May, 2011

Collective Choice

Justin Schwartz

COLLECTIVE CHOICE

Justin Schwartz

The John Marshall Law School, Chicago, IL USA

jkschw1@yahoo.com; 7schwartzj@jmls.edu

COLLECTIVE CHOICE

Collective choice involves the aggregation of individual preferences by some method such as voting to produce a social outcome. Analysis shows that it involves surprisingly intransigent paradoxes that seem to challenge the possibility of fair democratic decision-making. Among the most important of these is the Arrow Impossibility Theorem, the inspiration for the now vast field of social choice theory. That the theorem is valid cannot be disputed. The main objective of this informal, nontechnical discussion is to explicate its content and evaluate its significance. The less-than happy conclusion is that it states a deep problem for the theory of democratic politics that sixty years of intensive discussion has failed to dissipate.

The theorem was discovered by Kenneth J. Arrow, a Nobel Memorial Prize-winning economist. He showed that four minimal and obvious constraints seemingly necessary for a fair, democratic, and rational outcome cannot be satisfied by any method for aggregating preferences, such as majority rule, that generates a single ranking of preferences with a highest-ranked preference as the collective or social choice. The constraints can be stated as: (*D*), **Nondictatorship:** no individual's preferences unilaterally determine the outcome; (*P*) **the Weak Pareto Principle:** if any outcome is unanimously preferred by all to another outcome, the more preferred outcome must be higher ranked; (*I*) **the Indifference of Irrelevant Alternatives**: the social ranking of any set of preferences depends only on the actors' rankings of those preferences or "preference profiles," as they are termed in the literature. Finally, there is (*U*) **Unrestricted Domain**: the method will give a unique social choice among any logically coherent set of actors' preference profiles, however large, given that no choices or rankings are excluded from the start. Moreover, the actors' preferences and aggregate social outcomes are (*C*) **Collectively Rational**: they respect transitivity, so that if choice A is preferred to B, and B to C, then A is preferred to C; and

the outcomes are also "connected," so for any pair of alternatives, one is either preferred or the ranking is indifferent between them. In Arrow's presentation, *C* is part of *U*.

The Impossibility Theorem states that no social choice method involving at least three choices and a finite number of at least three individuals can satisfy all four constraints. It may produce no clear winner or rational outcome, generating paradox. To make the point vivid, a social choice method that satisfies *P*, *I,* and *U* (including *C*) can violate *D*, that is, be a dictatorship.

This startling result has teeth for at least three reasons. First, these are, plausibly, minimum, weak, and uncontroversial requirements for making collective choices in a fair and democratic way, so it is disturbing that they are mutually incompatible. Second, the Theorem is robust. It holds whether the actors are individuals or collective entities such as nations or organizations, and regardless of the specific method of aggregation, *e.g.,* simple majority rule, super-majoritarianism, Australian balloting, *etc.* The result holds regardless of the content of the preferences. Egoism is not required; it infects aggregation of altruistic preferences. The preferences need not be consequentialist, nor based on a common scale of values. They may incorporate incommensurable values as long as they can be ranked. Third, the theorem suggests that there is no such thing as the "common good," the "will of the people" or even the judgment of the majority. The paradoxes hold for any aggregation of individual preferences with more than three actors, be it national, subnational , or global. The theorem therefore raises the specter that no method of aggregating preferences can be rational, fair, and democratic.

Simple majority rule, for example, can be shown to violate C by generating "Condorcet cycling," named for the 18th century French mathematician who anticipated Arrow by discovering this special case of the theorem. Say the U.N., dealing with a Failed State, can vote to Negotiate (*N*),

Sanction (*S*), or Occupy (*O*). Rationally, the outcome should be transitive: If the U.N. prefers *N* to *S* and *S* to *O,* it should prefer *N* to *O.* But the outcome might be that each alternative will win against the other by a majority of a different composition, so that there is no clear winner. Furthermore, applying majority rule to the individual orderings here leads to intransitivity or cycling: *N* is preferred to *S,* and *S* to *O,* but *O* wins over *N*. This is irrational as well as inconclusive. The theorem shows that this result is general given the premises.

Responses are of three broad sorts. One is to try to relax at least one of the premises. A second is to urge that the results are of limited import because they do not arise much as a matter of fact in real politics. The third is to argue that the theorem is misleading or irrelevant because Arrow's framework fails to capture what democratic decision-making is about and how it works.

Attempts to evade Arrovian problems by relaxing the requirements have focused on *I*, Independence of Irrelevant Alternatives, *U*, Unrestricted Domain as well as *C* Collective Rationality, technically part of *U*. The other two conditions face little if any challenge. Nondictatorship, *D*, is as constitutive of fairness and democracy as anything could be. The Weak Pareto Principle, *P* insists no more than that unanimity be honored. What choice could count as collective and depart from that is hard to conceive.

The least intuitive condition is *I*, that social choice among some set of preferences depend only on individuals' aggregated ranking of those preferences and on nothing else. If the irrelevant alternatives that go into another social choice are genuinely irrelevant to the one at hand, they should not affect the ranking among preferences under consideration. But *I* also imposes a less obvious condition, the rejection of interpersonal comparisons about *how much* individuals prefer a given alternative. If *p* is both of our second choices, it does not matter under *I* that for me *p* is

abhorrent while for you it is merely distasteful. The social ranking depends on the ranking that each individual gives the outcomes and not on the relative intensity of preferences about those outcomes across individuals. The outcome turns wholly on ordinal rather than cardinal factors.

This does not seem particularly minimal. Cardinal ranking requiring interpersonal comparison of this sort is necessary for classical utilitarianism. Such comparisons are fundamental to welfare and development economics -- life expectancy and GDP per capita, for example, are cardinal (comparative) measures of welfare, and interpersonal comparison underlies the concept of diminishing marginal utility, *e.g.*, that a marginal increase of wealth means less to the rich than the poor, crucial to all these theories. Taking a major ethical theory and important collective choice method as well as several fundamental approaches to economics off the table *ab initio* seems like a strong rather than a weak starting point.

Three reasons are widely offered for the ordinal, noncomparative approach. First, there is doubt whether interpersonal comparisons are possible or meaningful in the absence of any generally accepted tests for determining whether, *e.g.,* I care twice as much about an additional year of expected life or an extra dollar than you do. Some of this skepticism is based on a discredited empiricist view of scientific method as merely a way of testing factual hypotheses about observations. Much work has been done, notably by economist Amartya Sen, a Harvard Nobel Memorial Laureate, and others, to show that such comparisons or proxies for them are technically possible. Some major American philosophers and decision theorists like Allen Gibbard and Donald Davidson contend that such comparisons might or must be made on the basis of personal experience and self-evaluation of our own preferences. Neither, however, offers a metric or avoids contentious normative hypotheses. Even if the interpersonal information exists and such comparisons could be made in principle, how it could be practicably ascertained

in a reliable way over many individuals and alternatives is not known. And social choice situations may create strategic incentives to misrepresent actual preferences and their intensities to obtain preferred outcomes. So Arrow's *I* may indeed be the weak, minimal, and best approach that it is claimed to be. John Rawls' and Robert Nozick's theories of justice are designed to respect it.

More deeply, even if reliable information about comparative intensity were available, it is not clear that it should matter in social choice. If, for instance, I rank more expected income higher than a longer expected life, and you vice versa, it seems worse than irrelevant for social choice if I care a lot about the additional money and you only a bit about the additional time. Counting intensity of preference would allow passionate minorities to dominate decision-making without winning votes or changing minds in the many cases when a majority choice is possible but less intensely favored. Arguably what should count in a vote is which choice is ranked higher, regardless of how intensely individuals care about lower ranked choices. Interpersonal comparative information, if available, could certainly be used in ethical or economic analysis, but such considerations can only be persuasive, not decisive, in democratic social choice.

The second "relaxation" target is *U*, Unrestricted Domain, that no possible set of preferences should be excluded as long as they satisfy *C*, Collective Rationality. Much effort has gone into domain restrictions. One, explored by Arrow, is illustrative of the ingenuity and barrenness of this sort of approach. It can be proven that majority decision will always be transitive if the range of allowed preference profiles is restricted to cases that are "single peaked," that is, if some alternative cannot be anyone's lowest choice, and the number of individuals is odd rather than even. This might arise in the Failed Nation problem posed above to the U.N. Suppose that "hawks" who support occupation (*O*) and "conciliators" who support negotiation (*N*) cannot

adopt the other's position as their second choice. If the number of U.N. members is odd, then, sanctions (*S*), advocated by "pragmatists," cannot be at the bottom of anyone's ranking.

The artificiality of this sort of result (and there are many more complicated results of this kind) is twofold. The single-peakedness solution works, mathematically, only with an odd number of individuals, and, politically, where consensus about alternatives is high. Such clever but ad hoc domain restrictions are impracticable, illiberal, and unmotivated. Merely avoiding Arrow Impossibility is no basis for excluding certain sets of preference profiles if people have them. Curiously, the philosopher Alfred MacKay, author of a major study of the theorem, advocates imposing single-peakedness on the grounds that the paradoxes are instances of an independently philosophically objectionable infinite regress and not merely as an unmotivated ad hoc way of avoiding paradox. But MacKay does not explain why it is democratically acceptable to impose a limited preference structure to avoid a philosophical problem that matters even less in real life than Arrow's paradoxes themselves, nor does he explain how to implement such a limitation.

A stronger rationale for domain restriction, one that people do care about, is that certain sets of preferences – for genocide or tyranny, say – are morally unworthy of consideration. The answer from a liberal democratic perspective is that there is no nondictatorial way of saying which preferences are not to count and so should be excluded from any domain of choice. It is important to see that choice under Arrovian conditions does not require preferences be taken as given ("exogenously") without critique or alteration, despite widespread claims to the contrary, like that of political scientist Jon Elster. Whether and how preferences may be subject to revision is different from whether aggregation should exclude any preferences ex ante because of their objectionable content. As long as we remain in Arrow's democratic world rather than

Clausewitz's realm of armed conflict or even the world of a benevolent (or otherwise) dictator, the range of acceptable alternatives must be decided ex post and politically.

A second sort of response to the Impossibility Theorem is that it is practically irrelevant as an empirical matter because divergence on outcomes is not, in the real world, great enough to generate Arrovian paradoxes. For example, Arrow notes that in a country like the U.S. with only two viable political parties, there may be only two choices. Or perhaps many of the proliferating profiles are marginal, such as a preference for a socialist party in the U.S., or "mainstream" preferences are highly similar and relatively consensual. This says, in effect, that the world conveniently provides domain restrictions without anyone's having to undemocratically eliminate alternatives in an ad hoc way. Where that is true, Arrow problems do not arise, but the assumption that the world is happily arranged to render them generally irrelevant is Panglossian. If the importance of Arrow's results turns on the probability of paradox in fact transpiring, that unhappy outcome is greater than empirical optimists suppose.

From a merely mathematical perspective, Professor Robert Abrams has shown that the probability of a paradoxical result, which is 6% among 216 possible preference profiles for the three-individual, three-alternatives case, does not rise above 10% regardless of the number of individuals involved if there are only three choices, but it fairly slowly approaches 100% as the number of choices approaches infinity. But 6 to 10% is a nontrivial probability of paradox – even a statistically significant one. Further, even if the world does restrict the domain de facto, one concern is whether it restricts them in a way that provides the *requisite* limitations. As single-peakedness illustrates, these can be quite peculiar, and that may be the simplest instance.

More deeply, systematic, widespread, and irreconcilable disagreement is a feature of any complex society, and moreso in a world of diverse societies. Rich and poor nations differ about bearing the cost of externalities like pollution and the unequal distribution of wealth generated by capitalism. The world community disagrees about how to handle challenges posed by Failed States, and who belongs in that category. Religious, racial, gender, and not least, class divisions are pervasive within and among nations. Points of consensus, even when expressed through democratic politics, often turn on bases of conflict with other groups. ("We can agree that fundamentalists/infidels/extremists are to be defeated.") These facts suggest that frequency of cycling and paradoxes may be greater than the cold mathematics suggest.

The final sort of reply is that the Impossibility Theorem misses the point of democratic politics. Since there is little agreement on what that is, the objection comes in many flavors. Consider three. An austere version is that our expectations about politics are lower than Arrow requires. Perhaps we do not need a complete social ranking of all alternatives. What matters is that we have a method that reliably produces a winner. However, here the bite of the theorem is, if such a method exists, it would require us to relax one of the four seemingly unavoidable conditions on social choice, which nonetheless remain compelling. It is not that just that the conclusion of the theorem is troublesome, but that its premises are more attractive than any alternative.

A richer objection of this kind is Elster's criticism that Arrow misconceives democracy by treating politics merely as a market-like mechanism for aggregating exogenously given preferences. Inspired by the philosophers Hannah Arendt and Jürgen Habermas, Elster, like Harvard Law Professor Cass Sunstein, advocates a deliberative democracy in which decision-making should be directed to making fair policy decisions by means of public debate that improves our preferences. Perhaps so. It might be noted that it is trivially obvious and not in the

least paternalistic to suggest that public discussion might change our preferences. Still, the appeal to deliberative democracy in this context misses the mark in two ways. First, as noted, nothing in Arrow excludes revision of preferences by public critique. Exogeneity of preference is not a premise in the Impossibility Theorem. Second, aggregation of the revised preferences is still required to reach a social choice. Since Elster does not challenge Arrow's premises, his deliberative democracy will face Arrow's paradoxes when it comes time for a vote.

A more radical critique in a similar spirit, offered by the philosopher Elizabeth Anderson and legal academic Richard Pildes does challenge the premises, specifically Collective Rationality. They maintain that politics expresses an understanding of vital choices based on incommensurable values. Inconsistencies are not "irrational" as long as the ultimate choice among incommensurable values is "not clearly inferior" to the alternatives. Cycling may serve important democratic ends by affirming the independent worth of incommensurable values that cannot all be realized at once. Inconsistent rankings may rationally reflect conflicting values generated by different norms governing multiple social roles that individuals occupy. The inability to attain a single complete consistent social choice ranking by a method that aggregates individual preferences is a virtue, not a paradoxical vice.

This thick description of a messy but vital political world is a refreshing contrast to the arid mathematical simplification of social choice theory. Nonetheless, this strategy seems to have subtly changed the subject in making a virtue of necessity. The "rationality" of inconsistent choices involves a comparative judgment of relative worth ("not clearly inferior") that presupposes the existence some common evaluative scale or at least a ranking that undermines a claim of radical incommensurability. The cycling Anderson and Pildes defend as functional derives from the very Arrovian assumptions they purport to reject. Moreover, while cycling may

be functional for democracy when conflicting interests and demands preclude a determinate choice, that does not mean that it is rational. To explain individual inconsistencies in preference due to competing demands of conflicting social roles is not to demonstrate their rationality, but to show how they arise. In the end, this approach offers a less than fully compelling persuasive redefinition of Arrow's paradoxes as "rational" because they may be the best that can be attained in the circumstances.

Collective choice, then, is still haunted by the specter of the Impossibility Theorem and similar social choice theoretic results, like Sen's Impossibility of a Paretan Liberal. Decades of attempts to circumvent, minimize, or dismiss these results have failed. The prospects for overcoming them seem dim. If the viability of democracy depended on showing that it could in principle always produce stable and rational results, it would be in trouble. Fortunately democracy, at least in the developed countries, is less debatable and more firmly rooted than any critique of it. No one is going to give it up because of the Arrow theorem. But the irrefrangibility of Arrovian paradox discourages complacency insofar as we wish to defend and understand it.

Prof. Justin Schwartz
The John Marshall Law School
Chicago, IL USA

See also

Ardent, Hannah; Capitalism; Climate Change; Class & Status, Collective Decision Problem; Common Good, Condorcet N., Marquis de; Consensus/Justification; Deliberative Democracy; Democratic Institutions; Diminishing Marginal Utility; Domination; Exploitation, Failed State; Genocide; Global Democracy; Habermas, Jürgen; Inequality; Rawls, John; Liberalism; Nozick, Robert; Pareto Optimality; Political Will→Common Good; Racism; Rational Choice; Socialism; Tyranny

Further Readings

Abrams, R. (1980) Foundations of Political Analysis: An Introduction to the Theory of Collective Choice: Columbia University Press, New York.

Arrow, K.J. (1963) Social Choice and Individual Values, 2nd ed. Yale University Press, New Haven (1st ed. 1951. Wiley, New York).

Bonner, J. (1986), Politics, Economics, and Welfare: An Elementary Introduction to Social Choice. Wheatsheaf Press, Brighton.

Copp. D., Hampton, J., & Roemer, J.E., eds., The Idea of Democracy. Cambridge University Press, Cambridge UK (contains, *inter alia*, contributions by C. Sunstein).

Elster, J. & Hylland, A., eds. (1986). Foundations of Social Choice Theory. Cambridge University Press, Cambridge UK (contains contributions by J. Elster, D. Davidson, and A. Gibbard) .

MacKay, A.F. (1980). Arrow's Theorem: The Paradox of Social Choice. A Case Study in the Philosophy of Economics. Yale University Press, New Haven.

Mill, J.S. (1962) Utilitarianism, On Liberty, Essay on Bentham, Together with Selected Writings of Jeremy Bentham and John Austin, M.Warnock, ed. Meridian, New York

Pildes, R.H. & Anderson, E.S. (1990) Slinging Arrows At Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics. 90 Colum. L. Rev. 2121-2214

Sen, A.K. (1970a) Collective Choice and Individual Welfare. North-Holland: Amsterdam.

_____. (1970b). The Impossibility of a Paretian Liberal. 78 Journal of Political Economy 152-157

_____ (1977), Social Choice Theory: A Re-examination. 45 Econometrica 53-89.