



“Summary of article by Peter J. Hammond: Social Choice: The Science of the Impossible?” in Frontier Issues in Economic Thought, Volume 3: Human Well-Being and Economic Goals. Island Press: Washington DC, 1997. pp. 107-109

Social Science Library: Frontier Thinking in Sustainable Development and Human Well-being

“Summary of article by Peter J. Hammond: Social Choice: The Science of the Impossible?”

The theory of social choice hardly existed before Kenneth Arrow’s pioneering work, particularly his “General Possibility Theorem.” This article describes the background and structure of Arrow’s theorem, and reviews some of the attempts that have been made to escape from Arrow’s discouraging conclusion.

THE BACKGROUND TO ARROW’S THEOREM

The fundamental concept of Arrow’s social choice theory is his social welfare function (SWF), a rule that determines a social preference ordering for any given profile of individual preference orderings. Arrow later suggested that it could be called a constitution, in order to distinguish it from Bergson’s SWF. The Bergson SWF is a real-valued function, defined on social states for only one profile of individual preferences; in contrast, the Arrow SWF is defined over all possible profiles of preferences, and its values are social preference orderings.

The simple voting procedure of majority rule appears to give rise to a natural SWF, but Arrow rediscovered the Condorcet paradox, named for its eighteenth-century author. Voting can give rise to intransitive or cyclical preferences: if individual 1's preferences, in order, are a,b,c, 2's preferences are b,c,a, and 3's preferences are c,a,b, then a majority prefers a to b, b to c, and c to a. On the other hand, a logically consistent but less attractive SWF can be created by selecting a dictator and following his decision on every issue.

Arrow’s work can be seen as asking whether there is a middle ground between these extremes. He included the requirement that the SWF must satisfy the Pareto criterion, a weak but familiar condition: in essence, the SWF must not overrule any unopposed vote.

INDEPENDENCE OF IRRELEVANT ALTERNATIVES

Is there a non-dictatorial SWF satisfying the Pareto criterion and generating a consistent social preference ordering from any profile of individual preferences? Arrow added one more condition to narrow the search, the most contentious aspect of his theorem: the independence of irrelevant alternatives. That is, the SWF’s ranking of a versus b must depend solely on individual preferences between a and b.

The motivation for this condition can be seen by examining a SWF that violates it. The Borda rule, named for another eighteenth-century writer, begins by assuming that each individual ranks

every conceivable social state (assumed to be finite in number), assigning 1 to the least preferred, 2 to the second-worst, and so on. All the individual rankings are then added to obtain a function that represents the social preference ordering. This is a consistent, non-dictatorial, Pareto-compatible SWF.

The Borda rule, however, is hopelessly unwieldy. Applying it to even a moderately complex state election can require each voter to evaluate more than a trillion possible combinations of outcomes. If the Borda rule is restricted to a smaller set of feasible choices, another rule is needed to decide the (often controversial) question of which alternatives to include. Arrow's requirement that the SWF be independent of irrelevant alternatives cuts through these and other conceptual muddles. Social preferences on any question must depend solely on individual preferences on the same question.

IS DICTATORSHIP INESCAPABLE?

Arrow's theorem proves that the only SWF defined over all logically possible patterns of individual preferences, that satisfies both the Pareto criterion and the independence of irrelevant alternatives, is one in which a single individual's preferences always prevail -- in short, a dictatorship.

Numerous escapes from this conclusion have been proposed. One of the least satisfactory solutions is to weaken the requirement that the SWF yields a consistent preference ordering over all social outcomes. This approach violates fundamental assumptions of collective rationality, and has not been widely accepted.

Another attempted escape involves abandoning the Pareto criterion. Yet this rejects the whole approach to social choice as based on individual preferences. If someone's preference for a over b is not opposed by any individual, how can the SWF fail to select a over b? Sen's 'liberal paradox' [see summary of Pattanaik article in this section] raises questions about the scope of the Pareto principle; but that paradox appears to rest on the inappropriate inclusion of envy, meddlesomeness, or even some forms of altruism and benevolence, i.e. other-directed preferences that are not relevant to an individual's personal welfare.

The only remaining option that maintains both collective rationality and the Pareto principle, while avoiding a dictatorship, is to allow dependence on irrelevant alternatives.

INTERPERSONAL COMPARISONS

The assumption of independence of irrelevant alternatives combines two important features. One is independence, requiring that only properties of the alternatives under consideration can count in the decision. The other feature is "ordinal non-comparability of utilities" -- that is, only the ordinal, non-comparable preferences of individuals can be considered in making the social decision. The latter feature is automatically satisfied by any Arrow SWF, since it relies solely on individual preference orderings.

Sen has argued that it is the exclusion of interpersonal comparisons, in particular, that leads to dictatorship in the proof of Arrow's theorem. To illustrate this point, Sen has introduced the notion of a "social welfare functional" (SWFL), a broader category than Arrow's SWF. Sen's SWFL derives a social preference ordering from individual utility functions, which may be ordinal or cardinal, interpersonally comparable or not. Any Arrow SWF is a Sen SWFL, but the converse is not true. The utilitarian sum of individual cardinal utility functions [see summary of Harsanyi article in this section] is a Sen SWFL, but not an Arrow SWF.

Sen has also proposed a modification of the assumption of independence of irrelevant alternatives. The corresponding condition for SWFLs may be called the "independence of irrelevant utilities": the social ordering of any set of outcomes depends only upon individuals' utilities obtained from those outcomes. Then there are many non-dictatorial SWFLs that satisfy the Pareto criterion and the independence of irrelevant utilities; one example is a maximin SWFL which is related to Rawls' concept of justice.

WHICH ALTERNATIVES ARE RELEVANT?

An additional problem arises in this escape from Arrow's theorem. When the interpersonal comparisons used in the Sen SWFL are placed on a rigorous theoretical foundation, the social decision process appears to include irrelevant alternatives. Application of the Rawls maximin rule to a particular social decision, for example, requires identification of the least well-off member of society; that identification process uses information that is irrelevant to (or independent of) the decision that is being made. In general, if social choice is to depend solely on individual preferences concerning the outcomes under consideration, then personal characteristics must be ignored. A potential solution to this problem is to broaden the definition of relevant alternatives to include personal characteristics as well as social states; the range of relevant alternatives could include those in which individuals exchange places in society and personal characteristics.

It remains to be seen how social choice theory can be reformulated to preserve as much science as one can while escaping the impossibility of non-dictatorship. (129)