

ECO 751, Public Economics
Fall 2007
Professor William Hoyt

Lecture 3: Social Welfare Functions

Arrow, Kenneth J. "A Difficulty in the Concept of Social Welfare," *Journal of Political Economy*, 58(1950), 328-46.

Samuelson, Paul A. "Social Indifference Curves," *The Quarterly Journal of Economics*, 70(1956), 1-22.

A Difficulty in the Concept of Social Welfare

1. Introduction

Arrow is interested in addressing the issue of whether it is possible to attribute "rationality" to the collective decisions of people when these decisions are made based on voting and the marketplace rather than convention or dictatorship.

The methods of dictatorship and convention are, or can be, rational in the sense that any individual can be rational in his choice. Can such consistency be attributed to collective modes of choice, where the wills of many people are involved?

It should be emphasized here that the present study is concerned only with the formal aspects of the foregoing question. That is, we ask if it is formally possible to construct a procedure for passing from a set of known individual tastes to a pattern of social decision making, the procedure in question being required to satisfy certain natural conditions.

An Illustration

- Let there be three alternatives: *A*, *B*, and *C*
- 3 voters: 1, 2, and 3.

Suppose that:

- 1 prefers *A* to *B* and *B* to *C* (and therefore *A* to *C*);
 - 2 prefers *B* to *C* and *C* to *A* (and therefore *B* to *A*); and
 - 3 prefers *C* to *A* and *A* to *B* (and therefore *C* to *B*).
- Then the majority prefer *A* to *B*, and a majority prefers *B* to *C*. Therefore the *community* prefers *A* to *B* and *B* to *C*. If the community is regarded as rational (has transitivity) then *A* should be preferred to *C*. But, in fact, a majority of the community prefer *C* to *A*. This simple way of determining collective choice fails to be rational.

2. *Definitions and Notation*

There is a basic set of mutually exclusive alternatives (consumption and/or production) with:

- x, y, z, \dots denoting the alternatives;
- S - subset of alternatives available to individual
- xRy — “ x is preferred or indifferent to y ” xPy — “ x is (strictly) preferred to y ” xIy — “ x is indifferent to y ”

Axiom I: For all x and y , either xRy or yRx .

Axiom II: For all x, y , and z either xRy and yRz imply xRz

These axioms allow an (weak) ordering of alternatives.

- Lemma:
- (a) For all x , xRx .
 - (b) If xPy , then xRy .
 - (c) If xPy , then xRy .
 - (d) If xIy and yIz , then xIz .
 - (e) For all x and y , either xRy or yPx .
 - (f) If xPy and yRz , then xPz .

3. *The Social Welfare Function*

In this section Arrow formally defines what is meant by a social welfare function (SWF) and then imposes the “reasonable” conditions that a social welfare function should be expected to have. These are essentially the same conditions we expect indifference curves to have for individuals.

Definition 3: By a “social welfare function” will be meant a process or rule which, for each set of individual orderings R_1, \dots, R_n for alternative social states (one ordering for each individual), states a corresponding social ordering of alternative social states, R .

The Conditions on Social Welfare

1. *Completeness*

- *Condition 1: The social welfare function is defined for every admissible pair of individual orderings R_1 and R_2 .*

If the individual preferences satisfy Axioms I and II and therefore are admissible, the social welfare function should be defined for these preferences. This is similar to the notion of completeness for individual preferences.

2. *Positive Association of Social and Individual Values*

- *Condition 2: If an alternative social state x rises or does not fall in the ordering of each individual without any other change in those orderings and x was preferred to another alternative y before the change in individual orderings then x is still preferred to y .*

This is essentially a statement that the SWF is non-decreasing in the utility of each and every individual.

3. *The Independence of Irrelevant Alternatives*

Condition 3: Let $R_1, R_2,$ and R'_1, R'_2 be two sets of individual orderings. If, for both individuals i and for all x and y in a given set of alternatives S , $xR_i y$ if and only if $xR'_i y$, then the social choice made from S is the same whether the individual orderings are $R_1, R_2,$ or R'_1, R'_2 .

As an example of the independence of irrelevant alternatives, Arrow gives an election slate. Let individual voters order the candidates and then through an election system a winner is chosen. Now suppose that one of the candidates dies (not the winner). Then if the election is held again, if the independence of irrelevant alternatives holds, the winner should not change.

Arrow also notes that the rank-order method of voting in which voters rank preferred candidate 1st, 2nd, 3rd, etc.) with higher rankings get more weight in the voting and the winner being the candidate with the greatest weighted sum of votes does not satisfy condition 3.

4. *The Condition of Citizens' Sovereignty*

Definition 4: A Social Welfare Function will be said to be imposed if for some pair of distinct alternatives x and y , xRy for any set of individual orderings $R_1, R_2,$ where R is the social ordering corresponding to R_1 and R_2 .

Essentially this means that no matter what the preferences of the community for x and y are, they do not affect the SWF — the ordering of x and y is independent of the preferences of the community.

Condition 4: The Social Welfare Function is not to be imposed.

5. *The Condition of Nondictatorship*

Definition 5: A social welfare function is said to be "dictatorial" if there exists an individual i such that for all x and y , $xP_i y$ implies xPy regardless of the orderings of individuals other than i , where P is the social preference relation corresponding to those orderings.

Condition 5: The social welfare function is not to be dictatorial (nondictatorship)

Summary of Conditions:

Then in summary we have the conditions on the social welfare function of:

- 1) Completeness;
- 2) Positive Association of Social and Individual Values;
- 3) Independence of Irrelevant Alternatives;
- 4) Citizens' Sovereignty (nonimposed social welfare function);
- 5) Nondictatorship.

4. The Possibility Theorem for Social Welfare Functions

Using a simple example with two individuals, 1 and 2, and three alternatives, 1-3, Arrow shows "that the supposition that there is a social welfare function satisfying those conditions [Conditions 1-5] leads to a contradiction." In other words, a social welfare function can not simultaneously satisfy all five conditions.

Example:

- 2 individuals (1 and 2);
- 2 commodities (1 and 2);
- 3 alternatives (1-3).

Preferences:

- 1) Ranking of alternatives by individual depends only on commodities he receives in alternative states.
- 2) More is better.

Alternative	Individual 1		Individual 2	
	Commodity 1	Commodity 2	Commodity 1	Commodity 2
1	5	1	5	9
2	4	2	6	8
3	3	3	7	7

Note that given the assumptions on preferences, a priori we can not order the alternatives for either individual so Condition 1 requires that the social welfare function be defined for all pairs of individual orderings R_1 and R_2 .

The Possibility Theorem

The contradiction in the Conditions 1-5 is now generated for this simple example.

Let x, y, z be the 3 alternatives and $x', y',$ and z' denote values for these alternatives.

We now draw some consequences (implications) of the 5 conditions for this example.

Consequence 1: If $x'P_1y'$ and $x'P_2y'$ then $x'Py'$.

Proof (sketch): Let there be orderings R_1' and R_2' and associated social preference that has $x'Py'$. Now have R_1'' formed from R_1' by raising valuation of x' to the top leaving evaluation of other 2 alone. Form R_2'' the same way. Then since we only increase everyone's valuation of x' it must be that $x'P''y'$ by Condition 2. So even if the valuation of x' and y' relative z' has changed by Condition 3 this does not matter as the (relative) valuation of z' is irrelevant.

Consequence 2: Suppose that for some x' and y' whenever $x'P_1y'$ and $y'P_2x'$, $x'Py'$. then, for that x' and y' , whenever $x'P_1y'$, $x'Py'$.

If individual 1's choice prevails against opposition of 2, then individual 1's choice prevails when 2 is indifferent between x' and y' .

Consequence 3: If $x'P_1y'$ and $y'P_2x'$, then $x'Iy'$.

If two individuals have exactly opposing interests, society must be indifferent between the two alternatives. Essentially Arrow obtains this by showing that if it does not hold then one of the two individuals is a dictator or the social ordering is imposed.

Proof: Suppose the consequence is false. Suppose for preferences R_1 and R_2 and alternatives x' and y' , we have $x'P_1y'$, $y'P_2x'$, but not $x'Iy'$. We will suppose that $x'Py'$ and show it leads to a contradiction as does $y'Px'$.

Let $x' = x$ and $y' = y$ (simplify notation). Then by assumption it must be the case that

Whenever xP_1y and yP_2x then xPy . (1)

Arrow then shows that (1) leads to a contradiction. Essentially what he shows is that if (1) (and Consequence 2) then either individual 1 is a dictator or the social ordering is imposed. He does this by considering (five) alternative orderings of x , y , and z to generate the contradiction.

a) Suppose 1's preferences are xP_1yP_1z and 2's are yP_2zP_2x . Since by (1) xPy and both prefer y to z then by Consequence 1, xPz . Since choice between x and z depends only on preferences for x and z (Condition 3) we have

Whenever xP_1z and zP_2x then xPz . (2)

b) Suppose 1's preferences are yP_1xP_1z and 2's are zP_2yP_2x . By Consequence 1, yPx and by (2) xPz and both prefer y to z then by Consequence 1, xPz . Then as before

Whenever yP_1z and zP_2y then yPz . (3)

c) Suppose 1's preferences are $y P_1 z P_1 x$ and 2's are $z P_2 x P_2 y$. By Consequence 1 and (3), $z P x$ and $y P z$ so $y P x$

$$\text{Whenever } y P_1 x \text{ and } x P_2 y \text{ then } y P x \quad (4)$$

d) Suppose 1's preferences are $z P_1 y P_1 x$ and 2's are $x P_2 z P_2 y$. By Consequence 1 and (4), $z P y$ and $y P x$ so $z P x$.

$$\text{Whenever } z P_1 x \text{ and } x P_2 z \text{ then } z P x. \quad (5)$$

e) Suppose 1's preferences are $z P_1 x P_1 y$ and 2's are $x P_2 y P_2 z$. By Consequence 1 and (5), $z P x$ and $x P y$ so $z P y$.

$$\text{Whenever } z P_1 y \text{ and } y P_2 z \text{ then } z P y. \quad (6)$$

What has been shown? Essentially what he shows is that it follows that if (1) (and (2)-(6) which directly follow) hold that individual 1's preference always determines the social ordering but this would be dictatorial so (1) can not hold and Consequence 3 must hold.

The Contradiction

Suppose for 1 has ordering x, y, z and 2 has z, x, y . By Consequence 1

$$x P y \quad (7)$$

(they both prefer x to y so society must). Since $y P_1 z$ and $z P_2 y$ it follows from Consequence 3 that

$$y I z \quad (8)$$

From (7) and (8) we have $x P z$ (transitivity) but $x P_1 z$ and $z P_2 x$ which implies $x I z$ by Consequence 3. But this is a contradiction since we can not have $x P z$ and $x I z$. Essentially what Arrow shows that we can not have positive association of values and the independence of irrelevant alternatives (Conditions 2 and 3) and transitivity when we also have nondictatorial or nonimposed social welfare.

Possibility Theorem. If there are at least three alternatives among which the members of the society are free to order in any way, then every social welfare function satisfying Conditions 2 and 3 and yielding a social ordering satisfying Axioms I and II (completeness and transitivity) must be either imposed or dictatorial.

“The Possibility Theorem shows that, if no prior assumptions are made about the nature of individual orderings, there is no method of voting which will remove the paradox of voting discussed in Part I (violation of irrelevant alternatives), neither plurality voting nor any scheme of proportional representation, no matter how complicated.”

Arrow restates the possibility theorem (p. 164) to state that if we are unwilling to make interpersonal comparisons of utility then the only way we can go from individual orderings to social orderings is to have imposed or dictatorial social welfare functions. Essentially then either interpersonal comparisons must be made or social welfare is defined outside of a voting system.

2. *Social Indifference Curves*

Samuelson, Paul A. "Social Indifference Curves," *The Quarterly Journal of Economics*, 70(1956), 1-22.

1. *Introduction and Summary*

In this article Samuelson is considering how appropriate it is to use social (community) indifference curves. However, the issue he is concerned with is not their relationship with voting but with market decisions. He is most directly addressing the use of social indifference curves in international economics as a means of characterizing the consumption decisions of nations (or regions).

After proving that in general community indifference curves do not exist, Samuelson then goes on to give conditions for when these curves do exist. Essentially it requires optimal redistribution of income within the community so that $dSWF^i/dI^i$, the marginal social utility of income, is equal for everyone in the community.

2. *Attempts to Justify the Use of Community Indifference Curves*

When can these indifference curves be justified?

- a) There is a single individual in the community;
- b) All individuals have identical tastes and endowments;
- c) The community is a totalitarian state.

Clearly these are very limiting and uninteresting cases.

3. *Proof of the Nonexistence of Community Indifference Curves*

Samuelson demonstrates that if we consider two dissimilar individuals then while we can express the indifference function for individual 1 as:

$$\frac{p_x}{p_y} = F^1(x^1, y^1) \tag{1a}$$

and for individual 2 as

$$\frac{p_x}{p_y} = F^2(x^2, y^2) \tag{1b}$$

The function $F'(\cdot, \cdot)$ is simply the marginal rate of substitution for individual i . Then these functions give the (compensated) demand relationships for the individuals. What Samuelson shows is that while you can have individual demand relationships given by (1) there is no collective community demand relationship

$$\frac{p_x}{p_y} = F'(x^1 + x^2, y^1 + y^2) \quad (2)$$

which given (1), that is given individuals freely choose consumption on the market, enables us to have a relationship between relative price and total demand in the community independent of how it is distributed among individuals. The difficulty arises because how the goods are allocated among the two individuals will change the equilibrium price ratio.

4. Regular Properties of Social Indifference Curves

When could it be the case that well-behaved (convex) social indifference curves can exist that will give a relationship between total demand and market prices and total income? Samuelson provides the following theorem:

Theorem: If each group member's demand and indifference contours have the conventional "regular" convexity and (b) if the social welfare function is defined to have similar regular convexity properties and (c) if within the group optimal lump-sum transfers are always made, then it follows:

(1) there will result observable demand totals that are functions of market prices and total income alone, and (2) that these demand functions will have all the Slutsky-Hicks or revealed preference properties of any single consumer's demand, and (3) there will exist a set of indifference contours relating the totals X, Y, \dots that has all the regular properties of any individual's contours and which we can pretend a single mind is maximizing.