Econ 502 - Microeconomic Theory - Spring Term 1985, second half

INTERPERSONAL COMPARISONS AND SOCIAL WELFARE FUNCTIONS

1. INTRODUCTION

Following Arrow's 'impossibility theorem', a large part of the theory of welfare economics has been concerned with the question of how individual preferences may be 'combined' to yield social preferences or choice rules. Independently, the more application-oriented branches of the subject have gone on using social welfare functions of the Bergson-Samuelson type, with individual utilities as their arguments. Some recent work has substantially clarified the confusion and reconciled the two approaches, by paying explicit attention to alternative forms of interpersonal comparisons. This is an expository summary of such work, based particularly on d'Aspremont and Gevers (1977) and Roberts (1980a).

This is meant to be a brief introductory account, with an eye towards applications. Therefore some subtle and theoretically important aspects are neglected. In particular, the following limitations should be noted:

(1) The research literature contains several alternative forms for the basic conditions, and seemingly small differences in these often lead to large variations in the consequences. I shall follow only one line of development. For fuller discussions, see Sen (1970, 1977), Roberts (1980a,b) and the references cited there.

(2) Rigorous theory of the subject supposes that individuals have complete, reflexive and transitive preference orderings, i.e. relations of the form 'x is at least as good as y'; social choice may have to be modelled using even more general kinds of relations. Economists are used to circumstances where preference relations are represented by numerical functions, i.e. functions f such that x is at least as good as y if and only if $f(x) \geq f(y)$. Debreu (1954) discusses the conditions under which such a representation is possible; an important exception is where the preferences are lexicographic. For ease of exposition, I shall work with numerically represented preferences. This is not too restrictive for individual preferences in most economic applications. However, the same is true for social preferences only in particular cases. I shall point out the most important instance where this matters.

2. DEFINITIONS AND BASIC PROPERTIES

The underlying structure of the problem is as follows: X denotes the set of social states; its typical elements are denoted by x, y etc. By a <u>social state</u> is meant a list of all matters pertinent to the problem; in economic contexts this will typically be all production and consumption allocations. With this in mind, I shall develop the exposition taking X to contain a continuum of states, but most of the issues of social choice arise in a similar way as soon as X has three or more elements.

The set of individuals is $N = \{1,2,\ldots n\}$; its typical elements are denoted by i,j etc. For geometric illustration I shall sometimes examine the case of two individuals.

$X * N$ is the set of all pairs (x,i) where x is a social state and i is an individual. U is the set of all real-valued functions on $X * N$, i.e. for u in U, x in X, and i in N, a unique real number u(x,i) is defined. For each i, this can be thought of as a function u(,i), and interpreted as a utility function representing a possible pattern of i's preferences over the social states. As far as individual choice is concerned, there is some arbitrariness in specifying such a representation. The individual's preferences and choices are not affected if the function is subjected to an arbitrary monotonic increasing transformation, i.e. the representation is ordinal. However, in particular contexts of social choice, the problem may offer us information beyond that contained in individual choice, and we may wish to use this in making value judgements with interpersonal comparisons. Then we must choose representations for different individuals so as to conform to these judgements. Therefore the function u involves two distinct aspects: Individual preferences and concerns of interpersonal comparisons. To emphasize this dual role, u(x,i) will be called the welfare level of individual i in state x. The function u itself will be called a pattern of preferences and comparisons, or often more simply a profile.

S is the set of real-valued functions on X; its typical element will be denoted by s. A function s will be interpreted as a numerical representation of a social preference ordering over the set of social states. Here we have an unavoidable ambiguity: if s represents a particular social preference relation, so does any monotonic increasing transform of s. We must therefore regard two such functions as equivalent. More precisely, s and s' are equivalent if, for all x and y in X, $s(x) > s(y)$ if and only if $s'(x) > s'(y)$.

By a social welfare functional is meant a function f whose independent variables range over a domain D contained in U, and whose values lie in S. Thus, for each profile in the domain, the function associates a social preference ordering. We must of course regard two such social welafre functionals f and f' as equivalent if, for all u in the domain, the outcomes $s = f(u)$ and $s' = f'(u)$ are equivalent in the sense defined above.

Social welfare functionals will be required to satisfy certain specified conditions. These are motivated by considerations of economic reasonableness or ethical beliefs. Three such conditions are defined, and very briefly discussed, in this section; others more explicitly concerned with value judgements follow in the next.

22

Condition (U): Unrestricted Domain

This requires D = U, i.e. that the functional should be capable of assigning a social preference relation to any conceivable pattern of individual preferences and interpersonal comparability.

A possible criticism of this condition is that it is too demanding in requiring a 'constitution' that lays down rules for resolving all conflicts of interest that could ever arise. It may be thought enough to have the ability to resolve only a limited set of issues, given some prior agreement limiting the domain of discussion. This argument provides one escape route if the various conditions prove to be mutually incompatible. Since the focus here is not on such 'impossibility theorems', unrestricted domain will not worry us. However, it should be noted that it permits all manner of external effects or meddlesomeness: each consumer's preferences can depend on all consumers' allocations in an arbitrary way.

Condition (P): The Pareto Principle

Let x,y be in X, u in U, and s = f(u). Then it is required that if (i) $u(x,i) \geq u(y,i)$ for all i implies $s(x) \geq s(y)$, (ii) $u(x,i) > u(y,i)$ for all i implies $s(x) > s(y)$.

A condition like this is very familiar to economists. However, slight differences in precise statements of it can have important consequences. The form used here can be described as moderately strong. It could be strengthened by requiring strict social preference as soon as at least one individual has a strict preference, i.e. strengthening (ii) to require $s(x) > s(y)$ when $u(x,i) \geq u(y,i)$ for all i and $u(x,i) > u(y,i)$ for at least one i. From the point of view of this exposition, that would cause a technical problem. Properly speaking, the social preference relation entailed by such a condition would be lexicographic and not representable by a numerical function s. This will be pointed out later where it matters.

On the other hand, the condition could be weakened by omitting part (i). The importance of that part is that it implies social indifference when all individuals are indifferent. For any one profile and any two states, then, an agreement of all welfare levels implies social indifference. This goes some way towards ruling out any role for non-welfare characteristics, i.e. features of the social states or individuals not contained in the welfare levels. In fact (ii) alone has almost as much force. Welfare comparisons cannot be overruled as soon as they are unanimous. It is only when welfare unanimity just fails, that non-welfare characteristics can be used for resolving ties, i.e. their role is at best lexically secondary.

For one particular profile, we may be fortunate and this may not matter. But when universal domain is allowed, so that any profile could arise, there are serious conflicts with other conceivable ethical desiderata such as liberalism or justice. See Sen (1970, ch.6) and a flood of following literature.

Condition (I): Independence of Irrelevant Alternatives

This requires that if two profiles u and u' agree over a subset A of X, then the social preferences that result from these profiles should also agree over the same subset, irrespective of any disagreement elsewhere. More formally, suppose $u(x,i) = u'(x,i)$ for all x in A and all i in N. Let $s = f(u)$ and $s' = f(u')$. Then the restrictions of s and s' to A should be equivalent, i.e. for all x,y in A, we should have $s(x) > s(y)$ if and only if $s'(x) > s'(y)$.

The distinctive feature of this condition is that it involves two profiles. As a result, it serves to extend the kind of irrelevance of non-welfare characteristics that the Pareto principle entailed for one profile, to all profiles in the domain. This consequence is an addition to the primary role of the condition, namely that of ruling out the influence of welfare characteristics of states other than those being compared. In all, this condition constrains the social choice mechanism to use relatively little information, namely that of welfare levels of the immediate states.

There is a line of argument that makes this condition the crucial difference between the approaches of Arrow and Bergson-Samuelson. See McManus (1975) for an example, and Sen (1977, section 9) for a discussion. In my view the aspects of interpersonal comparability are far more important; I shall comment on this later.

The next task is to make precise how the conjunction of conditions (P) and (I), together with (U), goes to deny any role to non-welfare characteristics in the social choice process. The resulting property is commonly called neutrality; I prefer to adopt Sen's perjorative description as a title:

Condition (W): Welfarism

Let x,y,x',y' be in X, and u,u' in U, such that for all i in N, we have $u(x,i) = u'(x',i)$ and $u(y,i) = u'(y',i)$. Then, on writing $s = f(u)$ and $s' = f(u')$, we require $s(x) > s(y)$ iff $s'(x') > s'(y')$.

Note how this condition combines aspects of Pareto-indifference and independence of irrelevant alternatives. It is trivial to verify that condition (W) implies (I), i.e. if a social welfare functional satisfies (W), it also satisfies (I). More importantly, we have the reverse connection:

Lemma: Conditions (U), (P) and (I) together imply (W).

Proof: This essentially consists of checking out the definition. Take x, y, x', y' in X, and u, u' in U, satisfying for all i in N, the equations

$$u(x,i) = u'(x',i) \quad , \quad u(y,i) = u'(y',i).$$

We must distinguish cases depending on how many of the states x, y, x' and y' are distinct. The simplest arises when they are all different, and serves to illustrate the line of reasoning and the role of the three conditions. Let

$$a_i = u(x,i) = u'(x',i) \ , \quad b_i = u(y,i) = u'(y',i)$$

Define another profile u'' as follows. Let $u''(x,i) = a_i$, $u''(y,i) = b_i$, $u''(x',i) = a_i$, $u''(y',i) = b_i$, i.e. u'' agrees with u on the states x, y and with u' on the states x', y'. Elsewhere, u'' can be defined arbitrarily. By condition (U), this is in the domain of f; let $s'' = f(u'')$.

Since u and u'' agree on the set $\{x,y\}$, by condition (I) we have $s(x) > s(y)$ if and only if $s''(x) > s''(y)$. Similarly, since u' and u'' agree on $\{x',y'\}$, we have $s'(x') > s'(y')$ if and only if $s''(x') > s''(y')$.

Now look at the profile u'' on its own. Since $u''(x,i) = u''(x',i)$ for all i, condition (P) yields $s''(x) = s''(x')$. Similarly, from $u''(y,i) = u''(y',i)$, we have $s''(y) = s''(y')$.

Putting these two lines of inference together, we have $s(x) > s(y)$ if and only if $s'(x') > s'(y')$, the desired conclusion.

The reasoning is slightly more complicated for cases where some of the states coincide, and requires more intermediate profiles to be constructed. For the most difficult case, see d'Aspremont and Gevers (1977, Lemma 2). ∎

The implication of condition (W) is that for social choice between any pair of states under any profile, all that matters is the pair of lists or vectors of the welfare levels. It is then possible to express the mechanism of social choice directly in terms of a preference relation, or its numerical representation, over the vectors of welfare levels. This is the content of the following:

Theorem: If a social welfare functional f satisfies conditions (U) and (W), then we can find a real-valued function W of n real variables with the following property: For x, y in X and u in U, set $s = f(u)$, and define vectors a and b with components $a_i = u(x,i)$, $b_i = u(y,i)$. Then we are to have $s(x) > s(y)$ if and only if $W(a) > W(b)$.

Proof: This essentially works backwards. Define a preference relation on the set of n-dimensional vectors as follows. Given two such vectors a and b, find u in U, and x, y in X, such that $a_i = u(x,i)$ and $b_i = u(y,i)$. This is always possible by condition (U). Then a is to be preferred to b if and only if $s(x) > s(y)$ where $s = f(u)$. By condition (W), this definition is independent of the particular choice of profile: if u', x' and y' are such that $a_i = u'(x',i)$ and $b_i = u'(y',i)$, we will have $s(x) > s(y)$ if and only if $s'(x') > s'(y')$. Thus the preference relation is well-defined. A numerical representation of it gives the desired function W. ∎

It is then natural to call W the Bergson-Samuelson social welfare function corresponding to the social welfare functional f. It should be noted that in order to ensure such a correspondence, we can impose <u>either</u> conditions (U) and (W) directly as in the theorem, <u>or</u> conditions (U), (P) and (I) which together imply the two required. The latter will be done in what follows.

Sometimes, when explicitly mentioned, another requirement will be imposed on the social welfare functional, and its corresponding Bergson-Samuelson function. The idea is that actual labels borne by individuals should have no influence on the social decision, i.e. if we permute the labels 1,2,... n among themselves, the same social preference rule should result. Stated more precisely, we have

<u>Condition (A): Anonymity</u>

Let p be a permuation of N, i.e. a one-to-one function onto N. If profiles u and u' are related by $u(x,i) = u'(x,p(i))$ for all x in X and all i in N, then it is required that $s = f(u)$ and $s' = f(u')$ are equivalent.

Let a and a' be vectors with components $a_i = u(x,i)$ and $a'_i = u(x',i)$. Then these vectors in their component lists contain the same numbers in different orders. If social choice is not to be affected by such a re-ordering of arguments in the Bergson-Samuelson function W corresponding to f, that function must be symmetric in its arguments. Under anonymity, social preference should depend only on the set of welfare levels, and not on which individual has which level.

3. ALTERNATIVE KINDS OF INTERPERSONAL COMPARISONS

The kinds of interpersonal comparisons we can make in this framework will depend on how sharpely we can distinguish one numerical representation of an individual's preferences from another one for his own preferences, or from a representation of another's preferences. These limitations may arise either because we do not have the necessry information, or because ethical or political considerations prevent us from using it.

For each i, a representation is $u(x,i)$ regarded as a function of x alone. For fixed i, our information may be ordinal, i.e. we may be unable to tell apart this function of x from any monotonically increasing transformation of it, or cardinal, i.e. we may be unable to distinguish this function from any increasing linear transforms while being able to recognise non-linear ones.

This may have to do with the kind of behaviour we can observe (under certainty or uncertainty) or with our judgement of the relevance of such observations for the problem of social choice. Or it may be that all we have to go on is information concerning individuals' income or wealth levels, forcing us to use these as proxies for welfare levels. Across individuals, we may be forced to maintain non-comparability by being unable to tell when different transforms are applied for different i, as when utility functions are inferred from behaviour. Or we may have particular kinds of comparability by being able or willing to insist that the transforms have certain common features across individuals, as when information pertains to income or wealth levels, and the units may be changed from pounds to pence in common for all individuals.

The general way to treat these possibilities is by dividing the totality of conceivable profiles, or the set U, into equivalence classes. These are a collection of mutually disjoint subsets of U which together exhaust all of U. Any pair of profiles which belong to the same equivalence class are supposed to be mutually indistinguishable. The implication for social choice is the requirement that if u and u' belong to the same equivalence class, then f(u) and f(u') should be equivalent representations of social preference.

Two points should be noted in this context. First, this requirement should be clearly distinguished from a condition restricting the domain. That would allow us to get away with a complete inability to assign a social preference to some profiles. Here we insist on the ability to evaluate any profile, but require the outcomes to coincide for certain profiles.

Secondly, the larger the equivalence class, the stronger the requirement imposed on the social welfare functional. This is because it is asked to take on equivalent outcomes for a wider collection of profiles. As a consequence, fewer social welfare functionals will be able to satisfy the requirement. In other words, the weaker our basis of information and comparison, the more profiles we will be obliged to regard as indistinguishable, and the more difficult the problem of finding social choice mechanisms will be.

The equivalence classes are defined by means of a group of transformations. This needs to be explained a little. Any two profiles in the same equivalence class can be obtained one from the other by applying one of a set of specified permissible transformations. Each transformation t is really a list or vector with $(t_1, t_2, \ldots t_n)$ as the component transformations, each component being a real-valued function of a real argument. Given u in U, then, $u' = t(u)$ is defined by

$$u'(x,i) = t_i(u(x,i)) \text{ for all } x \text{ in } X,\ i \text{ in } N.$$

The set of all transformations to be considered is denoted by T. In order to be compatible with the notion of equivalence classes, T must have a structure.

First, each profile is (trivially) in the same equivalence class as itself, therefore the identity transform which maps each profile u to itself must be in T. Secondly, the relation of being in the same equivalence class is symmetric between pairs of profiles, therefore the inverse of a transform in T should itself be in T. Finally, the relation is transitive, i.e. if u, u' are in the same equivalence class, and u', u'' are in the same one, then u, u'' must be in the same one. Therefore the composite of two transforms in T, i.e. the transform obtained by applying the two in succession, must itself be in T. A set with such a structure is called a group.

Different kinds of measurability and comparability of profiles yield different classifications of U into equivalence classes, corresponding to different groups of permissible transformations. Some such groups are described below, and named after the corresponding economic contexts of measurability and comparability.

ONC : Ordinal, non-comparable

Here T consists of all transforms $t = (t_1, t_2, \ldots t_n)$ where each t_i is monotonically increasing. Thus each individual's welfare can only be measured ordinally, and we cannot distinguish between different specifications of the orderings for different individuals.

OLC : Ordinal, level comparable

This includes all transforms $t = (t_o, t_o, \ldots t_o)$ where t_o is monotonically increasing. Individual welfares are still ordinal, but we can choose representations consistently across individuals, and can thus tell apart cases where different transforms are applied to different ones.

CNC : Cardinal, non-comparable

The permissible transforms are defined by the component functions

$$t_i(z) = \alpha_i + \beta_i \, z$$

where z is used as an all-purpose real variable, and α_i, β_i are constants, β_i being required to be positive. Thus each individual's welfare level is determined to within a linear transform, but this transform can be chosen differently for different ones.

CUC : Cardinal, unit comparable

This has the permissible transforms defined by $t_i(z) = \alpha_i + \beta \, z$, with β positive. Thus welfare levels for different individuals can have different origins, but must have the same scale. If an absolute scale can be specified, we may set $\beta = 1$, but that additional power makes no difference. The group of transforms with $t_i(z) = \beta_i \, z$ is essentially the same as this one, since $\log t_i(z) = \log \beta_i + \log z$ is of the same form.

CFC : Cardinal, fully comparable

Here we have $t_i(z) = \alpha + \beta z$ with β positive, i.e. the linear transform must be the same for all individuals. Now the ability to set $\beta = 1$ does entail a further restriction, but that case can be handled by taking logarithms in the next one, and is therefore not examined separately.

CRS : Cardinal, ratio scales

Now $t_i(z) = \beta z$ with β positive. Thus all origins of welfare levels are fixed. The scale can be varied, but must be common for all individuals.

PCM : Perfect comparability and measurability

The only permissible transform is the identity, so each equivalence class contains only one profile. In other words, any profile can be distinguished from any other.

Looking through the cases with cardinality, it may appear that the case of common origins but different scales, i.e. where

$$t_i(z) = \alpha + \beta_i z$$

has been forgotten. In fact this case is ruled out because this class of transforms is not a group. The inverse of such a transform is defined by

$$t_i^{-1}(z) = -(\alpha/\beta_i) + (1/\beta_i) z$$

which does not have the desired feature of common origins. (Exercise: verify that the cases defined above do give classes of transforms that are groups.)

The ultimate aim of the exercise is to see what kinds of social welfare functionals f are compatible with alternative kinds of measurability and comparability. This is easier to do by looking at the corresponding Bergson-Samuelson function W. We have the following crucial result:

THEOREM: Let W be compatible with a group of transformations T. Let a and b be n-dimensional vectors, and for t in T, let $a' = t(a)$, $b' = t(b)$. Then $W(a) = W(b)$ if and only if $W(t(a)) = W(t(b))$.

Proof: Let u, x, y be such that $a_i = u(x,i)$ and $b_i = u(y,i)$. This is possible by condition (U), and values of u elsewhere do not matter by condition (I). Write $u' = t(u)$, so that $t_i(a_i) = u'(x,i)$, $t_i(b_i) = u'(y,i)$. Let $s = f(u)$, $s' = f(u')$. Then we have

$W(a) = W(b)$ if and only if $s(x) = s(y)$ since W represents f

i.e. if and only if $s'(x) = s'(y)$ since u, u' are in one equivalence class

i.e. if and only if $W(t(a)) = W(t(b))$ since W represents f. ∎

This result can be given a simple geometric interpretation. The function W can be depicted by means of its contours (indifference cruves) in an n-dimensional space, with individual welfare levels shown on each axis. The theorem then says that if a and b lie on one contour, then t(a) and t(b) must lie on another contour. In other words, if we apply one of the group of permissible transforms to the variables on the axes, the whole indifference map of W should be carried into itself. Each contour will be moved, but its new position will coincide with the old position of another contour, and the overall picture of the contour map will remain unchanged.

The theorem helps us in establishing which functions are compatible with any specified group T. This is done by detecting which ones are not, and leaving a small class which can then be checked directly for compatibility. This is a matter of cunning choice of stransforms that will reveal violations of the above theorem. Geometrically, this is often easily done by constructing a permissible transform for which a shifted contour will intersect an original one, which then contradicts the Pareto principle. This programme is carried out below for the different cases of measurability and comparability defined earlier. To allow a simple geometric exposition, this is done for two individuals; some problems of generalisation are stated where they arise.

ORDINAL AND CARDINAL NON-COMPARABILITY.

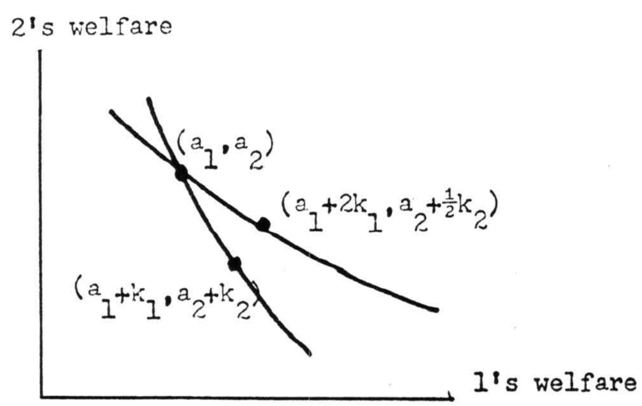
Suppose $W(a_1, a_2) = W(a_1+k_1, a_2+k_2)$. Because of the Pareto principle, k_1 and k_2 cannot be both positive or both negative.

Consider the transform, which is permissible under cardinal (and a fortiori under ordinal) non-comparability:

$$t_1(z) = 2\,(z - a_1) + a_1 \quad , \quad t_2(z) = \tfrac{1}{2}\,(z - a_2) + a_2$$

Applying the theorem, we find $W(a_1, a_2) = W(a_1+2k_1, a_2+\tfrac{1}{2}k_2)$. If both k_1 and k_2 are non-zero, this contradicts the Pareto principle.

The figure shows the case where k_1 is positive and k_2 negative. The point is that with the ability to apply independent transformations to the two axes (whether linear or non-linear), we can twist the indifference curves and then have the transformed contours intersect the original ones, which is not permissible.



Thus one of k_1 and k_2 must be zero, i.e. the contours of W must be either vertical or horizontal. In the former case, social preference depends only on individual 1's preference; in the latter, only on 2's. In other words, under non-comparability, social choice is simply dictatorship of one individual.

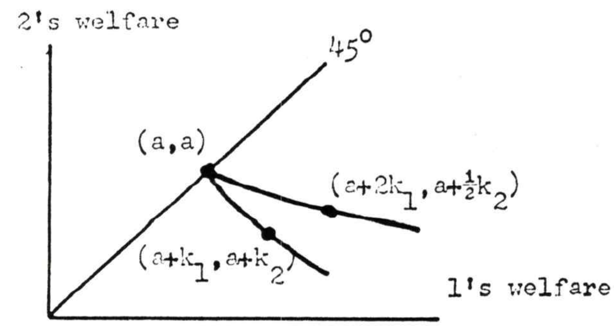
If we had imposed non-dictatorship as a prior desideratum on the problem, we would have an impossibility theorem: no social welfare functional would be compatible with conditions (U), (P), (I), non-dictatorship and non-comparability. In the ordinal case, this is Arrow's theorem; the corresponding result with cardinality is due to Sen. Anonymity would be even stronger than non-dictatorship, and would again yield impossibility theorems in these cases.

ORDINAL LEVEL-COMPARABILITY

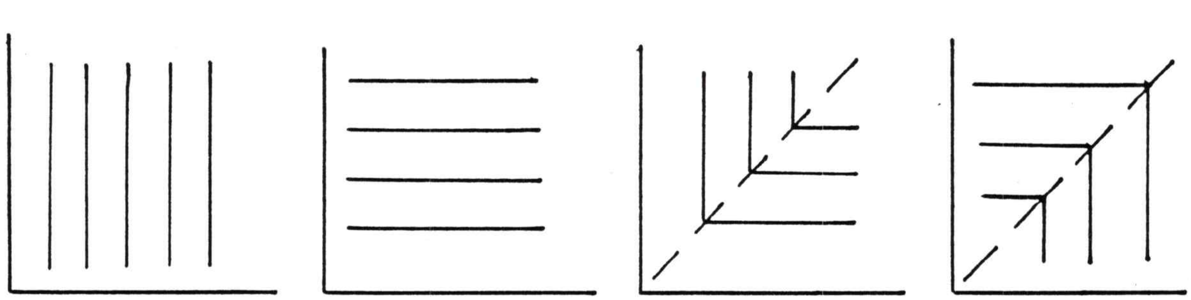
Suppose $W(a,a) = W(a+k_1, a+k_2)$ with $k_1 > 0$ and $k_2 < 0$. Consider the permissible transform, where each individual's welfare is transformed by

$$t_o(z) = \begin{cases} 2(z-a) + a & \text{for } z \geq a \\ \frac{1}{2}(z-a) + a & \text{for } z \leq a \end{cases}$$

Applying the theorem, we have $W(a,a) = W(a+2k_1, a+\frac{1}{2}k_2)$, which again contradicts the Pareto principle as shown in the figure. So one of k_1 and k_2 must be zero.



This differs from the previous case in that the reasoning is specific to one side of the 45° line. It could be applied independently to the other side, and a similar conclusion obtained. But it is now possible that k_1 is zero on one side and k_2 on the other, i.e. the contours of W may be vertical or horizontal throughout, or vertical on one side of the 45° line and horizontal on the other. The four cases that can arise are shown in the next figure.



(i) (ii) (iii) (iv)

The first two cases are those of dictatorship. We can rule them out by direct assumption, or by imposing a stronger condition like anonymity.

That leaves cases with L-shaped and inverse L-shaped contours, which can be represented by the numerical functions

$$W(a_1,a_2) = \min (a_1,a_2) \qquad \text{for case (iii)}$$
$$W(a_1,a_2) = \max (a_1,a_2) \qquad \text{for case (iv)}$$

In case (iii), social concern coincides with that of the individual with the lower welfare level of the two. This is a kind of dictatorship, but not that of a particular individual. Rather, it is dictatorship of a position in the ordering of welfare levels, namely the worst-off. In another light, we recognise this as a case at the extreme of egalitarian concern, or on yet another basis, as formally similar to Rawls' max-min criterion of justice.

Case (iv) is the dictatorship of the best-off, and at the extreme of inegalitarianism. Adjoining a condition of minimal equity will help us eliminate this, leaving (iii) as the only case compatible with ordinal level-comparability.

With more individuals, matters are somewhat more complicated. It remains true that social welfare coincides with the welfare of one individual, and which one is chosen depends on the order of welfare levels. An example is dictatorship of the k^{th} worst-off for some k. A strong equity condition, or a separability condition combined with a weaker equity condition, allow us to eliminate the possibilities other than the max-min.

Even then, if the worst-off person is indifferent between two states, we may wish to allow others' prefernces to have a role. The form of the Pareto principle used here precludes that and imposes indifference in such a social choice. If the principle were strengthened by requiring strict social preference as soon as one individual has a strict preference, then we would have to modify the max-min rule to its lexical extension: if the k worst-off individuals are indifferent, we look at the preference of the (k+1)st.

Details of all these points can be found in Hammond (1976), d'Aspremont and Gevers (1977) and Roberts (1980a,b).

CARDINAL UNIT COMPARABILITY

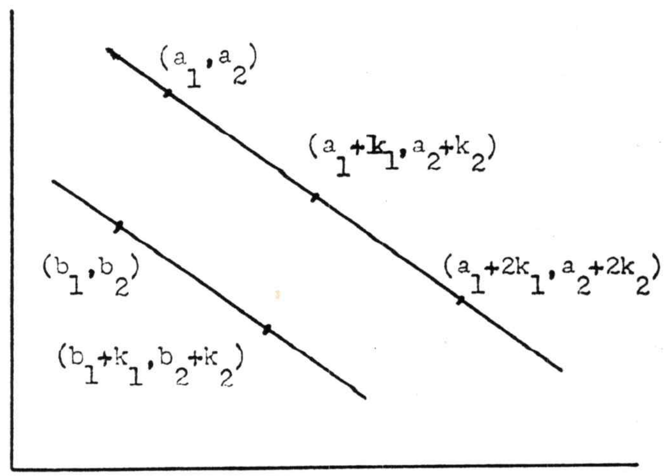
Suppose $W(a_1,a_2) = W(a_1+k_1,a_2+k_2)$. Try the transformation given by

$$t_1(z) = k_1 + z \quad , \quad t_2(z) = k_2 + z.$$

Applying the theorem, we find $W(a_1+k_1,a_2+k_2) = W(a_1+2k_1,a_2+2k_2)$. Therefore the welfare contours must be straight lines.

Next, take any numbers b_1 and b_2, and try the transform given by $t_1(z) = z + b_1 - a_1$, $t_2(z) = z + b_2 - a_2$. Applying the theorem, we have $W(b_1, b_2) = W(b_1 + k_1, b_2 + k_2)$. Therefore the lines are parallel.

The accompanying figure illustrates this. Another way to explain the result is that the indifference map has to translate into itself when it is subjected to any vertical or horizontal shift as the two welfare origins are changed; therefore it must consist of a family of parallel straight lines.

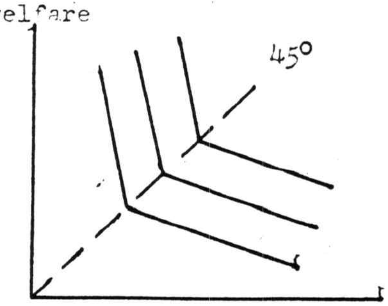


The numerical representation W can then be chosen to have the form $W(a_1, a_2) = c_1 a_1 + c_2 a_2$, where c_1 and c_2 are constants. In view of the Pareto principle, both must be non-negative, and at least one positive.

If anonymity is imposed, we further have $c_1 = c_2$, and then both can be set equal to one without loss of generality. This is the case of simple or pure utilitarianism.

CARDINAL FULL COMPARABILITY

Suppose $W(a,a) = W(a+k_1, a+k_2)$, and shift each individual's welfare by the transform $t_o(z) = 2(z-a) + a$. Applying the theorem, $W(a+k_1, a+k_2) = W(a+2k_1, a+2k_2)$, i.e. the welfare contours are straight lines. Next, using $t_o(z) = z + b - a$, we see that $W(b,b) = W(b+k_1, b+k_2)$, so the straight lines are parallel. However, the new feature is that the argument can be applied independently to the two sides of the 45° line, so the families of parallel lines can have different slopes on the two sides. The figure illustrates this.



With more individuals, similar arguments show that social welfare contours must be a family of parallel cones, with all their vertices on the line of equal welfares.

A numerical representation has been proved by Roberts (1980a). Let $a = (a_1, a_2, \dots a_n)$ be a typical vector of welfare levels. Write the average welfare as $\bar{a} = (a_1 + a_2 + \dots a_n)/n$, and let a be the vector of welfare deviations from the mean, i.e. its components $a_i = a_i - \bar{a}$. Then W can be chosen to have the form

$$W(a) = \bar{a} + g(a)$$

where g is a function homogeneous of degree one.

This can accomodate several commonly used forms. For example, g could depend on any central moments of the distribution of welfare levels, such as the standard deviation or skewness. Some care has to be taken to ensure homogeneity. Measures based on the Gini coefficient can also be cast in this form. Roberts suggests a class of measures

$$W(a) = \bar{a} + \gamma \min_i (a_i - \bar{a})$$

This is utilitarian for $\gamma = 0$, Rawlsian for $\gamma = 1$, and a weighted average of the two for intermediate values of γ.

CARDINAL RATIO SCALES

Here $W(a_1, a_2) = W(b_1, b_2)$ if and only if $W(ka_1, ka_2) = W(kb_1, kb_2)$ for any positive k. Thus the contours of social welfare must be scaled replicas of one another, i.e. homothetic. The numerical representation is also homothetic, and since only its ordinal form matters, it can be chosen to be a function homogeneous of degree one. A common example, yielding a family of social welfare functions that is quite rich for purposes of several applications, is the class of functions having a constant elasticity of substitution between individual welfare levels, i.e.

$$W(a_1, a_2, \dots a_n) = (a_1^{\delta} + a_2^{\delta} + \dots a_n^{\delta})^{1/\delta}$$

This is utilitarian for $\delta = 1$, and Rawlsian in the limit as δ goes to $-\infty$. Values of δ below 1 correspond to some egalitarian concern. Those above 1 are actually inegalitarian, the anti-Rawlsian limit occuring at infinity.

If our welfare information is limited to income or wealth, this may be thought to be the appropriate basis for social welfare judgements. The measure of inequality of income or wealth proposed by Atkinson (1970) fits naturally into this framework. It is based on a C.E.S. welfare function.

Any attempts to generalise Atkinson's index within the framework of ratio scale comparability must, however, remember the restriction of homotheticity.

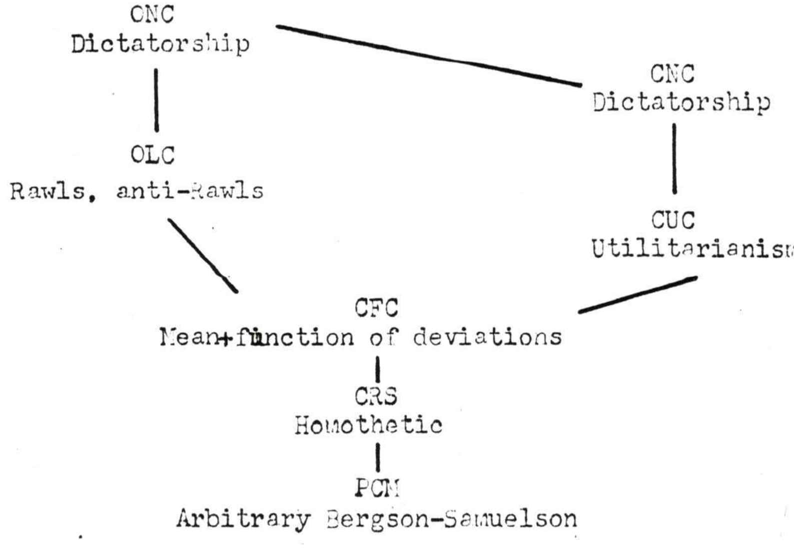
PERFECT MEASURABILITY AND COMPARABILITY

With no non-trivial transformations allowed, no restrictions can be placed, and any function $W(a_1, a_2)$ that is non-decreasing in the two arguments, and increasing in at least one of them at each point, is a welfare function compatible with this class.

This, therefore, appears to be the appropriate framework of information and comparability for the most general theory of welfare economics following the approach of Bergson and Samuelson. This view finds some support in the early expositions of Samuelson (1947, Ch.8). He asserts that social welfare is primarily defined over the allocations of goods and services, and that if the numerical representation of some individual's preferences by a utility function is altered by any transformation, the form of the social welfare function can be altered in a compensating way to preserve the basic judgements. To do this, of course, he needs to be able to distinguish between any two profiles, i.e. operate in a world of perfect measurability and comparability. This view of the Bergson-Samuelson approach is, however, at the opposite extreme from the one which has since become commonly accepted. For example, Sen (1977, section 9) declares the approach to be "like Arrow's ... rooted in ordinal non-comparability".

4. SUMMARY AND CONCLUSIONS

As the informational basis of measurability and comparability of individual welfares is strengthened, more types of social welfare functions become compatible. We have seen the earliest stages at which some common types become available; each remains open for further stages. The following diagram sums up the results:



This is useful in ensuring that, in theoretical as well as applied work, the type of welfare function being used is compatible with the informational basis that is available.

It should be emphasized again that the elementary account above merely begins to tap the existing literature. Sen (1977) and Roberts (1980a,b) provide several other points of discussion, and references to other work.

REFERENCES

Atkinson, A.B. (1970) "On the measurement of inequality", Journal of Economic Theory, reprinted in A.B. Atkinson (ed) Wealth, Income and Inequality, Penguin.

d'Aspremont, C. and Gevers, L. (1977) "Equity and the informational basis of collective choice", Review of Economic Studies.

Hammond, P.J. (1976) "Equity, Arrow's conditions, and Rawls' difference principle", Econometrica.

McManus, M. (1975) "Inter-tastes consistency in social welfare functions", in M. Parkin and A. Nobay (eds) Current Economic Problems.

Roberts, K.W.S. (1980a) "Interpersonal comparability adn social choice theory", Review of Economic Studies.

--------- (1980b) "Possibility theorems with interpersonally comparable welfare levels", Review of Economic Studies.

Samuelson, P.A. (1947) Foundations of Economic Analysis.

Sen, A.K. (1970) Collective Choice and Social Welfare.

----- (1977) "On weights and measures: Informational constraints in social welfare analysis", Econometrica.

23