in D_k for which the SCF chooses this alternative to the number of profiles in D_k.

from (2) and the neutrality of *f*,

$$P_{k'x} = P_{k''y} , \tag{4}$$

$$P_{k'y} = P_{k''x} , \tag{5}$$

$$P_{k'z} = P_{k''z} , \qquad \forall z \neq x, y ; \tag{6}$$

putting together (3), (4), (5) and (6),

$$P_{k'x} \geqslant P_{k''x} ,$$

$$P_{k'y} \leqslant P_{k''y} ,$$

$$P_{k'z} = P_{k''z} , \qquad \forall z \neq x, y ;$$

and as xk_0y the individual does not prefer $P_{k''}$ to $P_{k'}$. Q.E.D.

Returning to the proof of the theorem we turn to examine the modifications that are done in the true preference for manipulation. Every $k \neq k_0$, $k \in \Sigma$ can be constructed from k_0 by being the last element of a sequence $k_0, k_1, ..., k_l$ which satisfies for every $j = 1, 2, ..., l$, $k_j = T^1_{xy}(k_{j-1})\, xk_{j-1}y\, xk_0y$ for some $x, y \in A$. This is direct application of the theorem that every permutation is a product of transpositions. See Herstein (1964, ch. 2). Every manipulation is thus a sequence of modifications in the true preferences, none of them, according to the lemma, is deemed beneficial by the individual. This completes the proof of the theorem.

5. Discussion

The result proved in this paper can be given an intuitive explanation. The neutrality of the SCF coupled with the individual's ignorance about others' preferences face him, before he decides what preference to express, with equal chance for every alternative. (His ignorance is not enough, as if the SCF is not neutral the alternatives it favors will have a greater chance.) The individual will thus have no reason not to allocate his support among the alternatives according to his true preference. If the SCF is positively responsive, the higher an individual ranks an alternative in his expressed preference, the stronger support he gives to it. Therefore the individual will have no reason to rank in his expressed preference an alternative above one that he sincerely prefers.

The theorem assumed that individuals know the SCF. This condition can be weakened. Assume first that the SCF used is chosen from the set of neutral and positively responsive SCF's. It can be proved that if individuals are ignorant about the others' preferences and are uncertain (including ignorant) about which neutral and positively responsive SCF is used they will not manipulate. They face a probability distribution over states, in none of which, this paper has proved, is it beneficial to manipulate. Imperfect information about the SCF is thus compatible with non-manipulability. Yet, it is not sufficient for it (i.e., when not coupled with ignorance about others' preferences). To look at an example, assume individuals are ignorant about the SCF but know the others' preferences. An individual who faces a strong support for some alternative by the others assigns to it, knowing the SCF is neutral and positively responsive, a high probability of being chosen, which may lead him to manipulation.

Assume now that the SCF is not necessarily positively responsive, that it is chosen from the set of all neutral SCF's or from the set of all SCF's. In this case, which seems to be an unattractive one, ignorance about the SCF is sufficient for non-manipulability.

The case of ignorance is obviously an unrealistic one. Nevertheless, this extreme case serves to highlight the connection between information and manipulation. A further inquiry is necessary into the role information may have inside the domain between ignorance and perfect information, in addition to the crucial role information was proved to play on the corner of this domain.

References

Barbera, Salvador, 1977, The manipulation of social choice mechanisms that do not leave 'too much' to chance, Econometrica 45, 1573–1588.

Gibbard, Allan, 1973, Manipulation of voting schemes: A general result, Econometrica 41, 587–602.

Gibbard, Allan, 1978, Straightforwardness of game forms with lotteries as outcomes, Econometrica 46, 595–614.

Herstein, Israel N., 1964, Topics in algebra (Xerox College Publishing, Lexington).

Kelly, Jerry S., 1977, Strategy-proofness and social choice functions without singlevaluedness, Econometrica 45, 439–445.

Pasner, Elisha A. and Eugene Wesley, 1977, Stability of social choices in infinitely large societies, Journal of Economic Theory 14, 252–262.

Satterthwaite, Mark A., 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.