

Discussion Paper No. 175

MANIPULATION OF SCHEMES THAT MIX VOTING WITH CHANCE \*

by

ALLAN GIBBARD

January 1976

Department of Philosophy

University of Pittsburgh

\* I have been helped in revising this paper by conversations with Mark Satterthwaite, Thomas Schwartz, and Hugo Sonnenschein, and by letters from Peter Fishburn and Richard Zeckhauser.

I am grateful to the referee for remarkably detailed suggestions for shortening the proof of the main theorem.

## I. Introduction

An individual manipulates a system of voting if, by misrepresenting his preferences, he secures a result he prefers to the result that would obtain if he expressed his true preferences. For systems of pure voting, where chance plays no role in settling which alternative is adopted, the following result is known: such a scheme, if it is to preclude individual strategic manipulation, must either make someone dictator, or restrict the possible outcome to a fixed pair of alternatives. (See Gibbard, 1973, and Satterthwaite, 1975). This paper deals with systems of voting of a more general kind: systems by which a social decision is made through a combination of voting and chance. It will be shown that any such scheme, if it is to preclude individual strategic manipulation, must be a probability mixture of schemes, each of which either (1) accords a monopoly of influence to a single voter, or (2) restricts the final outcome to a fixed pair of alternatives. Schemes of the first kind I shall call unilateral; of the second kind, duple.

What is meant here by a combination of voting with chance? Suppose a decision is made in the following way: first, voting of some kind is used to pick out a set of one or more winning alternatives; then, in case there is more than one such winner, one of them is chosen by lot. Such a scheme, in effect, uses the way people vote to determine the probability each alternative has of being adopted. This I shall take as the defining feature of a scheme which combines voting with chance: on the basis of the way people vote, it assigns to each alternative a probability of being adopted.

This paper deals only with voting by rank order ballot: in the schemes to be considered here, voting consists in each voter's ranking the alternatives in

a professed order of preference. An individual is not allowed to express indifference between alternatives. The theorem in this paper applies to all systems of the kind I have characterized: to all systems by which voters' rank order ballots — no indifference allowed — determine the probability of each alternative's being adopted.

Systems of this kind will be called decision schemes, and they are defined, more precisely, as follows. Let there be a finite set of mutually exclusive alternatives, from which the community must select exactly one. Each voter ranks the alternatives on his ballot in professed order of preference. On the basis of these orderings, a probability of being adopted is assigned to each alternative, and the final choice is made by a suitable chance device. A decision scheme, then, is a function of the following kind. Let there be  $n$  voters, and let  $V$  be the set of mutually incompatible alternatives open to the community. Call an ordering of  $V$  with no ties a ranking and call an  $n$ -tuple of rankings a ranking  $n$ -tuple. Finally, let a lottery be an assignment of a probability to each alternative, with the probabilities adding up to one. A decision scheme is a function  $d$  whose domain is the set of all preference  $n$ -tuples, and whose values are lotteries.

How can manipulability be defined for decision schemes? A decision scheme is manipulable if there is a logically possible situation in which someone manipulates it; and an individual manipulates a decision scheme if, by misrepresenting his preferences, he secures a lottery he prefers to the lottery that would have obtained if he had expressed his true preferences. Whether he manipulates the scheme, then, depends on his preferences among lotteries. Now if an ordering of lotteries satisfies rationality conditions such as those of von Neumann and Morgenstern (1947, p. 26), then it can most

conveniently be given by an assignment of cardinal utilities to the alternatives. Whether individual  $k$  manipulates the scheme to his advantage, then, depends not only on the way everyone else votes, the way  $k$  votes, and the way  $k$  genuinely orders the alternatives; it depends further on the way  $k$  genuinely orders lotteries — on  $k$ 's cardinal utilities.

Manipulability, then, can be characterized as follows. In the first place,  $k$  manipulates decision scheme  $d$  if (a) where the actual votes are given by ranking  $n$ -tuple  $\langle P_1, \dots, P_n \rangle$  and  $k$ 's true utility scale is  $U$ ,  $k$ 's avowed ranking  $P_k$  is not the ranking of the alternatives given by scale  $U$ , and (b) if  $k$  had voted the ranking given by scale  $U$ , he would have secured a lottery of lower expected utility, as reckoned by  $U$ , than the lottery he actually secures. A decision scheme  $d$  is manipulable, then, if for some ranking  $n$ -tuple  $\langle P_1, \dots, P_n \rangle$ , for some person  $k$ , and for some utility scale  $U$ ,  $k$  manipulates  $d$ . If it is not manipulable, it will be called strategy - proof. These definitions are given explicitly in Section 4.

Unattractive examples of strategy - proof decision schemes are not hard to find. Here are three: Scheme 1: Put everyone's ballot in a hat, draw one at random, and choose the alternative which is ranked first on that ballot. (For a discussion of this scheme, see Gibbard, 1973, pp. 592-3, and Zeckhauser, 1973, pp. 938-40) Scheme 2: First collect the ballots. Next, put the names of the alternatives in a hat and select two at random. Then use the collected ballots to decide between those two alternatives by majority vote. This amounts to a decision scheme, since under it, the ballots cast determine the probability of each alternative's being adopted. Now if a voter misrepresents his preferences under this scheme, it can affect the outcome only to his

disadvantage. His misrepresentation can affect the outcome only if the following holds: for some pair of alternatives  $x$  and  $y$ , he prefers  $x$  to  $y$  but ranks  $y$  above  $x$  on his ballot, the names of  $x$  and  $y$  are drawn from the hat, and he swings the outcome from  $x$  to  $y$  by his vote — thus getting an outcome he likes less than the honest outcome. <sup>1/</sup> Scheme 3: A coin is flipped, and Scheme 1 is used if the coin lands heads; Scheme 2 if the coin lands tails.

It might have been hoped that there were strategy - proof decisions schemes more attractive than these: schemes, for instance, which select one or more optimal alternatives in a reasonable way on the basis of the way people vote, and then, in case there is more than one optimal alternative, choose the alternative actually to be adopted from among them by chance. The theorem in this paper shows, however, that all strategy - proof decision schemes are much like the unattractive schemes I have given as examples: all involve, in effect, selecting a ballot or a pair of alternatives by chance, and either ignoring all ballots but the one selected, or choosing somehow between the two selected alternatives. All, in other words, are probability mixtures of schemes, each of which is either unilateral or duple.

The precise statement and proof of this theorem is given in Section 4. Three corollaries are stated and proved in Section 5. The first is this: suppose a decision scheme guarantees Pareto optimal outcomes. Suppose, in other words, that no matter how people vote, if one alternative is unanimously outranked by another, then it gets a probability of zero. Suppose also that there are at least three alternatives, and that the decision scheme is strategy-proof. Then the decision scheme is a probability mixture of dictatorial schemes. <sup>2/</sup>

The second corollary is this. Suppose a decision scheme gives lotteries which are Pareto optimal ex ante, where a lottery is Pareto optimal ex ante if there is no other lottery which is unanimously preferred to it. Suppose, in other words, that no matter what each person's utility scale is, if each person votes the ranking of alternatives given by his utility scale, then the resulting lottery  $\rho$  has this property: there is no other lottery  $\rho'$  which ranks higher than  $\rho$  on everyone's utility scale. Suppose, as before, that there are at least three alternatives, and that the decision scheme is strategy - proof. Then the decision scheme is dictatorial — it is not, that is to say, merely a probability mixture of dictatorial schemes; it is itself dictatorial. This corollary extends to schemes which allow the expression of individual indifference.

The third corollary is simply the earlier theorem on non - chance voting schemes (Gibbard, 1973). The proof in this paper, then, constitutes a new proof of that earlier theorem.

## II. Background

The notion of manipulability used in this paper is a variant of the one formulated by Dummett and Farquharson (1961). Manipulability and closely related matters are discussed by Arrow (1951, p.7), Vickrey (1960), Murakami (1968, pp.74-81), Farquharson (1969), Sen (1970, pp. 192-6), and Pattanaik (1973, 1974, 1975, forthcoming). The theorem cited at the outset of this paper is proved independently, in quite different ways, by Gibbard (1973) and Satterthwaite (1975). A precise statement of the theorem is this: any scheme which uses rank order balloting in a non-chance way to select a single alternative is either manipulable, dictatorial (in that someone is guaranteed his first choice from among the possible outcomes), or restricted to no more than two possible outcomes. This result holds both for schemes which allow individual indifference to be expressed and for schemes which do not. A streamlined proof of the theorem is given by Schmeidler and Sonnenschein (1974). This earlier theorem does not apply to systems of voting which allow tied outcomes. In my discussion of that theorem (1973, pp. 592-3), I argued that it makes no sense to study the manipulability of schemes which allow ties unless one considers the system by which ties are to be broken. If ties are to be broken by chance, I argued, then the full system to be studied in effect yields outcomes which are lotteries among alternatives.

It was Zeckhauser (1969) who broached the study of voting with lotteries as alternatives. Fishburn (1972a, 1972b) studies the subject further. A subsequent paper by Zeckhauser (1973) is on virtually the topic of this paper. There Zeckhauser studies systems of voting which rely on individuals' self-interested balloting and may have lotteries as outcomes. He concludes (Theorem V, p. 945) that no such system can guarantee an outcome which is both ex ante

Pareto - optimal and, in a special sense, "non-dictatorial". (For the case of two voters, an outcome is "dictatorial" in Zeckhauser's sense if it is the first choice of one voter and the last choice of the other). Zeckhauser's result is logically independent of the one in this paper. It is stronger in one respect: the results here are confined to systems with rank order balloting, whereas Zeckhauser's is not. Zeckhauser, on the other hand, requires that the lotteries that serve as outcomes of the schemes he considers be Pareto-optimal ex ante, and, in his special sense, "non-dictatorial". The main theorem in this paper does not invoke Pareto optimality, and whereas the second corollary here is suggested by Zeckhauser's result, Zeckhauser's non-dictatorship condition is much stronger than the non-dictatorship condition in the corollary. Zeckhauser's result, then, neither directly entails the results in this paper nor is directly entailed by them.



### III. Individual Indifference

The main result in this paper fails when extended to systems that permit a voter to express indifference between alternatives. No doubt the easiest example of this failure is a serial dictatorship. Let a fixed "dictator" always get his first choice, and if more than one alternative ties as his first choice, let a "prime henchman" get his first choice from among the alternatives the dictator likes best. Let further ties be broken arbitrarily, say by selecting the tying alternative which is first on some predetermined list. Now a serial dictatorship is clearly strategy-proof: neither the dictator, the prime henchman, nor anyone else can ever gain by misrepresenting his preferences. If there are more than two alternatives, however, then the serial dictatorship is not equivalent to any probability mixture of unilateral or duple schemes, as the following considerations show.

Note at the outset that a serial dictatorship is not unilateral. A scheme is unilateral, in the sense the term has been given here, only if it accords a single voter — call him the ruler — a monopoly of influence, so that no matter how anyone votes, the ballots of all voters other than the ruler are ignored. Under a serial dictatorship, the ballot of the prime henchman is not invariably ignored, and hence a serial dictatorship is not itself unilateral.

Now a serial dictatorship is not a probability mixture which has any duple scheme as a part. If it were, then for at least one fixed pair of alternatives, the probability of the adopted alternative's being in that pair would have to be non-zero independently of how anyone voted. Under a serial dictatorship, both alternatives in any pair have probability zero of being adopted whenever neither is a first choice of the dictator. Thus if a serial dictatorship is a probability mixture of schemes each of which is unilateral or duple, then it is a probability mixture of unilateral schemes alone.

The only unilateral schemes that could be part of this mixture, though, are ones for which the dictator is ruler. Otherwise, there would be a non-zero chance that no matter what the dictator's ballot said, it would be ignored, so that for some way the dictator and others might vote, an alternative which was not the first choice of the dictator would be adopted with non-zero probability. Under a serial dictatorship, on the other hand, the probability that the dictator will fail to get his first choice (or one of his first choices in case he has no unique first choice) is always zero. Thus if a serial dictatorship were a probability mixture of unilateral and duple schemes, it would have to consist of a single unilateral scheme with the dictator as ruler — and we have already seen that it does not.

A serial dictatorship, then, is not a probability mixture of schemes which are unilateral or duple, and hence the theorem in this paper does not in general extend to systems which allow individuals to express indifference.

What, then, can be said about systems with ballots which do allow individual indifference to be expressed? What the theorem here tells us is this: if such a system  $s$  is strategy-proof, then there is a probability mixture  $m$  of unilateral and duple schemes which coincides with  $s$  whenever no one is indifferent between any pair of alternatives. For any ranking  $n$ -tuple  $\tilde{P}$  with no ties, in other words,  $m$  assigns the same prospect to  $\tilde{P}$  as does  $s$ .

The force of the theorem, then, extends to systems which allow individuals to express indifference. For the force of the theorem lies in the judgment that any probability mixture of unilateral and duple schemes is grossly defective as a way of making community decisions. What the theorem

says about systems with no individual indifference is, in effect, that non-manipulability can be had only in systems which are otherwise grossly defective. Now take a system which is non-manipulable and allows individual indifference. For all cases in which no one is indifferent between alternatives, the system is a fixed mixture of unilateral and dupe schemes. That in itself is a gross defect: for a significant class of combinations of individual preferences, the system stands ready to make the community choice in an unacceptable way. The theorem in this paper shows, then, that even in systems which permit a voter to express indifference, non-manipulability may be had only at an exorbitant price.

#### IV. Definitions and Proofs

Let  $V$  be a finite set, called the set of alternatives. Variables  $w, x, y,$  and  $z$  will have  $V$  as their range of values. A strict ordering of  $V$  is a binary relation  $P$  which, for all  $x, y,$  and  $z,$  satisfies:

Asymmetry:  $x P y \rightarrow \sim y P x.$

Negative transitivity:  $x P y \rightarrow (x P z \vee z P y).$

Such a relation is transitive, and may allow indifference between alternatives.

A ranking of  $V$  is a strict ordering of  $V$  which, for all  $x$  and  $y,$  satisfies:

Connectedness:  $x \neq y \rightarrow (x P y \vee y P x).$

A ranking  $n$ -tuple over  $V$  is an  $n$ -tuple  $\langle P_1, \dots, P_n \rangle$  of rankings of  $V.$

Ranking  $n$ -tuples will be represented by bold type (in typescript by squiggly underlining), on the pattern:

$$\underline{\underline{P}} = \langle P_1, \dots, P_n \rangle,$$

$$\underline{\underline{P}}^* = \langle P_1^*, \dots, P_n^* \rangle,$$

and the like.  $\underline{\underline{P}}$  and  $\underline{\underline{P}}'$  agree off  $k$  iff for all  $i \neq k, P_i' = P_i.$

$\underline{\underline{P}} /_k P$  is the preference  $n$ -tuple  $\underline{\underline{P}}'$  such that  $P_k' = P$  and  $\underline{\underline{P}}'$  agrees with  $\underline{\underline{P}}$  off  $k.$

We now define "proto-scheme", "scheme", and "decision scheme".

A measure over  $V$  is a function  $\rho$  which assigns a non-negative real

number,  $\rho(x)$ , to each member  $x$  of  $V$ . The sum  $\sum_x \rho(x)$  of these numbers is called the weight of the measure. A lottery is a measure of weight one.

A proto-scheme is a function  $d$  such that, for some positive integer  $n$ , called the number of voters of  $d$ , and for some finite set  $V$ , whose members are called alternatives of  $d$ , the domain of  $d$  is the set of all ranking  $n$ -tuples over  $V$ , and the values of  $d$  are measures over  $V$ . The value of  $d$  at  $\underline{p}$  will be written  $d_{\underline{p}}$ , and the probability  $d_{\underline{p}}$  assigns to an alternative  $x$  will be written  $d(x, \underline{p})$ . A scheme is a proto-scheme all of whose values have the same weight; this will be called the weight of the scheme. A decision scheme is a scheme of weight one. It thus assigns to each ranking  $n$ -tuple  $\underline{p}$  a lottery over  $V$ .

We consider, then, a fixed set  $V$  of alternatives and number  $n$  of voters. The variables will range as follows:

$w, x, y, z$ : alternatives in  $V$ .

$X, Y, Z$ : sets of alternatives, i.e. subsets of  $V$ .

$P, Q$ : rankings of  $V$ .

$\underline{P}, \underline{Q}$ : ranking  $n$ -tuples  $\langle P_1, \dots, P_n \rangle$  over  $V$ .

$b, c, d$ : Partial decision schemes for  $n$  voters and set  $V$  of alternatives.

Subscripts, superscripts, primes, and the like do not affect the range of variables.

A utility scale  $U$  over  $V$  is an assignment of real numbers to the members of  $V$ . Where  $U$  is a utility scale over  $V$  and  $\rho$  is a lottery over  $V$ , we define the expected utility  $U(\rho)$  of  $\rho$  on scale  $U$  in this

way:

$$U(\rho) = \sum_x U(x)\rho(x).$$

Utility scale  $U$  fits a strict ordering  $P$  iff for all  $x$  and  $y$ ,

$$U(x) > U(y) \leftrightarrow x P y.$$

A decision scheme  $d$  is potentially manipulable by  $k$  at  $\tilde{P}$  iff there are a utility scale  $U$  which fits  $P_k$  and a ranking  $P'_k$  of  $V$  such that where  $\tilde{P}' = \tilde{P} \setminus_k P'_k$ ,

$$U(d\tilde{P}') > U(d\tilde{P}).$$

$d$  is manipulable iff there are a voter  $k$  and a ranking  $n$ -tuple  $\tilde{P}$  such that  $d$  is potentially manipulable by  $k$  at  $\tilde{P}$ . Otherwise,  $d$  is strategy-proof.

We now give a number of definitions which will allow the theorem on strategy-proof decision schemes to be stated in a preliminary, weak version.

Definition 1. Scheme  $d$  is unilateral iff there is a  $k$  such that for all  $\tilde{P}$  and  $\tilde{P}'$ , if  $\tilde{P}'_k = \tilde{P}_k$ , then  $d\tilde{P}' = d\tilde{P}$ .

Definition 2. Scheme  $d$  is duple iff there are alternatives  $x$  and  $y$  such that for every other alternative  $z$ ,  $d(z, \tilde{P}) = 0$  for all  $\tilde{P}$ .

Definition 3. Scheme  $d$  is a probability mixture of schemes  $d_1, \dots, d_m$  iff there is a sequence  $\alpha_1, \dots, \alpha_m$ , with  $0 < \alpha_z \leq 1$  for each  $z \in \{1, \dots, m\}$  and  $\sum_{z=1}^m \alpha_z = 1$ , such that for each  $\tilde{P}$  and  $x$ ,

$$d(x, \tilde{P}) = \alpha_1 d_1(x, \tilde{P}) + \dots + \alpha_m d_m(x, \tilde{P}).$$

Where  $d$  is such a probability mixture, we shall write

$$d = \alpha_1 d_1 + \dots + \alpha_m d_m,$$

and where  $d = b + c$ , we shall write  $b = d - c$ .

Theorem (weak version). If  $d$  is a strategy-proof decision scheme, then  $d$  is a probability mixture of decision schemes each of which is either unilateral or duple.

This theorem can be strengthened to give conditions which are sufficient as well as necessary for a decision scheme's being strategy-proof. For any set  $X$  of alternatives and scheme  $d$ , we shall write  $d(X, P)$  for  $\sum_{x \in X} d(x, P)$ , the total probability assigned by measure  $dP$  to members of  $X$ .  $X$  heads ranking  $P_k$  iff for any  $x \in X$  and  $y \notin X$ ,  $x P_k y$ .

Definition 4. Proto scheme  $d$  is localized iff for every  $k, P, P'_k$ , and  $X$  such that  $X$  heads both  $P_k$  and  $P'_k$ ,  $d(X, P/P'_k) = d(X, P)$ .

A switch is a reversal of two adjacent alternatives in a ranking. A scheme is non-perverse if switching an alternative upward never decreases its probability.

Definition 5.  $x P_k y$  means that  $x P_k y$  &  $\sim(\exists z) (x P_k z \& z P_k y)$ . Where  $x P_k y$ ,  $P_k^y$  is the ranking which switches  $xy$  in  $P_k$  and permutes no other alternative,  $P_k^{ky} = P_k / P_k^y$ , and  $\epsilon_k^y(d, P)$ , the effect under  $d$  of  $k$ 's switching  $y$  upward, is  $d(y, P_k^{ky}) - d(y, P)$ . Scheme  $d$  is non-perverse iff for every  $P, k$ , and  $y$  such that  $\{y\}$  does not head  $P_k$ ,  $\epsilon_k^y(d, P) \geq 0$ .

Theorem. A decision scheme  $d$  is strategy-proof if and only if it is a probability mixture of decision schemes, each of which is localized, non-perverse, and either unilateral or duple.

The proof of the Theorem consists chiefly of five lemmas.

Definition 6. A proto-scheme  $d$  is pairwise responsive iff for every  $\tilde{P}$ ,  $k$ ,  $x$ ,  $y$ , and  $z$ , if  $x P_k^! y$  and  $z \notin \{x, y\}$ , then  $d(z, \tilde{P}^{ky}) = d(z, \tilde{P})$ .

Lemma 1. The following are equivalent.

- (a)  $d$  is a localized proto-scheme.
- (b)  $d$  is a pairwise responsive scheme.
- (c)  $d$  is a pairwise responsive proto-scheme, and for all  $x$ ,  $y$ ,  $\tilde{P}$ , and  $k$  such that  $x P_k^! y$ ,  
 $d(\{x, y\}, \tilde{P}^{ky}) = d(\{x, y\}, \tilde{P})$ .

Proof. Suppose  $d$  is a localized proto-scheme. Then since  $V$  heads any  $\tilde{P}$  and  $\tilde{P}^*$ ,  $d(V, \tilde{P}) = d(V, \tilde{P}^*)$ , and  $d$  is a scheme. Now suppose that  $x P_k^! y$ ,  $z \notin \{x, y\}$ , and  $W$  is the set of alternatives ranked above  $z$  in  $P_k$ . Then both  $W$  and  $W \cup \{z\}$  head both  $P_k$  and  $P_k^y$ . Thus since  $d$  is localized,  $k$ 's switching  $y$  upward changes neither the total probability of  $W$  nor the total probability of  $W \cup \{z\}$ . Thus it leaves the probability of  $z$  unchanged, and  $d$  is pairwise responsive. Thus (a) entails (b). For any pairwise responsive scheme, a switch of  $xy$  changes neither the total probability of  $V - \{x, y\}$  nor that of  $V$ ; thus it leaves that of  $\{x, y\}$  unchanged, and (b) entails (c). Now suppose (c); it follows that if  $x P_k^! y$  and  $\{x, y\} \subseteq Z$ , then  $d(Z, \tilde{P}^{ky}) = d(Z, \tilde{P})$ . If  $Z$  heads both  $P_k$  and  $P_k^!$ , then  $P_k^!$  can be formed from  $P_k$  by switches between members of  $Z$  and switches between non-members of  $Z$ , neither of which, we have seen,



change the total probability of  $Z$ . Thus  $d$  is localized, and (c) entails (a).

Lemma 2. A decision scheme  $d$  is strategy-proof iff  $d$  is localized and non-perverse.

Proof: Suppose that  $d$  is not localized, so that for some  $k$ , some  $\tilde{P}$  and  $\tilde{P}'$  that agree off  $k$ , and some  $X$  which heads both  $P_k$  and  $P'_k$ ,  $d(X, \tilde{P}') - d(X, \tilde{P}) = \epsilon > 0$ . Let  $U$  fit  $P_k$  and be such that for all  $x \in X$ ,  $1 \leq U(x) < 1 + \epsilon$ , and for all  $y \notin X$ ,  $0 \leq U(y) < \epsilon$ . Then, it can be calculated,

$$U(d\tilde{P}) < d(X, \tilde{P}) + \epsilon \leq U(d\tilde{P}'),$$

and so  $d$  is potentially manipulable by  $k$  at  $\tilde{P}$ .

If  $d$  is localized but perverse, then for some  $x, y$ , and  $k$ ,  $x \succ P_k ! y$  and  $k$ 's switching  $y$  upwards lowers the probability of  $y$  by some amount  $\epsilon > 0$ . By (c) of Lemma 1, the switch raises the probability of  $x$  by  $\epsilon$ , and changes no other probabilities. Hence if  $U$  fits  $P_k$ , so that  $U(x) > U(y)$ , then  $U(dP^{ky}) - U(dP) = \epsilon U(x) - \epsilon U(y) > 0$ , and so  $d$  is potentially manipulable by  $k$  at  $\tilde{P}$ .

Now suppose  $d$  is localized and non-perverse, and consider any  $k, \tilde{P}, P'_k$ , and  $U$  which fits  $P_k$ . Where  $\tilde{P}' = \tilde{P} /_k P'_k$ , we shall show that  $U(d\tilde{P}') \leq U(d\tilde{P})$ . Form  $P'_k$  from  $P_k$  by successive switches as follows: take the top alternative in  $P'_k$  and switch it from its position in  $P_k$  successively to the top, then take the second alternative in  $P'_k$  and switch it successively up from its position in  $P_k$  to its position in  $P'_k$ , and so forth. At each step, an alternative  $y$  is switched with an alternative which is above it in  $P_k$ . Since  $U$  fits  $P_k$ ,  $U(x) > U(y)$ , and so by (c) of Lemma 1 and the non-perversity of  $d$ , utility on scale  $U$  cannot be increased

by such steps. Hence  $U(d\tilde{P}') \leq U(d\tilde{P})$ . That proves the Lemma.

Definition 7.  $P_i \uparrow \{x,y\}$  is  $P_i$  restricted to  $\{x,y\}$ , and  $\tilde{P} \uparrow \{x,y\} = \langle P_1 \uparrow \{x,y\}, \dots, P_n \uparrow \{x,y\} \rangle$ . A scheme  $d$  is pairwise isolated iff for any  $k, \tilde{P}, \tilde{P}^*, x$ , and  $y$ , if  $x P_k ! y$ ,  $P_k^* = P_k$ , and  $\tilde{P}^* \uparrow \{x,y\} = \tilde{P} \uparrow \{x,y\}$ , then  $e_k^y(d, \tilde{P}^*) = e_k^y(d, \tilde{P})$ .  $d$  is decomposable iff for any fixed  $k, x$ , and  $y$  with  $x \neq y$ , there are functions  $\gamma$  and  $\delta$  such that for all  $\tilde{P}$  with  $x P_k ! y$ ,  $e_k^y(d, \tilde{P}) = \gamma(\tilde{P} \uparrow \{x,y\}) + \delta(P_k)$ .

Definition 8.  $P_{kxy}$  is  $P_k$  with  $x$  and  $y$  moved to the bottom, their ordering with respect to each other preserved, and the ordering of all other alternatives with respect to each other preserved.  $\tilde{P}_{xy}$  is  $\langle P_{1xy}, \dots, P_{nxy} \rangle$ ,  $\tilde{P}_{kx}$  is  $P_{kxx}$ , and  $\tilde{P}_{\sim x}$  is  $\tilde{P}_{\sim xx}$ .

Lemma 3. Let scheme  $d$  be localized. Then  $d$  is pairwise isolated and decomposable.

Proof that  $d$  is pairwise isolated: We first show that the switch of a pair by one person does not alter the effect of the switch of another pair by another person. Suppose  $j \neq k$ ,  $w P_j ! z$ ,  $x P_k ! y$ , and  $\{w,z\} \neq \{x,y\}$ .

Case 1:  $y \notin \{w,z\}$ .  $d$  is pairwise responsive and  $\tilde{P}^{jz}$  differs from  $\tilde{P}$  only in  $j$ 's switching  $wz$ ; thus  $d(y, \tilde{P}^{jz}) = d(y, \tilde{P})$ . Likewise,  $\tilde{P}^{jzky}$  differs from  $\tilde{P}^{ky}$  only in  $j$ 's switching  $wz$ ; thus  $d(y, \tilde{P}^{jzky}) = d(y, \tilde{P}^{ky})$

Hence

$$d(y, \tilde{P}^{jzky}) - d(y, \tilde{P}^{jz}) = d(y, \tilde{P}^{ky}) - d(y, \tilde{P}),$$

which is to say  $e_k^y(d, \tilde{P}^{jz}) = e_k^y(d, \tilde{P})$ .

Case 2:  $x \notin \{w, z\}$ . By an argument like that in Case 1,

$e_k^x(d, P_{\sim}^{jzky}) = e_k^x(d, P_{\sim}^{ky})$ . It follows from this and (c) of Lemma 1 that

$e_k^y(d, P_{\sim}^{jz}) = e_k^y(d, P_{\sim})$ . Now let  $x P_k! y$ ,  $P_k^* = P_k$ , and  $P_{\sim}^* \uparrow \{x, y\} = P_{\sim} \uparrow \{x, y\}$ .

Then  $P_{\sim}^*$  can be formed from  $P_{\sim}$  by a sequence of switches by voters other than  $k$ , none of which switches  $x$  with  $y$ . We have just seen that none of these changes the value of  $e_k^y$ , and thus  $e_k^y(d, P_{\sim}^*) = e_k^y(d, P_{\sim})$ . Thus  $d$  is pairwise isolated.

Proof that  $d$  is decomposable:

Take  $k, x$ , and  $y$  with  $x \neq y$ . For any  $P_{\sim}$  with  $x P_k! y$ , define

$\gamma(P_{\sim} \uparrow \{x, y\}) = e_k^y(d, P_{\sim, kxy})$ . Now let  $P_{\sim}$  and  $P_{\sim}^*$  be such that  $x P_k! y$  and  $P_k^* = P_k$ ; we shall show

$$e_k^y(d, P_{\sim}^*) - \gamma(P_{\sim}^* \uparrow \{x, y\}) = e_k^y(d, P_{\sim}) - \gamma(P_{\sim} \uparrow \{x, y\}). \quad (1)$$

Since  $d$  is pairwise isolated,  $e_k^y(d, P_{\sim})$  depends only on  $P_k$  and  $P_{\sim} \uparrow \{x, y\}$ ; thus we may suppose w.l.g. that everyone other than  $k$  ranks  $x$  and  $y$  last. Now form  $P_k^y$  from  $P_k$  by the following sequence of switches.

- (1\*) Progressively switch  $y$  to bottom.
- (2\*) Progressively switch  $x$  down to just above  $y$ .
- (3\*) Switch  $y$  with  $x$ .
- (4\*) Progressively switch  $y$  up to its original position.
- (5\*) Progressively switch  $x$  up to just below  $y$ . Call this sequence

$P_k^0, \dots, P_k^\mu$ , and consider the difference

$$d(y, P_{\sim} \downarrow_k P_k^z) - d(y, P_{\sim}^* \downarrow_k P_k^z) \quad (2)$$

as  $z$  goes from 0 to  $\mu$ . This difference changes only in step (3\*).

For the steps in (1\*) and (4\*) consist of switching  $y$  with various

alternatives  $z \notin \{x, y\}$ . Everyone except  $k$  ranks  $z$  above  $y$  in both  $\tilde{P}$  and  $\tilde{P}^*$ , and so since  $d$  is pairwise isolated, both terms of (2) change by the same amount, and (2) is unchanged. The steps in (2<sup>\*</sup>) and (5<sup>\*</sup>) consist of switching  $x$  with alternatives other than  $y$ ; since  $d$  is pairwise responsive, this changes neither term of (2). Now at step (3<sup>\*</sup>),  $x$  and  $y$  are switched in  $P_{kxy}$ . The change in (2) at step (3<sup>\*</sup>), then, is

$$e_k^y(d, \tilde{P}_{xy}) - e_k^y(d, \tilde{P}_{xy}^*). \quad (3)$$

This, then, is change in (2) from  $z = 0$  to  $z = \mu$ , that is,

$$[d, (y, \tilde{P}^{ky}) - d(y, \tilde{P}^{*ky})] - [d(y, \tilde{P}) - d(y, \tilde{P}^*)],$$

which is  $e_k^y(d, \tilde{P}) - e_k^y(d, \tilde{P}^*)$ .

From the equality of this with (3), (1) follows. Since the quantity in (1) depends only on  $P_k$ , let  $\delta(P_k)$  be this quantity; then  $e_k^y(d, \tilde{P}) = v(\tilde{P} \uparrow \{x, y\}) + \delta(P_k)$ , and  $d$  is decomposable.

Definition 9.  $k$ 's unilateral component of decision scheme  $d$  is the function  $d_k$  such that for all  $x$  and  $\tilde{P}$ ,

$$d_k(x, \tilde{P}) = \min_{\tilde{Q}} \{d(x, \tilde{Q} /_k P_k) - d(x, \tilde{Q} /_k P_{kx})\}.$$

Since the value of  $d_k(x, \tilde{P})$  depends only on  $x$  and  $P_k$ , this will be written  $d_k(x, P_k)$ .

Lemma 4. Let  $d$  be a strategy-proof decision scheme, and let  $d_k$  be  $k$ 's unilateral component of  $d$ . Then (a) if  $x P_k ! y$ , then

$$e_k^y(d_k, \tilde{P}) = \min_{\tilde{Q}} e_k^y(d, \tilde{Q} /_k P_k).$$

(b)  $d_k$  is a scheme which is unilateral, localized, and non-perverse.

Proof of (a): Let  $x P_k! y$ , let  $\alpha = d_k(y, P_k)$ , and let  $\beta = d_k(y, P_k^y)$ . Then  $\epsilon_k^y(d_k, \tilde{P}) = \beta - \alpha$ . Let  $\epsilon = \min_{\tilde{Q}} \epsilon_k^y(d, \tilde{Q} /_k P_k)$ ; we are to prove that  $\epsilon = \beta - \alpha$ .

By the definition of  $\beta$ , for some  $\tilde{Q}$ ,

$$\beta = d(y, \tilde{Q} /_k P_k^y) - d(y, \tilde{Q} /_k P_{ky}),$$

and by the definitions of  $\alpha$  and  $\epsilon$ ,

$$\alpha \leq d(y, \tilde{Q} /_k P_k) - d(y, \tilde{Q} /_k P_{kx});$$

$$\epsilon \leq d(y, \tilde{Q} /_k P_k^y) - d(y, \tilde{Q} /_k P_k).$$

Therefore by addition,  $\alpha + \epsilon \leq \beta$ .

We will have  $\alpha + \epsilon \geq \beta$  if there is a  $\tilde{Q}$  such that

$$d(y, \tilde{Q} /_k P_k^y) - d(y, \tilde{Q} /_k P_{ky}) = \alpha + \epsilon, \quad (4)$$

since  $\beta$  is the minimal value of this difference. Construct such a  $\tilde{Q}$  as follows. By the definition of  $\alpha$ , for some  $\tilde{P}^*$  with  $P_k^* = P_k$ ,  $\alpha = d(y, \tilde{P}^*) - d(y, \tilde{P}^* /_k P_{ky})$ . This difference is the sum of the effects, in context  $\tilde{P}^*$ , of  $k$ 's successively switching  $y$  from bottom to just below  $x$  in  $P_k$ . Since  $d$  is pairwise isolated and  $y$  is not switched with  $x$ , each of these effects is independent of where others besides  $k$  place  $x$  in their rankings. On the other hand, since  $d$  is pairwise isolated, the effect under  $d$  of  $k$ 's switching  $y$  with  $x$  is independent of the way others vote except for their ranking of  $x$  with respect to  $y$ . By the definition of  $\epsilon$ , for some  $\tilde{P}'$ ,  $P_k' = P_k$  and  $\epsilon = \epsilon_k^y(d, \tilde{P}')$ . Form  $\tilde{Q}$  from  $\tilde{P}^*$  by moving  $x$ , in the ranking of each voter  $i \neq k$ , to just above  $y$  in  $P_i^*$  or just below  $y$  in  $P_i^*$  according as  $x P_i' y$  or  $y P_i' x$ . Then, we have

seen,

$$\alpha = d(y, Q) - d(y, Q \underset{k}{/} P_{ky}),$$

$$\epsilon = e_k^y(d, Q) = d(y, Q \underset{k}{/} P_k^y) - d(y, Q),$$

and hence by addition, (4) holds.

Proof of (b): Since  $d$  is non-perverse,  $d(x, Q \underset{k}{/} P_k) - d(x, Q \underset{k}{/} P_{kx}) \geq 0$  for all  $Q$ . Hence its minimal value  $d_k(x, P_k)$  is non-negative, and  $d_k$  is a proto-scheme. Now let  $x P_k! y$ . By (a),

$$d_k(x, P_k) - d_k(x, P_k^y) = \min_Q e_k^x(d, Q \underset{k}{/} P_k^y);$$

$$d_k(y, P_k^y) - d_k(y, P_k) = \min_Q e_k^y(d, Q \underset{k}{/} P_k).$$

By (c) of Lemma 1, these two minima must be equal, and hence

$$d_k(\{x, y\}, P_k^y) = d_k(\{x, y\}, P_k).$$

Now let  $z \notin \{x, y\}$ , and take any  $Q$ . Since  $d$  is pairwise responsive,  $d(z, Q \underset{k}{/} P_k^y) = d(z, Q \underset{k}{/} P_k)$ . By (c) of Lemma 1,  $d(z, Q \underset{k}{/} P_{kz}^y) = d(z, Q \underset{k}{/} P_{kz})$ . Therefore

$$d(z, Q \underset{k}{/} P_k^y) - d(z, Q \underset{k}{/} P_{kz}^y) = d(z, Q \underset{k}{/} P_k) - d(z, Q \underset{k}{/} P_{kz}),$$

and so the minima are equal:  $d_k(z, P_k^y) = d_k(z, P_k)$ .

Thus (c) in Lemma 1 is satisfied for  $d_k$ , and so  $d_k$  is a localized scheme.

Finally, by (a), if  $x P_k! y$ , then for some  $Q$ ,  $e_k^y(d_k, P_k) = e_k^y(d, Q)$ , and since  $d$  is non-perverse, this is non-negative; therefore  $d_k$  is non-perverse.

$d_k$  is clearly unilateral, and the Lemma is proved.

Definition 10. A scheme  $d$  ignores external comparisons iff for any  $x, y, k, P$ , and  $P^*$ , if  $x P_k! y$  and  $P^* \uparrow \{x, y\} = P \uparrow \{x, y\}$ , then  $e_k^y(d, P^*) = e_k^y(d, P)$ .

Lemma 5. Suppose  $d$  is a strategy-proof decision scheme. For each  $k$ ,  
let  $d_k$  be  $k$ 's unilateral component of  $d$ , and define functions  $d_o, c,$   
and  $c_{yz}$  for each  $\{y, z\}$  such that  $y \neq z$  as follows.

$$d_o(x, P) = d(x, P_x);$$

$$c(x, P) = d(x, P) - d_o(x, P) - \dots - d_n(x, P);$$

$$c_{yz}(x, P) = c(x, P_{yz}) \text{ for } x \in \{y, z\} \text{ and } 0 \text{ for } x \notin \{y, z\}.$$

Then  $c$  ignores external comparisons. Each function  $d_o, \dots, d_n, c,$  and each  
 $c_{xy}$  is a localized, non-perverse scheme,  $d_o, \dots, d_n$  are unilateral, each  
 $c_{xy}$  is duple, and for all  $x$  and  $P,$

$$d(x, P) = d_o(x, P) + \sum_k d_k(x, P) + \sum_{yz} c_{yz}(x, P), \quad (5)$$

where  $\sum_{yz}$  sums over all pairs  $\{y, z\}$  with  $y \neq z.$

Proof.  $d_o$  is constant for each  $x$ , and thus, like each  $d_k$ , a unilateral,  
 localized, non-perverse scheme. Now consider  $c$ , and let  $x P_k! y$ .

$$e_k^y(d_i, P) = 0 \text{ for } i \neq k, \text{ and thus}$$

$$e_k^y(c, P) = e_k^y(d, P) - e_k^y(d_k, P). \text{ By Lemma 4(a),}$$

$$e_k^y(d_k, P) \leq e_k^y(d, P); \text{ hence } e_k^y(c, P) \geq 0$$

and  $c$  is non-perverse.  $c(x, P_x) = 0$  for all  $x$  and  $P$ , since  $d_o(x, P_x) = d(x, P_x)$  and  $d_i(x, P_x) = 0$  for  $i \neq 0$ . Thus since  $c$  is non-perverse,  $c(x, P) \geq 0$  for all  $x$  and  $P$ , and  $c$  is a proto-scheme. Since  $d, d_o, \dots, d_n$  are all localized,  $c$  is localized and hence a scheme.

$c$  ignores external comparisons. For let  $x P_k! y$ . Since  $d$  is decomposable, there are functions  $\gamma$  and  $\delta$  such that for any  $Q$  with  $x Q_k! y$ ,

$e_k^y(d, Q) = \gamma(Q \uparrow \{x, y\}) + \delta(Q_k)$ . Thus  $e_k^y(d, P) = \gamma(P \uparrow \{x, y\}) + \delta(P_k)$ .

Let  $Q$  minimize  $e_k^y(d, Q)$  for  $Q_k = P_k$ ; then by Lemma 4(a),

$e_k^y(d_k, P) = e_k^y(d, Q) = \gamma(Q \uparrow \{x, y\}) + \delta(P_k)$ . Therefore  $e_k^y(c, P) =$

$\gamma(P \uparrow \{x, y\}) - \gamma(Q \uparrow \{x, y\})$ .

Now let  $P^* \uparrow \{x, y\} = P \uparrow \{x, y\}$ ; by a like argument, where  $Q^*$  minimizes

$e_k^y(d, Q^*)$  for  $Q_k^* = P_k^*$ ,  $e_k^y(c, P^*) = \gamma(P^* \uparrow \{x, y\}) - \gamma(Q^* \uparrow \{x, y\})$ . Since  $Q$

minimizes  $e_k^y(d, Q) = \gamma(Q \uparrow \{x, y\}) + \delta(P_k)$  for  $Q_k = P_k$ ,  $Q$  minimizes

$\gamma(Q \uparrow \{x, y\})$  for  $x Q_k y$ , and since  $x P_k^* y$ ,  $Q^* = Q \downarrow_k P_k^*$  minimizes

$\gamma(Q^* \uparrow \{x, y\})$  for  $Q_k^* = P_k^*$ . For this  $Q^*$ ,  $Q^* \uparrow \{x, y\} = Q \uparrow \{x, y\}$ , and so

$e_k^y(c, P^*) = \gamma(P^* \uparrow \{x, y\}) - \gamma(Q^* \uparrow \{x, y\}) = \gamma(P \uparrow \{x, y\}) - \gamma(Q \uparrow \{x, y\}) = e_k^y(c, P)$ .

Thus  $c$  ignores external comparisons.

Since  $c(z, P) \geq 0$  for all  $z$  and  $P$ , from the way  $c_{xy}$  is defined,

$c_{xy}(z, P) \geq 0$ . Moreover,  $c_{xy}(V, P) = c(\{x, y\}, P_{xy})$ , and since  $c$  is localized,

this is constant for all  $P$ , and  $c_{xy}$  is a scheme.  $c_{xy}$  is pairwise

responsive and hence localized: an  $xy$  switch leaves all other probabilities

zero, and any other switch leaves all probabilities unchanged.  $c_{xy}$  is non-

perverse: if  $w P_k! z$  and  $\{w, z\} \neq \{x, y\}$ ,  $e_k^z(c_{xy}, P) = 0$ , and if  $x P_k! y$ ,

$e_k^y(c_{xy}, P) = e_k^y(c, P_{xy}) \geq 0$ , since  $c$  is non-perverse.

Finally,  $c(x, P) = \sum_{yz} c_{yz}(x, P)$  for  $y \neq z$ . We noted earlier that

$c(x, P_x) = 0$  for all  $x$  and  $P$ ; thus  $c_{yz}(x, P_x) = 0$ . For if  $x \notin \{y, z\}$ ,

then  $c_{yz}(x, P) = 0$  for all  $P$ , and  $c_{yz}(y, P_y) = c(y, (P_y)_{yz}) = 0$ , since

$(P_y)_{yz}$  has  $y$  uniformly on the bottom. Therefore

$c(x, P_x) = 0 = \sum_{yz} c_{yz}(x, P_x)$ . Now form  $P$  from  $P_x$  by successively shifting  $x$  upward in each ranking; call the resulting sequence  $P^0, \dots, P^u$ .

At each step from  $P^z$  to  $P^{z+1}$ , some  $k$  switches  $x$  with a  $w$  such that

$w P_k! x$ .



Now  $c_{\tilde{w}x}(x, \tilde{p}^{z+1}) - c_{\tilde{w}x}(x, \tilde{p}^z) = c(x, \tilde{p}_{\tilde{w}x}^{z+1}) - c(x, \tilde{p}_{\tilde{w}x}^z) = \epsilon_k^x(c, \tilde{p}_{\tilde{w}x}^z) = \epsilon_k^x(c, \tilde{p}^z)$

since  $c$  ignores external comparisons; thus

$c(x, \tilde{p}^{z+1}) - c(x, \tilde{p}^z) = c_{\tilde{w}x}(x, \tilde{p}^{z+1}) - c_{\tilde{w}x}(x, \tilde{p}^z)$ . For  $\{y, z\} \neq \{x, w\}$ ,  $\tilde{p}_{kyz}^{z+1} = \tilde{p}_{kyz}^z$ , and so  $c_{yz}(x, \tilde{p}^{z+1}) = c_{yz}(x, \tilde{p}^z)$ . Therefore  $c(x, \tilde{p}^{z+1}) - c(x, \tilde{p}^z) = \sum_{yz} c_{yz}(x, \tilde{p}^{z+1}) - \sum_{yz} c_{yz}(x, \tilde{p}^z)$ . Since  $c(x, \tilde{p}^0) = \sum_{yz} c_{yz}(x, \tilde{p}^0)$ , by induction  $c(x, \tilde{p}^z) = \sum_{yz} c_{yz}(x, \tilde{p}^z)$  for all  $\tilde{p}^z$ , and thus  $c(x, \tilde{p}) = \sum_{yz} c_{yz}(x, \tilde{p})$ . (5) follows immediately.

Proof of Theorem. In (5), drop schemes of weight zero, so that  $d$  is a sum  $b_1^* + \dots + b_m^*$  of schemes of positive weight  $\alpha_1, \dots, \alpha_m$  respectively. Let each decision scheme  $b_z = (1/\alpha_z) b_z^*$ ; then  $d = \alpha_1 b_1 + \dots + \alpha_m b_m$ , and each  $b_z$  is localized, non-perverse, and either unilateral or duple. Conversely, if decision scheme  $d$  is a probability mixture of this kind, then  $d$  is clearly localized and non-perverse; hence by Lemma 2,  $d$  is strategy-proof. That completes the proof of the theorem.

V. Corollaries

Definition 11.  $k$  is dictator for decision scheme  $d$  iff for every  $\tilde{P}$ ,  $x$ , and  $y$ , if  $x P_k y$  then  $d(y, \tilde{P}) = 0$ .  $d$  is dictatorial iff there is a dictator for  $d$ .

Definition 12. Lottery  $\rho$  is Pareto optimal ex post for ranking  $n$ -tuple  $\tilde{P}$  iff for any  $x$ , if there is a  $y$  such that  $y P_i x$  for all  $i$ , then  $\rho(x) = 0$ . Decision scheme  $d$  is Pareto optimific ex post iff for every  $\tilde{P}$ , lottery  $d\tilde{P}$  is Pareto optimal ex post for  $\tilde{P}$ .

Corollary 1 (Sonnenschein). Let decision scheme  $d$  be strategy-proof and Pareto optimific ex post. Let the set  $V$  of alternatives for  $d$  have at least three members. Then  $d$  is a probability mixture of dictatorial decision schemes.

Proof: Since  $d$  is strategy-proof,  $d$  is a probability mixture of decision schemes, each of which is unilateral or duple. Let

$$d = \alpha_1 d_1 + \dots + \alpha_m d_m,$$

where for each  $z \in \{1, \dots, m\}$ ,  $\alpha_z > 0$  and  $d_z$  is non-null and either unilateral or duple. Then no  $d_z$  is duple. For since  $d$  is Pareto optimific ex post, the alternatives in any pair  $\{x, y\}$  get a probability of zero whenever some alternative  $z$  is unanimously preferred to them. Thus for any pair  $\{x, y\}$ ,  $d_z(\{x, y\}, \tilde{P}) = 0$  whenever  $z P_i x$  and  $z P_i y$  for each  $i$ ; therefore

$d_z$  is not an  $xy$  duplicate scheme. Hence each  $d_z$  is unilateral.

Now let  $d_z$  be unilateral on the part of  $k$  and let  $x P_k y$ . Then, we shall show,  $d_z(y, P) = 0$ . For let  $\tilde{P}$  be such that  $P'_k = P_k$  and for all  $i$ ,  $x P_i y$ . Then since  $d$  is Pareto optimific ex post,  $d(y, P') = 0$ , and thus  $d_z(y, P') = 0$ . Since  $d_z$  is unilateral on the part of  $k$  and  $P'_k = P_k$ ,  $d_z(y, P) = 0$ . We have shown that for any  $\tilde{P}$ ,  $x$ , and  $y$ , if  $x P_k y$  then  $d_z(y, P) = 0$ . Thus  $k$  is dictator for  $d_z$ . We have shown that each  $d_z$  is dictatorial, and the Corollary is proved.

Definition 13. Lottery  $\rho$  is Pareto optimal ex ante for utility scales  $U_1, \dots, U_n$  iff there is no lottery  $\rho'$  such that for each  $i$ ,  $U_i(\rho') > U_i(\rho)$ . Decision scheme  $d$  is Pareto optimific ex ante iff for every ranking  $n$ -tuple  $\tilde{P}$  and every  $n$ -tuple  $\langle U_1, \dots, U_n \rangle$  of utility scales such that for each  $i$ ,  $P_i$  fits  $U_i$ , lottery  $d_{\tilde{P}}$  is Pareto optimal ex ante for  $U_1, \dots, U_n$ .

Corollary 2. Let decision scheme  $d$  be strategy-proof and Pareto optimific ex ante. Let the set  $V$  of alternatives for  $d$  have at least three members. Then  $d$  is dictatorial.

Proof: Note first that if  $d$  is Pareto optimific ex ante, then  $d$  is Pareto optimific ex post. For let preference rankings  $P_1, \dots, P_n$  fit  $U_1, \dots, U_n$  respectively, and suppose lottery  $\rho$  is not Pareto optimal ex post for  $P_1, \dots, P_n$ . Then for some pair of alternatives  $x$  and  $y$ ,  $y P_i x$  for all  $i$ , but  $\rho(x) \neq 0$ . Now let  $\rho'$  give  $x$ 's probability to  $y$ , so that

$\rho'(x) = 0$ ,  $\rho'(y) = \rho(x) + \rho(y)$ , and  $\rho'(z) = \rho(z)$  for all  $z \notin \{x, y\}$ .

Then for each  $i$ , since  $P_i$  fits  $U_i$ ,  $U_i(y) > U_i(x)$ , and so  $U_i(\rho') > U_i(\rho)$ . Therefore  $\rho$  is not Pareto optimal ex ante for  $U_1, \dots, U_n$ , and  $d$  is not Pareto optimific ex ante.

Suppose now that  $d$  is Pareto optimific ex ante, and therefore Pareto optimific ex post. Then  $d$  is a probability mixture of dictatorial decision schemes. Let

$$d = \alpha_1 d_1 + \dots + \alpha_n d_n,$$

where for each  $i$ ,  $\alpha_i \geq 0$  and  $i$  is dictator for  $d_i$ .

Suppose that  $d$  is not itself dictatorial, so that  $\alpha_i > 0$  for more than one  $i$ . Let  $\alpha_k > 0$ , let  $x, y$ , and  $z$  be distinct alternatives, and let the utility scales  $U_1, \dots, U_n$  be as follows.

$$U_k(x) = 1, \quad 1 > U_k(y) > \alpha_k, \quad U_k(z) = 0,$$

and for all  $w \notin \{x, y, z\}$ ,  $U_k(w) < 0$ .

For all  $i \neq k$ ,

$$U_i(z) = 1, \quad 1 > U_i(y) > 1 - \alpha_k, \quad U_i(x) = 0,$$

and for all  $w \notin \{x, y, z\}$ ,  $U_i(w) < 0$ .

For each  $i$ , let  $P_i$  fit  $U_i$ , so that  $\{x\}$  heads  $P_k$  and  $\{z\}$  heads  $P_i$  for all  $i \neq k$ . Then  $d(x, \tilde{P}) = \alpha_k$ ,  $d(z, \tilde{P}) = 1 - \alpha_k$ , and  $d(w, \tilde{P}) = 0$  for all  $w$  distinct from  $x$  and  $y$ . Therefore  $U_k(d\tilde{P}) = \alpha_k$ , and for  $i \neq k$ ,  $U_i(d, \tilde{P}) = 1 - \alpha_k$ .

Now let  $\hat{y}$  be the lottery that gives  $y$  as a sure thing. Then  $U_k(\hat{y}) > \alpha_k$  and for  $i \neq k$ ,  $U_i(\hat{y}) > 1 - \alpha_k$ . Therefore for all  $i$ ,

$U_i(\hat{y}) > U_i(d\tilde{p})$ , and so  $d\tilde{p}$  is not Pareto optimal ex ante for  $U_1, \dots, U_n$ . On the supposition that  $d$  is not dictatorial, we have shown that  $d$  is not Pareto optimal ex ante. Therefore  $d$  is dictatorial, and the Corollary is proved. 3/

Corollary 2 can be extended to schemes that allow individual indifference. Let a preference n-tuple over  $V$  be an  $n$ -tuple of strict orderings of  $V$ . Let an unrestricted decision scheme (UDS) be a function which, for some finite set  $V$  of alternatives and number  $n$ , takes as arguments all preference  $n$ -tuples over  $V$ , and takes as values lotteries over  $V$ . Manipulability is defined as before, with the term 'ranking' replaced by 'strict ordering'.

The following Assertion allows us both to extend Corollary 2 to UDS's, and to derive the old theorem on non-chance voting schemes. Where  $d$  is a decision scheme or UDS, a possible outcome for  $d$  is an alternative  $x$  such that for some  $\tilde{p}$  in the domain of  $d$ ,  $d(x, \tilde{p}) > 0$ . A weak dictator for  $d$  is a voter  $k$  such that for every  $\tilde{p}$ , where  $X$  is the set of possible outcomes which are first among possible outcomes in  $P_k$ ,  $d(X, \tilde{p}) = 1$ .  $d$  is weakly dictatorial iff there is a weak dictator for  $d$ .

Lemma 6. Let  $d$  be a strategy-proof UDS, and let  $d'$  be the decision scheme which is  $d$  with its domain restricted to ranking  $n$ -tuples. Then (i) any possible outcome of  $d$  is a possible outcome of  $d'$ , and (ii) a weak dictator for  $d'$  is weak dictator for  $d$ .

Proof of (i): Suppose  $x$  is a possible outcome of  $d$  but not of  $d'$ . Let  $\tilde{p}$  be a ranking  $n$ -tuple such that  $x$  ranks first in every  $P_i$ , and let  $\tilde{p}^*$  be a preference  $n$ -tuple such that  $d(x, \tilde{p}^*) > 0$ . Form a sequence of preference

n-tuples  $\underline{p}^0, \dots, \underline{p}^n$  as follows: let  $\underline{p}^0 = \underline{p}$ , and for each  $i$ , let  $\underline{p}^i = \underline{p}^{i-1} /_i P_i^*$ , so that  $\underline{p}^n = \underline{p}^*$ . Then since  $\underline{p}^0$  is a ranking n-tuple and  $x$  is not a possible outcome of  $d'$ , we have that  $d(x, \underline{p}^0) = 0$ , whereas  $d(x, \underline{p}^n) > 0$ . Take the least  $j$  such that  $d(x, \underline{p}^j) > 0$ , and let  $d(x, \underline{p}^j) = \epsilon$ . Then  $d(x, \underline{p}^{j-1}) = 0$ , and  $\underline{p}^{j-1} = P_j$ . Since  $P_j$  ranks  $x$  first, there is a utility scale  $U$  which fits  $P_j$ , such that  $U(x) = 1$  and for all  $y \neq x$ ,  $0 \leq U(y) < \epsilon$ . Since  $d(x, \underline{p}^j) = \epsilon$ ,  $U(x) = 1$ , and for all  $y \neq x$ ,  $U(y) \geq 0$ , we have that  $U(d\underline{p}^j) \geq \epsilon$ . Since  $d(x, \underline{p}^{j-1}) = 0$  and for all  $y \neq x$ ,  $U(y) < \epsilon$ , we have that  $U(d\underline{p}^{j-1}) < \epsilon$ . Thus  $U(d\underline{p}^j) > U(d\underline{p}^{j-1})$ , and since  $U$  fits  $P_j$  and  $\underline{p}^{j-1} = P_j$ ,  $d$  is potentially manipulable by  $j$  at  $\underline{p}^{j-1}$ .

Proof of (ii): From (i), the possible outcomes of  $d$  and  $d'$  are the same.

Now suppose  $k$  is weak dictator for  $d'$  but not for  $d$ . Then for some preference n-tuple  $\underline{p}$ , where  $X$  is the set of possible outcomes ranked first in  $P_k$ ,  $d(X, \underline{p}) \neq 1$ . For some  $x \in X$ , let  $P_k^*$  rank  $x$  first, and for every  $i \neq k$ , let  $P_i^*$  rank  $x$  last. Let  $\underline{p}^0 = \underline{p} /_k P_k^*$ . Then  $d(x, \underline{p}^0) \neq 1$ , for otherwise, on any utility scale  $U$  which fits  $P_k$ , we would have that  $U(d\underline{p}^0) = U(x)$  and  $U(d\underline{p}) < U(x)$ , so that  $U(d\underline{p}^0) > U(d\underline{p})$ . Thus  $d$  would be potentially manipulable by  $k$  at  $\underline{p}$ . Now form sequence  $\underline{p}^0, \dots, \underline{p}^n$  by letting  $\underline{p}^i = \underline{p}^{i-1} /_i P_i^*$  for each  $i$ , so that  $\underline{p}^n = \underline{p}^*$ . Then since  $\underline{p}^n$  is a ranking n-tuple with  $k$ , who is weak dictator for  $d'$ , ranking possible outcome  $x$  first, we have that  $d(x, \underline{p}^n) = 1$ . Take the least  $j$  such that  $d(x, \underline{p}^j) = 1$ ; then  $d(x, \underline{p}^{j-1}) < 1$ . We know that  $j \neq k$ , since from the way  $\underline{p}^j$  is defined,  $\underline{p}^k = \underline{p}^{k-1}$ . Therefore  $x$  is at the bottom of  $P_j^*$ , and so on any utility scale  $U$  which fits  $P_j^*$ ,  $U(d\underline{p}^{j-1}) > U(x)$ , whereas

$U(d_{\tilde{P}}^j) = U(x)$ . Thus  $U(d_{\tilde{P}}^{j-1}) > U(d_{\tilde{P}}^j)$ , and since  $P_j^j = P_j^*$  and  $U$  fits  $P_j^*$ ,  $d$  is potentially manipulable by  $j$  at  $\tilde{P}^j$ . That proves the Assertion.

Corollary 2'. Let UDS  $d$  be strategy-proof and Pareto optimific ex ante, and let  $d$  cover at least three alternatives. Then  $d$  is dictatorial.

Proof: Since  $d$  is Pareto optimific ex ante, all alternatives are possible outcomes, and so a weak dictator is dictator.  $d$  with its domain restricted to ranking  $n$ -tuples is dictatorial by Corollary 2, and thus by Assertion 17,  $d$  is dictatorial.

Definition 14. A voting scheme is a UDS  $v$  such that for every  $x$  and preference  $n$ -tuple  $\tilde{P}$ , either  $v(x, \tilde{P}) = 1$  or  $v(x, \tilde{P}) = 0$ .

Corollary 3. If a voting scheme is strategy-proof, then it is either duple or weakly dictatorial.

Proof: Let voting scheme  $v$  be strategy-proof, and let  $v'$  be  $v$  with its domain restricted to ranking  $n$ -tuples. Then by the Theorem,  $v'$  is a probability mixture  $\alpha_1 d_1 + \dots + \alpha_m d_m$ , where each  $\alpha_z$  is positive,  $\sum \alpha_z = 1$ , and each  $d_z$  is unilateral or duple. Now if  $v'(x, \tilde{P}) = 1$ , then for each  $z$ ,  $d_z(x, \tilde{P}) = 1$ , and if  $v'(x, \tilde{P}) = 0$ , then for each  $z$ ,  $d_z(x, \tilde{P}) = 0$ . Hence each  $d_z$  is identical with  $v'$ , and  $v'$  is either unilateral or duple.

If  $v'$  is duple -- that is, has at most two possible outcomes -- then by (i) of Assertion 17,  $v$  is duple.

Let  $v'$  be unilateral, with  $k$  as ruler. Then  $v'$  is weakly dictatorial. For let  $x$  be a possible outcome. Then for some ranking  $n$ -tuple  $\underline{p}^*$ ,  $v'(x, \underline{p}^*) > 0$ , and hence  $v'(x, \underline{p}^*) = 1$ . Thus since  $v'$  is unilateral,  $v'(x, \underline{p}) = 1$  wherever  $P_k = P_k^*$ . Suppose, then, that  $v'$  is not weakly dictatorial, so that for some alternative  $x$  and ranking  $n$ -tuple  $\underline{p}'$ ,  $x$  is first in  $P_k'$  but  $v'(x, \underline{p}') \neq 1$ . Where  $\underline{p} = \underline{p}' /_k P_k^*$ , we know that  $v'(x, \underline{p}) = 1$ . Thus for any utility scale  $U$  which fits  $P_k'$ ,  $U(v\underline{p}) > U(v\underline{p}')$ , and  $v'$  is potentially manipulable by  $k$  at  $\underline{p}$ . The supposition that  $v'$  is not weakly dictatorial has been shown false. It follows from (ii) of Assertion 17 that  $v'$  itself is weakly dictatorial. That proves the Corollary.



NOTES

1. Zeckhauser (1973, p. 939) describes an extension of the "random dictator system" as follows: "Provide each voter with  $q$  ballots for his first choice,  $r$  for his second,  $s$  for his third, etc., with  $q > r > s$ . The selection procedure is random as before." He then goes on to say, "Thus we find that only variants of the random dictator system will elicit ballots unique with respect to individuals' . . . ordinal preferences." (Being 'unique' in Zeckhauser's terminology is roughly the same as being 'strategy-proof' in mine). If by "variants of the random dictator system" he means systems of the form specified in passage I have quoted, then Scheme 2 is a counterexample to this claim.
2. I owe this corollary to Hugo Sonnenschein.
3. Aspects of this proof are suggested by arguments in Zeckhauser (1973).