Published in "Choice, Welfare, and Development"
A Festschrift in Honour of Amartya K. Sen

Edited by K. Basu, P. Pattanaik and K. Suzumura
Clarendon Press Oxford.

# 6

# Majority Rule, Social Welfare Functions, and Game Forms

ERIC S. MASKIN

## 1. Introduction

In his classic article, 'A Possibility Theorem on Majority Decisions' (Sen 1966), Amartya Sen showed that, in seeking domains of preferences on which the method of majority rule is transitive, one can, in effect, reduce the search to the case of three alternatives and three individuals. More specifically, he identified a condition—*value restriction*[1]—that is defined for triples of alternatives, and demonstrated that, provided the number of individuals is odd, majority rule is transitive regardless of how individuals' preferences are drawn from the domain $\mathfrak{R}$, if and only if $\mathfrak{R}$ satisfies value restriction for each triple (hence the reduction to three alternatives). Moreover, value restriction is the necessary and sufficient condition *regardless* of the number of individuals, and so we have transitivity with $n$ individuals ($n$ odd) for domain $\mathfrak{R}$ if and only if we have transitivity with three individuals for domain $\mathfrak{R}$ (hence the reduction to three individuals).

Reducing a social choice problem to the case of three alternatives and three individuals is, I believe, quite a powerful and general technique. In this paper I apply Sen's technique to establish simple proofs of two other well-known results from social choice theory: the Gibbard–Satterthwaite theorem on the impossibility of strategy-proof game forms, and the Arrow impossibility theorem for social welfare functions.

I then go on to establish a new result for a majority rule. Just as Sen showed that value restriction is a necessary condition on a domain of preferences for majority rule to be transitive, so one can establish that the same is true of *any* anonymous and neutral collective choice rule (CCR) that

It is a great pleasure to dedicate this paper to Amartya Sen. Not only was his article on majority rule the direct inspiration for the methods I use here, but, more generally, his work has long exerted a profound influence on my thinking. Indeed, my fascination with his beautiful monograph *Collective Choice and Social Welfare* (Sen 1970) was the principal reason I first tried my hand at social choice theory many years ago.

Research support from the NSF is gratefully acknowledged. I should like to thank D. Campbell, K. Suzumura, and S. Tadelis for helpful comments.

[1] The domain $\mathfrak{R}$ satifies *value restriction* for the triple $\{a, b, c\}$ if, for some $x \in \{a, b, c\}$, (i) for all $R \in \mathfrak{R}$, $x$ is not strictly preferred to both of the other alternatives; or (ii) for all $R \in \mathfrak{R}$, $x$ is not strictly between the other two alternatives; or (iii) for all $R \in \mathfrak{R}$, the other two alternatives are not both strictly preferred to $x$.

satisfies independence of irrelevant alternatives and the Pareto property. Hence, if $F$ is some such CCR and is transitive on domain $\mathfrak{R}$, then majority rule is transitive on $\mathfrak{R}$ as well (as long as the number of individuals is odd). Moreover, unless $F$ is itself majority rule, there exists a domain of preferences $\mathfrak{R}'$ on which $F$ is not transitive but majority rule is. Thus, among CCR's satisfying the above properties, majority rule is the one (and the *only* one) that is transitive on the widest class of domains of preferences. This conclusion gives us another characterization of majority rule to complement the one provided by May (1952).

## 2. The Gibbard–Satterthwaite Theorem

Let $A$ be a non-empty set (possibly infinite) of social alternatives. Let $N = \{1, \ldots, n\}$ be the set of players. Given abstract strategy spaces $S_1$, $S_2, \ldots, S_n$, an $n$-person *game-form* $g$ for $A$ is a mapping $g$: $S_1 \times \cdots \times S_n \to A$, where the mapping is *onto* $A$. Let $\mathfrak{R}_A$ be the class of all orderings of the elements of $A$. For $R_i \in \mathfrak{R}_A$ and $i \in N$, the strategy $\bar{s}_i \in S_i$ is said to be *dominant* for player $i$ with preference ordering $R_i$ if, for all $s_i \in S_i$, and $s_{-1} \in \underset{j \neq i}{\times} S_j$,[2]

$$g(\bar{s}_i, s_{-i}) \, R_i \, g(s_i, s_{-i}).$$

Given a subclass $\mathfrak{R} \subseteq \mathfrak{R}_A$, $g$ is said to be *strategy-proof* on $\mathfrak{R}$ if, for all $i \in N$ and all $R_i \in \mathfrak{R}$, player $i$ with ordering $R_i$ has a dominant strategy. Player $i \in N$ is a *dictator* for $g$ if, for all $a \in A$, there exists $s_i^a \in S_i$ such that, for all $s_{-i} \in \underset{j \neq i}{\times} S_j$

$$g(s_i^a, s_{-i}) = a.$$

If some player is a dictator for $g$, $g$ is *dictatorial.* The basic result on strategy-proof game forms is as follows.

> THEOREM (Gibbard–Satterthwaite). Suppose that $\#(A) \geqslant 3$ and $n \geqslant 2$.[3] Then, if an $n$-person game-form for $A$ is strategy-proof on $\mathfrak{R}_A$, it is dictatorial.

Gibbard (1973) proved this theorem by showing that the existence of a strategy-proof game-form (SPGF) on $\mathfrak{R}_A$ implies the existence of a social

[2] The notation $g(\bar{s}_i, s_{-i})$ is equivalent to $g(s_1, \ldots, s_{i-1}, \bar{s}_i, s_{i+1}, \ldots, s_n)$. For a coalition $C (\subseteq \{1, \ldots, n\})$, the notation $g(\bar{s}_C, s_{-C})$ is equivalent to $g(s_1^*, \ldots, s_n^*)$ where

$$s_i^* = \begin{cases} \bar{s}_i, & i \in C \\ s_i, & i \notin C \end{cases}$$

[3] $\#(A)$ denotes the cardinality of $A$.

welfare function (SWF) satisfying Arrow's conditions (see Section 3). Arrow's theorem states that such a SWF must be dictatorial, and it is a simple matter to translate the dictatorship of a SWF to that of the game-form. Satterthwaite (1975), and Schmeidler and Sonnenschein (1975), proved the result without appealing to the Arrow theorem, but they required rather lengthier arguments.

I offer quite a short proof. In the spirit of 'reductionism', we shall consider the theorem to be proved if we can establish that the existence of a non-dictatorial $n$-person SPGF on $\mathfrak{R}_A$ implies the existence of a three-person non-dictatorial SPGF on $\mathfrak{R}_B$ where $B \subseteq A$ and $\#(B) = 3$. Showing that there is no non-dictatorial SPGF in the three-person, three-alternative case can then readily be shown by purely mechanical calculation.

*Proof.* Consider an $n$-player non-dictatorial game-form $g$: $S_1 \times \cdots \times S_n \to A$ that is strategy-proof on $\mathfrak{R}_A$. We first claim that there exists a three-player non-dictatorial game-form that is strategy-proof on $\mathfrak{R}_A$. If $n = 2$, we can extend $g$ to three players by adding a dummy player who has no influence on the outcome. This extended game-form is clearly non-dictatorial and strategy-proof. If $n > 3$, then there exists a two-player coalition that is not decisive for $g$ (a coalition is decisive if, for all $a \in A$, the coalition has a strategy vector that results in $a$ regardless of the strategies of the complementary coalition), since, if $C$ is decisive, $N\backslash C$ is not decisive. Assume without loss of generality that $\{n-1, n\}$ constitutes a non-decisive coalition. Define the $(n-1)$-player game-form $g^{**}$: $S_1^{**} \times \cdots \times S_{n-1}^{**} \to A$ so that, for $i = 1, \ldots, n-2$, and $S_i^{**} = S_i$, and $S_{n-1}^{**} = S_{n-1} \times S_n$, and for all $(s_1, \ldots, s_{n-2}, s'_{n-1}) \in S_1^{**} \times \cdots \times S_{n-1}^{**}$,

$$g^{**}(s_1, \ldots, s_{n-2}, s'_{n-1}) = g(s_1, \ldots, s_{n-2}, s_{n-1}, s_n),$$

where $s'_{n-1} = (s_{n-1}, s_n)$. (Notice that $g^{**}$ is obtained from $g$ by 'collapsing' players $n-1$ and $n$ into a single player $n-1$). The game-form $g^{**}$ is non-dictatorial because $g$ is non-dictatorial and $\{n-1, n\}$ is non-decisive. Notice that, for $i = 1, \ldots, n-2$ and $R_i \in \mathfrak{R}_A$, if $\bar{s}_i$ is a dominant strategy for player $i$ with ordering $R_i$ in $g$, it remains a dominant strategy in $g^{**}$. Moreover, if, for $R \in \mathfrak{R}_A$, $\bar{s}_{n-1}$ and $\bar{s}_n$ are dominant strategies for players $n-1$ and $n$ with the same ordering $R$ in $g$, then the strategy $\bar{s}'_{n-1} = (\bar{s}_{n-1}, \bar{s}_n)$ is dominant for player $n-1$ with ordering $R$ in $g^{**}$. Hence $g^{**}$ is strategy-proof, and so we have shown that the existence of an $n$-player game-form that is non-dictatorial and strategy-proof implies the existence of an $(n-1)$-player game-form with the same properties. Thus our first claim is established.

Henceforth assume that $n = 3$. Because $g$ is non-dictatorial, we can choose a three-element subset $B = \{a, b, c\} \subseteq A$ such that, for all $i = 1, 2, 3$, there exists $x \in B$ that player $i$ cannot force. (Player $i$ can *force* $x$ in game-form $g$ if there exists $s_i^x$ such that, for all $s_{-i} \in \times_{j \neq i} S_j$, $g(s_i^x, s_{-i}) = x$.) Let

$\mathfrak{R}_A^B$ consist of all orderings in $A$ that rank the alternatives of $B$ strictly above all other alternatives. For each $i = 1, 2, 3$, let $S_i^B = \{s_i \in S_i \mid \exists\, R \in \mathfrak{R}_A^B$ such that $s_i$ is a dominant strategy for player $i$ with preference ordering $R$.

Consider the restriction $g^B$ of g to $S_1^B \times S_2^B \times S_3^B$. We claim that the range of $g^B$ is $B$. Suppose that there exist strategies $(s_1, s_2, s_3) \in S_1^B \times S_2^B \times S_3^B$ such that $g^B(s_1, s_2, S_3) \notin B$. By construction, each $s_i$ is dominant for player $i$ with some ordering $R_i \in \mathfrak{R}_A^B$. Now because $g$ is onto $A$, there exist strategies $(s'_1, s'_2, s'_3)$ such that $g(s'_1, s'_2, s'_3) \in B$. Because $s_1$ is dominant for player 1, and player 1 prefers any element in $B$ to anything in $A \backslash B$, $g(s_1, s'_2, s'_3) \in B$. Similarly, $g(s_1, s_2, s'_3)$, $g(s_1, s_2, s_3) \in B$, a contradiction. Hence the range of $g^B$ must be contained in $B$. But a similar argument shows that if, for $i = 1, 2, 3$ $s_i$ is dominant for $R_i$ and $R_i$ ranks $a$ above all other alternatives, then $g(s_1, s_2, s_3) = a$. Hence the range of $g^B$ equals $B$.

Now, $g^B$ is clearly strategy-proof on $\mathfrak{R}_B$ because it is strategy-proof on $\mathfrak{R}_A^B$. If it is also non-dictatorial, we are done. Hence, assume that some player, say player 1, is a dictator for $g^B$. But, by choice of $B$, there exists some alternative in $B$, say $a$, that player 1 cannot force in $g$. Thus, if $s_1^a \in S_1^B$ is a dominant strategy for player 1 with an ordering that ranks $a$ above all other alternatives, there exist $d \in A \backslash B$ and $(s_2^d, s_3^d) \in S_2 \times S_3$ such that $g(s_1^a, s_2^d, s_3^d) = d$. Let $B' = \{a, b, d\}$, and define $g^{B'}$ by analogy with $g^B$. Player 1 is not a dictator for $g^{B'}$ since, from the above argument, he cannot force $a$. But players 2 and 3 are not dictators either, since, from our hypotheses about player 1 in $g^B$, $g^B(s_1^*, s_2^*, s_3^*) = g^{B'}(s_1^*, s_2^*, s_3^*) = a$ whenever (i) $s_1^*$ is a dominant strategy for an ordering that ranks $a$ above all other alternatives, and (ii) $s_2^*$ and $s_3^*$ are dominant strategies for orderings that rank $a$ and $b$ above all other alternatives. Thus, $g^{B'}$ is a $3 \times 3$ non-dictatorial, strategy-proof game-form, as required. ■

### 3. The Arrow Impossibility Theorem

Let $A$ and $\mathfrak{R}_A$ be as in Section 2 except that $A$ is now restricted to be finite. For $\mathfrak{R} \subseteq \mathfrak{R}_A$, an $n$-person social welfare function (SWF) on $\mathfrak{R}$ is a mapping

$$f: \mathfrak{R}^n \to \mathfrak{R}_A.$$

Following Arrow (1951), we define the following properties of SWFs.

> PARETO PROPERTY. The SWF $f$ satisfies the Pareto property if, for all $a, b \in A$ and all $(R_1, \ldots, R_n) \in \mathfrak{R}^n$, $aPb$ provided that $ap_ib$ for all $i = 1, \ldots, n$ (where $aPb$ means '$a$ is strictly socially preferred to $b$, given $f(R_1, \ldots, R_n)$', and $P_i$ is the strict ordering for individual $i$ corresponding to $R_i$).

INDEPENDENT OF IRRELEVANT ALTERNATIVES (IIA). The SWF $f$ satisfies IIA if, for all $a$, $b \in A$ and all $(R_1, \ldots, R_n)$ and $(R'_1, \ldots, R'_n) \in \mathfrak{R}^n$, we have $aRb$ *iff* $aR'b$ (where $R$ and $R'$ are the social orderings corresponding to $(R_1, \ldots, R_n)$ and $(R'_1, \ldots, R'_n)$, respectively), provided that, for all $i \in N$, $aR_ib$ iff $aR'_ib$.

An individual $i \in N$ is a *dictator* on B for $f$ if, for all $a$, $b \in B$ and all $(R_1, \ldots, R'_n) \in \mathfrak{R}^n$, $aP_ib$ implies $aPb$.

NON-DICTATORSHIP. The SWF $f$ satisfies non-dictatorship if there is no dictator on $A$ for $f$.

We shall call a SWF $f$ on $\mathfrak{R}$ satisfying the Pareto property, IIA, and non-dictatorship an *Arrow social welfare function* (ASWF).

ARROW IMPOSSIBILITY THEOREM. If $n \geqslant 2$ and $\#(A) \geqslant 3$, there exists no $n$-person ASWF on $\mathfrak{R}_A$.

*Proof.* Once again, we shall be satisfied to reduce the question of existence to the three-person, three-alternative case. (See Suzumura (1988) for an alternative proof that reduces the problem to the three-person case.) We first show that, if $f$ is an $n$-person ASWF on $\mathfrak{R}$, then there exists a three-person ASWF on $\mathfrak{R}^*$, where $\mathfrak{R}^*$ consists of the strict orderings of $A$. If $n = 2$, then we can add a 'dummy' person (who has no effect on the social ordering) and trivially extend $f$ to the three-person case. Assume, therefore, that $n \geqslant 4$. Then there exists a two-person subset of players that is not *decisive* (a subset $M \subseteq N$ is decisive if, for all $a$, $b \in A$ $aPb$ wherever $aP_ib$ for all $i \in M$). Without loss of generality, suppose that the subset $\{n-1, n\}$ is not decisive. Define $f^*\colon \mathfrak{R}_A^{n-1} \to \mathfrak{R}_A$ such that, for all $(R_1, \ldots, R_{n-1}) \in \mathfrak{R}_A^{n-1}$,

$$f^*(R_1, \ldots, R_{n-1}) = f(R_1, \ldots, R_{n-1}, R_{n-1}).$$

That is, we are 'collapsing' individuals $n-1$ and $n$ in $f$ into a single individual to obtain $f^*$. Clearly, $f^*$ satisfies the Pareto property and IIA. Because $\{n-1, n\}$ is not decisive for $f$, individual $n-1$ is not a dictator for $f^*$. If nobody else is a dictator either, we are done. Suppose, therefore, that some other individual, say individual 1, *is* a dictator for $f^*$. Now, individual 1 is *not* a dictator for $f$. Thus for some $a$, $b \in A$ and some profile $(\hat{R}_1, \ldots, \hat{R}_n) \in (\mathfrak{R}_A^*)^n$ $a\hat{P}_1b$ but $a$ is not preferred to $b$ socially under $f(\hat{R}_1, \ldots, \hat{R}_n)$. Since 1 *is* a dictator for $f^*$, $\hat{R}_{n-1}$ and $\hat{R}_n$ cannot rank $a$ and $b$ the same way (if $\hat{R}_{n-1}$ and $\hat{R}_n$ rank $a$ and $b$ the same way, then $f^*$ determines the social ordering of $a$ and $b$). Thus there exists some other individual, say $n-2$, such that $\hat{R}_{n-2}$ ranks $a$ and $b$ the same as either $\hat{R}_{n-1}$ or $\hat{R}_n$. For simplicity, assume the former (i.e., $\hat{R}_{n-2}$ and $\hat{R}_{n-1}$ rank $a$ and $b$ the same way). Define $f^{**}\colon (\mathfrak{R}_A^*)^{n-1} \to \mathfrak{R}_A$ so that, for all $(R_1, \ldots, R_{n-1}) \in (\mathfrak{R}_A^*)^{n-1}$,

$$f^{**}(R_1, \ldots, R_{n-1}) = f(R_1, \ldots, R_{n-2}, R_{n-2}, R_{n-1}).$$

That is, we are 'collapsing' individuals $n-2$ and $n-1$ in $f$ into a single individual, $n-2$, to obtain $f^{**}$. Now individual 1 cannot be a dictator for $f^{**}$ (from IIA and our argument about $\hat{R}_1, \ldots, \hat{R}_n$)). Moreover, nobody else can be a dictator for $f^{**}$ either, because if 1 prefers $a$ to $b$ and everyone else prefers $b$ to $a$, the fact that 1 is a dictator for $f^*$ implies that $a$ is socially preferred to $b$ under $f^*$ and hence $f^{**}$. Thus $f^{**}$ is an ASWF, and so we may assume that $n = 3$.

Choose $a \in A$, and for any $R \in \mathfrak{R}_A$ let $R^a$ be the restriction of $R$ to $A \backslash a$. Define the three-person SWF $f^a$ to be the restriction of $f$ to $R_{A \backslash a}$. Because $f$ satisifies IIA, $f^a$ is well defined. Clearly, $f^a$ satisfies the Pareto property and IIA. Thus, if $f^a$ satisfies non-dictatorship, we have succeeded in reducing the cardinality of $A$ by 1. Assume, therefore, that some individual—say individual 1—is a dictator for $f^a$ on $A \backslash a$. Because 1 is not a dictator for $f$, however, there exists $b \in A \backslash a$ such that 1 is not a dictator on $(a, b)$ for $f$. Choose $c \in A \backslash \{a, b\}$. Consider $f^c$ (defined by analogy with $f^a$). Individual 1 is not a dictator for $f^c$ because 1 is not a dictator on $\{a, b\}$. Moreover, no other individual is a dictator for $f^c$ because, for any $d \in A \backslash \{a, b, c\}$, 1 is a dictator on {b,d} (since by assumption he is a dictator for $f^a$ on A\a). Therefore $f^c$ is an ASWF on $\mathfrak{R}_{A \backslash c}$. Proceeding iteratively, we may infer the existence of a three-person ASWF on $\mathfrak{R}_B$, where B is a three-alternative subset of A. ■

## 4. Majority Rule

An $n$-person *collective choice rule* (CCR) is a function that maps profiles of preferences $(R_1, \ldots, R_n)$ drawn from $\mathfrak{R}_A$ to a complete binary relation on $A$. This relation represents social preferences, but is not necessarily transitive. A CCR is *anonymous* if it is invariant with respect to permutations of the individuals' labels $\{1, \ldots, n\}$; it is *neutral* if it is invariant with respect to permutations of the alternatives' labels $\{a, b, \ldots\}$. The most familiar CCR is the *method of majority rule* (MMR): if $R^M$ is the majority rule social relation corresponding to the profile $((R_1, \ldots, R_n)$, then, for all $a$ and $b$, $aR^Mb$ if and only if $\#\{i \mid aP_ib\} \geqslant \#\{i \mid bP_ia\}$.

May (1952) characterized MMR by establishing that it is the unique CCR defined on $\mathfrak{R}_A$ (see Campbell (1982) and (1988) for extensions to the case of restricted domains; see Bordes (1976) and Campbell (1980) for alternative characterizations) satisfying anonymity, neutrality, the Pareto property (see Section 3), and *positive responsiveness*. (If there is a shift in someone's preference ordering in favour of alternative $a$ relative to $b$, and $a$ was previously no worse than $b$ socially, then $a$ must now be strictly socially preferred to $b$; note that positive responsiveness implies IIA.) I shall provide an alternative characterization. We know from Arrow's theorem (Section 3) that there exists no anonymous SWF satisfying IIA and the Pareto property on the unrestricted domain of preferences $\mathfrak{R}_A$. Thus, for example, MMR is intransitive for the celebrated Condorcet profile, that is, it is not a social welfare function for the unrestricted domain. But we may ask for which *restricted* domains of preferences $\mathfrak{R}$ a given CCR constitutes a SWF, i.e. on which domains it is transitive. This is precisely the question that Sen (1966) answered for MMR: (provided that the number of individuals is odd) MMR is a SWF on domain $\mathfrak{R}$ if and only if $\mathfrak{R}$ satisfies value restriction (see fn. 1). I will show that, among CCRs satisfying anonymity, neutrality, the Pareto property, and IIA, MMR is transitive on the widest class of domains of preferences (and is the unique such CCR).

> THEOREM. Let $F$ be an $n$-person ($n$ odd) CCR satisfying anonymity, neutrality, the Pareto property, and IIA. Suppose that $\mathfrak{R}$ is a domain of strict preferences on which $F$ is transitive, i.e. a social welfare function. Then MMR is also an SWF on $\mathfrak{R}$. Moreover, unless F is itself MMR, there exists some domain $\mathfrak{R}'$ on which MMR is a SWF but $F$ is not.

*Proof.* Consider a domain $\mathfrak{R}$ on which a CCR $F$ satisfying the hypotheses of the theorem is a SWF for some $n$ odd. We claim that $\mathfrak{R}$ satisfies value restriction (VR), and so will conclude, from Sen (1966), that MMR is a SWF on this domain too.

If $\mathfrak{R}$ fails to satisfy VR, then for some triple of alternatives $\{a, b, c\}$ there exists a triple of 'cyclic' preferences $R(abc)$, $R(bca)$, $R(cab) \in \mathfrak{R}$ such that $aP(abc)\,bp(abc)\,c$, $bP(bca)\,cP(bca)\,a$, and $cP(cab)\,aP(cab)\,b$. To see this, note that if VR fails then, for some $\{a, b, c\}$, there exists $R \in \mathfrak{R}$ such that $R$ ranks $a$ highest (among $\{a, b, c\}$). Because the labels '$b$' and '$c$' are arbitrary, we might as well suppose that $R = R(abc)$ Now, similarly, there exist $R'$, $R'' \in \mathfrak{R}$ such that $R'$ ranks $b$ highest and $R''$ ranks $c$ highest (again, among $\{a, b, c\}$). Now, either $R$, $R'$, and $R''$ together form a triple of cyclic preferences (there are two such cycles), or else two of the three orderings belong to the same cycle and the third belongs to the other cycle. In the former case our assertion is established. Thus, we might as well assume that $R' = R(bca)$ and $R'' = R(cba)$. Because VR fails, there exists $r \in \mathfrak{R}$ that ranks $a$ strictly between $b$ and $c$. Therefore, we might as well suppose that $\hat{R} = R(bac)$. (If $\hat{R} = R(cab)$, then $R$, $R'$ and $\hat{R}$ belong to the same cycle.) Similarly, the failure of VR implies that there exists $\hat{\hat{R}} \in \mathfrak{R}$ that ranks $b$ strictly below $a$ and $c$. As noted above, we might as well assume that $\hat{\hat{R}} \neq R(cab)$. But then $\hat{\hat{R}} = R(acb)$, and so the cycle $R(acb)$, $R(bac)$, $R(cba)$ belongs to $\mathfrak{R}$, as asserted. Henceforth, therefore, assume that $R(abc)$, $R(bca)$, $R(cab) \in \mathfrak{R}$.

Consider a profile $(R_1, \ldots, R_n)$ in which $aP_1b$ and $bP_ia$ for all $i \neq 1$. Let $R^F$ be the corresponding social ordering under $F$. There are three possible cases: (i) $bP^Fa$, (ii) $aI^Fb$ society is indifferent between $a$ and $b$, and (iii) $aP^Fb$. We will consider each of these in turn. Suppose first that $aP^Fb$. From IIA, anonymity, and neutrality, we infer from $bP^Fa$ that. for any alternatives

$x$ and $y$ and any profile in which all individuals but one prefer $x$ to $y$, $x$ is socially preferred to $y$. Consider the profile $(\hat{R}_1, \ldots, \hat{R}_n)$ such that $\hat{R}_1 = R(abc)$, $\hat{R}_2 = R(bca)$, and $\hat{R}_3 = \cdots = \hat{R}_n = R(cab)$. Let $\hat{R}^F$ be the corresponding social ordering. Now, everybody but individual 2 prefers $a$ to $b$. Hence, from anonymity and IIA, we have $a\hat{P}^F b$. Similarly, everybody but individual 1 prefers $c$ to $a$. Therefore, from neutrality, $c\hat{P}^F a$. We infer, from transitivity, that $c\hat{P}^F b$. Notice, moreover, that all individuals but 1 and 2 prefer $c$ to $b$. IIA, anonymity, and neutrality therefore imply that, whenever all but two individuals prefer alternative $x$ to $y$, $x$ is socially preferred to $y$.

Consider the profile $(\hat{\hat{R}}_1, \ldots, \hat{\hat{R}}_n)$ such that $\hat{\hat{R}}_1 = R(abc)$, $\hat{\hat{R}}_2 = \hat{\hat{R}}_3 = R(bca)$ and $\hat{\hat{R}}_4 = \ldots = \hat{\hat{R}}_n = R(cab)$. Arguing as above, we conclude that $c\hat{\hat{P}}^F b$, where $\hat{\hat{R}}^F$ is the social ordering corresponding to the profile. Continuing in the same way, we conclude that, for any $m < n$, any alternatives $x$ and $y$, and any profile in which all but $m$ individuals prefer $x$ to $y$, $x$ is socially preferred to $y$. But this is contradictory, because if all but $m$ individuals prefer $x$ to $y$, then all but $n - m$ individuals prefer $y$ to $x$, and so socially $y$ should be preferred to $x$. Thus, case (i), where $bP^F a$, is impossible.

We can derive the same contradiction in case (iii), where $aP^F b$. Suppose, therefore, that $aI^F b$. Arguing as above, we can show that, for all $m < n$, if all but $m$ individuals prefer $a$ to $b$, then society is indifferent between $a$ and $b$. Now, consider the profile $\bar{R}_1, \ldots, \bar{R}_n)$, in which $\bar{R}_1 = \cdots = \bar{R}_{n-1} = R(abc)$ and $\bar{R}_n = R(bca)$. From the above conclusion, we have $a\bar{I}^F b$ and $a\bar{I}^F c$, and so, from transitivity, $b\bar{I}^F c$. But everyone in this profile prefers $b$ to $c$, a contradiction of the Pareto property.

We conclude that $\mathfrak{R}$ must satisfy VR after all. Hence, from Sen (1966), MMR is a SWF on $\mathfrak{R}$.

We next turn to the second assertion of the theorem. Suppose that $F$ is an anonymous and neutral $n$-person CCR that satisfies IIA and the Pareto principle. Suppose, moreover, that $F \neq$ MMR. Hence there exist a profile $(R_1, \ldots, R_n)$ and alternatives $a$ and $b$ such that $n - m$ individuals (where $n - m > m$) prefer $a$ to $b$ and yet $a$ is not socially preferred to $b$. Let $A = \{a, b, c\}$ and $\mathfrak{R}' = \{R(cab), R(bca), R(cba)\}$. Notice that $\mathfrak{R}'$ satisfies VR ($a$ is never on top), and so MMR is transitive on $\mathfrak{R}'$. Consider the profile $\hat{R}_1, \ldots, \hat{R}_n)$ in which $\hat{R}_1 = \cdots = \hat{R}_m = R(cab)$, $\hat{R}_{m+1} = \cdots = \hat{R}_{2m} = R(bca)$, and $\hat{R}_{2m+1} = \cdots = \hat{R}_n = R(cba)$. Notice that $n - m$ individuals prefer $b$ to $a$, and so, from the above assumption about $F$, $a\hat{R}^F b$, where $\hat{R}^F$ is the social ordering corresponding to the profile. Similarly $b\hat{R}^F c$, and so, from transitivity, $a\hat{R}^F c$. But everyone in the profile prefers $c$ to $a$, and so, from the Pareto property, $c\hat{P}^F a$, a contradiction. We conclude that $F$ is not transitive on $\mathfrak{R}'$. ■

To illustrate this theorem, consider the *Pareto extension rule* (PER), which, after majority rule, is probably the best known CCR satisfying the theorem's hypotheses. PER is the rule defined so that, for all $(R_1, \ldots, R_n)$ and all a, $b \in A$, $aR^P b$ (where $R^P$ is the social preference relation corresponding to $(R_1, \dots R_n)$) if and only if $b$ does not strictly Pareto-dominate $a$. It can readily be shown that PER is transitive on domain $\mathfrak{R}$ consisting of strict orderings if and only if $\mathfrak{R}$ satisfies *quasi-agreement* (QA): for all $B = \{a, b, c\} \subseteq A$, there exists $x \in B$ such that either (i) for all $R \in \mathfrak{R}$, $x$ is ranked higher than the other two alternatives in $B$; or (ii) for all $R \in \mathfrak{R}$, $x$ is ranked in between the other two alternatives in $B$; or (iii) for all $R \in \mathfrak{R}$, $x$ is ranked below the other two alternatives in $A$. But QA is clearly strictly stronger than VR. In particular, for the six strict orderings of the alternatives $\{a, b, c\}$, VR requires the deletion of no more than two (one ordering from each triple of cyclic preferences). Thus, for example, $\{R(abc), R(bca), R(acb), R(cba)\}$ constitutes a domain satisfying VR ($a$ is never in between). In contrast, QA requires the deletion of at least four orderings. For example, $\{R(abc), R(acb)\}$ constitutes a maximal domain satisfying QA; the addition of any of the four other orderings would lead to a violation.

To conclude, I shall examine the roles of the neutrality and anonymity assumptions in the above theorem, since they are not required in Arrow's treatment of social welfare functions. If we are willing to dispense with neutrality, then it becomes possible to define SWFs on domains strictly bigger than those satisfying value restriction. For example, let $A = \{a, b, c\}$, and define the CCR $F^*$ so that, for all profiles $(R_1, \dots, R_n)$ and all $x, y \in \{a, b, c\}$ $xP^*y$ if and only if $xP(abc)y$ unless $yP_ix$ for all $i$, in which case $yP^*x$. It is not hard to see that $F^*$ is transitive on the domain consisting of all strict orderings but $R(cab)$, i.e. on a domain including one more ordering than any value-restricted domain.

If we drop anonymity, it is easy to construct a CCR other than MMR that satisfies all other hypotheses of the theorem as well as non-dictatorship and, moreover, is transitive on any value-restricted domain. For example, for $n > 3$, consider the CCR that determines social preferences according to majority rule applied to individuals 1, 2, and 3 alone, so that the remaining individuals are 'dummies' having no affect on the social ranking. Such a CCR is clearly transitive whenever the same is true of ordinary majority rule, and thus it demonstrates that the theorem is false if we simply substitute non-dictatorship for anonymity. I conjecture, however, that the theorem can be restored if we replace anonymity with the *conjunction* of non-dictatorship and a 'no dummy' assumption (so that every individual has some influence on the social ranking). This conjecture, however, must be left to future work.

## References

Arrow, K. (1951), *Social Choice and Individual Values*. John Wiley, New York.

Bordes, G. (1976), 'Consistency, Rationality and Collective Choice', *Review of Economic Studies*, 43: 447–57.

Campbell, D. (1980), 'Algorithms for Social Choice Functions', *Review of Economic Studies*, 47: 617–27.

—— (1982), 'On the Derivation of Majority Rule', *Theory and Decision*, 14: 133–40.

—— (1988), 'A Characterization of Simple Majority Rule for Restricted Domains', *Economics Letters*, 28: 307–10.

Gibbard, A. (1973), 'Manipulation of Voting Schemes: A General Result', *Econometrica*, 41: 587–601.

May, K. (1952), 'A Set of Independent, Necessary, and Sufficient Conditions for Simple Majority Decision', *Econometrica*, 20: 680–4.

Satterthwaite, M. (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.

Schmeidler, D. and Sonnenschein, H. (1975), 'The Possibility of a Nonmanipulable Social Choice Function', mimeo, Northwestern University.

Sen, A. (1966), 'A Possibility Theorem on Majority Decisions', *Econometrica*, 34: 491–9.

—— (1970), *Collective Choice and Social Welfare*, Holden Day, San Francisco.

Suzumura, K. (1988), 'Reduction of Social Choice Problems: A Simple Proof of Arrow's General Possibility Theorem', *Hitotsubashi Journal of Economics*, no. 2, 219–21.