The Impossibility of a Paretian Liberal

The Impossibility of a Paretian Liberal

Amartya Sen

Delhi School of Economics and Harvard University

I. Introduction

The purpose of this paper is to present an impossibility result that seems to have some disturbing consequences for principles of social choice. A common objection to the method of majority decision is that it is illiberal. The argument takes the following form: Given other things in the society, if you prefer to have pink walls rather than white, then society should permit you to have this, even if a majority of the community would like to see your walls white. Similarly, whether you should sleep on your back or on your belly is a matter in which the society should permit you absolute freedom, even if a majority of the community is nosey enough to feel that you must sleep on your back. We formalize this concept of individual liberty in an extremely weak form and examine its consequences.

II. The Theorem

Let R_i be the ordering of the ith individual over the set X of all possible social states, each social state being a complete description of society including every individual's position in it. There are n individuals. Let R be the social preference relation that is to be determined.

DEFINITION 1: *A collective choice rule* is a functional relationship that specifies one and only one social preference relation R for any set of n individual orderings (one ordering for each individual).

A special case of a collective choice rule is one that Arrow (1951) calls a social welfare function, namely, a rule such that R must be an ordering.

DEFINITION 2: *A social welfare function* is a collective choice rule, the range of which is restricted to orderings.

A weaker requirement is that each R should generate a "choice function," that is, in every subset of alternatives there must be a "best"

For comments and criticisms I am grateful to Kenneth Arrow, Peter Diamond, Milton Friedman, Tapas Majumdar, Stephen Marglin, and Thomas Schelling.

alternative, or, in other words, there must be some (but not necessarily only one) alternative that is at least as good as all the other alternatives in that subset. This may be called a "social decision function."

DEFINITION 3: *A social decision function* is a collective choice rule, the range of which is restricted to social preference relations that generate a choice function.

It was shown in Sen (1969) that the conditions that were proven to be inconsistent by Arrow (1951, 1963) in his justly famous "impossibility theorem" in the context of a social welfare function are in fact perfectly consistent if imposed on a social decision function. The impossibility theorem to be presented here holds, however, for social decision functions as well.

Arrow's condition of collective rationality (Condition 1′) can be seen to be merely a requirement that the domain of the collective choice rule should not be arbitrarily restricted.

CONDITION U (Unrestricted Domain): Every logically possible set of individual orderings is included in the domain of the collective choice rule.

Arrow used a weak version of the Pareto principle.

CONDITION P: If every individual prefers any alternative x to another alternative y, then society must prefer x to y.

Finally, we introduce the condition of individual liberty in a very weak form.

CONDITION L (Liberalism): For each individual i, there is at least one pair of alternatives, say (x, y), such that if this individual prefers x to y, then society should prefer x to y, and if this individual prefers y to x, then society should prefer y to x.[1]

The intention is to permit each individual the freedom to determine at least one social choice, for example, having his own walls pink rather than white, other things remaining the same for him and the rest of the society.[2]

The following impossibility theorem holds.

THEOREM I: There is no social decision function that can simultaneously satisfy Conditions U, P, and L.

[1] The term "liberalism" is elusive and is open to alternative interpretations. Some uses of the term may not embrace the condition defined here, while many uses will. I do not wish to engage in a debate on the right use of the term. What is relevant is that Condition L represents a value involving individual liberty that many people would subscribe to. Whether such people are best described as liberals is a question that is not crucial to the point of this paper.

[2] Even this informal statement, which sounds mild, is much more demanding than Condition L. If the individual's preference over a personal choice (like choosing the color of his wall) is to be accepted by the society, other things remaining the same, then this gives the individual rights not only over one pair, which is all that is required by Condition L, but over many pairs (possibly an infinite number of pairs) varying with the "other things." If it is socially all right for me to have my walls either pink or white as I like in a social state where you smoke cigars, it should be socially all right for me to do the same where you indulge yourself in ways other than smoking cigars. Even this is not required by Condition L, which seems to demand very little.

In fact, we can weaken the condition of liberalism further. Such freedom may not be given to all, but to a proper subset of individuals. However, to make sense the subset must have more than one member, since if it includes only one then we might have a dictatorship. Hence, we demand such freedom for at least two individuals.

CONDITION L* (Minimal Liberalism): There are at least two individuals such that for each of them there is at least one pair of alternatives over which he is decisive, that is, there is a pair of x, y, such that if he prefers x (respectively y) to y (respectively x), then society should prefer x (respectively y) to y (respectively x).

The following theorem is stronger than Theorem I and subsumes it.

THEOREM II: There is no social decision function that can simultaneously satisfy Conditions U, P, and L^*.

PROOF: Let the two individuals referred to in Condition L^* be 1 and 2, respectively, and the two pairs of alternatives referred to be (x, y) and (z, w), respectively. If (x, y) and (z, w) are the same pair of alternatives, then there is a contradiction. They have, therefore, at most one alternative in common, say $x = z$. Assume now that person 1 prefers x to y, and person 2 prefers w to $z(=x)$. And let everyone in the community including 1 and 2 prefer y to w. There is in this no inconsistency for anyone, not even for 1 and 2, and their respective orderings are: 1 prefers x to y and y to w, while 2 prefers y to w and w to x. By Condition U this should be in the domain of the social decision mechanism. But by Condition L^*, x must be preferred to y, and w must be preferred to x $(=z)$, while by the Pareto principle, y must be preferred to w. Thus, there is no best element in the set $(x=z, y, w)$ in terms of social preference, and every alternative is worse than some other. A choice function for the society does not therefore exist.

Next, let x, y, z, and w, be all distinct. Let 1 prefer x to y, and 2 prefer z to w. And let everyone in the community including 1 and 2 prefer w to x and y to z. There is no contradiction for 1 or 2, for 1 simply prefers w to x, x to y, and y to z, while 2 prefers y to z, z to w, and w to x. By Condition U this configuration of individual preferences must yield a social choice function. But by Condition L^* society should prefer x to y and z to w, while by the Pareto principle society must prefer w to x, and y to z. This means that there is no best alternative for this set, and a choice function does not exist for any set that includes these four alternatives. Thus, there is no social decision function satisfying Conditions U, P, and L^*, and the proof is complete.[3]

[3] We can strengthen this theorem further by weakening Condition L^* by demanding only that 1 be decisive for x against y, but not vice versa, and 2 be decisive for z against w, but not vice versa, and require that $x \neq z$, and $y \neq w$. This condition, too, can be shown to be inconsistent with Condition U and P, but the logical gain involved in this extension does not, alas, seem to be associated with any significant increase of relevance that I can think of.

III. An Example

We give now a simple example of the type of impossibility that is involved in Theorem II by taking a special case of two individuals and three alternatives. There is one copy of a certain book, say *Lady Chatterly's Lover*, which is viewed differently by 1 and 2. The three alternatives are: that individual 1 reads it (x), that individual 2 reads it (y), and that no one reads it (z). Person 1, who is a prude, prefers most that no one reads it, but given the choice between either of the two reading it, he would prefer that he read it himself rather than exposing gullible Mr. 2 to the influences of Lawrence. (Prudes, I am told, tend to prefer to be censors rather than being censored.) In decreasing order of preference, his ranking is z, x, y. Person 2, however, prefers that either of them should read it rather than neither. Furthermore, he takes delight in the thought that prudish Mr. 1 may have to read Lawrence, and his first preference is that person 1 should read it, next best that he himself should read it, and worst that neither should. His ranking is, therefore, x, y, z.

Now if the choice is precisely between the pair (x, z), i.e., between person 1 reading the book and no one reading it, someone with liberal values may argue that it is person 1's preference that should count; since the prude would not like to read it, he should not be forced to. Thus, the society should prefer z to x. Similarly, in the choice exactly between person 2 reading the book (y) and no one reading it (z), liberal values require that person 2's preference should be decisive, and since he is clearly anxious to read the book he should be permitted to do this. Hence y should be judged socially better than z. Thus, in terms of liberal values it is better that no one reads it rather than person 1 being forced to read it, and it is still better that person 2 is permitted to read the book rather than no one reading it. That is, the society should prefer y to z, and z to x. This discourse could end happily with the book being handed over to person 2 but for the fact that it is a Pareto inferior alternative, being worse than person 1 reading it, in the view of both persons, i.e., x is Pareto superior to y.

Every solution that we can think of is bettered by some other solution, given the Pareto principle and the principle of liberalism, and we seem to have an inconsistency of choice. This is an example of the type of problem that is involved in Theorems I and II.

IV. Relevance

The dilemma posed here may appear to be somewhat disturbing. It is, of course, not necessarily disturbing for every conceivable society, since the conflict arises with only particular configurations of individual preferences. The ultimate guarantee for individual liberty may rest not on rules for social choice but on developing individual values that respect each other's

personal choices. The conflict posed here is concerned with societies where such a condition does not hold and where pairwise choice based on liberal values may conflict with those based on the Pareto principle. Like Arrow's "General Possibility Theorem," here also the Condition of Unrestricted Domain is used.

However, unlike in the theorem of Arrow, we have not required transitivity of social preference. We have required neither transitivity of strict preference, nor transitivity of indifference, but merely the existence of a best alternative in each choice situation.[4] Suppose society prefers x to y, and y to z, and is indifferent between z and x. Arrow would rule this out, since there is an intransitivity; but we do not, for here alternative x is "best" in the sense of being at least as good as both the other alternatives. Our requirements are, in this respect, very mild, and we still have an impossibility.

Second, we have not imposed Arrow's much debated condition of "the independence of irrelevant alternatives."[5] Many people find the relaxation of this condition to be an appealing way of escaping the Arrow dilemma. This way out is not open here, for the theorem holds without imposing this condition.

The Pareto principle is used here in a very weak version, as in Arrow. We do not necessarily require that if someone prefers x to y and everyone regards x to be at least as good as y, then x is socially better. We permit the possibility of having collective choice rules that will violate this provided everyone strictly preferring x to y must make x socially better than y.

[4] It may appear that one way of solving the dilemma is to dispense with the social choice function based on a binary relation, that is, to relax not merely transitivity but also *acyclicity*. A choice function that need not correspond to any binary relation has undoubtedly a wider scope. But then Condition P and Condition L would have to be redefined, for example, (1) x should not be chosen when y is available, if everyone prefers y to x, and (2) for each individual there is a pair (x_i, y_i) such that if he prefers x_i (respectively y_i) to y_i (respectively x_i), then y_i (repectively x_i) should not be chosen if x_i (respectively y_i) is available. Thus redefined, the choice set for the set of alternatives may be rendered empty even without bringing in acyclicity, and the contradiction will reappear. This and other possible "ways out" are discussed more fully in my forthcoming book (Sen, in press, chap. 6).

[5] Using the condition of the independence of irrelevant alternatives, A. Gibbard, in an unpublished paper, has recently proved the following important theorem: Any social decision function that must generate social preferences that are all transitive in the strict relation (quasi-transitive) and which must satisfy Conditions U, P, nondictatorship, and the independence of irrelevant alternatives, must be an oligarchy in the sense that there is a unique group of individuals each of whom, by preferring x to y, can make the society regard x to be at least as good as y, and by all preferring x to y can make the society prefer x to y, irrespective of the preferrences of those who are not in this group. Gibbard's Theorem is disturbing, for the conditions look appealing but the resultant oligarchy seems revolting, and it is a major extension of the problem posed by Arrow (1951, 1963). Gibbard argues against the simultaneous insistence on a binary relation of social preference generating a choice function and on the condition of the independence of irrelevant alternatives. We have not imposed the latter.

Nevertheless it turns out that a principle reflecting liberal values even in a very mild form cannot possibly be combined with the weak Pareto principle, given an unrestricted domain. If we do believe in these other conditions, then the society cannot permit even minimal liberalizm. Society cannot then let more than one individual be free to read what they like, sleep the way they prefer, dress as they care to, etc., *irrespective* of the preferences of others in the community.

What is the moral? It is that in a very basic sense liberal values conflict with the Pareto principle. If someone takes the Pareto principle seriously, as economists seem to do, then he has to face problems of consistency in cherishing liberal values, even very mild ones.[6] Or, to look at it in another way, if someone does have certain liberal values, then he may have to eschew his adherence to Pareto optimality. While the Pareto criterion has been thought to be an expression of individual liberty, it appears that in choices involving more than two alternatives it can have consequences that are, in fact, deeply illiberal.

References

Arrow, K. J. *Individual Values and Social Choice*. New York: Wiley, 1951; 2d ed., 1963.

Sen, A. K. "Quasi-transitivity, Rational Choice and Collective Decisions." Discussion paper no. 45, Harvard Institute of Economic Research 1968. *Rev. Econ. Studies* 36, no. 3 (July 1969):381–93.

———. *Collective Choice and Social Welfare*. San Francisco: Holden-Day; and Edinburgh: Oliver & Boyd, in press.

[6] The difficulties of *achieving* Pareto optimality in the presence of externalities are well known. What is at issue here is the *acceptability* of Pareto optimality as an objective in the context of liberal values, given certain types of externalities.