

On The Proof of the Possibility Theorem for  
Social Welfare Functions

January 25, 1949

By: Kenneth J. Arrow

A number of readers have found the proof of the principal theorem in an earlier paper of the author's<sup>1</sup> to be needlessly obscure. An amplified version of the proof is here presented. Opportunity is also taken to correct proof of a lemma in the earlier part of the paper; the lemma in question is not used in the proof of the possibility theorem.

The following parts of Section 3 of Possibility will be presupposed: pages 3-4 down through Lemma 1, and Definition 4 on page 5. I will also refer freely to Conditions 1-6 on pages 9-11. It will be shown that Conditions 1-6 are inconsistent; this may be restated to say that if Conditions 1-3 and Condition 6 are satisfied, either Condition 4 or Condition 5 must be violated. This is precisely the statement of Theorem 2 (Possibility, page 14), when the meaning of Conditions 4-6 is taken into account.

To avoid confusion, the terms "weak ordering," "weak ordering relation," or simply, "ordering," or "ordering relation," will be used to denote the relation, "preferred or indifferent to." A knowledge of the weak ordering tells for all pairs of alternatives whether or not they are indifferent, and if not, which is preferred

---

1. The Possibility of a Universal Social Welfare Function, RAD (L)-289, 1948, Project Rand, referred to below as Possibility.

to which. The concept of ordering, as used here, is then equivalent to the ordinary term, "preference scale," or "indifference map," though avoiding the continuity assumptions usually implicit in the latter concept. The term, "preference relation," will be used here only in the sense of strict preference; if  $P$  is a preference relation, then  $x P y$  means that  $x$  is preferred to  $y$  and not merely indifferent.

By a social welfare function is meant a method of assigning to each possible set of individual ordering relations  $R_1, \dots, R_n$  a corresponding social ordering,  $R$ . The social welfare function, it should be made clear, is the method of making the assignment, and not any particular social ordering. In the following proof, we assume a given social welfare function satisfying Conditions 1-6 and show that the assumption leads to a contradiction.

1. Restatement of the Proof of the Inconsistency of Conditions 1-6

In what follows,  $V$  will stand for a set of individuals. In particular,  $V_0$  will stand for the null set of individuals (the set containing no individuals),  $V'$  will be a set containing a single individual, and  $V''$  will be the set of all individuals.

Definition 10. The set  $V$  is said to be decisive for  $x$  against  $y$  if  $x \succ y$  and  $x P_i y$  for all sets of individual ordering relations such that  $x P_i y$  for all  $i$  in  $V$ .

This definition may be explained as follows: Let  $\bar{R}$  stand for the set of individual ordering relations  $R_1, \dots, R_n$ . The condition that  $x P_i y$  for all  $i$  in  $V$  restricts the  $\bar{R}$ 's under consideration by restricting the range of variation of those components of  $\bar{R}$  whose subscripts are in  $V$  to ordering relations have the given property with respect

to  $x$  and  $y$ . To each  $\bar{R}$ , a given social welfare function assigns a social preference scale  $R$ ; according to this scale, we may have, in general,  $x P y$ , or  $x I y$ , or  $y P x$ . Suppose that it so happens that for all  $\bar{R}$  consistent with the condition that  $x P_i y$  for all  $i$  in  $V$ , the resultant  $R$  is such that  $x P y$ ; then we can say that  $V$  is decisive for  $x$  against  $y$ . Intuitively, the concept of decisive set can be explained as follows: a set of individuals is decisive if, whenever they all prefer  $x$  to  $y$ , society prefers  $x$  to  $y$  regardless of what preferences any individuals may have concerning any alternatives other than  $x$  or  $y$ .

It should be emphasized that the question of whether or not a given set of individuals is decisive with respect to a given pair of alternatives,  $x$  and  $y$ , is determined by the social welfare function and does not vary with the actual orderings of individuals at any given time.

Consequence 1. If there is some set of individual ordering relations  $R_1, \dots, R_n$  such that  $x P_i y$  for all  $i$  in  $V$  and  $y P_i x$  for all  $i$  not in  $V$  and such that the corresponding social preference relation yields the outcome,  $x P y$ , then  $V$  is decisive for  $x$  against  $y$ .

$R_1, \dots, R_n$  is a particular set of individual ordering relations having the property that,

$$x P_i y \text{ for all } i \text{ in } V, \tag{1}$$

$$y P_i x \text{ for all } i \text{ not in } V. \tag{2}$$

By hypothesis, we have that for this set of individual orderings, the corresponding social preference relation is such that,

$$x P y. \tag{3}$$

Now, let  $R_1^i, \dots, R_n^i$  be any set of individual orderings subject only to the condition that,

$$x P_i^i y \text{ for all } i \text{ in } V. \tag{4}$$

To show that  $V$  is decisive, it is necessary, according to the definition, to show that for every such set  $R_1^i, \dots, R_n^i$ , the corresponding social ordering,  $R^i$ , is such that  $x P^i y$ .

For any given  $R_1^i, \dots, R_n^i$  satisfying (4), construct a third set of individual preference scales  $R_1^{ii}, \dots, R_n^{ii}$ , as follows: Let  $V_1$  be the set of individuals not in  $V$  for whom  $x P^i y$ ; let  $V_2$  be the set of individuals for whom  $x I_1^i y$ . By (4),  $V_2$  does not overlap  $V$ . For  $i$  not in  $V_1$  or  $V_2$ , define  $R_1^{ii}$  to be the same as  $R_1^i$ ; in particular, this holds for all  $i$  in  $V$ , since  $V$  does not overlap  $V_1$  or  $V_2$ . That is,

$$x' R_1^{ii} y' \text{ if and only if } x' R_1^i y' \text{ for } i \text{ not in } V_1 \text{ or } V_2. \tag{5}$$

For  $i$  in  $V_1$ , define  $x' R_1^{ii} y'$  to hold if and only if one of the following conditions holds:

$$x' \neq x, y' \neq x, x' R_1^i y'; \tag{6}$$

$$x' \neq x, y' \neq x, x' P_i^i y; \tag{7}$$

$$x' = x, y R_1^i y'. \tag{8}$$

For  $i$  in  $V_2$ , define  $x' R_1^{ii} y'$  to hold if and only if one of the following conditions holds:

$$x' \neq x, y' \neq x, x' R_1^i y'; \tag{9}$$

$$x' \neq x, y' = x, x' R_1^i y; \tag{10}$$

$$x' = x, y R_1^i y'. \tag{11}$$

---

2. The definitions of the new orderings  $R_1^{ii}$  in the two cases ( $i$  in  $V_1$  and  $i$  in  $V_2$ ) as given by (6-11) are detailed ways of writing  
1 2

It must first be shown that the relations  $R_i^n$  so defined are in fact weak orderings, i.e., that they satisfy axioms I and II

(Possibility, page 4). For  $i$  not in  $V_1$  or  $V_2$ , this follows from (5), since  $R_i$  by assumption is an ordering relation for each  $i$ .

First, consider  $i$  in  $V_1$ . Since  $R_i$  is a weak ordering relation, for all  $x'$  and  $y'$ , we may assert that either  $x' R_i y'$  or  $y' R_i x'$ .

---

2. (cont. from pg. 3) rather simple transformations. For  $i$  in  $V_1$ , we simply take the original ordering and make only the following transformation: move  $x$  from its previous position below  $y$  to a position just above  $y$ , i.e., so that  $x$  is preferred both to  $y$  and to everything to which  $y$  is preferred or indifferent but inferior to every alternative which is actually preferred to  $y$ . For  $i$  in  $V_2$ , the only alteration needed to transform  $R_i$  to  $R_i^n$  is to move  $x$  from its previous position below  $y$  to a position indifferent to  $y$ . These alterations are of the simple type which occur in the hypothesis of Condition 2; since all that has happened in passing from the orderings  $R_i$  to the orderings  $R_i^n$  is that the position of  $x$  has either been raised or remained unaltered, and since society preferred  $x$  to  $y$  in the first situation, it is not unreasonable to postulate that society still prefers  $x$  to  $y$ . But the orderings  $R_i^n$  were so designed that for each individual the relative position of  $x$  and  $y$  is the same as in the scales  $R_i^j$ ; if we accept the principle of Condition 3 that all that matters in the social choice between  $x$  and  $y$  is the choices of individuals as between  $x$  and  $y$ , then the social choice between  $x$  and  $y$  when the individual choices are given by the orderings  $R_i^1, \dots, R_i^j$  must be the same as when the individual choices are given by the orderings  $R_i^1, \dots, R_i^j$ , i.e.,  $x$ .

Then, by (6), for any  $x' \neq x$ ,  $y' \neq x$ , either  $x' R_1'' y'$  or  $y' R_1'' x'$ .

Also, since  $R_1$  is an ordering, for all  $x'$  and  $y'$ , either  $x' P_1 y'$  or  $y' R_1 x'$ ; in particular, for all  $x'$ , either  $x' P_1 y$  or  $y R_1 x'$ . From (7) and (8), this means that for all  $x'$ , either

$x' R_1'' x$  or  $x R_1'' x'$ . Together with the previous result, we can

say that for all  $x'$  and  $y'$ , either  $x' R_1'' y'$  or  $y' R_1'' x'$ , and

therefore  $R_1''$ , with  $i$  in  $V_1$ , satisfies axiom I.

It must also be shown that  $R_1''$  is transitive (satisfies axiom II).

We will suppose that  $x' R_1'' y'$  and  $y' R_1'' z'$  and seek to prove that

$x' R_1'' z'$ . If  $x' \neq x$ ,  $y' \neq x$ ,  $z' \neq x$ , this follows from (6) and

the fact that  $R_1$  is a weak ordering relation and therefore transitive.

If  $x' = x$ ,  $y' \neq x$ ,  $z' \neq x$ , then from (8) and (6) we have

$y R_1 y'$  and  $y' R_1 z'$ , so that  $y R_1 z'$ , from which it follows that

$x' R_1'' z'$  by (8).

If  $x' \neq x$ ,  $y' \neq x$ ,  $z' = x$ , then from (7) and (8), we have

$x' P_1 y$  and  $y R_1 z'$ , so that  $x' R_1 z'$ , and therefore,  $x' R_1'' z'$  by (6).

If  $x' \neq x$ ,  $y' = x$ ,  $z' = x$ , then from (6) and (7),  $x' R_1 y'$

and  $y' P_1 y$ , so that  $x' P_1 y$ , and therefore  $x' R_1'' z'$  by (7).

If  $x' = x$ ,  $y' = x$ , then  $x' = y'$ , and therefore  $y' R_1'' z'$  implies

$x' R_1'' z'$ .

If  $x' = x$ ,  $z' = x$ , then we seek to prove that  $x R_1'' x$ . If  $i$  is

in  $V_1$ ,  $i$  is not in  $V$ , by definition, and therefore  $y P_1 x$ , by (2).

Hence,  $y R_1 x$ , and therefore,  $x R_1'' x$ , by (8). Since  $x' = x$ ,  $z' = x$ ,

$x' R_1'' z'$ .

If  $y' = x$ ,  $z' = x$ , then  $y' = z'$ , and therefore  $x' R_1'' y'$  implies

$x' R_1'' z'$ . All possible combinations of  $x'$ ,  $y'$ ,  $z'$  have been considered,

and we may say that for all  $x'$ ,  $y'$ ,  $z'$ ,  $x' R_1'' y'$  and  $y' R_1'' z'$  imply

$x' R_1^n z'$ .  $R_1^n$  is therefore transitive and hence is a weak ordering relation for  $i$  in  $V_1$ .

Now consider  $i$  in  $V_2$ . Again, it must be shown that  $R_1^n$  satisfies axioms I and II. As before, it follows from (7), since  $R_1$  satisfies axiom I (is a connected relation), that for  $x' \neq x$ ,  $y' \neq x$ , either  $x' R_1^n y'$  or  $y' R_1^n x'$ . Also since, in particular, for all  $x'$ , either  $x' R_1 y$  or  $y R_1 x$ , it follows from (8) and (9) that, for all  $x'$ , either  $x' R_1^n x$  or  $x R_1^n x'$ . Hence, for all  $x'$  and  $y'$ , either  $x' R_1 y'$  or  $y' R_1 x'$ , and therefore  $R_1^n$  satisfies axiom I for  $i$  in  $V_2$ .

Now consider the question of transitivity. We suppose that  $x' R_1^n y'$  and  $y' R_1^n z'$  and seek to prove that  $x' R_1^n z'$ . For  $x' \neq x$ ,  $y' \neq x$ ,  $z' \neq x$ , this follows from (9) and the fact that  $R_1$  is transitive.

If  $x' = x$ ,  $y' \neq x$ ,  $z' \neq x$ , then from (11) and (9), we have  $y R_1 y'$  and  $y' R_1 z'$ , so that  $y R_1 z'$  and therefore  $x' R_1^n z'$  by (11).

If  $x' \neq x$ ,  $y' = x$ ,  $z' \neq x$ , then from (10) and (11), we have  $x' R_1 y$  and  $y R_1 z'$ , so that  $x' R_1 z'$ , and therefore  $x' R_1^n z'$  by (9).

If  $x' \neq x$ ,  $y' \neq x$ ,  $z' = x$ , then from (9) and (10), we have  $x' R_1 y'$  and  $y' R_1 y$ , so that  $x' R_1 y$ , and therefore  $x' R_1^n z'$  by (10), since  $z' = x$ .

If  $x' = x$ ,  $y' = x$ , then  $x' = y'$ , and  $y' R_1^n z'$  implies  $x' R_1^n z'$ .

If  $x' = x$ ,  $z' = x$ , then we seek to prove that  $x R_1^n x$ . If  $i$  is in  $V_2$ , then  $i$  is not in  $V$ , as noted above, and therefore  $y P_1 x$ , by (2). This implies  $y R_1 x$ , which is equivalent to  $x R_1^n x$  by (11), if we replace  $y'$  there by  $x$ .

If  $y' = x$ ,  $z' = x$ , then  $y' = z'$ , so that  $x' R_1^n y'$  implies  $x' R_1^n z'$ . All possible combinations of  $x'$ ,  $y'$ ,  $z'$  have been considered, and we may say that for all  $x'$ ,  $y'$ ,  $z'$ ,  $x' R_1^n y'$  and  $y' R_1^n z'$ . Hence,  $R_1^n$  satisfies axiom II for all  $i$  in  $V_2$ . We have now shown that for all  $i$ ,  $R_1^n$  is a weak ordering relation.

From (5), (6), and (9),

$$x' R_1 y' \text{ if and only if } x' R_1^n y' \text{ for } x' \neq x, y' \neq x. \quad (12)$$

Suppose  $x R_1 y'$ . If  $i$  is not in  $V_1$  or  $V_2$ , then  $x R_1^n y'$  by (5). If  $i$  is in  $V_1$  or  $V_2$ , then  $i$  is not in  $V$ ,  $y P_1 x$  by (2). From  $y R_1 x$  and  $x R_1 y'$ , we have  $y R_1 y'$ , and  $x R_1^n y'$  by (8) or (11), whichever is appropriate.

$$x R_1 y' \text{ implies } x R_1^n y'. \quad (13)$$

Now suppose that  $x P_1 y'$ . That is,  $x R_1 y'$  but not  $y' R_1 x$ . From the first statement,  $x R_1^n y'$ , by (13). Suppose also that  $y' R_1^n x$ . If  $i$  is not in  $V_1$  or  $V_2$ , then  $y' R_1 x$ , by (5), contrary to hypothesis. If  $i$  is in  $V_1$ , then  $y' P_1 y$  from (7); if  $i$  is in  $V_2$ , then  $y' R_1 y$  from (10). For  $i$  in either  $V_1$  or  $V_2$ ,  $i$  is not in  $V$ , as noted above, and therefore  $y P_1 x$  by (2). Hence, for  $i$  in either  $V_1$  or  $V_2$ ,  $y' R_1 x$ , contrary to hypothesis. Therefore, the supposition that  $y' R_1^n x$  cannot hold for any  $i$ . The statement,  $x P_1 y'$  implies  $x R_1^n y'$  but not  $y' R_1^n x$ . That is,

$$x P_1 y' \text{ implies } x P_1^n y'. \quad (14)$$

But the relations between the set of ordering relations  $R_1, \dots, R_n$ , and the set of ordering relations  $R_1^n, \dots, R_n^n$ , as contained in (12-14), are precisely those between  $R_1, \dots, R_n$  and  $R_1^i, \dots, R_n^i$  postulated in the hypothesis of Condition 2 (monotonicity). Indeed, from (3) and (12-14),



it is seen that all of the hypothesis of Condition 2 is satisfied, and therefore, the conclusion follows. If  $P^n$  is the social preference relation corresponding to  $R_1^n, \dots, R_n^n$ , then we may write,

$$x P^n y. \quad (15)$$

Now compare the sets of ordering relations  $R_1^i, \dots, R_n^i$  and  $R_1^n, \dots, R_n^n$  as to how they order the elements of the set  $[x, y]$ , consisting of the two elements  $x$  and  $y$ . From (4) and the definition of  $V_1$ , we note that,

$$x P_i^i y \text{ for all } i \text{ in } V \text{ or } V_1. \quad (16)$$

By the definition of  $V_2$ ,

$$x I_i^i y \text{ for all } i \text{ in } V_2. \quad (17)$$

$$y P_i^i x \text{ for all } i \text{ not in } V, V_1, \text{ or } V_2. \quad (18)$$

For  $i$  in  $V$ , it follows from (5) and (1) that,

$$x P_i^n y \text{ for } i \text{ in } V. \quad (19)$$

Since  $y R_1^i y$ , it follows from (8) that  $x R_i^n y$  for  $i$  in  $V_1$ . On the other hand, since it is not true that  $y P_i^i y$ , it follows from (7) that it is not true that  $y R_i^n x$  for  $i$  in  $V_1$ .

$$x P_i^n y \text{ for } i \text{ in } V_1. \quad (20)$$

For  $i$  in  $V_2$ , it follows from the fact that  $y R_1^i y$  and from (10) and (11) that both  $x R_i^n y$  and  $y R_i^n x$  hold.

$$x I_i^n y \text{ for } i \text{ in } V. \quad (21)$$

For  $i$  not in  $V, V_1, \text{ or } V_2$ , it follows from (5) and (2) that,

$$y P_i^n x \text{ for } i \text{ not in } V, V_1, \text{ or } V_2. \quad (22)$$

Comparing (16) with (19) and (20), (17) with (21), and (18) with (22), we see that for each  $i$ ,  $R_i^i$  and  $R_i^n$  order the elements of the set  $[x, y]$  in exactly the same way. Hence, if  $C^i(S)$  and  $C^n(S)$  are the

social choices made out of the set  $S$  when the individual preference scales are  $R_1^i, \dots, R_n^i$  and  $R_1^{ii}, \dots, R_n^{ii}$ , respectively, then  $C'([x, y])$  and  $C''([x, y])$  contain exactly the same elements by Condition 3 (independence of irrelevant alternatives). But, by (15),  $C''([x, y])$  contains the single element  $x$ , and therefore so does  $C'([x, y])$ . Hence, if  $P'$  is the social preference relation corresponding to the set of individual ordering relations  $R_1^i, \dots, R_n^i$ , then,

$$x P' y. \tag{23}$$

But  $R_1^i, \dots, R_n^i$  were any set of individual preference scales subject only to (1), and (23) has been shown to hold for any such set. Therefore, by Definition 10,  $V$  is decisive for  $x$  against  $y$ . Q.E.D.

This extremely long-winded proof is really simple in principle. The purpose in introducing the auxiliary ordering relations  $R_1^{ii}$  was to permit a comparison with the original ordering relations  $R_1^i$  which would satisfy the hypotheses of Condition 2. At the same time, as far as the choice between alternatives  $x$  and  $y$  is concerned, the relations  $R_1^{ii}$  are essentially equivalent to the relations  $R_1^i$ ; this is shown by the latter part of the proof.

A brief summary of the proof can be given if we pretend that the whole universe of possible alternatives consists of  $x$  and  $y$ ; Condition 3 guarantees us, in effect, that in trying to find the social choice as between  $x$  and  $y$ , this pretense will not lead to any trouble. In this case, let  $R_1^i$  be defined as follows:  $x P_1^i y$  for all  $i$  in  $V$ ,  $y P_1^i x$  for all  $i$  not in  $V$ . By hypothesis,  $x P y$ . Let  $R_1^i, \dots, R_n^i$  be any other set of individual preference scales in  $[x, y]$  such that  $x P_1^i y$  for all  $i$  in  $V$ . If  $i$  belongs to  $V$ , then  $R_1^i$  is identical

to  $R_i^1$  in  $[x, y]$ , so that the hypotheses of Condition 2 are satisfied. If  $i$  does not belong to  $V$ , then the hypotheses of Condition 2 are satisfied vacuously. Therefore,  $x P^1 y$ ; since this result holds for any  $R_1^1, \dots, R_n^1$  satisfying the condition that  $x P^1 y$  for all  $i$  in  $V$ ,  $V$  is decisive for  $x$  against  $y$  by Definition 10.

The meaning of this consequence may be formulated somewhat as follows: Imagine an observer seeing individuals write down their individual orderings and hand them to the central authorities who then form a social ordering based on the individual orderings in accordance with the social welfare function. Suppose further that this observer notices that for a specific pair of alternatives  $x$  and  $y$ , every individual in a certain set  $V$  of individuals preferred  $x$  to  $y$ , while everybody not in  $V$  preferred  $y$  to  $x$ , and that the resultant social ordering ranks  $x$  higher than  $y$ . Then, the observer is entitled to say, without looking at any other aspects of the individual and social orderings, that  $V$  is a decisive set for  $x$  against  $y$ , i.e., that if tastes change but in such a way that all the individuals in  $V$  still prefer  $x$  to  $y$  (though they might have changed their rankings for all other alternatives and though the individuals not in  $V$  might have changed their scales completely), then the social ordering will still rank  $x$  higher than  $y$ .

Consequence 2. For every  $x$  and  $y$  such that  $x \neq y$ , there is a decisive set for  $x$  against  $y$ .

Proof: If we interchange  $x$  and  $y$  in Definition 8, then Condition 4 says that there exists a set of individual orderings  $R_1, \dots, R_n$  such that not  $y R x$ , where  $R$  is the social ordering corresponding to the set of individual ordering relations  $R_1, \dots, R_n$ . That is to say,  $x P y$ .

Let  $V$  be the set of individuals such that  $x R_1 y$ ,  $V_1$  the set of individuals such that  $x I_1 y$ . For  $i$  not in  $V_1$ , define  $R'_i$  to be the same as  $R_i$ :

$$\text{For all } x' \text{ and } y', x' R'_i y' \text{ if and only if } x' R_i y', \text{ for } i \text{ not in } V_1. \quad (24)$$

For  $i$  in  $V_1$ , define  $x' R'_i y'$  to hold if and only if one of the following hold:

$$x' = x \text{ or } y' \neq x, x' R_i y', \quad (25)$$

$$x' \neq x, y' = x, x' P_i x. \quad (26)$$

From (24) and the assumption that  $R_i$  is a weak ordering relation for each  $i$ , it follows that  $R'_i$  is a weak ordering for  $i$  not in  $V_1$ . Now suppose  $i$  in  $V_1$ . It must first be shown that  $R'_i$  obeys axiom I. Since  $R_i$  is a weak ordering, for all  $x'$  and  $y'$ , either  $x' R_i y'$  or  $y' R_i x'$ ; therefore, by (25), for all  $x' \neq x, y' \neq x$ , either  $x' R'_i y'$  or  $y' R'_i x'$ . Also, for all  $x'$ , either  $x' P_i x$  or  $x R_i x$ ; from (25) and (26), this implies either  $x' R'_i x$  or  $x R'_i x$ . Hence, in every case,  $R'_i$  is connected (satisfies axiom I).

To show transitivity in the case  $i$  is in  $V_1$ , assume that  $x' R'_i y'$  and  $y' R'_i z'$ . From (25) and (26), if  $x' R'_i y'$ , then  $x' R_i y'$  in any case. Similarly,  $y' R_i z'$ . Hence,  $x' R_i z'$ ; if  $x' = x$  or  $z' \neq x$ , then  $x' R'_i z'$ , by (25), and transitivity is demonstrated.

Now consider the remaining case,  $x' \neq x, z' = x$ . If  $y' = x$ , then  $y' = z'$ , so that the statement  $x' R'_i z'$  follows immediately from the hypothesis

<sup>3</sup>For  $i$  in  $V_1$ , the original ordering  $R$  made  $x$  and  $y$  indifferent.  $R'$  modifies the ordering  $R$  only by moving  $x$  just ahead of  $y$ . For  $i$  not in  $V_1$ , the orderings  $R$  and  $R'$  are the same. Since the only changes made move  $x$  up on the individual scales, the previous social preference for  $x$  over  $y$  remains unchanged. But the new orderings fall into the pattern of the hypothesis of Consequence 1 and hence can be used to establish the decisiveness of  $V$ .

$x' R_1' y'$ . If  $y' \neq x$ , then from (25) and (26), we have  $x' R_1' y'$  and  $y' P_1' x$ , which imply  $x' P_1' x$  and therefore, by (26),  $x' R_1' x$  or  $x' R_1' z'$ . Hence,  $R_1'$  is transitive, and is therefore a weak ordering relation.

From (24) and (25),

If  $x' \neq x$ ,  $y' \neq x$ , then  $x' R_1' y'$  if and only if  $x' R_1' y'$ . (27)

Also from (24) and (25),

$x R_1' y'$  implies  $x R_1' y'$ . (28)

Now suppose  $x P_1' y'$ . Then  $x R_1' y'$  but not  $y' R_1' x$ . From the first statement and (28), we have  $x R_1' y'$ . From the second, we cannot have  $y' = x$ , since otherwise not  $x R_1' x$ , which is absurd. Suppose also  $y' R_1' x$ ; then, since  $y' \neq x$ , it follows from (26) that  $y' P_1' x$  for  $i$  in  $V_1$ , which contradicts the statement that not  $y' R_1' x$ . Hence, for such  $i$ ,  $x P_1' y'$  implies  $x P_1' y'$ . The same result holds for  $i$  not in  $V_1$ , by (24).

$x P_1' y'$  implies  $x P_1' y'$ . (29)

Since, as noted earlier,  $x P y$ , all the hypotheses of Condition 2 are satisfied, according to (27-29); we have, if  $P'$  is the social preference relation corresponding to the individual orderings  $R_1', \dots, R_n'$ ,

$x P' y$ . (30)

If  $i$  is in  $V$  but not in  $V_1$ , then  $x P_1' y$  by definition of  $V$  and  $V_1$ , and therefore, by (24),

$x P_1' y$  for  $i$  in  $V$  but not in  $V_1$ . (31)

If  $i$  is in  $V_1$ , then, by definition of  $V_1$ ,  $x I_1' y$ . Hence,

$x R_1' y$  for  $i$  in  $V_1$ , (32)

$y I_1' x$  for  $i$  in  $V_1$ . (33)

From (32) and (25),  $x R_i' y$ . From (33) and (26), not  $y R_i' x$ , so that  $x P_i' y$  for all  $i$  in  $V_1$ . Combine this result with (31) to yield,

$$x P_i' y \text{ for } i \text{ in } V. \tag{34}$$

For  $i$  not in  $V$ , we have, by definition of  $V$ ,  $y P_i' x$ . By (24),

$$y P_i' x \text{ for } i \text{ not in } V. \tag{35}$$

From (30), (34) and (35), we find that there is a set of individual orderings such that the set of individuals  $V$  can, by all preferring  $x$  to  $y$ , cause society to prefer  $x$  to  $y$ , even though all other individuals prefer  $y$  to  $x$ . By Consequence 1, it follows that  $V$  is a decisive set for  $x$  against  $y$ .

Consequence 3. For every  $x$  and  $y$  such that  $x \neq y$ ,  $V''$  is decisive for  $x$  against  $y$ .

Recall that  $V''$  is the set of all individuals. Consequence 3 says that if everybody prefers  $x$  to  $y$ , then society prefers  $x$  to  $y$ , regardless of any other comparisons of alternatives made by individuals.

Proof: By Consequence 2, there is a set  $V$  which is decisive for  $x$  against  $y$ . Let  $R_1, \dots, R_n$  be a set of individual preference scales such that  $x P_i' y$  for all  $i$ . That is,

$$x P_i' y \text{ for all } i \text{ in } V'', \tag{36}$$

$$y P_i' x \text{ for all } i \text{ not in } V''. \tag{37}$$

The last statement is vacuously true, since there are no  $i$  not in  $V''$ . In particular,  $x P_i' y$  for all  $i$  in  $V$ . By Definition 10 (the definition of a decisive set),

$$x P y, \tag{38}$$

where  $P$  is the social preference relation corresponding to the set of individual ordering relations  $R_1, \dots, R_n$ . But, by (35-38), this set of individual ordering relations satisfies the hypothesis of Consequence 1, so that we may infer that  $V''$  is a decisive set.

Consequence 4. If  $V'$  is decisive for either  $x$  against  $y$  or  $y$  against  $z$ ,  $V'$  is decisive for  $x$  against  $z$ , where  $x$ ,  $y$ , and  $z$  are distinct alternatives.

Recall that  $V'$  is a set consisting of a single individual. The consequence asserts that if a single individual is decisive for a given alternative  $x$  against any other alternative, he is decisive for  $x$  against any alternative, and that if he is decisive for any alternative against a given alternative  $z$ , he is decisive for any alternative against  $z$ . This is the first consequence in which some paradoxes begin to appear.

Proof: (1) Assume that  $V'$  is decisive for  $x$  against  $y$ . We seek to prove that  $V'$  is decisive for  $x$  against any  $z$ .

Let the individual in  $V'$  be given the number 1. Let  $R_1, \dots, R_n$  be a set of individual ordering relations satisfying the conditions,

$$x P_1 y, \tag{39}$$

$$y P_i z \text{ for all } i, \tag{40}$$

$$z P_i x \text{ for } i \neq 1. \tag{41}$$

From (39),  $x P_i y$  for all  $i$  in  $V'$ ; therefore, by Definition 10,

$$x P y, \tag{42}$$

where  $F$  is the social preference relation corresponding to the set of individual orderings  $R_1, \dots, R_n$ . From (40),  $y P_i z$  for all  $i$  in  $V''$ , so that, from Consequence 3 and the definition of a decisive set,

$$y P z. \tag{43}$$

By Condition 1, the social ordering relation satisfies axioms I and II. Therefore, from (42) and (43),

$$x P z. \tag{44}$$

But, from (39) and (40),  $x P_i y$  and  $y P_i z$ , so that  $x P_i z$ , or,

$$x P_i z \text{ for all } i \text{ in } V'. \tag{45}$$

(41) may be written,

$$z P_i x \text{ for all } i \text{ not in } V'. \tag{46}$$

By (44)-(46), the hypotheses of Consequence 1 are satisfied, so that  $V'$  must be decisive for  $x$  against  $z$ . That is, there is one set of individual ordering relations in which all the individuals in  $V'$  (in this case, one individual) prefer  $x$  to  $z$  while all other individuals prefer  $z$  to  $x$ , and the social welfare function is such as to yield a social preference for  $x$  as against  $z$ . This suffices, by Consequence 1, to establish that  $V'$  is decisive for  $x$  against  $z$ .

(2) Now assume that  $V'$  is decisive for  $y$  against  $z$ . Let the individual in  $V'$  have the number 1, and let  $R_1, \dots, R_n$  be a set of individual ordering relations such that,

$$x P_i y \text{ for all } i, \tag{47}$$

$$y P_1 z, \tag{48}$$

$$z P_i x \text{ for } i \neq 1. \tag{49}$$

Then, as in part (1), (47) implies that  $x P y$ , while (48) implies that  $y P z$ , so that  $x P z$ . But, from (47) and (48),  $x P_1 z$ , which, in conjunction with (49), shows that the hypotheses of Consequence 1 are satisfied, and therefore,  $V'$  is decisive for  $x$  against  $z$  again.

Consequence 5. There exists no individual  $i$  and no alternatives  $x$  and  $y$  such <sup>that</sup>  $x P_i y$  implies  $x P y$  regardless of the ordering relations of



all other individuals.

This consequence states that no individual can be a dictator for even one pair of alternatives; i.e., there is no individual such that, with the given social welfare function, the community automatically prefers a certain  $x$  to a certain  $y$  whenever the individual in question does so.

Proof: Suppose the consequence false. Then there is an individual, whom we may designate by  $l$ , and a pair of alternatives  $x$  and  $y$  such that  $x P_l y$  implies  $x P y$ . Let  $V'$  be the set consisting of the single individual  $l$ ; then, by Definition 10,

$V'$  is decisive for  $x$  against  $y$ . (50)

Let  $y'$  be any alternative distinct from  $x$  and  $y$ . Then, from (50) and Consequence 4,  $V'$  is decisive for  $x$  against  $y'$ . Since, by (50), this statement is still true for  $y' = y$ , we may say,

$V'$  is decisive for  $x$  against any  $y' \neq x$ . (51)

For a fixed  $y' \neq x$ , let  $x'$  be an alternative distinct from  $x$  and  $y'$ . This choice is possible by Condition 6 (that there are three alternatives). Then, from (51) and Consequence 4,  $V'$  is decisive for  $x'$  against  $y'$ . By (51), this statement still holds if  $x' = x$ .

$V'$  is decisive for  $x'$  against  $y'$ , provided  $x' \neq y'$ ,  $y' \neq x$ . (52)

Choose any  $x' \neq x$ , and a particular  $y''$  distinct from both  $x$  and  $x'$ . This choice is possible by Condition 6. Then (52) holds; since  $x'$ ,  $y''$ ,  $x$  are distinct, it follows from Consequence 4, if we substitute  $x'$  for  $x$ ,  $y''$  for  $y$ , and  $x$  for  $z$ , that

$V'$  is decisive for  $x'$  against  $x$ , provided  $x' \neq x$ . (53)

(52) and (53) together can be written,

$V'$  is decisive for any  $x'$  against any  $y'$ , provided  $x' \neq y'$ . (54)

But, by Definition 10, (54) says that for all  $x'$  and  $y'$  (distinct),  $x' P y'$  whenever  $x' P_1 y'$ . By Definition 9, this means that the social welfare function is dictatorial, which, however, is excluded by Condition 5. Hence, the supposition that the consequence is false leads to a contradiction with one of the conditions. Q.E.D.

It will now be shown that Conditions 1-6 lead to a contradiction. Use will be made of the preceding five consequences of the conditions. Let  $S$  be a set composed of three distinct alternatives; the existence of such a set is guaranteed by Condition 6. For each possible ordered pair  $x', y'$  such that  $x'$  and  $y'$  both belong to  $S$  and  $x' \neq y'$  (there are six such ordered pairs), there is at least one set of individuals which is decisive for  $x'$  against  $y'$  by Consequence 2. Consider all sets of individuals who are decisive for some  $x'$  in  $S$  against some  $y'$  in  $S$  and distinct from  $x'$ . Among these sets, choose the one with the fewest number of individuals; if this condition does not uniquely specify the set, choose any of those decisive sets which does not have more members in it than some other decisive set. E.g., if, among all the sets which are decisive for some  $x'$  in  $S$  against some (distinct)  $y'$  in  $S$ , there is one with two members and all the others have more than two members, choose that one; on the other hand, if there are two sets decisive for some  $x'$  in  $S$  against some  $y'$  in  $S$  which have three members each while all the other decisive sets have more than three members, choose any one of the three-member sets. Designate the chosen set by  $V_1$ . It is decisive for some alternative in  $S$  against some other one in  $S$ ; by suitable labeling, we may say that  $V_1$  is decisive for  $x$  against  $y$ .  $S$  contains just one alternative other than  $x$  and  $y$ ; call that alternative  $z$ . Let the number of members of  $V_1$  be  $k$ ; designate the members of  $V_1$  by the numbers

$1, \dots, k$ , and number the remaining members  $k + 1, \dots, n$ . Let  $V'$  contain the single individual 1,  $V_2$  the individuals  $2, \dots, k$ , and  $V_3$  individuals  $k + 1, \dots, n$ . From the construction of  $V_1$ , we may conclude,

$$V_1 \text{ is decisive for } x \text{ against } y, \quad (55)$$

any set which is decisive for some alternative in  $S$  against some other alternative in  $S$  contains at least  $k$  members. (56)

By construction,  $V_2$  contains  $k-1$  members. Hence, from (56),

$V_2$  is not decisive for any alternative in  $S$  against any other alternative in  $S$ . (57)

Consequence 5 is equivalent to stating that if  $V'$  contains exactly one member, then  $V'$  is not decisive for any alternative against any other alternative. (58)

Let  $R_1, \dots, R_n$  be a set of individual ordering relations such that, for  $i$  in  $V'$ ,  $x P_i y$ , and  $y P_i z$ , (59)

for  $i$  in  $V_2$ ,  $z P_i x$ , and  $x P_i y$ , (60)

for  $i$  in  $V_3$ ,  $y P_i z$ , and  $z P_i x$ . (61)

From (59), (60), and the definitions of  $V_1$ ,  $V_2$ , and  $V'$ ,  $x P_i y$  for all  $i$  in  $V_1$ . From (55),

$$x P y, \quad (62)$$

where  $P$  is the social preference relation corresponding to  $R_1, \dots, R_n$ .

From (60), and the fact that  $R_i$  is a weak ordering relation and hence transitive,

$$z P_i y \text{ for all } i \text{ in } V_2. \quad (63)$$

From (59) and (61),

$$y P_i z \text{ for all } i \text{ not in } V_2. \quad (64)$$

Suppose  $z P y$ . Then from (63), (64), and Consequence 1, it would follow that  $V_2$  was decisive for  $y$  against  $z$ ; but this contradicts (57). Hence, we must say, not  $z P y$ , or,

$$y R z, \tag{65}$$

where  $R$  is the social ordering relation corresponding to  $R_1, \dots, R_n$ , the relation from which the preference relation  $P$  was derived. By Condition 1, the relation  $R$  is a weak ordering relation, having all the usual properties assigned to preference scales, including that of transitivity. Hence, from (62) and (65),

$$x P z. \tag{66}$$

From (59), it follows from the transitivity of  $R_i$  that,

$$x P_i z \text{ for } i \text{ in } V', \tag{67}$$

while from (60) and (61),

$$z P_i x \text{ for } i \text{ not in } V'. \tag{68}$$

From (66-68) and Consequence 1, it follows that  $V'$  is decisive for  $x$  against  $z$ . But this contradicts (58). Thus, we have shown that Conditions 1-6 taken together lead to a contradiction. Put another way, if we assume that our social welfare function satisfies Conditions 1-3 and we further suppose that Condition 6 holds, that there are more than two alternatives which society is to rank, then either Condition 4 or Condition 5 must be violated. Condition 4 states that the social welfare function is not conventional; Condition 5 states that it is not dictatorial.

Theorem 2. If the number of alternatives exceeds two, every social welfare function satisfying Conditions 1-3 is either conventional or dictatorial.

2. A Correct Proof for Lemma 11.

I wish to take this opportunity of correcting the statement and proof of Lemma 11 on page 8 of Possibility. The notation will be changed slightly.

The situation we wish to formalize is one in which some but not all of the choices made by a chooser are known in advance. Of some pairs  $x, y$  it is known that  $x$  is preferred to  $y$ ; of some it is known that  $x$  is indifferent to  $y$ . This can be reworded by saying that for some pairs of alternatives,  $x$  is known to be preferred or indifferent to  $y$ ; the case where  $x$  is known to be indifferent to  $y$  is covered by saying that  $x$  is known to be preferred or indifferent to  $y$  and  $y$  is known to be preferred or indifferent to  $x$ . We further assume that if  $x$  is known to be preferred or indifferent to  $y$  but  $x$  is not known to be indifferent to  $y$ , then  $x$  is known to be preferred to  $y$ . Clearly, the relation of "known preference or indifference" is a partial (weak) ordering, as defined in Possibility, p. 8. The relation between this partial ordering and the complete ordering of all choices can be expressed in the following definition.

Definition 11. If  $R$  and  $Q$  are relations,  $R$  is said to be compatible with  $Q$ , if the following conditions hold:

- (a)  $R$  is a weak ordering and  $Q$  is a partial ordering;
- (b) for all  $x$  and  $y$ , if  $x Q y$ , then  $x R y$ ;
- (c) for all  $x$  and  $y$ , if  $x Q y$  but not  $y Q x$ , then  $x R y$  but not  $y R x$ .

Before restating and proving Lemma 11 several preliminary lemmas will be presented.

Lemma A. Let  $X'$  be a space divided into three mutually exclusive and exhaustive sets,  $\underline{S}$ ,  $S$  and  $\bar{S}$ . Let  $R$  be a relation between pairs of

elements in  $X'$  having the following properties:

- (a)  $R$  is a weak ordering within each of the three sets;
- (b) if  $x$  and  $y$  belong to different sets, then  $x R y$  holds if and only if  $x$  belongs to  $\bar{S}$  and  $y$  does not or if  $x$  belongs to  $S$  and  $y$  to  $\underline{S}$ .

Then  $R$  is a weak ordering on the entire space  $X'$ .

Essentially, we have divided the space into three parts, such that every element in the first is superior to every element in the second and third, while every element in the second is superior to every element in the third. Further, within each category, the elements are ranked consistently. Then, we conclude that all elements are ranked consistently.

Proof: To prove that  $R$  is a weak ordering, we have to show that  $R$  is connected and transitive, i.e., satisfies axioms I and II.

(1)  $R$  is connected: If  $x$  and  $y$  are in the same set, then either  $x R y$  or  $y R x$ , by hypothesis (a) and I.

If  $x$  is in  $\bar{S}$  and  $y$  is not, then  $x R y$ , by (b) of the hypothesis.

If  $x$  is in  $S$  and  $y$  in  $\bar{S}$ , then  $y R x$ , by (b).

If  $x$  is in  $S$  and  $y$  in  $\underline{S}$ , then  $x R y$ , by (b).

If  $x$  is in  $\underline{S}$  and  $y$  is not, then  $y R x$ , by (b).

Since the three sets exhaust  $X'$ , we have shown that for all  $x$  and  $y$ , either  $x R y$  or  $y R x$ .

(2)  $R$  is transitive: We assume  $x R y$  and  $y R z$ , and seek to show that  $x R z$ .

If  $x$  belongs to  $\bar{S}$  and  $z$  does not, then certainly  $x R z$  by (b). If  $z$  also belongs to  $\bar{S}$ , then, by (b),  $y R z$  can only be true if  $y$  also belongs to  $\bar{S}$ . Since  $R$  is a weak ordering in  $\bar{S}$  by (a), it is transitive by II, so that  $x R z$ .

If  $x$  belongs to  $S$ , then  $x R y$  can only hold if  $y$  belongs to  $S$  or

$\underline{S}$  by (b). If  $y$  belongs to  $\underline{S}$ , then by the same argument,  $y R z$  implies that  $z$  belongs to  $\underline{S}$  or  $\underline{S}$ . If  $z$  also belongs to  $\underline{S}$ , then  $x R z$ , by (a) and II; if  $z$  belongs to  $\underline{S}$ , then  $x R z$  by (b). If  $y$  belongs to  $\underline{S}$ , then  $y R z$  requires that  $z$  belongs to  $\underline{S}$  by (b), so that  $x R z$  by (b).

If  $x$  belongs to  $\underline{S}$ , then by (b),  $x R y$  implies  $y$  belongs to  $\underline{S}$ , and then  $y R z$  implies that  $z$  belongs to  $\underline{S}$ . Then  $x R z$  by (a) and II.

All cases have been covered so that  $R$  is transitive, and therefore a weak ordering on the entire space  $X'$ , Q.E.D.

Lemma B. Let the space  $X$  be divided into two mutually exclusive subsets  $X'$  and  $S^*$ . Let  $f(x)$  be a function on  $S^*$  into  $X'$ , i.e., a rule which assigns to each element of  $S^*$  a unique corresponding element of  $X'$ . Let  $R$  be a relation on  $X$  having the following properties:

- (a)  $R$  is a weak ordering on  $X'$ ;
- (b) if  $x$  belongs to  $S^*$  and  $y$  to  $X'$ , then  $x R y$  if and only if  $f(x) R y$ ;
- (c) if  $x$  belongs to  $X'$  and  $y$  to  $S^*$ , then  $x R y$  if and only if  $x R f(y)$ ;
- (d) if  $x$  and  $y$  both belong to  $S^*$ , then  $x R y$  if and only if  $f(x) R f(y)$ .

Then  $R$  is a weak ordering on the entire space  $X$ .

The meaning of this is that there is assumed to be an ordering of a particular space  $X'$ ; then some additional elements are introduced, and they are given the same place in the order as a corresponding element in  $X'$ . Then it is asserted that this method yields a consistent ranking.

Proof: Again, it is to be shown that  $R$  is connected and transitive.

The simplest procedure is to define  $f(x)$  for all elements in  $X$  by setting  $f(x) = x$  for  $x$  in  $X'$ . It is already defined for  $x$  in  $S^*$ . Then,

with the aid of the hypothesis,

For all  $x$  in  $X$ ,  $f(x)$  is in  $X'$ . (1)

Further, from the extended definition of  $f(x)$  and (b)-(d),

For all  $x$  and  $y$  in  $X$ ,  $x R y$  if and only if  $f(x) R f(y)$ . (2)

(2) can easily be shown by considering all possible cases; thus, if  $x$  and  $y$  both belong to  $X'$ , then  $f(x)$  and  $f(y)$  are the same as  $x$  and  $y$ , respectively, so that (2) is trivial. The other cases follow from (b)-(d), respectively.

However, we know  $R$  to be connected and transitive in  $X'$ , by (a). From (1), for all  $x$  and  $y$ , either  $f(x) R f(y)$  or  $f(y) R f(x)$ , so that by (2),  $x R y$  or  $y R x$ . Similarly, from (a), II, and (1), for all  $x, y$ , and  $z$ ,  $f(x) R f(y)$  and  $f(y) R f(z)$  imply  $f(x) R f(z)$ ; from this, the transitivity of  $R$  follows by (2).

Lemma C. Let  $X$  be a space divided into two mutually exclusive subsets,  $X'$  and  $S^*$ . Suppose

- (a)  $Q$  is a partial ordering on  $X$ ;
- (b)  $f(x)$  is a function on  $S^*$  to  $X'$ ;
- (c)  $x Q f(x)$  and  $f(x) Q x$  for all  $x$  in  $S^*$ ;
- (d)  $R$  is compatible with  $Q$  on  $X'$ ;
- (e) if  $x$  belongs to  $S^*$  and  $y$  to  $X'$ , then  $x R y$  if and only if  $f(x) R y$ ;
- (f) if  $x$  belongs to  $X'$  and  $y$  to  $S^*$ , then  $x R y$  if and only if  $x R f(y)$ ;
- (g) if  $x$  and  $y$  both belong to  $S^*$ , then  $x R y$  if and only if  $f(x) R f(y)$ .

Then  $R$  is compatible with  $Q$  on  $X$ .

In this lemma, it is assumed that part of the space  $X'$  is ordered



and in a way compatible with some choices known in advance. The remaining elements each known to be indifferent to one of the elements of  $X'$ , and they are placed in the order in the same position as the element to which they are indifferent. Then it is asserted that these operations create a consistent ordering of the entire space.

Proof: By (d) and Definition 11(a),  $R$  is a weak ordering on  $X'$ . In conjunction with (b) and (e)-(g), this shows, by Lemma B, that

$R$  is a weak ordering on  $X$ . (1)

As in the proof of Lemma B, extend the definition of  $f(x)$  by setting  $f(x) = x$  for  $x$  in  $X'$ . Then, again, as in (1) and (2) under the proof of Lemma B,

for all  $x$  in  $X$ ,  $f(x)$  is in  $X'$ , (2)

for all  $x$  and  $y$ ,  $x R y$  if and only if  $f(x) R f(y)$ . (3)

From (a),  $x Q x$  for all  $x$ , and therefore,  $x Q f(x)$  and  $f(x) Q x$  for all  $x$  in  $X'$ ; in conjunction with (c), we have

for all  $x$ ,  $x Q f(x)$  and  $f(x) Q x$ . (4)

Suppose  $x Q y$ . By (4),  $f(x) Q x$  and  $y Q f(y)$ . By (a),  $Q$  is a partial ordering and therefore transitive. From  $f(x) Q x$  and  $x Q y$  follows  $f(x) Q y$ ; from this last and  $y Q f(y)$  follows  $f(x) Q f(y)$ , so that  $x Q y$  implies  $f(x) Q f(y)$ . Since by (4), also  $x Q f(x)$  and  $f(y) Q y$ ,  $f(x) Q f(y)$  implies  $x Q y$ , by a similar argument.

For all  $x$  and  $y$ ,  $x Q y$  if and only if  $f(x) Q f(y)$ . (5)

Suppose  $x Q y$ . By (5),  $f(x) Q f(y)$ . From (2), (d) and Definition 11(b),  $f(x) R f(y)$ , so that  $x R y$  by (3).

For all  $x$  and  $y$ ,  $x Q y$  implies  $x R y$ . (6)

Now suppose  $x Q y$  and not  $y Q x$ . By (5),  $f(x) Q f(y)$  and not  $f(y) Q f(x)$ . By (2), (d), and Definition 11(c),  $f(x) R f(y)$  and not

$f(y) R f(x)$ . By (3),  $x R y$  and not  $y R x$ .

For all  $x$  and  $y$ ,  $x Q y$  and not  $y Q x$  imply  $x R y$  and not  $y R x$ . (7)

By (1), (a), (6), and (7), Definition 11 is satisfied so that  $R$  is compatible with  $Q$ .

Lemma 11. Let  $X$  be a space of alternatives,  $Q$  a partial ordering on  $X$ , and  $S$  a subset of  $X$  having the following properties:

(a) for all  $x$  and  $y$  in  $S$ , not  $x Q y$ ;

(b) there exists at least one weak ordering  $T'$  which is compatible with  $Q$ ;

then for every weak ordering  $T''$  on  $S$ , there is a weak ordering  $R$  on the whole of  $X$  compatible with  $Q$  such that for  $x$  and  $y$  in  $S$ ,  $x R y$  if and only if  $x T'' y$ .<sup>4</sup>

This means the following: suppose, of all possible pairs of elements in  $X$ , the choices among some pairs is fixed in advance, and in a consistent way (so that if  $x$  is fixed in advance to be chosen over  $y$  and  $y$  fixed in advance to be chosen over  $z$ , then  $x$  is fixed in advance to be chosen over  $z$ ). Suppose, however, there is a set  $S$  of elements in  $X$  such that the choice between no pair of them is prescribed in advance. Suppose further there is some way of setting up an ordering relation in  $X$  which will be compatible with the choices prescribed in advance, i.e., will yield the same choice as that prescribed in advance whenever the choice was so prescribed. Then the lemma states that, given any ordering of the elements in  $S$ , there is a way of ordering all the elements of  $X$  which will be compatible both with the given ordering in  $S$  and with the choices made in advance. In other words, if we know that there is some ordering on  $X$ , and we know some of the choices implied by that ordering but the known choices do not give any direct information as to choices between

---

4. An error in the original formulation of this lemma was pointed out to the author by J. C. C. McKinsey, Project Rand.

elements in a subset  $S$  of  $X$ , then there is also no indirect information as to choices in  $S$ , i.e., the ordering on  $X$  is compatible with any ordering on  $S$ . This statement is intuitively obvious; nevertheless, a proof may be in order.

Proof: Let  $\underline{S}$ ,  $S^*$  be defined as follows:  $\underline{S}$  is the set of elements for which there exists an element  $y$  in  $S$  such that  $y \leq x$  and not  $x \leq y$ ;  $S^*$  is the set of elements  $x$  such that there is an element  $y$  in  $S$  for which  $x \leq y$  and  $y \leq x$ .

Suppose there is an element  $x$  in  $S$  which is also in either  $\underline{S}$  or  $S^*$ . In either case, there is an element  $y$  in  $S$  such that  $y \leq x$ ; but since both  $y$  and  $x$  in  $S$ , this contradicts (a) of the hypothesis. Therefore,  $S$  has no elements in common with either  $\underline{S}$  or  $S^*$ .

$$S \cap \underline{S} = 0, S \cap S^* = 0. \tag{1}$$

Here,  $S \cap \underline{S}$ , for example, stands for the set of elements common to both  $S$  and  $\underline{S}$ , and  $0$  denotes the set consisting of no elements. The first part of (1) states that  $S$  and  $\underline{S}$  have no elements in common. Suppose there is an element  $x$  in both  $\underline{S}$  and  $S^*$ . Since  $x$  is in  $\underline{S}$ , there is an element  $y$  of  $S$  such that  $y \leq x$ ; since  $x$  is in  $S^*$ , there is an element  $z$  of  $S$  such that  $x \leq z$ . From these two statements, it follows that  $y \leq z$ , since  $\leq$  is a partial ordering by hypothesis and therefore transitive. But since  $y$  and  $z$  are both in  $S$ , it is impossible that  $y \leq z$ , by (a); hence, there can be no elements common to  $\underline{S}$  and  $S^*$ .

$$\underline{S} \cap S^* = 0. \tag{2}$$

Finally, let  $\bar{S}$  consist of all elements not in  $S$ ,  $\underline{S}$ , or  $S^*$ . Clearly,

$$S \cap \bar{S} = 0, \underline{S} \cap \bar{S} = 0, S^* \cap \bar{S} = 0. \tag{3}$$

$$S \cup \underline{S} \cup S^* \cup \bar{S} = X. \tag{4}$$

Here,  $S \cup \underline{S} \cup S^* \cup \bar{S}$  denotes the set of elements in at least one of the named sets; (4) says that every element of  $X$  is in at least one of them.

Finally, let

$$X' = S \cup \underline{S} \cup \bar{S} \tag{5}$$

so that

$$X = X' \cup S^*, \tag{6}$$

while, from (5) and (1-3),

$$X' \cap S^* = \emptyset. \tag{7}$$

We will now define a relation R, first only for elements of X'.

For x and y in X', let x R y hold if and only if one of the following holds:

$$x \text{ and } y \text{ in } \bar{S} \text{ and } x T' y; \tag{8}$$

$$x \text{ in } \bar{S} \text{ and } y \text{ in } S \text{ or } \underline{S}; \tag{9}$$

$$x \text{ and } y \text{ in } S \text{ and } x T'' y; \tag{10}$$

$$x \text{ in } S \text{ and } y \text{ in } \underline{S}; \tag{11}$$

$$x \text{ and } y \text{ in } \underline{S} \text{ and } x T' y. \tag{12}$$

Here, T' is the ordering on X, whose existence is guaranteed by (b) of the hypothesis, and T'' is any ordering on S (satisfying axioms I and II, of course). Since T' and T'' are orderings in their respective domains, it follows from (8), (10), and (12) that

$$R \text{ is a weak ordering within each of the three sets } S, \underline{S}, \text{ and } \bar{S}. \tag{13}$$

From (1), (3), and (5), X' is divided into the three mutually exclusive subsets  $\underline{S}$ , S, and  $\bar{S}$ . By (13), (9), and (11), hypotheses (a) and (b) of Lemma A are satisfied, so that

$$R \text{ is a weak ordering on } X'. \tag{14}$$

We now wish to show that R is compatible with Q on X'. First, suppose x Q y. Consider various possible combinations of x and y.

If x and y belong to  $\bar{S}$ , then x Q y implies x T' y by (b) and Definition 11(b), and therefore x R y, by (8).

If  $x$  belongs to  $\bar{S}$  and  $y$  does not, then  $x R y$  by (9).

If  $x$  belongs to  $S$ , then  $x Q y$  implies that  $y$  belongs to  $\underline{S}$  or  $S^*$  by definition of those sets, but since we are only considering elements in  $X'$ ,  $y$  must belong to  $\underline{S}$ , by (7). Then  $x R y$  by (11).

If  $x$  belongs to  $\underline{S}$ , then, by definition, there is an element  $z$  in  $S$  such that  $z Q x$ . From  $z Q x$  and  $x Q y$  follows  $z Q y$ , since  $Q$  is a partial ordering by hypothesis and therefore transitive. As  $z$  is in  $S$ ,  $y$  is in  $\underline{S}$  or  $S^*$  by definition; but since we are only considering elements in  $X'$ ,  $y$  is in  $\underline{S}$  by (7). By (b) and Definition 11(b),  $x Q y$  implies  $x T' y$ ; since  $x$  and  $y$  are both in  $\underline{S}$ ,  $x R y$  by (12).

For all  $x$  and  $y$  in  $X'$ ,  $x Q y$  implies  $x R y$ . (15)

Now suppose both  $x Q y$  and not  $y Q x$ . Consider again various possible combinations of  $x$  and  $y$ .

If  $x$  and  $y$  both belong to  $\bar{S}$ , then  $x Q y$  and not  $y Q x$  imply  $x T' y$  and not  $y T' x$  by (b) and Definition 11(c), and therefore  $x R y$  and not  $y R x$ , by (8).

If  $x$  belongs to  $\bar{S}$  and  $y$  does not, then  $x R y$  by (9) while not  $y R x$  by (8-12).

If  $x$  belongs to  $S$ , as before,  $y$  belongs to  $\underline{S}$ , so that  $x R y$  and not  $y R x$  by (8-12).

If  $x$  belongs to  $\underline{S}$ , then, as before,  $x Q y$  implies that  $y$  is in  $\underline{S}$ . By <sup>(b) and</sup> Definition 11(c),  $x Q y$  and not  $y Q x$  imply  $x T' y$  and not  $y T' x$ , and, since  $x$  and  $y$  are both in  $\underline{S}$ ,  $x R y$  and not  $y R x$ , by (12).

For all  $x$  and  $y$  in  $X'$ ,  $x Q y$  and not  $y Q x$  imply  $x R y$  and not  $y R x$ . (16)

By (14-16) and hypothesis, Definition 11 is satisfied.

$R$  is compatible with  $Q$  on  $X'$ . (17)

We will now extend  $R$  to the whole of  $X$ . For  $x$  in  $S^*$ , choose a corresponding element  $f(x)$  in  $S$  such that

$$x \ Q \ f(x) \ \text{and} \ f(x) \ \downarrow \ x; \tag{18}$$

this can always be done by the definition of  $S^*$ . Since  $f(x)$  is in  $S$ , it follows from (5) that

$$f(x) \ \text{is a function on } S^* \ \text{to } X'. \tag{19}$$

Now define  $x \ R \ y$ , where  $x$  or  $y$  or both is in  $S^*$ , to hold if and only if one of the following conditions holds:

$$x \ \text{in } S^*, \ y \ \text{in } X', \ \text{and} \ f(x) \ R \ y; \tag{20}$$

$$x \ \text{in } X', \ y \ \text{in } S^*, \ \text{and} \ x \ R \ f(y); \tag{21}$$

$$x \ \text{and} \ y \ \text{in } S^*, \ \text{and} \ f(x) \ R \ f(y). \tag{22}$$

By (6-7), the hypothesis on  $Q$ , (19), (18), (17), and (20-22), the various hypotheses of Lemma C are fulfilled. Hence,

$$R \ \text{is compatible with } Q \ \text{everywhere on } X. \tag{23}$$

For  $x$  and  $y$  in  $S$  (10) guarantees that  $x \ R \ y$  if and only if  $x \ T'' \ y$ . In conjunction with (23), this proves the lemma.<sup>5</sup>

The motivation of the above proof can be stated rather simply. We know that the space  $X$  can be ordered in a way compatible with a particular partial ordering  $Q$ . We wish to order the space in a way compatible both with  $Q$  and with any order whatever on  $S$ , where  $Q$  tells us nothing about the relative position of any two elements in  $S$ . Our procedure is to first separate out all the elements which must necessarily, according to our knowledge of  $Q$ , follow at least one of the elements of  $S$ , and

---

5. In the hypothesis of Lemma 11, just proved, there appeared the assumption that there exists at least one ordering of  $X$ , namely  $T'$ , which is compatible with the preassigned partial ordering  $Q$ . This assumption is superfluous, since, as D. Blackwell, Howard University, has proved to the author, a partial ordering  $Q$  can always be extended to a complete ordering. The proof involves the rather delicate mathematical tool of transfinite induction. In the present work, since the space involved is the totality of alternatives which are the objects of choice by individuals and by society, and since, in the applications, the partial ordering which is discussed is one which is assumed to begin with to be compatible with an actual individual preference scale, there is no harm in making the superfluous assumption in order to simplify the proof.

also the elements known to be indifferent to at least one element of  $S$ . The first type of elements are made to be inferior to all the elements of  $S$ ; within themselves, the inferior elements are ordered by the original ordering which is known to be consistent with  $Q$ . All elements not in  $S$  and not known to be inferior or indifferent to any element of  $S$  are made to be superior to all elements of  $S$ ; within themselves, these elements which are superior to the elements of  $S$  are ordered by the original ordering. Within themselves, the elements of  $S$  can be ordered in any way. Finally, the elements indifferent to elements of  $S$  are given the same place in the order as the latter. It is intuitively clear that this method of reordering the elements of  $X$  so as to be compatible with the partial ordering  $Q$  and with any given ordering of  $S$  will in fact accomplish its purpose and be a true ordering relation. However, it does not seem possible to give a rigorous proof less detailed than the above.

### 3. Proof of Theorem 3.

As noted on page 16 of Possibility, we may argue that we know on a priori grounds something about the possible individual preference scales. In that case, we would not ask that our social welfare function be defined for all logically possible individual ordering relations but only those consistent with our a priori knowledge. One example is the assumption very frequently made in welfare economics that each individual evaluates social alternatives only in terms of the commodities he receives under each and also that in comparing the commodity bundles he receives under two social alternatives, he will prefer one bundle over another if the first yields him at least as much of every commodity as the second and more of at least one commodity. These assumptions constitute in effect a preassigned partial ordering of the social alternatives

with which the actual order must be compatible; an individual is indifferent as between any two social alternatives which yield him the same commodity bundle and he will prefer one social alternative over another if the first yields him at least as much of every commodity as the second and more of at least one commodity. The relation so defined is clearly transitive; but it is not a total ordering, since it does not tell us how the individual compares two social alternatives, one of which yields him more of one commodity than the second while the second yields him more of a second commodity than the first. Also note that the partial ordering of the social alternatives involved is different for different individuals, since each individual ranks the alternatives according to what he gets out of them, and an alternative which yields more of each commodity to one individual than a second alternative may yield less of each commodity than the second alternative to a second individual.

Formally, we may say that there are  $n$  partial orderings  $Q_1, \dots, Q_n$ , and it is known in advance that the individual ordering relations  $R_1, \dots, R_n$  are compatible with the partial orderings  $Q_1, \dots, Q_n$ , respectively. I.e.,

$$\text{For each } i, x Q_i y \text{ implies } x R_i y. \quad (1)$$

$$\text{For each } i, x Q_i y \text{ and not } y Q_i x \text{ imply } x R_i y \text{ and not } y R_i x. \quad (2)$$

If, before constructing our social welfare function, we know that (1) and (2) hold with known  $Q_1, \dots, Q_n$ , we might feel it superfluous to require our social welfare function to be so defined as to yield a social ordering for a set of individual orderings  $R_1, \dots, R_n$  which do not satisfy (1) and (2). Hence, Condition 1 (Possibility, page 9) can be rephrased as follows:



Condition 1'. For every set of individual orderings  $R_1, \dots, R_n$  compatible with  $Q_1, \dots, Q_n$  respectively, the corresponding social ordering  $R$  is a weak ordering relation.

The proof of Theorem 2 made use of the requirement that the social welfare function was defined for any constellation of individual tastes which could be represented by individual ordering relations, since at several points in the proof new sets of individual ordering relations were constructed out of given ones and the corresponding social ordering relations discussed; it was assumed that these new social ordering relations existed if it could only be shown that the new individual relations were in fact ordering relations. Therefore, it might seem possible that replacing Condition 1 by Condition 1' upset the conclusion of Theorem 2.

This, however, is not so in the usual context of economic policy discussions. Closer inspection of the proof of Theorem 2 shows that all that is really required is that for some set of three alternatives the range of individual ordering relations for which the social welfare function is defined is sufficiently wide so that any individual can order those particular three alternatives in any way. To put it formally, let an admissible set of individual ordering relations be a set for which the social welfare function defines a corresponding social ordering; if Condition 1' is accepted, an admissible set of individual orderings is one such that  $R_i$  is compatible with  $Q_i$  for all  $i$ . Let  $S$  be a set of three alternatives. Theorem 2 remains valid if we can show that, for any given  $n$  orderings  $T_1^n, \dots, T_n^n$  of  $S$ , there is an admissible set of individual orderings  $R_1, \dots, R_n$ , such that for each  $i$ ,  $R_i$  coincides with  $T_i^n$  on  $S$ .

In the present case, let  $S$  be a set of three alternatives about the ordering in which none of the known partial orderings tell us anything; i.e., suppose that for each  $i$  and each pair  $x, y$  of elements in  $S$ , not  $x Q_i y$ . Under the individualistic hypotheses discussed above, an example of such a set would be a set of three alternative distributions of fixed stocks of commodities such that for each pair of distributions, one gave to any given individual more of one commodity while the second distribution gave more of another commodity.

To such a set  $S$ , we may apply Lemma 11. Let  $T_1^w, \dots, T_n^w$  be any set whatever of  $n$  orderings of  $S$ . Since the totality of all alternatives is assumed to be capable of being ordered in a manner compatible with any one of the partial orderings (by Condition 1', it is assumed that the known partial orderings  $Q_i$  are compatible with the actual total orderings), let  $T_i^w$  be any particular ordering of all the alternatives, both in  $S$  and out, which is compatible with the partial ordering  $Q_i$ , for each  $i$ . All the hypotheses of Lemma 11 are satisfied, if we add the subscript  $i$  to all relations. Hence for each  $i$ , there is a total ordering  $R_i$  which is compatible with both the preassigned partial ordering  $Q_i$  and with the arbitrary ordering  $T_i^w$  of the set  $S$ . This set of individual orderings is admissible under Condition 1". By the last remark in the second preceding paragraph, the conclusion of Theorem 2 remains valid.

Theorem 3. If  $Q_1, \dots, Q_n$  are a set of weak partial orderings for which there exists a set  $S$  containing at least three alternatives such that for all  $i$  and all  $x$  and  $y$  in  $S$  not  $x Q_i y$ , then every social welfare function satisfying Condition 1' and Conditions 2 and 3 is either conventional or dictatorial.

Although the motivation for developing this theorem was the analysis of the individualistic case, the conclusion is applicable to any attempt to restrict the range of individual orderings by requiring them to be consistent with any preassigned partial orderings. Thus, if we modified the previous individualistic ordering by saying that between two alternatives which yielded the same commodity bundle to a given individual, the individual was not indifferent but chose on the basis of some measure of income equality, while however, he still preferred any alternative which gave him at least as much of all commodities and more of at least one, the above negative conclusions as to the possibility of a social welfare function would still be applicable, since we could still form the set  $S$  which satisfies the hypotheses of Theorem 3.

ERRATA OF COWLES COMMISSION DISCUSSION PAPER: ECONOMICS: 245

- p. 4 In equation (7), " $y' \neq x$ " should be " $y' = x$ ."
- p. 6 Line 15; same as preceding correction.
- p. 7 Line 4; replace "(7)" by "(9)".
- p. 7 Line 7; replace "(8) and (9)" by "(10) and (11)".
- p. 9 In equation (21), add subscript 2 to " $v$ ".
- p. 12 In n. 3, put subscript 1 on  $R$  and  $R'$  wherever they occur.
- p. 12 Line 12; after " $x' R'_1 y'$ ", insert, "or  $y'$ ."
- p. 22 Line 4; replace " $X$ " by " $S$ ."